UC Davis
UC Davis Electronic Theses and Dissertations

Title

Metabolomic markers of exclusive breastfeeding in infants

Permalink

https://escholarship.org/uc/item/7kz2g9cq

Author

Wang, Aidong

Publication Date

2021

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

Metabolomic Markers of Exclusive Breastfeeding in Infants

By

AIDONG WANG
DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Food Science

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

Carolyn M. Slupsky, Chair

Bo Lönnerdal

Christopher W. Simmons

Committee in Charge

2021

Abstract

Breast milk is the gold standard for infant nutrition because in addition to nutrients, it also provides bioactive components that benefit the health of the developing infant. Infants under six months are recommended to be exclusively breastfed (EBF). Infant feeding practices determined by parental recall methods have been reported to have discrepancies when compared with results measured by the deuterium dose-to-mother (DTM) technique. However, the DTM method is limited by its high workload of sample collection and its validity outside the sample collection window. Finding markers of infant feeding practice using a readily accessible biological sample (such as urine or feces) is a promising novel approach to determine infant feeding practices.

Breast milk is dynamic and its composition is impacted by many factors among which maternal secretor status and Lewis blood type impacts the levels of several human milk oligosaccharides (HMOs). Previous research has demonstrated the influence of maternal HMO phenotype on the infant gut microbiome. However, these studies used recall methods to determine infant feeding practice and the sample size used were relatively small. Therefore, it is important to illustrate this impact in a larger infant population whose feeding practice were determined objectively.

This dissertation investigates the impact of maternal HMO phenotype on the milk metabolome, infant metabolism and gut microbiome, and seeks to develop metabolic biomarkers of feeding practice from samples of infant feces and urine. Chapter 1 reviews the compositional differences between human milk and infant formula, as well as the recent efforts to narrow the gap between the two. Chapter 2 demonstrates the impact of maternal secretor status and Lewis blood type on breast milk metabolome. Fold change analysis showed that the non-secretor (Se-) Lewis negative (Le-) milk had major differences in free fatty acids, free amino acids, and metabolites related to energy metabolism. Chapter 3 shows the impact of maternal secretor status on the infant fecal

metabolome and gut microbiome. Maternal secretor status did not alter the within-community (alpha) diversity, between-community (beta) diversity, or the relative abundance of bacterial taxa at the genus level. However, differences in infant gut microbial fermentation products were observed with succinate, amino acids and their derivatives, and 1,2-propanediol. Chapters 4 and 5 develop and validate metabolic biomarkers for the discrimination between EBF and nonexclusively breastfed (non-EBF) infants using fecal and urine samples, respectively. Metabolites with excellent prediction performance (area under the receiver operating curve (AUC) above 0.8) were identified as biomarkers, and these included HMOs, short chain fatty acids, amino acids (AAs) and organic acids in infant feces, and AA derivatives, nicotinamide-adenine dinucleotide degradation products, and metabolites from dietary sources and/or host-microbial co-metabolism in infant urine.

This work provides new insights into the impact of maternal secretor status and Lewis blood type on the breast milk metabolome, as well as a further knowledge on how maternal secretor status influences the EBF infant gut metabolome and microbiome. The results in this study pave the road to a deeper understanding of how the choice of feeding practices impacts infant metabolism and the gut microbiome, and demonstrates the power of utilizing metabolic biomarkers as a novel approach to determine infant feeding practice.

Dedicated to my loving family for providing me unconditional love and support

Acknowledgement

First and foremost, I would like to express my sincere gratitude to Dr. Carolyn Slupsky, my advisor, for providing me a tremendous amount of help, guidance, encouragement, and support along the way. You have always been a wonderful mentor who inspires me in both my research and life. Your insightful and patient guidance has helped me overcome a lot of challenges and build confidence especially when I was filled with self-doubt.

I am deeply indebted to Dr. Bo Lonnerdal and Dr. Christopher Simmons – members of my Dissertation Committee and Qualifying Exam Committee. Thank you for generously providing me help, feedback, and advice along my doctoral research and dissertation writing.

I would also like to thank other members of my Qualifying Exam Committee: Dr. Gary Smith, Dr. Dr. Bruce German, and Dr. David Mills. Your generous advice improved my research and knowledge.

Thank you to Dr. Gary Smith and Dr. Alison Mitchell for helping me stay on course of my PhD study as my Academic Advisor.

I would like to extend my sincere thanks to Dr. Lisa Houghton – collaborator of my major research project. Your insightful advice and timely support on providing the needed information on the project helped move my research forward.

I would also like to thank all the Slupsky Lab members. It has been a wonderful experience to get to know and work together with everyone in the lab. I really appreciate the inspiring discussions and helpful advice from everyone in the lab.

Contents

Abstract ii

Acknowledgement v

Chapter 1 Breast milk and infant formula: a comparative view into their composition and health impact 1

Overview 1

Major compositional difference between breast milk and infant formula 2

Protein and amino acids 3

Lipids and fatty acids 4

Human milk oligosaccharides 7

Efforts to narrow the gap between breast milk and infant formula 7

HMOs and prebiotics 8

MFGM 9

Probiotics 10

Conclusion 12

Reference 14

Chapter 2 Difference in milk metabolome between secretor status and Lewis blood type . 21

Introduction 22

Materials and methods 24

Milk sample preparation 24

NMR data acquisition and processing 25

HMO phenotype determination 26

Statistical analysis 26

Results 27

Discussion 37

References 41

Chapter 3 Impact of milk secretor status on fecal metabolome and microbiota of breastfed infants 44

Introduction 45

Materials and methods 47

Breastfeeding status determination 47

Breast milk sample collection and metabolite extraction 47

Fecal sample collection and metabolite extraction 48

NMR acquisition, data processing and quantification 49

Milk secretor status phenotype determination 49

Fecal microbial DNA extraction and library preparation 50

Analysis of 16S rRNA gene amplicon sequences 51

Statistical analysis 52

Results 53

Discussion 60

Conclusion 66

References 68

Chapter 4 Infant fecal metabolites as potential biomarkers for exclusive breastfeeding 72

Introduction 73

Materials and methods 75

Breastfeeding status and maternal secretor status determination 76

Fecal metabolite measurement 76

Fecal microbial DNA extraction, library preparation and 16S rRNA gene sequencing 78

Statistical analysis 79

Results 80

Discussion 91

Conclusion 96

References 97

Chapter 5 Infant urine metabolic profile as a source of potential biomarkers for exclusive breastfeeding 101

Introduction 102

Materials and methods 103

Feeding practice, breast milk intake, and maternal secretor status determination 104

Urine metabolite measurement 105

Statistical analysis 105

Results 106

Discussion 112

Conclusion 116

References 117

Appendix: Multi-micronutrient and choline supplementation during pregnancy impacted children urine metabolome at age of 2 to 4 years 120

Introduction 121

Methods 122

Urine sample collection and metabolite extraction 122

Urine metabolite measurements 122

Statistical analysis 123

Results 124

Discussion 127

Conclusion 128

References 130

Chapter 1 Breast milk and infant formula: a comparative view into their composition and health impact

Overview

Exclusive breastfeeding (EBF) provides the optimal nutrition for newborn infants. Human milk provides infants a variety of short-term benefits including protection against infectious disease, mortality, otitis media, leukemia, and allergies (Young 2017), as well as long-term benefits such as lowering the risk of obesity and diabetes (Binns, Lee, and Low 2016). Suboptimal breastfeeding, defined as feeding any food or beverage other than human milk, is responsible for 45% of neonatal infectious deaths, 30% of diarrheal deaths, and 18% of acute respiratory deaths among under-five children in developing countries (World Health Organization 2009). The World Health Organization (WHO) recommends infants under six months to be exclusively breastfed (EBF) (WHO, 2019). EBF is especially important in resource-poor areas because breast milk is a reliable hygienic food source containing both nutrients and water (Kakute et al. 2005). Despite this, the global prevalence of EBF remains low. Surveys and recent studies report that the EBF rate at 6 months ranges from 24% to 32% in developing countries, and is as low as ~13% in developed countries (WHO 2009; Jones et al. 2011).

For nonexclusively breastfed (non-EBF) and non-breastfed infants, infant formula is the most common food source, which can be introduced to infants as early as postpartum day one (Rosenberg et al. 2008). Other than infant formula, common complementary foods consumed by non-EBF and non-breastfed infants under six months of age include solids such infant cereal, mashed vegetables or fruits, and porridge (which may include egg and meat), as well as liquids

including water, tea, and juice (Kronborg, Foverskov, and Væth 2015; Duong, Binns, and Lee 2005; Kuchenbecker et al. 2015; Samady et al. 2020; Hornsby et al. 2021).

There are many factors impacting maternal choice and capability to exclusively breastfeed her infant. Maternal education level and age have been reported to be positively associated with EBF (Jones et al. 2011). Children from two-parent families are more likely to be ever EBF than children from other family structures (Jones et al. 2011). Barriers to successful EBF among employed mothers include policy, support and environment at her workspace (El-Houfey et al. 2017). Extending paid maternal leave has been reported to significantly increase the EBF rate in low- and middle-income countries (Chai, Nandi, and Heymann 2018). Recommendation from health professionals is another factor impacting parental action of feeding practice. A recent survey conducted among 563 pediatric practitioners showed that 31.8% and 42.5% of the practitioners recommended mothers to introduce complementary food at 4 months of age for EBF and non-EBF infants, respectively (Samady et al. 2020).

Although EBF is the optimal feeding practice, there are occasions that EBF is not an option. Therefore it is important to understand the differences between breast milk and infant formula and how infants are affected by parental choice of feeding practice. This chapter summarizes the current understanding of the compositional difference between breast milk and infant formula, the potential impact of this difference on infant health, as well as efforts to improve infant formula to better resemble human milk.

Major compositional difference between breast milk and infant formula

Most common infant formulas are derived from bovine milk. For infants with allergies to bovine milk, infant formula with alternative protein sources such as hydrolyzed bovine milk protein and

soy protein are available (Martin, Ling, and Blackburn 2016). The constituents of infant formula are designed to resemble components of human milk, and are highly regulated by national and international standards to support the healthy development of infants (Food and Drug Administration (FDA) 2020; Koletzko et al. 2005). However, there are differences in multiple components between human milk and infant formula.

Protein and amino acids

One of the major differences between human milk and infant formula is protein content. Human milk is reported to have between 1.35-1.8 g of protein / 150 mL, while the bovine milk ranges between 2.7 g to 3.0 g of protein / 150 mL (Thompkinson and Kharb 2007). The nutrient requirements published by the United States Food and Drug Administration (FDA) for infant formula state that there should be between 1.8-4.5 g of protein / 150 mL, or 1.8-4.5 g/100 kcal. (The 150 mL volume was adapted from the instructions for diluting infant formula to better directly compare with human milk; 150 mL properly dissolved infant formula provides 100 kcal of energy (Timby et al. 2014; Koletzko et al. 2005). For infant formula based on soy protein and hydrolyzed bovine protein, the minimum protein content is 2.25 g of protein / 150 mL (Koletzko et al. 2005). Human milk contains all the essential amino acids (AAs) for optimal growth of human infants. Compared to bovine milk, human milk has a lower protein content (Dewey et al. 1996). The major consideration regarding the protein level in infant formula is the provision of adequate essential amino acids (AAs) to similar levels found in human milk. Per gram of protein, the levels of certain AAs including cysteine, lysine, threonine and tryptophan are lower in bovine milk compared to human milk (Thompkinson and Kharb 2007). Soy milk has also been shown to have significantly lower levels of a broad range of free AAs (histidine, isoleucine, lysine, phenylalanine, threonine

and valine) but higher levels of methionine and tyrosine compared to human milk (Agostoni et al. 2000).

Another major difference between human milk and bovine milk is the ratio of casein to whey proteins. In human milk, the ratio of casein to whey is ~40:60, while in bovine milk it is 80:20. For infant formula, the FDA requires the minimum amount of protein (1.8 g/150 mL) to be adjusted by the casein content (if the casein content is 75% of the total protein, the minimum protein required will be 1.8/0.75 g/150 mL) (FDA 2020). The casein family of proteins is composed of α-s1-casein, α-s2-casein, β-casein, and κ-casein. The predominant casein subclass is β-casein in human milk and α-s1-casein in bovine milk (Park 2017). The whey composition of human milk is also different from bovine milk. Human milk has been reported to have higher concentrations of α-lactalbumin, lactoferrin, immunoglobulins A, G and M compared to bovine milk (Goldsmith et al. 1983; Butler 1994). Additionally, β-lactoglobulin has been shown to be the major whey protein in bovine milk, whereas it is only detected in trace amounts in human milk (Goldsmith et al. 1983; Butler 1994). Interestingly, bovine α-s1-casein, β-casein and κ-casein, as well as β-lactoglobulin have been reported to be responsible for bovine milk allergies (Hochwallner et al. 2014).

Lipids and fatty acids

Triglycerides compose 98% of the milk lipids and contribute approximately 50% of the total energy intake of infants (Manson and Weaver 1997). They are present as globules covered by the milk fat globule membrane (MFGM) in both human and bovine milk at comparable amounts: ~ 5.6 g/150 mL (Lindmark Månsson 2008; Thompkinson and Kharb 2007). MFGM, which is not present in formulas utilizing fats from plant sources, is composed of various proteins and lipids with beneficial functions (Lee et al. 2018). For infant formula, the FDA requires the total fat content to be between 3.3-6.0 g / 150 mL (FDA 2020), and the European Society for Paediatric

Gastroenterology Hepatology and Nutrition (ESPGHAN) recommends 4.4-6 g / 150 mL (Koletzko et al. 2005). The FDA standard only specifies a minimum requirement of linoleic acid (300 mg/150 mL) with no requirements of other fatty acids (FDA 2020). In contrast, the ESPGHAN recommendation states that the ratio of linoleic acid to α-linoleic acid should be between 5:1 to 15:1, the maximum portion of lauric plus myristic acids should contribute to 20% of total fat, trans fatty acids should not exceed 3% of total fat, and erucic acid should be 1% of total fat (Koletzko et al. 2005).

Fats from dairy sources were widely used in infant formula by utilizing whole milk as ingredients in the early 20th century; however, its use diminished as the infant formula industry developed (Delplanque et al. 2015; Innis 2011). Currently, in most formulas, dairy fats are replaced by vegetable fats (Medicine 2004) to achieve a higher amount of unsaturated fatty acids, to avoid contamination and unpleasant odors caused by milk processing, and to decrease cost (Innis 2011; Fomon 2001). Multiple plant oils including coconut oil, corn oil, soybean oil, palm oil, sunflower oil, safflower oil, and rapeseed oil have been used to generate the lipid fraction of infant formula, and it is common to use a mixture of several oil types (Berger, Fleith, and Crozier 2000; Mendonça et al. 2017; Zou, Pande, and Akoh 2016). Although bovine milk fat is less common, it can be found in some formulad (Sun et al. 2016). Multiple studies have shown a positive effect on cognition when including bovine-derived MFGM as part of infant formula (Gurnida et al. 2012; Timby et al. 2014), as well as a positive impact on gut microbial function (He et al. 2019).

Human milk possesses over 200 different fatty acids, with the dominant fatty acids being oleic acid, palmitic acid and linoleic acid (Jensen et al. 1990; Lindmark Månsson 2008). Approximately 98% of the fatty acids are long-chain (C > 10) (LCFA) with ~40% of them saturated. The remaining 2% of the fatty acids are medium-chain (6 < C <10) (MCFA). In bovine milk, LCFAs

comprise ~90% of the total fatty acid content, while the proportion of MCFAs is 6–7%, and the short chain fatty acid (SCFA) butyrate 3–4% and there are over 400 fatty acids present (Jensen et al. 1990; Lindmark Månsson 2008), with saturated fatty acids accounting for about 67% of the total. Oleic acid and palmitic acid are the major fatty acids in bovine milk. Butyrate is widely detected in bovine milk while in human milk and vegetable oils, it is present in trace amounts. Vegetable oils have a less diverse profile of fatty acids, and they tend to miss those between C4:0 to C12:0 (Dorni et al. 2018). To mimic the fatty acid composition of human milk, multiple vegetable oils are usually blended to generate infant formula (Zou, Pande, and Akoh 2016).

Besides fatty acid composition, there are also differences in the triglyceride structure when comparing human milk, bovine milk, and vegetable oils. In human milk, palmitic acid is present approximately 70% of the time at the sn-2 position of the triglyceride, with other long-chain saturated fatty acids (LCSFAs), including lauric and arachidic acids, esterified at this position ~30% of the time (López-López et al. 2002). Stearic acid, another LCSFA, is present only 10% of the time at sn-2 position (López-López et al. 2002). Long chain unsaturated fatty acids (LCUFAs) including oleic, linoleic acid are predominantly esterified at the sn-1 and sn-3 positions (~12% and ~22% esterified at sn-2 position, respectively) (López-López et al. 2002). In bovine milk, 40–45% of the total amount of palmitic acid is present at the sn-2 position of the triglyceride, with oleic acid attached to the sn-1 or sn-3 position (Bourlieu et al. 2015). In infant formula, the amount of saturated fatty acids located at the sn-1 and/or sn-3 is higher than in human milk (Tu et al. 2017). It has been suggested that the position of fatty acids within the triglyceride may be associated with stool frequency and consistency. Specifically, LCSFAs esterified at the sn-1 and sn-3 positions have been reported to form calcium fatty acids soaps which could contribute to constipation while

LCSFAs located at the sn-2 position can help form softer stools (Mehrotra, Sehgal, and Bangale 2019).

Human milk oligosaccharides

Human milk oligosaccharides (HMOs), the third most abundant component of human milk (after lactose and lipids), are generally absent in infant formula. Beneficial effects associated with HMOs include serving as prebiotics to *Bifidobacterium*, decoys for pathogens, modulators of intestinal epithelial cell responses, and immune modulators (Bode 2012). HMOs are comprised of over 200 different structures built with five building blocks: Galactose (Gal), Glucose (Glc), N-acetylglucosamine (GlcNAc), fucose (Fuc), and sialic acid (Neu5Ac). Fucosylated, sialylated, and nonfucosylated neutral HMOs account for 35–50%, 12–14%, and 42–55% of the total HMOs in term human milk, respectively (Totten et al. 2012; Ninonuevo et al. 2006). Oligosaccharide analysis of bovine milk revealed that ~40 different structures are present. When compared to HMOs, bovine milk oligosaccharides (BMOs) contain shorter oligomeric chains (Tao et al. 2008). In bovine milk, sialylated oligosaccharides composes nearly 70% of total oligosaccharides. Moreover, while HMOs are built on a lactose core, BMOs are built on either a lactose core or lactose amines (Tao et al. 2008). To better mimic the prebiotic function of HMOs, several plant-based and synthesized structures including galactooligosaccharides (GOS), polydextrose (PDX), lactulose (LOS), inulin, and fructooligosaccharides (FOS) have been added to infant formula (Vandenplas, Zakharova, and Dmitrieva 2015; Roberfroid 2007).

Efforts to narrow the gap between breast milk and infant formula

Health benefits such as lower morbidity, less early antibiotic exposure, and a gut microbiome with a higher abundance of beneficial microbes in infants consuming breast milk over those who were

fed infant formula have been widely reported (Parizkova et al. 2020; Young 2017; Binns, Lee, and Low 2016; He et al. 2019; Roger et al. 2010; Zhao et al. 2021). Such differences have driven the evolution of infant formula in terms of adjusting the content of existing ingredients and including novel components with proven health benefits.

A major difference between formula-fed and breastfed infants is their gut microbiome. Infants are exposed to human microbes potentially already prior to birth (Del Chierico et al. 2015; Moles et al. 2013). The infant gut microbiome gradually matures towards an adult-like configuration within three years after birth (Yatsunenko et al. 2012). When compared to breastfed infants, formula-fed infants have been reported to have a significantly altered gut microbiome (Bäckhed et al. 2015; Bezirtzoglou, Tsiotsias, and Welling 2011; He et al. 2019). Overall, EBF infants were reported to have higher taxa from the protective bacterial class *Actinobacteria,* and a lower alpha diversity while formula-fed infants had higher levels of the proinflammatory bacterial class *γ-Proteobacteria* and a more diverse gut microbiota (Bäckhed et al. 2015; Bezirtzoglou, Tsiotsias, and Welling 2011; He et al. 2019). *Bifidobacterium* species (*B. bifidum*, *B. breve*, *B. longum* subsp. *longum*, and *B. longum* subsp. *infantis*), the most dominant microbes in the infant gut, were shown to be at a lower relative abundance in the gut of formula-fed infants in multiple studies (He et al. 2019; Roger et al. 2010; Zhao et al. 2021; Makino et al. 2015). *Bifidobacterium* species have been reported to exert several beneficial effects in infants such as preventing necrotizing enterocolitis (NEC) in preterm infants (Zhu et al. 2019), as well as reducing acute gastroenteritis (Taipale et al. 2016) and respiratory tract infections (Hojsak et al. 2015).

HMOs and prebiotics

HMOs are prebiotics that selectively benefit the growth of some *Bifidobacterium* species among which *Bifidobacterium subspecies infantis* (*B. infantis*) is capable of consuming a complete profile

of HMOs (Underwood et al. 2014; Zivkovic et al. 2011). To mimic the prebiotic function of HMOs, multiple plant-based and synthesized oligosaccharide structures have been added to infant formula. The main target has been *Bifidobacterium* (Marín-Manzano et al. 2020; Scalabrin et al. 2012; Lu, Yeung, and Yeung 2018; Oswari et al. 2019). Non-fructosylated α-galactooligosaccharides (α-GOS) from pea was shown to be fully metabolized by microbes from infant fecal slurries, and was shown to promote the growth of *Bifidobacterium longum* subsp. *longum* and *Bifidobacterium catenulatum/pseudo-catenulatum* (Marín-Manzano et al. 2020). Term infants consuming infant formula containing polydextrose (PDX) and GOS had a higher total *Bifidobacteria*, *B. longum* and *B. infantis* than infants fed formula without PDX (Scalabrin et al. 2012). Under anaerobic conditions, both lactulose and fructooligosaccharides (FOS) were shown to support the growth of several *Bifidobacterium* species (*B. animalis*, *B. bifidum*, *B. infantis*, and *B. lactis*) (Lu, Yeung, and Yeung 2018). Inulin supplemented to infant formula at a concentration of 0.4 g/100 mL significantly increased the relative abundance of *Bifidobacteria* and *Lactobacilli* in infants aged 3 to 5 months (Oswari et al. 2019). Recently, attributed to the advances in large-scale production of HMO structures, it is possible to supplement 2'fucosyllactose (2'FL) and lacto-N-neotetraose (LNnT) in infant formula (Bych et al. 2019; Vandenplas et al. 2018). Infants fed formula containing 2'FL and LNnT were reported to have softer stools, fewer episodes of night-time wake-ups, fewer parental reports of bronchitis, and lower incidences of lower respiratory tract infections, use of antipyretics, and use of antibiotics when compared to infant fed formula without the two HMOs (Puccio et al. 2017).

MFGM

MFGM, containing multiple bioactive components, is historically removed during infant formula production from bovine milk. A recent study conducted *in vitro* growth experiments of *B. infantis*

DSMZ20090 with key components from MFGM and reported that lactadherin, sialic acid and phospholipid significantly promoted the growth of *B. infantis* DSMZ20090 (Zhao et al. 2021). In this study, it was shown that a series of genes coding enzymes/proteins (2-hydroxyacid dehydrogenase, 4-hydroxy-tetrahydrodipicolinate synthase, and ABC transporter ATP-binding protein) involved in energy and carbohydrate metabolism were enriched, which may be responsible for stimulating the growth of *B. infantis* DSMZ20090 (Zhao et al. 2021). Infant formula supplemented with bovine MFGM has been reported to protect the infant against pathogens such as *Clostridium difficile* or *Listeria monocytogenes* (Timby et al. 2015). In a rat model, provision of a formula supplemented with bovine MFGM restored gut microbiota, intestinal growth, Paneth and goblet cell numbers, and tight junction protein patterns to that of rat pups who consumed mother's milk, while rats fed formula without MFGM showed significant deficits in intestinal development (Bhinder et al. 2017). In the same study, it was observed that MFGM afforded protection against *Clostridium difficile* toxin induced inflammation (Bhinder et al. 2017).

Probiotics

Breast milk has long been recognized to contain bacteria even when collected under aseptic conditions (Dudgeon and Jewesbury 1924). Traditionally the breast milk bacteria had been believed to come from maternal skin and/or the infant oral cavity (West, Hewitt, and MURPHY 1979). Recently it has been suggested that breast milk microorganisms could be originated from both internal and external sources (Fernández et al. 2013). The internal origin of breast milk microorganisms can be supported by the discovered microbes in colostrum collected before the first infant feeding (Damaceno et al. 2017), the vertical transmission of pathogenic bacteria from

mother to the infant (Lawrence and Lawrence 2004), and the isolated probiotic bacteria in the breast milk of the lactating mothers who orally consumed probiotics (Abrahamsson et al. 2009).

Is has been shown that milk intake of 800 mL/day can provide between 1×10^5 and 1×10^7 bacteria daily (Heikkilä and Saris 2003). A broad profile of bacteria has been isolated from breast milk and the dominant microbes includes *Staphylococci*, *Streptococci*, *Lactic acid bacteria* (LAB) and *Propionibacteria* (Martín et al. 2003; Heikkilä and Saris 2003; Gavin and Ostovar 1977). *Staphylococcus, Corynebacteria,* and *Propionibacteria* are common breast milk facultative anaerobic microbes originated from adult skin (Grice et al. 2009). Interestingly, breast milk has also been shown to possess *Bifidobacteria* which is strict anaerobic, and several species including *B. longum, B. animalis, B. bifidum* and *B. catenulatum* have been detected (Gueimonde et al. 2007). But it remains controversial that if breast milk is a source of *Bifidobacterium* in the infant gut (Martín et al. 2009).

Thus, including probiotics in infant formula is another way to potentially lessen the gap between human milk and infant formula. Multiple studies have focused on the benefits of supplementing the probiotic *Bifidobacterium animalis ssp lactis CNCMI-3446* in infant formula (Bakker-Zierikzee et al. 2005; Velaphi et al. 2008; Weizman, Asli, and Alsheikh 2005; Weizman and Alsheikh 2006; 2006; Storm et al. 2019). Other studies focused on the impact of adding *Streptococcus thermophilus* (Béghin et al. 2021) and *Lactobacillus helveticus* (Chouraqui, Van Egroo, and Fichot 2004) to infant formula. Administration of these probiotics was shown to be well tolerated and the major benefits observed were a reduction in the risk of nonspecific gastrointestinal infections, reduced risk of antibiotic use, and lower frequency of colic and/or irritability (Braegger et al. 2011). Probiotic use in preterm infants showed benefits in reducing NEC and late-onset sepsis, and improving feed related outcomes (time to full feeds) (Bermudez-

Brito et al. 2012; Hidalgo-Cantabrana et al. 2017; Hill et al. 2014; Plaza-Diaz et al. 2019). There are several probiotic microbes (*L. rhamnosus* GGATCC53103 or the combination of *B. infantis* Bb-02, *B. lactis* Bb-12, and *Str. thermophilus* TH-4) recommended by the ESPGHAN Committee on Nutrition to reduce NEC rates (van den Akker et al. 2020).

Conclusion

Although breastfeeding is the optimal food for infants, there are various occasions where breastfeeding is not an option. Infant formula serves as a substitute for human milk and ensures normal growth and development of infants during this critical window of time. There are major differences in the composition between human milk and infant formula including protein composition, lipid structures and composition, HMOs and other bioactive constituents in human milk. Advances in infant formula development are driven by the growth of understanding of human milk. As the beneficial effects of the bioactive components in human milk are proven, multiple missing components such as HMOs (or other probiotic structures), MFGM and probiotics have been included into infant formula. However, significant challenges to fully understand the components of human milk remain and the mechanism of how feeding practices impact health development of infants requires further research.

Human milk components are influenced by many factors among which maternal secretor status and Lewis blood type determine the HMO profile (Kunz et al. 2000). However, it remains unknown if maternal secretor status and Lewis blood type will impact the human milk metabolome. The second chapter of this dissertation looks into differences in the milk metabolome among different maternal secretor status and Lewis blood phenotypes with an emphasis on the milk metabolome of phenotypically non-secretor Lewis negative women. Maternal secretor status

has been reported to impact the infant gut microbiome, but the results have been inconsistent from studies conducted at different locations and with infants fed using different methods (EBF, partial breastfed or formula-fed), especially regarding the impact of maternal secretor status on infant gut *Bifidobacteria* (Lewis et al. 2015; Korpela et al. 2018; Smith-Brown et al. 2016; Underwood et al. 2015; Wang et al. 2015). Therefore, the third chapter characterizes the influence of maternal secretor status on the infant gut microbiome and metabolome in EBF infants.

A further understanding of the mechanisms of how human milk components impact the health development of infants relies on an accurate measurement of infant feeding practice. However, routinely used recall methods such as recall since birth, recall over the past 24 h, and report of current feeding practice tend to overestimate the EBF rate when compared to the deuterium oxide dose-to-mother (DTM) method (Mazariegos, Slater, and Ramirez-Zea 2016; Noel-Weiss, Boersma, and Kujawa-Myles 2012). The DTM method provides accurate measurement of infant feeding practice but its utility is limited due to the high workload of sample collection and its validity outside the sample collection window. Therefore the development of a novel approach with an accurate assessment of infant feeding practice using easily collectable samples is needed. The fourth and fifth chapters characterize the metabolic and gut microbial differences between EBF and non-EBF infants utilizing their fecal and urine samples, respectively. These two chapters also identify metabolic biomarkers with a predictive capability of infant feeding practices using random forest model coupled with a backwards selection algorithm.

Reference

WHO. Exclusive Breastfeeding for Optimal Growth, Development and Health of Infants. Geneva: World Health Organization; 2019.

U.S. Food and Drug Administration. 2020. 21CFR107.100, Infant Formula Nutrient Requirements Nutrition Specifications.

Abrahamsson, Thomas R, Gabriela Sinkiewicz, Ted Jakobsson, Mats Fredrikson, and Bengt Björkstén. 2009. "Probiotic Lactobacilli in Breast Milk and Infant Stool in Relation to Oral Intake during the First Year of Life." *Journal of Pediatric Gastroenterology and Nutrition* 49 (3). doi: 10.1097/MPG.0b013e31818f091b.

Agostoni, Carlo, Brunella Carratù, Concetta Boniglia, Enrica Riva, and Elisabetta Sanzini. 2000. "Free Amino Acid Content in Standard Infant Formulas: Comparison with Human Milk." *Journal of the American College of Nutrition* 19 (4): 434–38. doi: 10.1080/07315724.2000.10718943.

van den Akker, Chris HP, Johannes B. van Goudoever, Raanan Shamir, Magnus Domellöf, Nicholas D. Embleton, Iva Hojsak, Alexandre Lapillonne, Walter A. Mihatsch, Roberto Berni Canani, and Jiri Bronsky. 2020. "Probiotics and Preterm Infants: A Position Paper by the ESPGHAN Committee on Nutrition and the ESPGHAN Working Group for Probiotics and Prebiotics." *Journal of Pediatric Gastroenterology and Nutrition*. doi: 10.1097/MPG.0000000000002655.

Bäckhed, Fredrik, Josefine Roswall, Yangqing Peng, Qiang Feng, Huijue Jia, Petia Kovatcheva-Datchary, Yin Li, et al. 2015. "Dynamics and Stabilization of the Human Gut Microbiome during the First Year of Life." *Cell Host & Microbe* 17 (5): 690–703. doi: 10.1016/j.chom.2015.04.004.

Bakker-Zierikzee, Astrid M., Martine S. Alles, Jan Knol, Frans J. Kok, Jules J. M. Tolboom, and Jacques G. Bindels. 2005. "Effects of Infant Formula Containing a Mixture of Galacto- and Fructo-Oligosaccharides or Viable Bifidobacterium Animalis on the Intestinal Microflora during the First 4 Months of Life." *The British Journal of Nutrition* 94 (5): 783–90. doi: 10.1079/bjn20051451.

Béghin, Laurent, Sebastian Tims, Mieke Roelofs, Carole Rougé, Raish Oozeer, Thameur Rakza, Gaetano Chirico, et al. 2021. "Fermented Infant Formula (with Bifidobacterium Breve C50 and Streptococcus Thermophilus O65) with Prebiotic Oligosaccharides Is Safe and Modulates the Gut Microbiota towards a Microbiota Closer to That of Breastfed Infants." *Clinical Nutrition* 40 (3): 778–87. doi: 10.1016/j.clnu.2020.07.024.

Berger, Alvin, Mathilde Fleith, and Gayle Crozier. 2000. "Nutritional Implications of Replacing Bovine Milk Fat with Vegetable Oil in Infant Formulas." *Journal of Pediatric Gastroenterology and Nutrition* 30 (2): 115–30. doi: 10.1097/00005176-200002000-00006.

Bermudez-Brito, Miriam, Julio Plaza-Díaz, Sergio Muñoz-Quezada, Carolina Gómez-Llorente, and Angel Gil. 2012. "Probiotic Mechanisms of Action." *Annals of Nutrition and Metabolism* 61 (2): 160–74. doi: 10.1159/000342079

Bezirtzoglou, Eugenia, Arsenis Tsiotsias, and Gjalt W. Welling. 2011. "Microbiota Profile in Feces of Breast- and Formula-Fed Newborns by Using Fluorescence in Situ Hybridization (FISH)." *Anaerobe* 17 (6): 478–82. doi: 10.1016/j.anaerobe.2011.03.009.

Bhinder, Ganive, Joannie M Allaire, Cyrielle Garcia, Jennifer T Lau, Justin M Chan, Natasha R Ryz, Else S Bosman, et al. 2017. "Milk Fat Globule Membrane Supplementation in Formula Modulates the Neonatal Gut Microbiome and Normalizes Intestinal Development." *Scientific Reports* 7 (March): 45274–45274. doi: 10.1038/srep45274.

Binns, Colin, MiKyung Lee, and Wah Yun Low. 2016. "The Long-Term Public Health Benefits of Breastfeeding." *Asia-Pacific Journal of Public Health* 28 (1): 7–14. doi: 10.1177/1010539515624964.

Bode, Lars. 2012. "Human Milk Oligosaccharides: Every Baby Needs a Sugar Mama." *Glycobiology* 22 (9): 1147–62. doi: 10.1093/glycob/cws074.

Bourlieu, Claire, Karima Bouzerzour, Stéphanie Ferret-Bernard, Cindy Le Bourgot, Sophie Chever, Olivia Ménard, Amélie Deglaire, et al. 2015. "Infant Formula Interface and Fat Source Impact on Neonatal

Digestion and Gut Microbiota." *European Journal of Lipid Science and Technology* 117 (10): 1500–1512. doi: 10.1002/ejlt.201500025.

Braegger, Christian, Anna Chmielewska, Tamas Decsi, Sanja Kolacek, Walter Mihatsch, Luis Moreno, Malgorzata Piescik, et al. 2011. "Supplementation of Infant Formula with Probiotics and/or Prebiotics: A Systematic Review and Comment by the ESPGHAN Committee on Nutrition." *Journal of Pediatric Gastroenterology and Nutrition* 52 (2). doi: 10.1097/MPG.0b013e3181fb9e80.

Butler, J. E. 1994. "Passive Immunity and Immunoglobulin Diversity." In *Indigenous Antimicrobial Agents of Milk. Recent Developments, Uppsala (Sweden), 31 Aug-1 Sep 1993*. FIL-IDF. Secretariat general.

Bych, Katrine, Marta Hanna Mikš, Ted Johanson, Markus Jondelius Hederos, Louise Kristine Vigsnæs, and Peter Becker. 2019. "Production of HMOs Using Microbial Hosts — from Cell Engineering to Large Scale Production." *Food Biotechnology • Plant Biotechnology* 56 (April): 130–37. doi: 10.1016/j.copbio.2018.11.003.

Chai, Yan, Arijit Nandi, and Jody Heymann. 2018. "Does Extending the Duration of Legislated Paid Maternity Leave Improve Breastfeeding Practices? Evidence from 38 Low-Income and Middle-Income Countries." *BMJ Global Health* 3 (5): e001032–e001032. doi: 10.1136/bmjgh-2018-001032.

Chouraqui, Jean-Pierre, Louis-Dominique Van Egroo, and Marie-Claire Fichot. 2004. "Acidified Milk Formula Supplemented With Bifidobacterium Lactis: Impact on Infant Diarrhea in Residential Care Settings." *Journal of Pediatric Gastroenterology and Nutrition* 38 (3). doi: 10.1097/00005176-200403000-00011.

Damaceno, Quésia S., Jaqueline P. Souza, Jacques R. Nicoli, Raquel L. Paula, Gabriela B. Assis, Henrique C. Figueiredo, Vasco Azevedo, and Flaviano S. Martins. 2017. "Evaluation of Potential Probiotics Isolated from Human Milk and Colostrum." *Probiotics and Antimicrobial Proteins* 9 (4): 371–79. doi: 10.1007/s12602-017-9270-1.

Del Chierico, Federica, Pamela Vernocchi, Andrea Petrucca, Paola Paci, Susana Fuentes, Giulia Praticò, Giorgio Capuani, et al. 2015. "Phylogenetic and Metabolic Tracking of Gut Microbiota during Perinatal Development." *PLOS ONE* 10 (9): e0137347. doi: 10.1371/journal.pone.0137347.

Delplanque, Bernadette, Robert Gibson, Berthold Koletzko, Alexandre Lapillonne, and Birgitta Strandvik. 2015. "Lipid Quality in Infant Nutrition: Current Knowledge and Future Opportunities." *Journal of Pediatric Gastroenterology and Nutrition* 61 (1): 8. doi: 10.1097/MPG.0000000000000818.

Dewey, K. G., G. Beaton, C. Fjeld, B. Lönnerdal, and P. Reeds. 1996. "Protein Requirements of Infants and Children." *European Journal of Clinical Nutrition* 50 Suppl 1 (February): S119-47; discussion S147-150.

Dorni, Charles, Paras Sharma, Gunendra Saikia, and T. Longvah. 2018. "Fatty Acid Profile of Edible Oils and Fats Consumed in India." *11th IFDC Special Issue 'Food Composition and Public Health Nutrition* 238 (January): 9–15. doi: 10.1016/j.foodchem.2017.05.072.

Duan, Yifan, Zhenyu Yang, Jianqiang Lai, Dongmei Yu, Suying Chang, Xuehong Pang, Shan Jiang, et al. 2018. "Exclusive Breastfeeding Rate and Complementary Feeding Indicators in China: A National Representative Survey in 2013." *Nutrients* 10 (2). doi: 10.3390/nu10020249.

Dudgeon, Leonard S., and Reginald C. Jewesbury. 1924. "The Bacteriology of Human Milk." *Journal of Hygiene* 23 (1): 64–76. doi: 10.1017/S0022172400008470.

Duong, Dat V, Colin W Binns, and Andy H Lee. 2005. "Introduction of Complementary Food to Infants within the First Six Months Postpartum in Rural Vietnam." *Acta Paediatrica* 94 (12): 1714–20. doi: 10.1111/j.1651-2227.2005.tb01843.x.

El-Houfey, Amira A., Khaled Saad, Ahmed M. Abbas, Safaa R. Mahmoud, and Mohsen Wadani. 2017. "Factors That Influence Exclusive Breastfeeding: A Literature Review." *Int. J. Nurs. Didact* 7: 24–31. doi: 10.15520/ijnd.2017.vol7.iss11.264.24-31

Fernández, Leónides, Susana Langa, Virginia Martín, Antonio Maldonado, Esther Jiménez, Rocío Martín, and Juan M. Rodríguez. 2013. "The Human Milk Microbiota: Origin and Potential Roles in Health and Disease." *SI: Human Microbiome and Health* 69 (1): 1–10. doi: 10.1016/j.phrs.2012.09.001.

Fomon, Samuel J. 2001. “Infant Feeding in the 20th Century: Formula and Beikost.” *The Journal of Nutrition* 131 (2): 409S-420S. doi: 10.1093/jn/131.2.409S

Gavin, A., and K. Ostovar. 1977. “Microbiological Characterization of Human Milk.” *Journal of Food Protection* 40 (9): 614–16.

Goldsmith, Sara J., James S. Dickson, Harold M. Barnhart, Romeo T. Toledo, and Ronald R. Eiten-Miller. 1983. “IgA, IgG, IgM and Lactoferrin Contents of Human Milk during Early Lactation and the Effect of Processing and Storage.” *Journal of Food Protection* 46 (1): 4–7. doi: 10.4315/0362-028X-46.1.4.

Grice, Elizabeth A., Heidi H. Kong, Sean Conlan, Clayton B. Deming, Joie Davis, Alice C. Young, Gerard G. Bouffard, Robert W. Blakesley, Patrick R. Murray, and Eric D. Green. 2009. “Topographical and Temporal Diversity of the Human Skin Microbiome.” *Science* 324 (5931): 1190–92.

Gueimonde, M., K. Laitinen, S. Salminen, and E. Isolauri. 2007. “Breast Milk: A Source of Bifidobacteria for Infant Gut Development and Maturation?” *Neonatology* 92 (1): 64–66. doi: 10.1159/000100088.

Gurnida, Dida A., Angela M. Rowan, Ponpon Idjradinata, Deddy Muchtadi, and Nanan Sekarwana. 2012. “Association of Complex Lipids Containing Gangliosides with Cognitive Development of 6-Month-Old Infants.” *Early Human Development* 88 (8): 595–601. doi: 10.1016/j.earlhumdev.2012.01.003.

He, Xuan, Mariana Parenti, Tove Grip, Bo Lönnerdal, Niklas Timby, Magnus Domellöf, Olle Hernell, and Carolyn M. Slupsky. 2019. “Fecal Microbiome and Metabolome of Infants Fed Bovine MFGM Supplemented Formula or Standard Formula with Breast-Fed Infants as Reference: A Randomized Controlled Trial.” *Scientific Reports* 9 (1): 11589. doi: 10.1038/s41598-019-47953-4.

Heikkilä, Mari P., and P. E. J. Saris. 2003. “Inhibition of Staphylococcus Aureus by the Commensal Bacteria of Human Milk.” *Journal of Applied Microbiology* 95 (3): 471–78. doi: 10.1046/j.1365-2672.2003.02002.x.

Hidalgo-Cantabrana, Claudio, Susana Delgado, Lorena Ruiz, Patricia Ruas-Madiedo, Borja Sánchez, and Abelardo Margolles. 2017. “Bifidobacteria and Their Health-Promoting Effects.” *Microbiology Spectrum* 5 (3): 5.3. 21. doi: 10.1128/microbiolspec.BAD-0010-2016.

Hill, Colin, Francisco Guarner, Gregor Reid, Glenn R. Gibson, Daniel J. Merenstein, Bruno Pot, Lorenzo Morelli, Roberto Berni Canani, Harry J. Flint, and Seppo Salminen. 2014. “The International Scientific Association for Probiotics and Prebiotics Consensus Statement on the Scope and Appropriate Use of the Term Probiotic.” *Nature Reviews Gastroenterology & Hepatology* 11 (8): 506–14.

Hochwallner, Heidrun, Ulrike Schulmeister, Ines Swoboda, Susanne Spitzauer, and Rudolf Valenta. 2014. “Cow’s Milk Allergy: From Allergens to New Forms of Diagnosis, Therapy and Prevention.” *Methods (San Diego, Calif.)* 66 (1): 22–33. doi: 10.1016/j.ymeth.2013.08.005.

Hojsak, Iva, Višnja Tokić Pivac, Ana Močić Pavić, Agneza Marija Pasini, and Sanja Kolaček. 2015. “Bifidobacterium Animalis Subsp. Lactis Fails to Prevent Common Infections in Hospitalized Children: A Randomized, Double-Blind, Placebo-Controlled Study.” *The American Journal of Clinical Nutrition* 101 (3): 680–84. doi: 10.3945/ajcn.114.102004.

Hornsby, Paige P., Mark R. Conaway, Tegan J. Medico, Kelly K. Gurka, and Ann Kellams. 2021. “Timing of Introduction of Complementary Foods and Beverages to Infants of Low-Income Women.” *Breastfeeding Medicine* 16 (7): 547–52. doi: 10.1089/bfm.2020.0352.

Innis, Sheila M. 2011. “Dietary Triacylglycerol Structure and Its Role in Infant Nutrition.” *Advances in Nutrition* 2 (3): 275–83.

Jensen, Robert G., Ann M. Ferris, Carol J. Lammi-Keefe, and Robin A. Henderson. 1990. “Lipids of Bovine and Human Milks: A Comparison.” *Journal of Dairy Science* 73 (2): 223–40.

Jones, Jessica R., Michael D. Kogan, Gopal K. Singh, Deborah L. Dee, and Laurence M. Grummer-Strawn. 2011. “Factors Associated With Exclusive Breastfeeding in the United States.” *Pediatrics* 128 (6): 1117. doi: 10.1542/peds.2011-0841.

Kakute, Peter Nwenfu, John Ngum, Pat Mitchell, Kathryn A. Kroll, Gideon Wangnkeh Forgwei, Lillian Keming Ngwang, and Dorothy J. Meyer. 2005. “Cultural Barriers to Exclusive Breastfeeding by

Mothers in a Rural Area of Cameroon, Africa." *Special Issue - International Health* 50 (4): 324–28. doi: 10.1016/j.jmwh.2005.01.005.

Koletzko, Berthold, Susan Baker, Geoff Cleghorn, Ulysses Fagundes Neto, Sarath Gopalan, Olle Hernell, Quak Seng Hock, et al. 2005. "Global Standard for the Composition of Infant Formula: Recommendations of an ESPGHAN Coordinated International Expert Group." *Journal of Pediatric Gastroenterology and Nutrition* 41 (5): 584–99. doi: 10.1097/01.mpg.0000187817.38836.42.

Korpela, Katri, Anne Salonen, Brandon Hickman, Clemens Kunz, Norbert Sprenger, Kaarina Kukkonen, Erkki Savilahti, Mikael Kuitunen, and Willem M. de Vos. 2018. "Fucosylated Oligosaccharides in Mother's Milk Alleviate the Effects of Caesarean Birth on Infant Gut Microbiota." *Scientific Reports* 8 (1): 13757. doi: 10.1038/s41598-018-32037-6.

Kronborg, Hanne, Else Foverskov, and Michael Væth. 2015. "Breastfeeding and Introduction of Complementary Food in Danish Infants." *Scandinavian Journal of Public Health* 43 (2): 138–45. doi: 10.1177/1403494814567171.

Kuchenbecker, J., I. Jordan, A. Reinbott, J. Herrmann, T. Jeremias, G. Kennedy, E. Muehlhoff, B. Mtimuni, and M. B. Krawinkel. 2015. "Exclusive Breastfeeding and Its Effect on Growth of Malawian Infants: Results from a Cross-Sectional Study." *Paediatrics and International Child Health* 35 (1): 14–23. doi: 10.1179/2046905514Y.0000000134.

Kunz, C., S. Rudloff, W. Baier, N. Klein, and S. Strobel. 2000. "Oligosaccharides in Human Milk: Structural, Functional, and Metabolic Aspects." *Annual Review of Nutrition* 20 (1): 699–722. doi: 10.1146/annurev.nutr.20.1.699.

Lawrence, Robert M., and Ruth A. Lawrence. 2004. "Breast Milk and Infection." *Current Controversies in Perinatal Medicine IV* 31 (3): 501–28. doi: 10.1016/j.clp.2004.03.019.

Lee, Hanna, Emily Padhi, Yu Hasegawa, Jules Larke, Mariana Parenti, Aidong Wang, Olle Hernell, Bo Lönnerdal, and Carolyn Slupsky. 2018. "Compositional Dynamics of the Milk Fat Globule and Its Role in Infant Development." *Frontiers in Pediatrics* 6: 313. doi: 10.3389/fped.2018.00313.

Lewis, Zachery T., Sarah M. Totten, Jennifer T. Smilowitz, Mina Popovic, Evan Parker, Danielle G. Lemay, Maxwell L. Van Tassell, Michael J. Miller, Yong-Su Jin, and J. Bruce German. 2015. "Maternal Fucosyltransferase 2 Status Affects the Gut Bifidobacterial Communities of Breastfed Infants." *Microbiome* 3 (1): 13.

Lindmark Månsson, Helena. 2008. "Fatty Acids in Bovine Milk Fat." *Food & Nutrition Research* 52 (1): 1821.

López-López, A., MC López-Sabater, C. Campoy-Folgoso, M. Rivero-Urgell, and AI Castellote-Bargalló. 2002. "Fatty Acid and Sn-2 Fatty Acid Composition in Human Milk from Granada (Spain) and in Infant Formulas." *European Journal of Clinical Nutrition* 56 (12): 1242–54. doi: 10.1038/sj.ejcn.1601470.

Lu, Evelyn, Marie Yeung, and Chi Kong Yeung. 2018. "Comparative Analysis of Lactulose and Fructooligosaccharide on Growth Kinetics, Fermentation, and Antioxidant Activity of Common Probiotics." *Food and Nutrition Sciences*.

Makino, H., R. Martin, E. Ishikawa, A. Gawad, H. Kubota, T. Sakai, K. Oishi, et al. 2015. "Multilocus Sequence Typing of Bifidobacterial Strains from Infant's Faeces and Human Milk: Are Bifidobacteria Being Sustainably Shared during Breastfeeding?" *Beneficial Microbes* 6 (4): 563–72. doi: 10.3920/BM2014.0082.

Manson, William G., and Lawrence T. Weaver. 1997. "Fat Digestion in the Neonate." *Archives of Disease in Childhood-Fetal and Neonatal Edition* 76 (3): F206–11.

Marín-Manzano, María D., Oswaldo Hernandez-Hernandez, Marina Diez-Municio, Cristina Delgado-Andrade, Francisco J. Moreno, and Alfonso Clemente. 2020. "Prebiotic Properties of Non-Fructosylated α-Galactooligosaccharides from PEA (Pisum Sativum L.) Using Infant Fecal Slurries." *Foods* 9 (7). doi: 10.3390/foods9070921.

Martin, Camilia R, Pei-Ra Ling, and George L Blackburn. 2016. "Review of Infant Feeding: Key Features of Breast Milk and Infant Formula." *Nutrients* 8 (5): 279. doi: 10.3390/nu8050279.

Martín, Rocío, Esther Jiménez, Hans Heilig, Leonides Fernández, María L. Marín, Erwin G. Zoetendal, and Juan M. Rodríguez. 2009. "Isolation of Bifidobacteria from Breast Milk and Assessment of the Bifidobacterial Population by PCR-Denaturing Gradient Gel Electrophoresis and Quantitative Real-Time PCR." *Applied and Environmental Microbiology* 75 (4): 965–69.

Martín, Rocío, Susana Langa, Carlota Reviriego, Esther Jimínez, María L. Marín, Jordi Xaus, Leonides Fernández, and Juan M. Rodríguez. 2003. "Human Milk Is a Source of Lactic Acid Bacteria for the Infant Gut." *The Journal of Pediatrics* 143 (6): 754–58.

Mazariegos, Monica, Christine Slater, and Manuel Ramirez-Zea. 2016. "Validity of Guatemalan Mother's Self-Reported Breast-Feeding Practices of 3-Month-Old Infants." *Food and Nutrition Bulletin* 37 (4): 494–503. doi: 10.1177/0379572116654644.

Medicine, National Academy of Sciences Institute of. 2004. *Infant Formula: Evaluating the Safety of New Ingredients*. National Academies Press.

Mehrotra, Vipul, Sunil Kumar Sehgal, and Nikhil Rajeev Bangale. 2019. "Fat Structure and Composition in Human Milk and Infant Formulas: Implications in Infant Health." *Clinical Epidemiology and Global Health* 7 (2): 153–59. doi: 10.1016/j.cegh.2018.03.005.

Mendonça, Marcio Antonio, Wilma Maria Coelho Araújo, Luiz Antonio Borgo, and Ernandes de Rodrigues Alencar. 2017. "Lipid Profile of Different Infant Formulas for Infants." *PloS One* 12 (6): e0177812.

Moles, Laura, Marta Gómez, Hans Heilig, Gerardo Bustos, Susana Fuentes, Willem de Vos, Leónides Fernández, Juan M. Rodríguez, and Esther Jiménez. 2013. "Bacterial Diversity in Meconium of Preterm Neonates and Evolution of Their Fecal Microbiota during the First Month of Life." *PLOS ONE* 8 (6): e66986. doi: 10.1371/journal.pone.0066986.

Ninonuevo, Milady R., Youmie Park, Hongfeng Yin, Jinhua Zhang, Robert E. Ward, Brian H. Clowers, J. Bruce German, et al. 2006. "A Strategy for Annotating the Human Milk Glycome." *Journal of Agricultural and Food Chemistry* 54 (20): 7471–80. doi: 10.1021/jf0615810.

Noel-Weiss, Joy, Sonya Boersma, and Sonya Kujawa-Myles. 2012. "Questioning Current Definitions for Breastfeeding Research." *International Breastfeeding Journal* 7 (1): 9. doi: 10.1186/1746-4358-7-9.

Oswari, Hanifah, Ariani Dewi Widodo, Frieda Handayani, Mohammad Juffrie, Tonny Sundjaya, Jacques Bindels, and Badriul Hegar. 2019. "Dosage-Related Prebiotic Effects of Inulin in Formula-Fed Infants." *Pediatric Gastroenterology, Hepatology & Nutrition* 22 (1): 63–71. doi: 10.5223/pghn.2019.22.1.63.

Parizkova, Petra, Nika Dankova, Pavel Frühauf, Jitka Jireckova, Jiri Zeman, and Martin Magner. 2020. "Associations between Breastfeeding Rates and Infant Disease: A Survey of 2338 Czech Children." *Nutrition & Dietetics* 77 (3): 310–14. doi: 10.1111/1747-0080.12532.

Park, Young W. 2017. "Goat Milk – Chemistry and Nutrition." In *Handbook of Milk of Non-Bovine Mammals*, 42–83. John Wiley & Sons, Ltd. doi: 10.1002/9781119110316.ch2.2.

Plaza-Diaz, Julio, Francisco Javier Ruiz-Ojeda, Mercedes Gil-Campos, and Angel Gil. 2019. "Mechanisms of Action of Probiotics." *Advances in Nutrition* 10 (suppl_1): S49–66.

Puccio, Giuseppe, Philippe Alliet, Cinzia Cajozzo, Elke Janssens, Giovanni Corsello, Norbert Sprenger, Susan Wernimont, Delphine Egli, Laura Gosoniu, and Philippe Steenhout. 2017. "Effects of Infant Formula with Human Milk Oligosaccharides on Growth and Morbidity: A Randomized Multicenter Trial." *Journal of Pediatric Gastroenterology and Nutrition* 64 (4): 624–31. doi: 10.1097/MPG.0000000000001520.

Roberfroid, Marcel. 2007. "Prebiotics: The Concept Revisited." *The Journal of Nutrition* 137 (3): 830S-837S. doi: 10.1093/jn/137.3.830S.

Roger, Laure C., Adele Costabile, Diane T. Holland, Lesley Hoyles, and Anne L. McCartney. 2010. "Examination of Faecal Bifidobacterium Populations in Breast- and Formula-Fed Infants during the First 18 Months of Life." *Microbiology*. Microbiology Society.

Rosenberg, Kenneth D., Carissa A. Eastham, Laurin J. Kasehagen, and Alfredo P. Sandoval. 2008. "Marketing Infant Formula through Hospitals: The Impact of Commercial Hospital Discharge

Packs on Breastfeeding." *American Journal of Public Health* 98 (2): 290–95. doi: 10.2105/AJPH.2006.103218.

Samady, Waheeda, Emily Campbell, Ozge Nur Aktas, Jialing Jiang, Alexandria Bozen, Jamie L. Fierstein, Alanna Higgins Joyce, and Ruchi S. Gupta. 2020. "Recommendations on Complementary Food Introduction among Pediatric Practitioners." *JAMA Network Open* 3 (8): e2013070–e2013070. doi: 10.1001/jamanetworkopen.2020.13070.

Scalabrin, Deolinda M. F., Susan H. Mitmesser, Gjalt W. Welling, Cheryl L. Harris, John D. Marunycz, D. Carey Walker, Nico A. Bos, Satu Tölkkö, Seppo Salminen, and Jon A. Vanderhoof. 2012. "New Prebiotic Blend of Polydextrose and Galacto-Oligosaccharides Has a Bifidogenic Effect in Young Infants." *Journal of Pediatric Gastroenterology and Nutrition* 54 (3): 343–52. doi: 10.1097/MPG.0b013e318237ed95.

Smith-Brown, Paula, Mark Morrison, Lutz Krause, and Peter S. W. Davies. 2016. "Mothers Secretor Status Affects Development of Childrens Microbiota Composition and Function: A Pilot Study." *PLOS ONE* 11 (9): e0161211. doi: 10.1371/journal.pone.0161211.

Storm, Heidi M., Julie Shepard, Laura M. Czerkies, Brian Kineman, Sarah S. Cohen, Heidi Reichert, and Ryan Carvalho. 2019. "2′-Fucosyllactose Is Well Tolerated in a 100% Whey, Partially Hydrolyzed Infant Formula with Bifidobacterium Lactis: A Randomized Controlled Trial." *Global Pediatric Health* 6: 2333794X19833995.

Sun, Cong, Xiaoqiang Zou, Yunping Yao, Jun Jin, Yuan Xia, Jianhua Huang, Qingzhe Jin, and Xingguo Wang. 2016. "Evaluation of Fatty Acid Composition in Commercial Infant Formulas on the Chinese Market: A Comparative Study Based on Fat Source and Stage." *International Dairy Journal* 63: 42–51.

Taipale, Teemu J., Kaisu Pienihäkkinen, Erika Isolauri, Jorma T. Jokela, and Eva M. Söderling. 2016. "Bifidobacterium Animalis Subsp. Lactis BB-12 in Reducing the Risk of Infections in Early Childhood." *Pediatric Research* 79 (1–1): 65–69. doi: 10.1038/pr.2015.174.

Tao, N., E.J. DePeters, S. Freeman, J.B. German, R. Grimm, and C.B. Lebrilla. 2008. "Bovine Milk Glycome." *Journal of Dairy Science* 91 (10): 3768–78. doi: 10.3168/jds.2008-1305.

Thompkinson, D.K., and Suman Kharb. 2007. "Aspects of Infant Food Formulation." *Comprehensive Reviews in Food Science and Food Safety* 6 (4): 79–102. doi: 10.1111/j.1541-4337.2007.00020.x.

Timby, Niklas, Erik Domellöf, Olle Hernell, Bo Lönnerdal, and Magnus Domellöf. 2014. "Neurodevelopment, Nutrition, and Growth until 12 Mo of Age in Infants Fed a Low-Energy, Low-Protein Formula Supplemented with Bovine Milk Fat Globule Membranes: A Randomized Controlled Trial." *The American Journal of Clinical Nutrition* 99 (4): 860–68. doi: 10.3945/ajcn.113.064295.

Timby, Niklas, Olle Hernell, Outi Vaarala, Merit Melin, Bo Lönnerdal, and Magnus Domellöf. 2015. "Infections in Infants Fed Formula Supplemented With Bovine Milk Fat Globule Membranes." *Journal of Pediatric Gastroenterology and Nutrition* 60 (3).

Totten, Sarah M., Angela M. Zivkovic, Shuai Wu, UyenThao Ngyuen, Samara L. Freeman, L. Renee Ruhaak, Momodou K. Darboe, J. Bruce German, Andrew M. Prentice, and Carlito B. Lebrilla. 2012. "Comprehensive Profiles of Human Milk Oligosaccharides Yield Highly Sensitive and Specific Markers for Determining Secretor Status in Lactating Mothers." *Journal of Proteome Research* 11 (12): 6124–33. doi: 10.1021/pr300769g.

Tu, Anqi, Qiang Ma, Hua Bai, and Zhenxia Du. 2017. "A Comparative Study of Triacylglycerol Composition in Chinese Human Milk within Different Lactation Stages and Imported Infant Formula by SFC Coupled with Q-TOF-MS." *Food Chemistry* 221 (April): 555–67. doi: 10.1016/j.foodchem.2016.11.139.

Underwood, Mark A., Stephanie Gaerlan, Maria Lorna A. De Leoz, Lauren Dimapasoc, Karen M. Kalanetra, Danielle G. Lemay, J. Bruce German, David A. Mills, and Carlito B. Lebrilla. 2015. "Human Milk Oligosaccharides in Premature Infants: Absorption, Excretion, and Influence on the Intestinal Microbiota." *Pediatric Research* 78 (August): 670.

Underwood, Mark A., J. Bruce German, Carlito B. Lebrilla, and David A. Mills. 2014. “Bifidobacterium Longum Subspecies Infantis: Champion Colonizer of the Infant Gut.” *Pediatric Research* 77 (October): 229.

Vandenplas, Yvan, Bernard Berger, Virgilio P. Carnielli, Janusz Ksiazyk, Hanna Lagström, Manuel Sanchez Luna, Natalia Migacheva, et al. 2018. “Human Milk Oligosaccharides: 2′-Fucosyllactose (2′-FL) and Lacto-N-Neotetraose (LNnT) in Infant Formula.” *Nutrients* 10 (9). doi: 10.3390/nu10091161.

Vandenplas, Yvan, Irina Zakharova, and Yulia Dmitrieva. 2015. “Oligosaccharides in Infant Formula: More Evidence to Validate the Role of Prebiotics.” *The British Journal of Nutrition* 113 (9): 1339–44. doi: 10.1017/S0007114515000823.

Velaphi, Sithembiso C., Peter A. Cooper, Keith D. Bolton, Mantoa Mokhachane, Ramatsimele M. Mphahlele, Elaine Beckh-Arnold, Liemo Monaheng, and Elisabeth Haschke-Becher. 2008. “Growth and Metabolism of Infants Born to Women Infected with Human Immunodeficiency Virus and Fed Acidified Whey-Adapted Starter Formulas.” *Nutrition* 24 (3): 203–11. doi: 10.1016/j.nut.2007.11.002.

Wang, Mei, Min Li, Shuai Wu, Carlito B Lebrilla, Robert S Chapkin, Ivan Ivanov, and Sharon M Donovan. 2015. “Fecal Microbiota Composition of Breast-Fed Infants Is Correlated with Human Milk Oligosaccharides Consumed.” *Journal of Pediatric Gastroenterology and Nutrition* 60 (6): 825–33. doi: 10.1097/MPG.0000000000000752.

Weizman, Zvi, and Ahmed Alsheikh. 2006. “Safety and Tolerance of a Probiotic Formula in Early Infancy Comparing Two Probiotic Agents: A Pilot Study.” *Journal of the American College of Nutrition* 25 (5): 415–19. doi: 10.1080/07315724.2006.10719554.

Weizman, Zvi, Ghaleb Asli, and Ahmed Alsheikh. 2005. “Effect of a Probiotic Infant Formula on Infections in Child Care Centers: Comparison of Two Probiotic Agents.” *Pediatrics* 115 (1): 5. doi: 10.1542/peds.2004-1815.

West, P. A., J. H. Hewitt, and OLIVE M. MURPHY. 1979. “The Influence of Methods of Collection and Storage on the Bacteriology of Human Milk.” *Journal of Applied Bacteriology* 46 (2): 269–77.

World Health Organization. 2009. “Global Health Risks : Mortality and Burden of Disease Attributable to Selected Major Risks.”

Yatsunenko, Tanya, Federico E. Rey, Mark J. Manary, Indi Trehan, Maria Gloria Dominguez-Bello, Monica Contreras, Magda Magris, et al. 2012. “Human Gut Microbiome Viewed across Age and Geography.” *Nature* 486 (7402): 222–27. doi: 10.1038/nature11053.

Young, B.E. 2017. “Chapter 2 - Breastfeeding and Human Milk: Short and Long-Term Health Benefits to the Recipient Infant.” In *Early Nutrition and Long-Term Health*, edited by Jose M. Saavedra and Anne M. Dattilo, 25–53. Woodhead Publishing. doi: 10.1016/B978-0-08-100168-4.00002-1.

Zhao, Junying, Wei Yi, Bin Liu, Yaohua Dai, Tiemin Jiang, Shuxing Chen, Jianwu Wang, et al. 2021. “MFGM Components Promote Gut Bifidobacterium Growth in Infant and in Vitro.” *European Journal of Nutrition*, July. doi: 10.1007/s00394-021-02638-5.

Zhu, Xiu-Li, Xiao-Gang Tang, Fan Qu, Yu Zheng, Wen-Hao Zhang, and Yu-Qiao Diao. 2019. “Bifidobacterium May Benefit the Prevention of Necrotizing Enterocolitis in Preterm Infants: A Systematic Review and Meta-Analysis.” *International Journal of Surgery* 61 (January): 17–25. doi: 10.1016/j.ijsu.2018.11.026.

Zivkovic, Angela M., J. Bruce German, Carlito B. Lebrilla, and David A. Mills. 2011. “Human Milk Glycobiome and Its Impact on the Infant Gastrointestinal Microbiota.” *Proceedings of the National Academy of Sciences* 108 (Supplement 1): 4653. doi: 10.1073/pnas.1000083107.

Zou, Long, Garima Pande, and Casimir C. Akoh. 2016. “Infant Formula Fat Analogs and Human Milk Fat: New Focus on Infant Developmental Needs.” *Annual Review of Food Science and Technology* 7 (1): 139–65. doi: 10.1146/annurev-food-041715-033120.

Chapter 2 Difference in milk metabolome between secretor status and Lewis blood type

The functional role of milk for the developing neonate is an area of great interest, and a significant amount of research has been done. However, a lot of work remains to fully understand the complexities of milk, and the variations imposed through genetics. It has previously been shown that both secretor (Se) and Lewis blood type (Le) status impacts the human milk oligosaccharide (HMO) content of human milk. While some studies have compared the non-HMO milk metabolites of Se+ and Se- women, none have reported on the non-HMO milk metabolome of Se- and Le- mothers. To determine the differences in the non-HMO milk metabolome between Se-Le- mothers and other HMO phenotypes (Se+Le+, Se+Le-, and Se-Le+), 10 milk samples from 10 lactating mothers were analyzed using nuclear magnetic resonance (NMR) spectroscopy. Se or Le HMO phenotypes were assigned based on the presence and absence of 6 HMOs generated by the Se and Le genes. After classification, 58 milk metabolites were compared among the HMO phenotypes. Principal component analysis (PCA) identified clear separation between Se-Le- milk and the other milks. Fold change analysis demonstrated that the Se-Le- milk had major differences in free fatty acids, free amino acids, and metabolites related to energy metabolism. The results suggest that the milk metabolome of mothers with the Se-Le- phenotype differs in its non-HMO metabolite composition from mothers with other HMO phenotypes.

This chapter was originally published in *Frontiers in Nutrition*:

Wang, Aidong, Petya Koleva, Elloise du Toit, Donna T Geddes, Daniel Munblit, Susan L Prescott, Merete Eggesbø, et al. 2021. "The Milk Metabolome of Non-Secretor and Lewis Negative Mothers." *Frontiers in Nutrition* 7 (February): 576966–576966. doi: 10.3389/fnut.2020.576966.

Introduction

Human milk is the gold standard for infant nutrition as it provides essential nutrients for infant growth, as well as bioactive components such as human milk oligosaccharides (HMOs). While the variation of HMOs among different maternal HMO phenotypes has been widely studied (Thurl et al. 2010a; Stahl et al. 2001; Elwakiel et al. 2018; Newburg et al. 2004), the impact of the maternal HMO phenotypes on other low-molecular-weight milk metabolites remains unclear. Metabolites other than oligosaccharides are thought to play important roles in infant health. For example, milk glutamate has been shown to impact appetite and growth (Ventura, Beauchamp, and Mennella 2012), biogenic amines have been reported to provide protection against infectious disease (Gregory and Walker 2013), taurine has been recognized to contribute to neonatal brain development (Tochitani 2017), and creatine appears to be essential for normal neural development (Edison et al. 2013). An understanding of how these metabolites change with HMO phenotype may be important to further understanding of the function of these metabolites in milk.

Maternal HMO phenotypes are determined by the activity of two genes: the secretor (Se) gene fut2, coding for α-1,2-fucosyltransferase (FUT2), and the Lewis (Le) gene fut3, coding for α-1,3/1,4-fucosyltransferase (FUT3). FUT2 and FUT3 are responsible for the fucosylation of milk oligosaccharides. There are five monosaccharides upon which all HMOs are built: D-glucose, D-galactose, N-acetylglucosamine (GlcNAc), L-fucose and sialic acid (Neu5Ac) (Kunz et al. 2000). At the core of the HMO structure is lactose, which can be sialylated to form α2-3 (e.g., 3′sialyllactose, 3'SL) or α2-6 (e.g., 6′sialyllactose, 6'SL) linkages to sialic acid, or fucosylated to form α1-2 (e.g., 2′FL), or α1-3 (e.g., 3FL) linkages to fucose. To form more complex HMOs, lactose can be elongated through a β1-3 linkage to lacto-N-biose (type I) or a β1-6 linkage to N-acetyllactosamine (type II). Lactose or the formed polylactosamine backbone can then be

sialylated and/or fucosylated to create an additional 200 different oligosaccharide structures (Ninonuevo et al. 2006). FUT2 synthesizes 2'FL or lacto-N-fucopentose I (LNFP I) by attaching a fucose to lactose or lacto-N-tetraose (LNT) respectively. FUT3 synthesizes lacto-N-difucohexaose I (LDFH I) and lactodifucotetraose (LDFT) from LNFP I and 2'FL respectively by attaching an additional fucose. FUT3 can also directly transfer fucose to LNT, lactose, and lacto-N-neotetraose (LNnT) to form lacto-N-neotetraose II (LNFP II), 3FL, and lacto-N-neotetraose III (LNFP III), respectively (Morrow et al. 2003). Additionally, the α-1,3-fucosyltransferases encoded by fut4, 5, 6, 7 and/or 9, which are Se- and Le-independent, also play roles in attaching fucose to lactose, and thus 3FL and LNFP III can sometimes be observed in milk from Lewis negative (Le-) women (Oriol, Le Pendu, and Mollicone 1986; Kunz et al. 2017). It has been speculated that FUT1 also participates in HMO fucosylation, as α-1-2-fucosylated HMOs have been observed in milk from Se- women (Newburg, Ruiz-Palacios, and Morrow 2005). In human milk from Se+/Le+ women, 35-50% of the HMOs are fucosylated, 12-14% are sialylated and 42-55% are nonfucosylated neutral (Totten et al. 2012a).

While the Se and Le genes are important to generate a variety of HMOs in both free and conjugated forms, many individuals have polymorphisms in one or both of these genes making them non-functional. In European and American populations, the Le- frequency is between 4% - 6%, and 20% of the population are Se-, making Se-Le- extremely rare. In contrast, in certain African populations, over 30% of the population are Le- and approximately 38% are Se- (Nordgren et al. 2014; Barnicot and Lawler 1953; Koda et al. 2001), which makes the probability of having Se-Le- mothers higher. The importance of functional Se and Le genes in infant development is an area of active research. One study showed that maternal secretor status appeared to be important for preventing diarrhea, as although the gut microbiota measured through 16S rRNA sequencing did

not differ between infants of Se+ and Se- mothers, the prevalence of diarrhea was higher among infants of Se- mothers (Paganini et al. 2019). Moreover, when these infants were provided iron supplements, infants of Se- mothers were more likely to experience a decrease in the abundance of Bifidobacterium and an increase in pathogens compared to infants of Se+ mothers (Paganini et al. 2019). However, supplementation with galactooligosaccharides appeared to ameliorate the impact of iron supplementation (Paganini et al. 2019).

Studies comparing non-HMO milk metabolites from mothers who were phenotypically Se+ to Se- demonstrated no differences between groups (Dessì et al. 2018; Praticò et al. 2014). We have previously reported on the milk metabolome at day 90 (Smilowitz et al. 2013) and over the first month of lactation (Spevacek et al. 2015) in Se+Le+ and Se-Le+ women. We observed no significant difference in non-HMO metabolites between the two groups. To date, no studies have compared the non-HMO metabolites of milk from phenotypically Se-Le- mothers to any other phenotype. We hypothesized that the non-HMO milk metabolome from Se-Le- women would be similar to the other phenotypes one month after delivery. This study provides preliminary data on the comparison of the milk metabolic profile between women with the Se-Le- phenotype and other phenotypes.

Materials and methods

Milk sample preparation

In this pilot study, to maximize the homogeneity of subjects (Ten-Doménech et al. 2020), human milk samples were collected one month postpartum from 10 randomly-selected healthy women (age 29.8 ± 4.8, pre-pregnancy BMI 25.0 ± 2.9) in Cape Town, South Africa, who gave birth to term infants (50% male) through vaginal delivery, and practiced exclusive breastfeeding prior to

sample collection. The exclusion criteria included antibiotic or probiotic treatment during the last trimester of pregnancy and the breastfeeding period. Ethical approval for this study was provided by the University of Cape Town's Human Research Ethical Committee (HREC REF: 306/2014). Mature milk samples from mothers were collected after obtaining their consent. Women were asked to wash their hands, their nipple, and surrounding breast area with soap, then soak the breast area with chlorhexidine to reduce contamination by skin microbes, followed by washing with sterile water. A small volume of milk was collected manually or with an electric breast pump into a sterile collection bottle after discarding the first few drops. Time since last feed was not recorded. After collection, samples were transported on ice and stored at −20 °C until further processing. This study is a subset of a larger study on the relationship of milk short chain fatty acids and atopy (Stinson et al. 2020).

Milk samples were prepared as previously described (Gay et al. 2018). Briefly, samples were thawed on ice, mixed, then 500 μL of each sample was filtered through Amicon Ultra 0.5 mL 3-kDa cutoff spin filters (Millipore Sigma, Burlington, MA, USA) at 10,000× g for 15 min at 4 ºC to remove lipids and protein, as the study was interested in low-molecular-weight polar metabolites. 350 μL of filtrate was mixed with 70 μL of deuterium oxide and 60 μL of standard buffer solution (consisting of 585 mM NaHPO4 (pH 7.0), 11.667 mM disodium-2,2-dimethyl-2-silapentane-5-sulfonate (DSS, internal standard), and 0.47% NaN_3 in H_2O) in a 1.5 mL Eppendorf tube (Gay et al. 2018). 460 μL of the mixture was transferred to a nuclear magnetic resonance (NMR) tube for subsequent NMR spectral analysis.

NMR data acquisition and processing

^1H NMR spectra were acquired at 25 ºC using the first transient of the Varian tnnoesy pulse sequence on a Varian 500 MHz Inova spectrometer equipped with a 5-mm HCN cold probe. Water

suppression pulses were calibrated to achieve a bandwidth of 80 G. Spectra were collected with 128 transients and 8 steady-state scans using a 4 s acquisition time (48,000 complex points) and a 1 s recycle delay. Before spectral analysis, all free induction decays were zero-filled to 64,000 data points and line broadened to 0.5 Hz. The methyl singlet produced by DSS internal standard was used for chemical shift referencing (set to 0 ppm) and for quantification. Spectra were manually processed and profiled using Chenomx NMR Suite version 8.1 (Chenomx Inc., Edmonton, AB, Canada).

HMO phenotype determination

The HMO phenotype was determined based on the presence or absence of specific milk oligosaccharides in the NMR spectra, which were identified and quantified from an NMR spectral library created through the analytical preparation of 10 commercially available HMOs as previously described (Smilowitz et al. 2013). In this study, the limit of detection was set to 20 μM for these compounds based on the ability to clearly observe spectral peaks of these HMOs over noise in the spectra generated from the Varian 500 MHz Inova spectrometer. Detection of both 2'FL and LNFP I in milk resulted in phenotype assignment as Se+, otherwise Se-. When LNFP II, 3FL, LDFT, and LNFP III were visible in the NMR spectra, the phenotype was assigned as Le+, otherwise Le-.

Statistical analysis

Statistical computing and graphical generation were performed using the R (version 3.5.2) programing environment. Prior to principal component analysis (PCA), generalized log transformation (defined as $\log_2(1+y)$ where y is the metabolite concentration) was applied to all

metabolomics data. PCA was computed using the *prcomp* function in the *stats* package of R without scaling the transformed data, and the first two components were plotted.

Metabolomics data without log transformation was used to perform log2_Fold calculation according the following equation.

$$\text{log2\_Fold} = \log_2\left(\frac{Mean\ concentration\ of\ Metabolite\ 1\ in\ Se-Le-\ or\ Se-Le+\ group}{Mean\ concentration\ of\ Metabolite\ 1\ in\ Se+\ group}\right)$$

Briefly, the mean concentration of each metabolite was first calculated for the Se-Le-, Se-Le+, and Se+ phenotypes (the Se+Le+ and Se+Le- samples were combined since there was only one Se+Le- sample). The mean concentration of each metabolite in the Se-Le- (or Se-Le+) groups was divided by the mean concentration of the same metabolite in the Se+ group to calculate the ratio between Se-Le- (or Se-Le+) and Se+ phenotypes. To ensure metabolites were expressed in the same range, log2 transformation was applied. To decrease the chance of false discovery using FDR-corrected p-values (since most metabolites were significantly different using this method), we considered a log2 fold change cut off of ±1.5 as an indication of significance.

Results

In total, 10 milk samples were collected from South African women 1 month after term delivery, of which 60% (n=6) were Mixed Race, 20% (n=2) were Black, and 20% (n=2) were Caucasian. None of the women had atopic disease. An NMR spectrum annotated with HMO peaks is shown in Figure 2.1A. Multiple peaks of each HMO could be identified, with some overlapping with other metabolites in milk. The HMO phenotypes of the subjects was estimated by assessing the presence or absence of specific HMOs in the milk samples (Table 2.1), with examples of the NMR spectrum corresponding to each of the HMO phenotypes shown in Figure 1B. Samples where both

2’FL and LNFP I could be measured were assigned as Se+, while samples where these two HMOs could not be detected were designated Se-. No sample was detected with only one of the two HMOs. Se+ samples with the presence of LNFP II, 3FL, LDFT, and LNFP III were assigned as Se+Le+, otherwise they were assigned as Se+Le-. Se- samples with detectable levels of LNFP II, 3FL, and LNFP III were classified as Se-Le+, and for those without these three HMOs as Se-Le-. Out of 10 samples analyzed, three samples were designated Se-Le-, as none of the six targeted HMOs was detected in any of these samples. Additionally, the area under the peak for the three FUT 3-catalyzed HMOs (LNFP II, LNFP III and 3FL) were higher in milk from Se-Le+ mothers compared to milk from Se+Le+ mothers (Figure 2.1B).

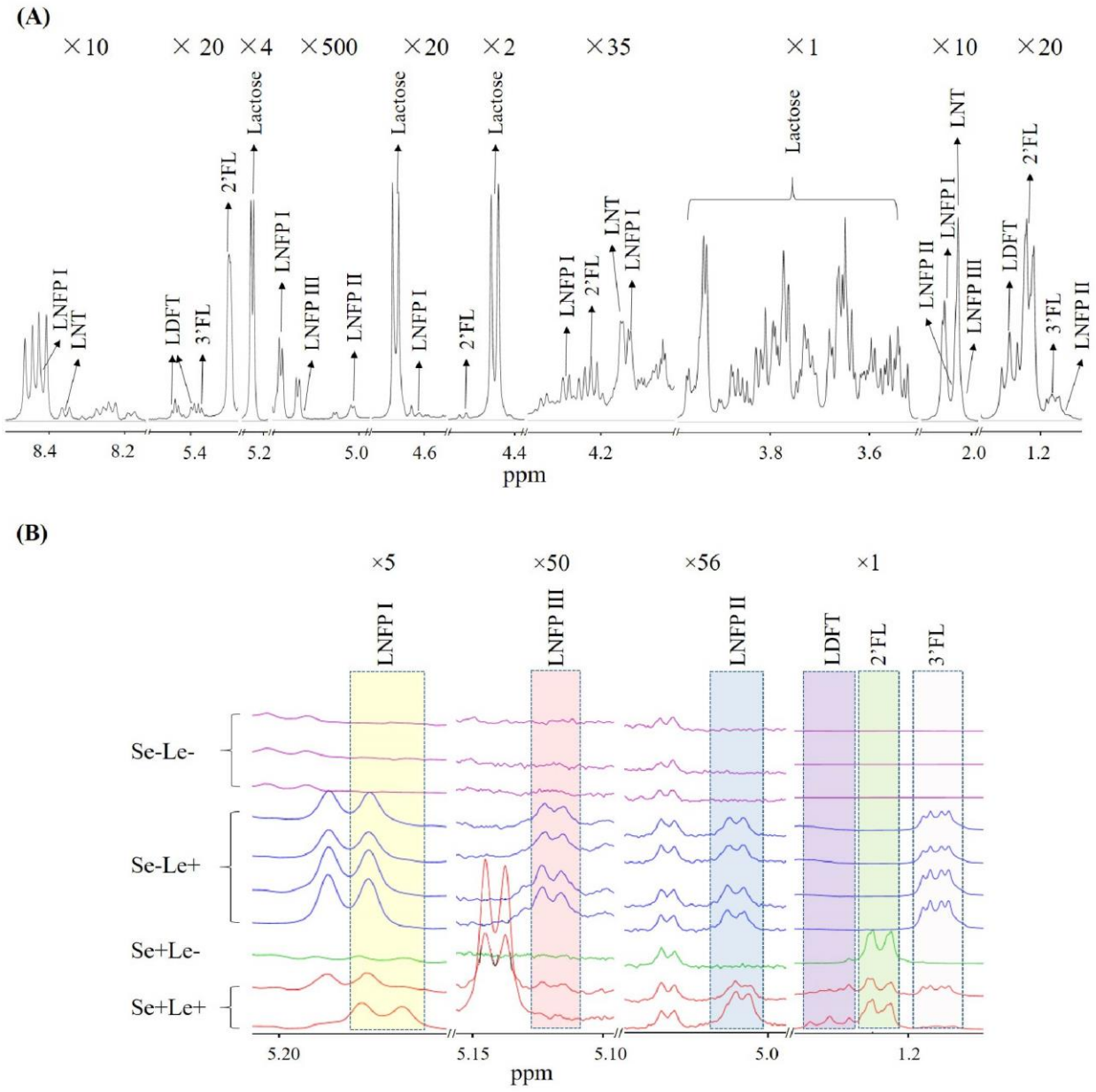


Figure 2.1. Identification of HMOs in human milk NMR spectra. A) Multiple peaks of each HMO are shown in ten different chemical shift regions at various vertical scales to illustrate characteristic peaks associated with identified HMOs. Magnification is indicated at the top of each segment. B) Comparison of NMR spectra of milk between individuals with putative differences in Se and Le status.

Table 2.1. Proposed synthetic pathways of the principal fucosyloligosaccharides used to identify Secretor (Se) and Lewis (Le) gene types based on presence/absence in the 10 milk samples.

Subject	Starting structure	LNT	Lactose	LNT	LNnT	Lactose	2'FL	
	HMO phenotype	Se+		Le+				HMO phenotype
	HMO synthesized	LNFP I	2'FL	LNFP II	LNFP III	3FL	LDFT	
1		bld	bld	bld	bld	bld	bld	Se-Le-
2		√	√	√	√	√	√	Se+Le+
3		bld	bld	√	√	√	bld	Se-Le+
4		bld	bld	√	√	√	bld	Se-Le+
5		bld	bld	bld	bld	bld	bld	Se-Le-
6		bld	bld	√	√	√	bld	Se-Le+
7		bld	bld	√	√	√	bld	Se-Le+
8		bld	bld	bld	bld	bld	bld	Se-Le-
9		√	√	bld	bld	bld	bld	Se+Le-
10		√	√	√	√	√	√	Se+Le+

LNFP I: lacto-N-fucopentose I, LNFP II: lacto-N-fucopentose II, LNFP III: lacto-N-fucopentose III, LNT: lacto-N-tetraose, LNnT: lacto-N-neotetraose, LDFH: lacto-N-difucohexaose, LDFT: lactodifucotetraose, 2'FL: 2'fucosyllactose, 3FL: 3-fucosyllactose,
Se: secretor
Le: Lewis
√: Detected. HMO concentrations over 20 uM were considered detected.
bld: Below limit of detection. HMO concentrations below 20 uM were considered below detection limit.

To evaluate whether the milk metabolites were different among the HMO phenotypes, 58 quantified polar metabolites (excluding the HMOs resulting from FUT2 and FUT3) were examined and compared. Figure 2.2A shows a principal component analysis (PCA) of milk metabolites of women from the identified HMO phenotypes. Separation along PC1, which explained 48.7% of the variance, revealed a difference between the Se-Le- group and all other groups. Along PC2, which explained 15.0% of the variance, separation based on Se status was observed. As there was only one sample identified as Se+Le-, and it did not separate from the

Se+Le+ samples in the PCA plot (Figure 2.2A), it was combined with the Se+Le+ samples (Se+ samples) in further analyses.

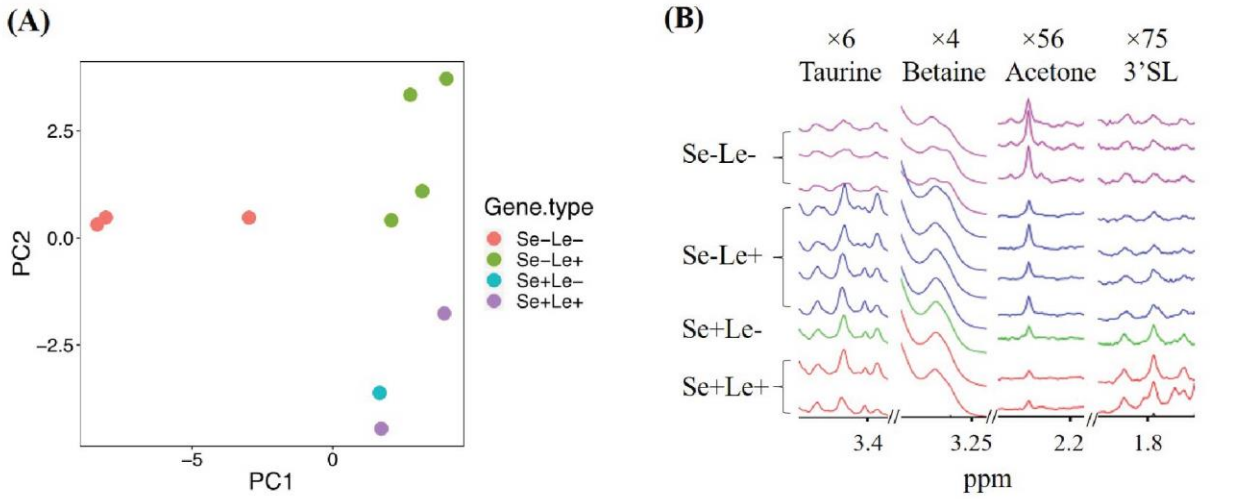


Figure 2.2. Comparison of metabolic profiles of each group excluding oligosaccharide concentrations. (A) Principal component analysis of metabolites not used for HMO phenotype assignment. (B) Metabolites with a different trend between secretor status compared to the literature**.** Magnification is indicated at the top of each segment.

In order to further compare milk metabolites among groups, the fold/ratio of metabolite concentrations in milk from Se-Le- and Se-Le+ mothers relative to milk from Se+ mothers were calculated (Figure 2.3). In terms of the oligosaccharides and their metabolites, 3'galactosyllactose, 3'SL, fucose, and LNnT were between 2 and 10-fold lower in milk samples from Se-Le- and Se-Le+ compared to Se+ mothers. Galactose was 6 and 1 times higher in milk samples from Se-Le- and Se-Le+ mothers respectively compared to samples from Se+ mothers. For metabolites associated with energy metabolism, samples from Se-Le- milk were approximately 4 times higher in creatine phosphate, 12 times higher in creatine, 4 times higher in creatinine, 5 times higher in

citrate, 6 times higher in pyruvate, and 10 times higher in succinate compared to Se+ milk, while these metabolites were similar in concentration between milk from Se-Le+ and Se+ mothers.

Milk from Se-Le- mothers also differed with respect to free amino acid concentrations compared to milk from Se+ and Se-Le+ mothers. Arginine, asparagine, glycine, leucine, isoleucine, lysine, and tyrosine were 2-4 fold higher in milk samples from Se-Le- compared to Se+. Interestingly, the fold difference of these amino acids in milk between Se-Le+ and Se+ samples was less than 2-fold. Carnitine was higher in milk from both the Se-Le- (~4 fold) and Se-Le+ (~2 fold) groups compared to Se+. Alanine, glutamate, glutamine, taurine, and betaine were all between 2- and 5-fold lower in the Se-Le- group compared to the Se+ group, while they were similar in concentration between the Se-Le+ and Se+ groups. Aspartate was also 2-fold lower in milk from both Se-Le- and Se-Le+ samples compared to Se+.

Free fatty acids and associated metabolites such as acetate, choline, and sn-glycero-3-phosphocholine were 12, 2 and 2-fold higher respectively in the Se-Le- group compared to samples from Se+ mothers. Azelate, butyrate, caprate, and caprylate were also 8, 10, 6 and 9-fold higher respectively in the Se-Le+ group compared to Se+ samples. Additionally, butyrate was 3 and 10-fold higher in Se-Le- and Se-Le+ groups compared to Se+. O-phosphocholine was lower in the Se-Le- group (10-fold) compared to Se+. Representative peaks of taurine, betaine, acetone and 3'SL are shown in Figure 2.2B. Metabolite concentrations for each subject are shown in Table 2.2.

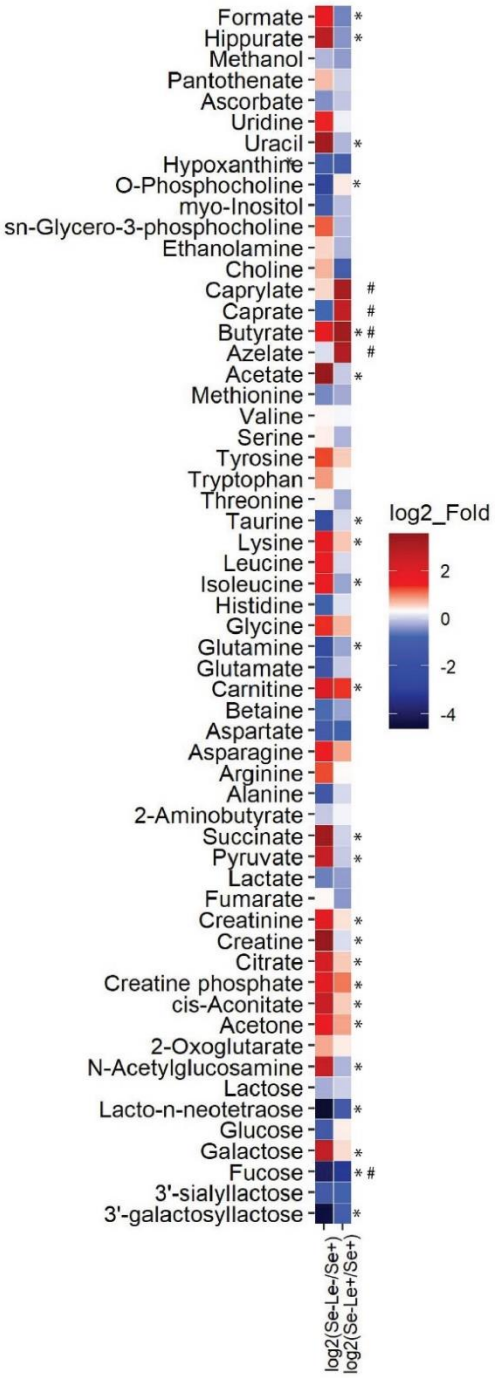


Figure 2.3. Fold difference of metabolite concentrations in milk from Se-Le- and Se-Le+ mothers relative to Se+. The mean concentration of each metabolite was calculated for all groups and the means of the Se-Le- and Se-Le+ groups were divided by the mean of the metabolite concentration from the Se+ groups to determine the ratio relative to the Se+ groups. The ratio values were then log2 transformed. Log2_fold values over 1.5 or below -1.5 are indicated in the figure. * log2_fold over 1.5 or below -1.5 when comparing Se-Le- to Se+ samples. # log2_fold over 1.5 or below -1.5 when comparing Se-Le+ to Se+ samples.

Table 2.2. Milk metabolites concentrations of the subjects.

Sample ID	1	5	8	3	4	6	7	9	2	10
HMO phenotypes	Se-Le-			Se-Le+				Se+Le-	Se+Le+	
Metabolites (uM)										
2'fucosyllactose	0.0	0.0	0.0	0.0	0.0	0.0	0.0	7259.1	5766.1	3392.7
2-Aminobutyrate	8.7	10.7	8.0	6.7	8.6	10.7	17.9	13.2	3.7	11.1
2-Oxoglutarate	77.3	79.4	112.1	77.5	74.9	69.2	57.8	68.6	41.9	45.4
3-fucosyl.lactose	0.0	0.0	0.0	7877.3	7510.0	4687.8	5293.1	0.0	602.3	3377.6
3'galactosyllactose	7.7	4.9	25.5	187.4	168.9	158.7	46.6	27.4	786.9	17.8
3'sialyllactose	117.0	124.5	126.1	175.0	183.7	114.5	126.5	242.0	288.1	286.3
Acetate	360.0	370.9	80.5	14.1	20.7	31.6	25.5	23.1	28.9	18.2
Acetone	22.6	24.0	16.6	12.9	11.3	15.7	10.1	9.2	5.8	6.2
Alanine	72.8	74.1	56.9	122.4	161.0	139.6	280.8	220.9	46.5	235.0
Arginine	90.2	92.8	30.5	53.5	36.7	28.9	32.4	48.3	29.9	11.3
Ascorbate	91.6	91.6	62.8	94.4	159.4	81.7	88.3	140.1	122.9	67.6
Asparagine	29.5	29.5	31.1	5.7	9.1	38.0	26.5	11.9	6.8	15.2
Aspartate	23.2	23.2	47.2	17.1	42.5	39.9	44.1	101.7	29.1	67.2
Azelate	12.1	12.1	24.6	28.6	108.5	172.7	161.9	10.3	9.3	25.6
Betaine	65.0	54.5	60.3	79.9	73.2	73.7	90.8	101.2	103.6	88.4
Butyrate	52.4	39.5	92.8	32.8	63.9	339.8	233.1	7.5	4.6	38.7
Caprate	8.3	8.3	16.8	69.3	74.1	97.8	224.9	9.6	24.2	22.0
Caprylate	27.5	27.5	38.8	97.6	86.6	250.3	331.6	7.4	20.6	36.7
Carnitine	58.8	54.5	30.8	24.6	23.5	29.7	42.1	19.1	5.5	11.1
Choline	550.2	568.5	315.8	179.6	173.6	118.1	84.2	221.7	415.5	230.0
Citrate	11444.8	11615.2	11628.0	3788.6	3529.9	3754.1	4038.3	2074.1	3994.9	1350.1
Creatine	555.9	678.8	824.7	98.6	86.4	34.3	33.2	46.4	77.4	52.7
Creatine phosphate	4.4	5.8	179.4	29.5	32.2	39.3	40.8	11.4	15.7	24.3
Creatinine	220.8	230.7	81.6	78.6	75.8	53.3	59.7	45.7	49.4	48.8
Ethanolamine	188.5	231.4	73.1	109.2	127.4	74.1	89.2	122.9	136.8	78.2

(continued)

Table 2.2 continued

Sample ID	1	5	8	3	4	6	7	9	2	10
HMO phenotypes		Se-Le-				Se-Le+		Se+Le-	Se+Le+	
Metabolites (uM)										
Formate	145.8	144.2	38.1	21.4	27.4	27.4	22.8	26.0	69.3	11.2
Fucose	18.2	27.9	29.4	73.8	69.7	14.0	21.7	437.7	229.1	531.2
Fumarate	1.8	1.2	22.4	4.3	5.2	3.7	7.1	8.5	5.2	6.1
Galactose	592.4	482.8	204.6	160.3	52.0	106.2	80.0	11.6	170.7	27.3
Glucose	493.4	275.3	709.5	951.3	1233.8	1494.1	2289.9	1447.0	386.0	1526.5
Glutamate	305.0	351.0	473.9	902.8	1022.0	996.1	1177.8	1736.1	199.6	1193.0
Glutamine	60.5	87.7	24.7	110.6	167.5	204.0	436.6	430.1	27.4	389.2
Glycine	590.5	590.5	174.9	112.5	266.6	403.4	383.2	180.3	228.0	123.1
Hippurate	86.0	88.5	121.5	16.1	9.4	7.1	11.9	7.1	11.2	25.5
Histidine	6.9	7.5	4.7	11.0	7.2	20.6	12.1	23.5	5.6	6.1
Hypoxanthine	2.7	3.7	6.2	2.6	4.2	5.8	4.2	6.7	16.6	2.9
Isoleucine	24.2	26.5	8.8	9.0	7.2	2.5	4.0	11.2	3.1	6.6
Lacto-N-fucopentaose II	3.4	1.0	1.2	363.5	306.2	388.7	563.9	1.9	324.1	175.4
Lactate	180.6	191.5	130.5	235.3	198.9	128.6	207.7	186.5	356.6	189.0
Lacto-N-fucopentaose I	0.0	0.0	0.0	0.0	0.0	0.0	0.0	503.8	1810.8	64.2
Lacto-N-fucopentaose III	14.0	12.0	0.0	243.5	199.2	179.3	172.7	0.0	35.0	106.4
Lacto-N-neotetraose	28.7	15.2	30.4	335.3	352.8	288.4	87.7	424.4	1368.3	46.7
Lacto-N-tetraose	16.5	32.2	6.7	671.9	725.6	719.0	604.5	332.5	3215.8	492.5
Lactodifucotetraose	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	445.9	823.9
Lactose	177325.2	173715.2	191447.1	201532.0	204280.0	214368.3	222308.8	215690.1	189847.9	225892.4
Leucine	59.3	63.7	17.1	20.7	19.3	16.2	14.4	21.5	9.3	19.7
Lysine	57.6	58.2	12.3	21.5	23.5	17.8	15.5	16.3	13.2	8.6
Methanol	86.1	84.0	97.0	82.5	71.8	75.3	80.6	84.3	86.8	125.5
Methionine	17.9	18.7	4.3	15.0	13.9	17.4	17.9	22.0	14.2	21.1
N-Acetylglucosamine	784.9	899.2	779.0	83.5	89.9	166.9	168.2	45.0	338.6	46.8
O-Phosphocholine	37.8	41.4	137.0	518.0	575.0	797.8	834.0	349.8	543.3	614.0

(continued)

Table 2.2 continued

Sample ID	1	5	8	3	4	6	7	9	2	10
HMO phenotypes		Se-Le-			Se-Le+			Se+Le-	Se+Le+	
Metabolites (uM)										
Pantothenate	22.5	16.4	21.7	13.2	14.7	10.5	12.6	13.6	5.6	18.6
Pyruvate	93.9	92.7	51.9	15.3	11.7	10.1	15.0	8.3	21.4	10.3
Serine	903.2	1143.4	841.8	524.7	720.2	344.3	940.1	487.0	993.8	684.3
Succinate	160.7	165.5	30.5	12.1	11.6	13.7	12.0	9.6	15.8	11.0
Taurine	123.0	124.8	89.4	620.7	632.2	403.3	339.5	535.4	382.1	513.5
Threonine	28.9	17.9	8.5	8.4	18.2	10.9	10.5	11.6	13.9	17.1
Tryptophan	27.4	21.1	2.8	9.0	10.7	12.3	13.4	10.6	8.2	9.4
Tyrosine	39.6	41.0	8.3	24.1	24.6	12.4	13.4	14.7	8.2	14.0
Uracil	69.9	70.6	22.9	4.8	0.0	15.7	0.0	7.0	10.3	0.0
Uridine	6.5	2.0	63.9	8.2	6.7	23.9	4.4	15.0	7.5	5.4
Valine	35.4	38.0	19.9	32.1	31.2	27.7	23.3	33.3	7.5	31.3
cis-Aconitate	60.8	66.5	19.1	9.6	12.7	17.2	12.4	10.6	5.8	9.2
myo-Inositol	391.8	400.5	231.1	692.6	566.9	786.8	749.6	334.8	1366.7	554.0
sn-Glycero-3-phosphocholine	1257.2	1274.1	1476.9	463.4	461.8	621.3	621.0	803.8	475.7	503.1

Discussion

Significant research has been undertaken to understand the impact of maternal secretor status and Lewis blood type on the milk glycome (Kunz et al. 2017) and subsequent influence on infant health including their gut microbiota (Lewis et al. 2015), susceptibility to rotavirus (Nordgren et al. 2014), allergy to bovine milk (Seppo et al. 2017), and weight during the first 6 months (Berger et al. 2019). But no research studies have focused on the metabolic profile of milk from the Se-Le- population due to its low prevalence.

In the current study, we found that in Se-Le- samples, all of the 6 fucosylated HMOs used to determine HMO phenotypes were below the detection limit of the instrument. 3FL and LNFP III were previously reported to be present in milk from Le- women, which could potentially be due to the activity of FUT4, 5, 6, 7, and 9 enzymes (Thurl et al. 2010). In the current study, neither of these HMOs was detected in Le- women, which may be due to the difference in detection methods. Mass Spectrometry can measure down to the picomolar level, whereas for spectra obtained from the Varian 500 MHz spectrometer used in this study, the limit of detection of these metabolites was 20 μM. It remains to be determined if oligosaccharides present in milk below 20 μM would have a significant impact on infant health.

Other HMOs and related metabolites such as 3'galactosyllactose, 3'SL, fucose, galactose, and GlcNAc also trended different within the HMO phenotypes, indicating other factors influencing the glycome of human milk (Figure 2.3). 3'SL (Figure 2.2B), lower in both Se-Le- and Se-Le+ groups compared to Se+ in this study, was reported to be similar in concentration in milk from Se+ and Se- women (Smilowitz et al. 2013; Totten et al. 2012) or even ~20-56% higher in milk from Se- compared to Se+ women (Paganini et al. 2019; McJarrow et al. 2019). 6'-sialyllactose, which

is not reported in the current study, was demonstrated to be significantly higher in milk from non-secretor women (Totten et al. 2012). Further studies are needed to investigate if this is due to a preference of α-2, 6-sialylation / α-2, 3-sialylation or simply a difference amongst populations.

Pyruvate, citrate, cis-aconitate, and succinate, which are metabolites involved in the tricarboxylic acid (TCA) cycle, were higher in Se-Le- milk. Increased TCA cycling could indicate greater energy provision, and previous studies have speculated that a higher level of TCA intermediates in bovine milk compared to human milk may be to enhance growth (Qian et al. 2016; Scano et al. 2016).

HMO biosynthesis was suggested to be an extension of lactose biosynthesis which occurs in the Golgi of the mammary gland epithelial cells (Rudloff et al. 2006). Therefore, inactivity of both α-1,2- and α-1,3/1,4- fucosylltransferases in Se-Le- women might profoundly impact mammary gland metabolism, and thus impact milk composition. Alanine, taurine, glutamine and glutamate are the most abundant free amino acids in human milk (Smilowitz et al. 2013; Zhang et al. 2013), and these were all lower in the milk from the Se-Le- group compared to milk from the Se+ group. Higher free glutamate in bovine milk infant formula has been reported to decrease its intake (Ventura, Beauchamp, and Mennella 2012). It could be that a lower level of glutamate in milk from Se-Le- women could increase milk intake by the infant to compensate for the low and less diverse HMO content. Branched chain amino acids (leucine and isoleucine) and lysine were higher in Se-Le- compared to Se+ milk. A similar pattern of free amino acids in human milk was seen in a previous study comparing high and low growth rate groups of premature infants (Alexandre-Gouabau et al. 2019), where a higher content of insulinotrophic amino acids and tyrosine was associated with faster infant growth.

Choline in the Se-Le- group was almost double the level in Se+ samples, while phosphocholine was one-tenth the level. A previous study showed a negative correlation between choline and phosphocholine in human milk (Moukarzel et al. 2017), and a similar correlation was observed in this study. The origin of choline in milk is not completely understood. One study reported that breast milk choline is related to maternal choline intake and genetic polymorphisms (Fischer et al. 2010), while another study showed no difference in milk choline content based on maternal diet (Perrin et al. 2020). Indeed, the betaine level in Se-Le- milk was 2-fold lower than that in Se+ samples, suggesting a possible lower conversion of choline to betaine. It could be that the difference in milk choline (and other metabolites) in the Se-Le- group compared to Se+ group could result in differences in milk lipid synthesis (Rudolph et al. 2007). Indeed, choline is an essential precursor of phosphatidylcholine and sphingomyelin, which are essential components of biological membranes and precursors for intracellular messengers such as ceramide and diacylglycerol (Blusztajn 1998). This would imply that the milk fat globule would be different in Se-Le- mothers since maternal phenotype will impact conjugated glycolipids in addition to HMOs (Morrow et al. 2005). Differences in the milk fat would need to be assessed in a separate study.

Here we showed differences in non-HMO metabolites between phenotypically Se-Le- mothers and Se-Le+, and Se+ mothers. These differences included metabolites related to energy metabolism, amino acids, and fatty acids. The current study is limited by the small sample size and the rarity of Se-Le- HMO phenotype. Factors such as the completeness of milk expression, time since last feed, time of the day during sample collection, and information on mother's diet were not collected; however, the impact of these factors on milk composition is negligible compared to the impact of genetics. Nonetheless, this study shows that the Se and Le status of the mother has an important role to play in the composition of non-oligosaccharide milk metabolites. Further research involving

larger sample sizes should be done to confirm the findings, investigate the impact on milk lipid and proteins, and investigate potential biological consequences of Se-Le- milk on infant gut microbial succession and metabolism. This will help further unravel the link between human milk and infant health.

References

Alexandre-Gouabau, Marie-Cécile, Thomas Moyon, Agnès David-Sochard, François Fenaille, Sophie Cholet, Anne-Lise Royer, Yann Guitton, et al. 2019. "Comprehensive Preterm Breast Milk Metabotype Associated with Optimal Infant Early Growth Pattern." *Nutrients* 11 (3): 528. doi: 10.3390/nu11030528.

Barnicot, N. A., and Sylvia D. Lawler. 1953. "A Study of the Lewis, Kell, Lutheran and P Blood Group Systems and the ABH Secretion in West African Negroes." *American Journal of Physical Anthropology* 11 (1): 83–90.

Berger, Paige, Jasmine Plows, Roshonda Jones, Tanya Alderete, Chloe Yonemitsu, Ji Hoon Ryoo, Lars Bode, and Michael Goran. 2019. *Human Milk Oligosaccharides and Infant Weight in the First 6 Months of Life (P11-053-19)*. Oxford University Press.

Blusztajn, Jan Krzysztof. 1998. "Choline, a Vital Amine." *Science* 281 (5378): 794. doi: 10.1126/science.281.5378.794.

Dessì, Angelica, Despina Briana, Sara Corbu, Stavroula Gavrili, Flaminia Cesare Marincola, Sofia Georgantzi, Roberta Pintus, Vassilios Fanos, and Ariadne Malamitsi-Puchner. 2018. "Metabolomics of Breast Milk: The Importance of Phenotypes." *Metabolites* 8 (4). doi: 10.3390/metabo8040079.

Edison, Erica E., Margaret E. Brosnan, Khalid Aziz, and John T. Brosnan. 2013. "Creatine and Guanidinoacetate Content of Human Milk and Infant Formulas: Implications for Creatine Deficiency Syndromes and Amino Acid Metabolism." *British Journal of Nutrition* 110 (6): 1075–78. doi: 10.1017/S000711451300010X.

Elwakiel, M., J. A. Hageman, W. Wang, I. M. Szeto, J. B. van Goudoever, K. A. Hettinga, and H. A. Schols. 2018. "Human Milk Oligosaccharides in Colostrum and Mature Milk of Chinese Mothers: Lewis Positive Secretor Subgroups." *Journal of Agricultural and Food Chemistry* 66 (27): 7036–43. doi: 10.1021/acs.jafc.8b02021.

Fischer, Leslie M, Kerry Ann da Costa, Joseph Galanko, Wei Sha, Brigitte Stephenson, Julie Vick, and Steven H Zeisel. 2010. "Choline Intake and Genetic Polymorphisms Influence Choline Metabolite Concentrations in Human Breast Milk and Plasma." *The American Journal of Clinical Nutrition* 92 (2): 336–46. doi: 10.3945/ajcn.2010.29459.

Gay, Melvin, Petya Koleva, Carolyn Slupsky, Elloise Toit, Merete Eggesbo, Christine Johnson, Ganesa Wegienka, Naoki Shimojo, Dianne Campbell, and Susan Prescott. 2018. "Worldwide Variation in Human Milk Metabolome: Indicators of Breast Physiology and Maternal Lifestyle?" *Nutrients* 10 (9): 1151.

Gregory, Katherine E, and W Allan Walker. 2013. "Immunologic Factors in Human Milk and Disease Prevention in the Preterm Infant." *Current Pediatrics Reports* 1 (4): 10.1007/s40124-013-0028–2. doi: 10.1007/s40124-013-0028-2.

Koda, Yoshiro, Hidenori Tachida, Hao Pang, Yuhua Liu, Mikiko Soejima, Abbas A. Ghaderi, Osamu Takenaka, and Hiroshi Kimura. 2001. "Contrasting Patterns of Polymorphisms at the ABO-Secretor Gene (FUT2) and Plasma α(1,3)Fucosyltransferase Gene (FUT6) in Human Populations." *Genetics* 158 (2): 747.

Kunz, C., S. Rudloff, W. Baier, N. Klein, and S. Strobel. 2000. "Oligosaccharides in Human Milk: Structural, Functional, and Metabolic Aspects." *Annual Review of Nutrition* 20 (1): 699–722. doi: 10.1146/annurev.nutr.20.1.699.

Kunz, Clemens, Christina Meyer, Maria Carmen Collado, Lena Geiger, Izaskun García-Mantrana, Bibiana Bertua-Ríos, Cecilia Martínez-Costa, Christian Borsch, and Silvia Rudloff. 2017a. "Influence of Gestational Age, Secretor, and Lewis Blood Group Status on the Oligosaccharide Content of Human Milk." *Journal of Pediatric Gastroenterology and Nutrition* 64 (5): 789–98.

Lewis, Zachery T., Sarah M. Totten, Jennifer T. Smilowitz, Mina Popovic, Evan Parker, Danielle G. Lemay, Maxwell L. Van Tassell, et al. 2015. "Maternal Fucosyltransferase 2 Status Affects the Gut

Bifidobacterial Communities of Breastfed Infants." *Microbiome* 3 (1): 13. doi: 10.1186/s40168-015-0071-z.

McJarrow, Paul, Hadia Radwan, Lin Ma, Alastair K.H. MacGibbon, Mona Hashim, Hayder Hasan, Reyad S. Obaid, et al. 2019. "Human Milk Oligosaccharide, Phospholipid, and Ganglioside Concentrations in Breast Milk from United Arab Emirates Mothers: Results from the MISC Cohort." *Nutrients* 11 (10). doi: 10.3390/nu11102400.

Morrow, Ardythe L., David S. Newburg, Guillermo M. Ruiz-Palacios, Jareen Meinzen-Derr, Maria de Lourdes Guerrero, Mekibib Altaye, and Prasoon Chaturvedi. 2003. "Innate Protection Conferred by Fucosylated Oligosaccharides of Human Milk against Diarrhea in Breastfed Infants." *Glycobiology* 14 (3): 253–63. doi: 10.1093/glycob/cwh020.

Morrow, Ardythe L., Guillermo M. Ruiz-Palacios, Xi Jiang, and David S. Newburg. 2005. "Human-Milk Glycans That Inhibit Pathogen Binding Protect Breast-Feeding Infants against Infectious Diarrhea." *The Journal of Nutrition* 135 (5): 1304–7. doi: 10.1093/jn/135.5.1304.

Moukarzel, Sara, Lynda Soberanes, Roger A. Dyer, Susan Albersheim, Rajavel Elango, and Sheila M. Innis. 2017. "Relationships among Different Water-Soluble Choline Compounds Differ between Human Preterm and Donor Milk." *Nutrients* 9 (4): 369.

Newburg, David S., Guillermo M. Ruiz-Palacios, Mekibib Altaye, Prasoon Chaturvedi, Jareen Meinzen-Derr, Maria de Lourdes Guerrero, and Ardythe L. Morrow. 2004. "Innate Protection Conferred by Fucosylated Oligosaccharides of Human Milk against Diarrhea in Breastfed Infants." *Glycobiology* 14 (3): 253–63. doi: 10.1093/glycob/cwh020.

Newburg, David S., Guillermo M. Ruiz-Palacios, and Ardythe L. Morrow. 2005. "HUMAN MILK GLYCANS PROTECT INFANTS AGAINST ENTERIC PATHOGENS." *Annual Review of Nutrition* 25 (1): 37–58. doi: 10.1146/annurev.nutr.25.050304.092553.

Ninonuevo, Milady R., Youmie Park, Hongfeng Yin, Jinhua Zhang, Robert E. Ward, Brian H. Clowers, J. Bruce German, et al. 2006. "A Strategy for Annotating the Human Milk Glycome." *Journal of Agricultural and Food Chemistry* 54 (20): 7471–80. doi: 10.1021/jf0615810.

Nordgren, Johan, Sumit Sharma, Filemon Bucardo, Waqas Nasir, Gökçe Günaydın, Djeneba Ouermi, Leon W. Nitiema, Sylvia Becker-Dreps, Jacques Simpore, and Lennart Hammarström. 2014. "Both Lewis and Secretor Status Mediate Susceptibility to Rotavirus Infections in a Rotavirus Genotype–Dependent Manner." *Clinical Infectious Diseases* 59 (11): 1567–1573.

Oriol, R., J. Le Pendu, and R. Mollicone. 1986. "Genetics of ABO, H, Lewis, X and Related Antigens." *Vox Sanguinis* 51 (3): 161–71.

Paganini, Daniela, Mary A. Uyoga, Guus A.M. Kortman, Jos Boekhorst, Sacha Schneeberger, Simon Karanja, Thierry Hennet, and Michael B. Zimmermann. 2019. "Maternal Human Milk Oligosaccharide Profile Modulates the Impact of an Intervention with Iron and Galacto-Oligosaccharides in Kenyan Infants." *Nutrients* 11 (11). doi: 10.3390/nu11112596.

Perrin, Maryanne T, Roman Pawlak, Lindsay H Allen, and Daniela Hampel. 2020. "Total Water-Soluble Choline Concentration Does Not Differ in Milk from Vegan, Vegetarian, and Nonvegetarian Lactating Women." *The Journal of Nutrition* 150 (3): 512–17. doi: 10.1093/jn/nxz257.

Praticò, Giulia, Giorgio Capuani, Alberta Tomassini, Maria Elisabetta Baldassarre, Maurizio Delfini, and Alfredo Miccheli. 2014. "Exploring Human Breast Milk Composition by NMR-Based Metabolomics." *Natural Product Research* 28 (2): 95–101. doi: 10.1080/14786419.2013.843180.

Qian, Linxi, Aihua Zhao, Yinan Zhang, Tianlu Chen, Steven H Zeisel, Wei Jia, and Wei Cai. 2016. "Metabolomic Approaches to Explore Chemical Diversity of Human Breast-Milk, Formula Milk and Bovine Milk." *International Journal of Molecular Sciences* 17 (12): 2128. doi: 10.3390/ijms17122128.

Rudloff, Silvia, Susanne Obermeier, Christian Borsch, Gottfried Pohlentz, Rudolf Hartmann, Herbert Brösicke, Michael J. Lentze, and Clemens Kunz. 2006. "Incorporation of Orally Applied (13)C-Galactose into Milk Lactose and Oligosaccharides." *Glycobiology* 16 (6): 477–87. doi: 10.1093/glycob/cwj092.

Rudolph, Michael C., James L. McManaman, TzuLip Phang, Tanya Russell, Douglas J. Kominsky, Natalie J. Serkova, Torsten Stein, Steven M. Anderson, and Margaret C. Neville. 2007. "Metabolic Regulation in the Lactating Mammary Gland: A Lipid Synthesizing Machine." *Physiological Genomics* 28 (3): 323–36. doi: 10.1152/physiolgenomics.00020.2006.

Scano, Paola, Antonio Murgia, Martina Demuru, Roberto Consonni, and Pierluigi Caboni. 2016. "Metabolite Profiles of Formula Milk Compared to Breast Milk." *Food Research International (Ottawa, Ont.)* 87 (September): 76–82. doi: 10.1016/j.foodres.2016.06.024.

Seppo, Antti E., Chloe A. Autran, Lars Bode, and Kirsi M. Järvinen. 2017. "Human Milk Oligosaccharides and Development of Cow's Milk Allergy in Infants." *Journal of Allergy and Clinical Immunology* 139 (2): 708-711. e5.

Smilowitz, Jennifer T., Aifric O'sullivan, Daniela Barile, J. Bruce German, Bo Lönnerdal, and Carolyn M. Slupsky. 2013. "The Human Milk Metabolome Reveals Diverse Oligosaccharide Profiles." *The Journal of Nutrition* 143 (11): 1709–1718.

Spevacek, Ann R., Jennifer T. Smilowitz, Elizabeth L. Chin, Mark A. Underwood, J. Bruce German, and Carolyn M. Slupsky. 2015. "Infant Maturity at Birth Reveals Minor Differences in the Maternal Milk Metabolome in the First Month of Lactation." *The Journal of Nutrition* 145 (8): 1698–1708.

Stahl, B., S. Thurl, J. Henker, M. Siegel, B. Finke, and G. Sawatzki. 2001. "Detection of Four Human Milk Groups with Respect to Lewis-Blood-Group-Dependent Oligosaccharides by Serologic and Chromatographic Analysis." *Advances in Experimental Medicine and Biology* 501: 299–306. doi: 10.1007/978-1-4615-1371-1_37.

Stinson, Lisa F., Melvin C. L. Gay, Petya T. Koleva, Merete Eggesbø, Christine C. Johnson, Ganesa Wegienka, Elloise du Toit, et al. 2020. "Human Milk From Atopic Mothers Has Lower Levels of Short Chain Fatty Acids." *Frontiers in Immunology* 11: 1427. doi: 10.3389/fimmu.2020.01427.

Ten-Doménech, Isabel, Victoria Ramos-Garcia, José David Piñeiro-Ramos, María Gormaz, Anna Parra-Llorca, Máximo Vento, Julia Kuligowski, and Guillermo Quintás. 2020. "Current Practice in Untargeted Human Milk Metabolomics." *Metabolites* 10 (2): 43. doi: 10.3390/metabo10020043.

Thurl, Stephan, Manfred Munzert, Jobst Henker, Günther Boehm, Beate Müller-Werner, Jürgen Jelinek, and Bernd Stahl. 2010a. "Variation of Human Milk Oligosaccharides in Relation to Milk Groups and Lactational Periods." *British Journal of Nutrition* 104 (9): 1261–1271.

Tochitani, Shiro. 2017. "Functions of Maternally-Derived Taurine in Fetal and Neonatal Brain Development." In *Taurine 10*, edited by Dong-Hee Lee, Stephen W. Schaffer, Eunkyue Park, and Ha Won Kim, 17–25. Dordrecht: Springer Netherlands.

Totten, Sarah M., Angela M. Zivkovic, Shuai Wu, UyenThao Ngyuen, Samara L. Freeman, L. Renee Ruhaak, Momodou K. Darboe, J. Bruce German, Andrew M. Prentice, and Carlito B. Lebrilla. 2012. "Comprehensive Profiles of Human Milk Oligosaccharides Yield Highly Sensitive and Specific Markers for Determining Secretor Status in Lactating Mothers." *Journal of Proteome Research* 11 (12): 6124–6133.

Ventura, Alison K, Gary K Beauchamp, and Julie A Mennella. 2012. "Infant Regulation of Intake: The Effect of Free Glutamate Content in Infant Formulas." *The American Journal of Clinical Nutrition* 95 (4): 875–81. doi: 10.3945/ajcn.111.024919.

Zhang, Zhiying, Alicia S Adelman, Deshanie Rai, Julia Boettcher, and Bo Lőnnerdal. 2013. "Amino Acid Profiles in Term and Preterm Human Milk through Lactation: A Systematic Review." *Nutrients* 5 (12): 4800–4821. doi: 10.3390/nu5124800.

Chapter 3 Impact of milk secretor status on fecal metabolome and microbiota of breastfed infants

Abstract

Maternal secretor status has been shown to be associated with the presence of specific fucosylated human milk oligosaccharides (HMOs), and the impact of maternal secretor status on infant gut microbiota measured through 16S rRNA gene sequencing has previously been reported. None of those studies have confirmed exclusive breastfeeding, nor investigated the impact of maternal secretor status on gut microbial fermentation products. The present study focused on exclusively breastfed (EBF) infants, with exclusive breastfeeding validated through the dose-to-mother (DTM) technique, and the impact of maternal secretor status on the infant fecal microbiome and metabolome. Maternal secretor status did not alter the within-community (alpha) diversity, between-community (beta) diversity, or the relative abundance of bacterial taxa at the genus level. However, infants fed milk from secretor (Se+) mothers exhibited a lower level of fecal succinate, amino acids and their derivatives, and a higher level of 1,2-propanediol when compared to infants fed milk from non-secretor (Se-) mothers. Interestingly, for infants consuming milk from Se+ mothers, there was a correlation between the relative abundance of *Bifidobacterium* and *Streptococcus*, and between each of these genera and fecal metabolites that was not observed in infants receiving milk from Se- mothers. Our findings indicate that secretor status of the mother impacts gut microbial metabolism.

Introduction

Although the total concentration of oligosaccharides in human milk has low biological variability, on any given day the concentrations of individual human milk oligosaccharides (HMOs) exhibit high inter-individual variation (Smilowitz et al. 2013; Spevacek et al. 2015). Variation in the concentration of individual oligosaccharides is driven by maternal genetics, stage of lactation, as well as other unknown factors (Azad et al. 2018). The secretor (Se) gene, *fut2*, codes for α-1,2-fucosyltransferase 2 (FUT2), which is responsible for producing HMOs such as 2'fucosyllactose (2'FL), lacto-N-fucopentose I (LNFP I), and lactodifucotetraose (LDFT) (Thurl et al. 2010). For women with a functional FUT2 enzyme (Se+), 2'FL is the most abundant HMO in their breast milk, whereas for women with a non-functional FUT2 enzyme (Se-), 2'FL has been shown to be below detection limits (Totten et al. 2012; Smilowitz et al. 2013). Some studies have reported that for mothers with Se- status, α-1,2-fucosyltransferase 1 (FUT1) can synthesize 2'FL, albeit at low levels (Newburg, Ruiz-Palacios, and Morrow 2005). Nonetheless, breast milk from Se- mothers has been reported to have lower concentrations of fucosylated HMOs, total HMOs, and higher levels of non-fucosylated neutral HMOs when compared to breast milk from Se+ women (Totten et al. 2012; Kunz et al. 2017; Smilowitz et al. 2013)

The differences in HMO profiles between women with a Se+ and Se- phenotype poses the question as to whether there may be selective advantages for the infant. As HMOs are important substrates for microbial fermentation, the impact of secretor status on infant gut microbial composition has been the focus of most studies. One such study reported that in premature infants fed Se+ milk, a trend toward lower levels of *Proteobacteria* and higher levels of *Firmicutes* was observed in the fecal microbiome compared to premature infants fed Se- milk (Underwood et al. 2015). Another study showed that Se+ milk consumption was associated with a higher abundance and faster

colonization of *Bifidobacterium* in the gut of term breastfed infants (Lewis et al. 2015). The impact of maternal secretor status on gut microbial composition was reported to persist to the age of 2-3 years, when children at this age were reported to have higher *Bifidobacterium* and lower *Bacteroides* (Smith-Brown et al. 2016). However, these associations have not been observed in all studies. One study reported no impact of maternal secretor status on infant gut microbiota when partially or exclusively breastfed (EBF) infants were born vaginally, but a shift in microbial composition was observed in caesarean-born infants (Korpela et al. 2018). Another study reported a negative correlation between gut milk 2'FL and *Bifidobacterium* in EBF infants (Wang et al. 2015).

Although studies have investigated the impact of maternal secretor status on infant microbial composition, few have reported whether it alters the gut microbial fermentation capability. It is of great interest to understand if this difference in the milk metabolome based on maternal secretor status has an impact on production of microbial fermentation products. While previous reports investigating the impact of maternal secretor status on infant gut microbiome (Lewis et al. 2015; Smith-Brown et al. 2016; Korpela et al. 2018; Wang et al. 2015) have taken into account infant feeding practices, determination of EBF status was solely based on mother's self-reporting and was not confirmed through objective techniques, such as the dose-to-mother (DTM) deuterium-oxide method. A previous study investigating the EBF rate of Guatemalan mothers showed that while the self-reported EBF rates were 50% (report of current feeding practice) and 61% (by 24-hour recall), the EBF rate measured via the DTM method was only 36% (Medoua et al. 2012). A separate study reported that based on the DTM technique, 75% of self-reported EBF infants were fed food other than breast milk (Mazariegos, Slater, and Ramirez-Zea 2016). It is important to confirm EBF status with an objective technique such as the DTM method to reduce the

confounding impact of food and water intake on infant gut microbiome. The present study specifically targeted EBF infants whose EBF status were confirmed by the DTM method in a less-developed area in West Java, Indonesia and evaluated the influence of maternal secretor status on the infant gut microbial composition and fecal metabolome.

Materials and methods

The current study included a total of 160 mother–infant pairs with infant post-natal age from 2 to 5.5 months from the Sumedang district in the province of West Java, Indonesia. The inclusion criteria for infants included gestational age and health status, and was described in a previous study (Liu et al. 2019). Ethical approval for the study was granted by University of Otago Human Research Ethics Committee New Zealand (H15/125) and the Health Research Ethics Committee Faculty of Medicine Universitas Padjadjaran, Bandung (081), Indonesia.

Breastfeeding status determination

The determination of EBF status was based on the dose-to-mother (DTM) method described previously (Liu et al. 2019; Leong et al. 2021). Briefly, mothers were provided a dose of deuterium oxide, and saliva samples were collected from mother-infant pairs over the course of 14 days. Using these data, infant daily water intake from sources other than breast milk was calculated and then compared to the cut off value of 86.6 g/d (Liu et al. 2019) to identify EBF status.

Breast milk sample collection and metabolite extraction

Full expression of morning milk samples was collected using breast pumps (Harmony, Medela, Baar, Switzerland) after instructing the mothers to take strict precautions to avoid all sources of contamination. After gentle mixing and aliquoting, milk samples were stored at -80 °C until analysis.

For metabolomics analysis, milk samples were thawed on ice, vortexed, and centrifuged at 12k rcf, at 4 ℃ for 5 min to separate milk lipids. A total of 350 μL of the aqueous phase was transferred to a pre-washed 3 kDa Amicon filter (Amicon ultra centrifugal filter, Millipore, Billerica, MA) to remove lipids and proteins. To 207 μL of filtrate, 23 μL of internal standard (5 mM 3-(trimethylsilyl)-1-propanesulfonic acid-d6 (DSS-d6) in 99.8% D_2O (to serve as a lock) and 0.2% NaN_3 (to inhibit bacterial growth)) was added. To minimize pH-based peak movement in the NMR spectra, the pH of each sample was adjusted to 6.85 ± 0.07 by adding small amounts of NaOH or HCl. A total of 180 μL of the mixture was transferred to 3mm Bruker NMR tubes (Bruker, Brillerica, MA) and stored at 4 ℃ until spectral acquisition.

Fecal sample collection and metabolite extraction

Fresh fecal samples were collected into Eppendorf tubes directly from the nappy and stored at -80 °C until analysis. To prepare for microbiome and metabolome analysis, fecal samples were thawed on ice and fecal metabolites were extracted as preciously described (He et al. 2019a). Briefly, 250 mg of fecal material was combined with 1.5 mL of ice-cold Dulbecco's phosphate buffered saline (DPBS, 1X, pH 7.4) for metabolites extraction. After vortexing and centrifugation (14 kcrf, 4 ℃, for 5 min), the supernatant was filtered through a syringe filter (0.22 μm pore size, Millex-GP syringe filter, Millipore, Billerica, MA) followed by an Amicon filter (3kDa). Samples were prepared as above for NMR analysis. The pellet was collected and saved at 4 ℃ for DNA extraction (described below).

To estimate the water content of each sample, approximately 75 mg of feces was weighed into a 2 mL screw-cap tube and lyophilized (Labconco FreeZone 4.5 L Freeze Dry System, Labconco,

Kansas City, MO). The weight of the tube was analytically determined before and after drying and used below for calculating the amount of water in the extracted sample.

NMR acquisition, data processing and quantification

^{1}H NMR spectra were acquired at 298K using a NOESY ^{1}H presaturation experiment ('noesypr1d') on a Bruker Avance 600 MHz NMR spectrometer (Bruker BioSpin, Germany) equipped with a SampleJet autosampler (Bruker BioSpin, Germany) as previously described (He et al. 2019b). After manual phasing and baseline correction in Chenomx processor, metabolites were quantified using Chenomx Profiler (Chenomx NMR Suite v8.3, Chenomx Inc, Edmonton, Alberta, Canada) based on the established method of targeted profiling (Weljie et al. 2006). The resulting metabolite concentrations were corrected based water content as described in our previous study (He et al. 2019a), where fecal water estimate was calculated for each sample as follows:

$$\text{Fecal water estimate} = \left(1 - \frac{Dry\ weight}{Wet\ weight}\right) \times Weight\ of\ extracted\ sample \quad (1)$$

Milk secretor status phenotype determination

Milk secretor status determination was based on the presence or near absence of 2'FL in the NMR spectra, which was identified and quantified from an NMR spectral library created through the analytical preparation of commercially available HMO standards as previously described (Smilowitz et al. 2013; Wang et al. 2021). In this study, when [2'FL] ≥ 200 μM, milk secretor status was assigned as Se+, while milk samples with [2'FL] < 200 μM were assigned as Se-.

Fecal microbial DNA extraction and library preparation

The pellet of each fecal sample from above was used for DNA extraction according to the Human Microbiome Project (HMP) protocol (Wesolowska-Andersen et al. 2014) with minor modifications (He et al. 2019a) using the MoBio PowerLyzer PowerSoil DNA isolation kit (MoBio, Carlsbad, CA). The DNA purity was determined spectrophotometrically using a NanoDrop 2000C Spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA). One negative control sample using PCR-grade water (MoBio, Carlsbad, CA) was prepared for each batch.

The V4 hypervariable region of the 16S rRNA gene was targeted using a two-step PCR protocol. In step 1, the V4 region was amplified using F515/R806 primers modified to contain an Illumina overhang sequence and a 0-5 bp spacer on the 5' end. The modified F515 forward primer sequence was: 5'-*TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG[spacer]GTGCCAGCMGCCGCG GTAA*-3', and the modified R806 reverse primer sequence was: 5'-*GTCTCGTGGGCTCGGAGAT GTGTATAAGAGACAG[spacer]GGACTACHVGGGTWTCTAAT*-3'. Before PCR, the forward and reverse primers were diluted to 10 μM. The PCR reactions were performed in 15 μL reaction volumes containing 4 μL DNA template, 0.75 μL DMSO (Fisher Scientific, Waltham, MA), 3 μL 5X KAPA HiFi Buffer (KAPA Biosystems, Woburn, MA), 0.45 μL dNTP Mix (10 mM), 0.3 μL KAPA HiFi HotStart Polymerase (KAPA Biosystems, Woburn, MA), 5.7 μL PCR-grade water (MoBio, Carlsbad, CA), and primers (0.4 μL for each). The amplified DNA products in this step were diluted 1:10 using PCR water and mixed well. In step 2, an 8 bp index was used to multiplex the samples in both the forward and reverse primers. The forward indexing primer sequence was: *AATGATACGGCGACCACCGAGATCTACACXXXXXXXXTCGTCGGCAGCGTC,* and the reverse indexing primer sequence was: *CAAGCAGAAGACGGCATACGAGATXXXXXXXXGTCT*

CGTGGGCTCGG (X indicates the positions of the 8-bp indices). The indexing primers were diluted to 5 μM before using. The PCR reactions were performed in duplicate in 20 μL reaction volumes containing 1 μL diluted DNA template from step 1, 1 μL DMSO (Fisher Scientific, Waltham, MA), 1 μL 5X KAPA HiFi Buffer (KAPA Biosystems, Woburn, MA), 0.6 μL dNTP Mix (10 mM), 0.4 μL KAPA HiFi HotStart Polymerase (KAPA Biosystems, Woburn, MA), 9 μL PCR-grade water (MoBio, Carlsbad, CA), and primers (2 μL for each). In both steps 1 and 2, the PCR reactions consisted of an initial denaturation at 95 °C for 5 min followed by a 10-cycle program of 20 s at 98 °C for denaturation, 15 s at 55 °C for annealing, 60 s at 72 °C for primer extension and a final extension of 72 °C for 10 min.

Amplified PCR products from step 2 were quality checked by gel electrophoresis. The band intensity (around 430 bp) was visualized using SYBR safe DNA stain (Invitrogen) and its quantity (in ng/μL) was estimated using a molecular ladder with known concentration (BioRad EZ ladder 1 kb) through ImageLab software (v5.2.1, BioRad, Hercules, CA). Amplicons were pooled in equimolar ratios and purified using the QIAquick PCR Purification Kit (QIAGEN, Hilden, Germany) using a modified protocol from the manufacturer's instruction (He, Parenti, Grip, Domellöf, et al. 2019). A purified amplicon library was quality checked by Bioanalyzer and submitted to the UC Davis Genome Center DNA Technologies Core for 300 bp paired-end sequencing on the Illumina MiSeq platform.

Analysis of 16S rRNA gene amplicon sequences

Sequencing reads were demultiplexed by the UC Davis Genome center DNA Technologies Core after sequencing based on the index sequences provided upon library submission. Sequence reads without a corresponding barcode and primer sequence were discarded. Spacers and the V4 region

primers were removed using the *cutadapt* function in *cutadapt* module (version 1.8.3). Reads were then split into forward (R1) and reverse (R2) reads before feeding into the *fastqc* and *multiqc* function to evaluate reading quality. DADA2 (version 1.12.1) was then used to filter, trim, merge reads and assign taxonomy. Briefly, forward and reverse reads were trimmed to 200 and 150 bases with a maxEE of 2, respectively. After error rate learning and sample inference, paired reads were merged and sequences with a length between 250 and 258 bp were kept. Taxonomy was assigned using the SILVA 16S rRNA database version 138 formatted for DADA2 (Callahan et al. 2016; Quast et al. 2013; Yilmaz et al. 2014). The ASV table and taxonomy table were then exported from the DADA2 pipeline for further analysis in R (version 4.0.3).

Statistical analysis

Statistical analysis and graphical generation were performed using the R programing environment version 4.0.3. The ASV table generated from the DADA2 pipeline was used to calculate the relative abundance of each genus and α-diversity matrices using the *phyloseq* package. Generalized log transformation (defined as $log(y + 1))$ was applied to all metabolomics data.

Principal coordinate analysis (PCoA) of microbiome data was computed using the *pcoa* package based on the Bray distance while PCoA of metabolome data was done based on Euclidean distances. The centroids of each cluster (centroid of mass) were calculated as the average of PC1 and PC2 for each group. The ellipses were constructed based on multivariate normal distribution at a 95% confidence level.

The differential analysis for the microbiome was computed at the genus level using Analysis of composition of microbiomes (ANCOM), and infant age was set as a covariate. The false discovery rate (FDR, type I error) was controlled using the Benjamini-Hochberg procedure. For

metabolomics data, the significance between secretor status was evaluated using the Mann-Whitney Test (*wilcox.test* function) and p-values from Mann-Whitney test were then adjusted to reduce FDR (*p.adjust(, method = 'fdr')*). The overall level of significance was set at $p < 0.05$.

The *cliff.delta* function from the *effsize* package was used to evaluate the effect size between the Se+ and Se- group using Cliff's delta (δ) statistics. The 95 % confidence interval of each computed Cliff's delta was further estimated. The threshold of negligible, small, and large δ were assessed according to Romano *et al* (Romano et al. 2006) where $|\delta|<0.147$ corresponds to negligible, $0.147<|\delta|<0.33$ corresponds to small, $0.33 <|\delta|<0.474$ corresponds to medium, and $|\delta|>0.474$ large.

The Spearman correlation coefficient (R) was computed using *cor*(, *method = "spearman"*) to evaluate the strength of a correlation. All plots were generated using *ggplot2*.

Results

The present study aimed to compare the gut microbiota and fecal metabolome between EBF infants receiving milk from phenotypically Se+ and Se- mothers. Only samples from those mother-infant pairs where exclusive breastfeeding was confirmed through the DTM method, and from which infant fecal and maternal breast milk samples were collected were used in this study. After data generation, a few fecal samples were excluded because of suspected sequencing errors (n=2) or suspected urine contamination (n=5, determined by the presence of urea in the NMR spectrum). In the end, data from a total of 160 mother-infant pairs were included.

Although the maternal secretor genotype was not determined, milk secretor phenotype was defined as Se+: [2'FL] ≥ 200 μM; and Se-: [2'FL] < 200 μM (Figure 3.1 a). Lewis blood type associated HMOs (3-fucosyllactose 3FL, lacto-N-fucopentose II (LNFP II) and lacto-N-fucopentose III

(LNFP III)) were also measured (Figure 3.1 d-f). As described previously (A. Wang et al. 2021), based on these data, none of the mothers were classified as phenotypically Lewis negative.

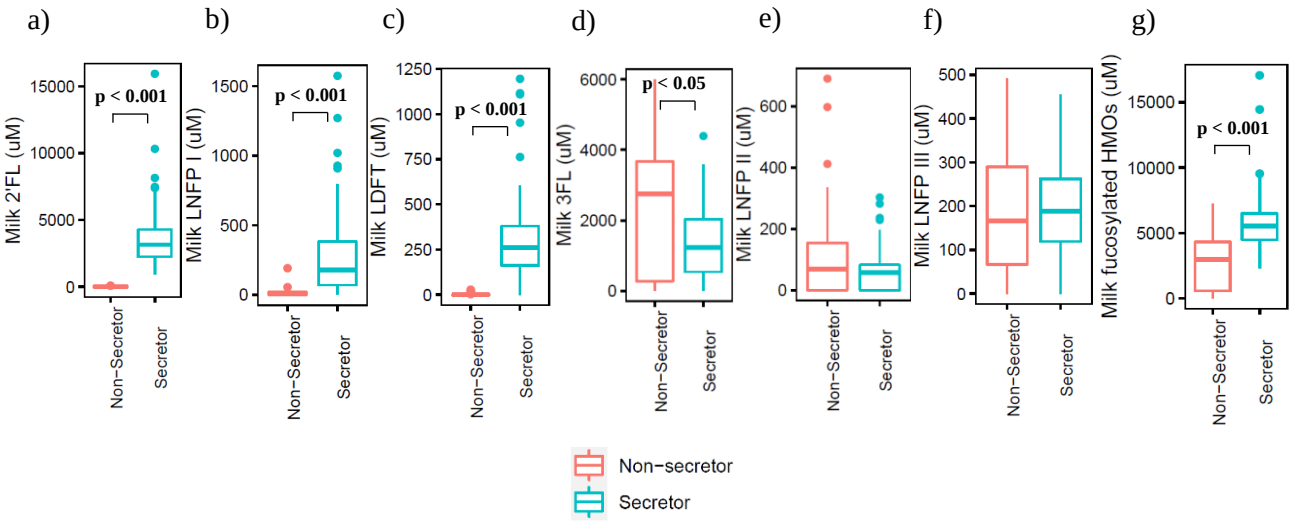


Figure 3.1. Representative fucosylated milk oligosaccharides. a) 2'fucosyllactose (2'FL), b) lacto-N-fucopentose I (LNFP I), c) lactodifucotetraose (LDFT), d) 3-fucosyllactose (3FL), e) lacto-N-fucopentose II (LNFP II), f) lacto-N-fucopentose III (LNFP III), g) the summation of a)-f)

Characteristics of the study population. As shown in Table 3.1, 113 of the mother-infant pairs were assigned to the Se+ group while 47 were assigned to the Se- group. Maternal age and body mass index (BMI), infant length, weight, *z*-score, and age were comparable between the two groups. Morbidity data reflected the prevalence of disease over a two-week time period prior to sample collection. A total of 8.0% and 14.2% of infants in the Se+ group were reported with diarrhea and cough respectively, while the prevalence of these conditions in Se- group was 17.0% and 34.0%, respectively. Prevalence of vomiting, fever, and tachypnea was similar between the two groups.

Table 3.1. Clinical characteristics of mother and infant subjects.

Characteristics	Milk secretor status phenotype	
	Secretor	**Non-secretor**
Sample size, n	113	47
Mother		
Age, year	26.2 (6.0)	25.2 (5.8)
BMI, kg/m^2	24.5 (3.6)	24.0 (3.3)
Infant		
Sex, %female	60.2%	46.8%
Length, cm	58.8 (6.0)	59.4 (3.1)
Weight, kg	6.1 (0.9)	6.4 (1.0)
z-score	0.4 (1.1)	0.6 (1.0)
Age, month	3.5 (1.1)	3.5 (1.1)
Vomit, n (%) #	8 (7.1%)	3 (6.4%)
Fever, n (%) #	19 (16.8%)	8 (17.0%)
Diarrhea, n (%) #	9 (8.0%)	5 (17.0%)
Cough, n (%) #	16 (14.2%)	16 (34.0%)
Tachypnea, n (%) #	4 (3.5%)	2 (4.3%)

BMI, body mass index. Mother age, BMI, length, weight, *z*-score, and infant age were shown as mean (standard deviation). # Morbidity data reflect the prevalence of disease over a two week time period before sample collection.

Impact of milk secretor status on the infant microbiome. To explore the impact of maternal secretor phenotype on infant fecal microbiota, 16S rRNA gene sequencing was performed. A PCoA showed no clear separation in the microbiome between infants receiving Se+ and Se- milk (Figure 3.2 a). Chao1 and Shannon alpha diversity showed no significant differences between the Se+ and Se- groups via Mann-Whitney test (Chao1 $p = 0.12$ and Shannon $p = 0.21$, Figure 3.2 b). Figure 3.2 c and 3.2 d shows those genera with over 1% relative abundance for each of the infants from both groups, and Figure 3.2 e shows a comparison of the average relative abundance of each genus. As expected, *Bifidobacterium* comprised the major genus, and the abundance was similar

between infants consuming milk from Se+ and Se- mothers. Other genera present in significant amounts included *Bacteroides, Collinsella, Streptococcus,* and *Veillonella,* and no difference between the relative abundance of these taxa or other taxa with over 1% relative abundance was observed in the feces of infants consuming milk from Se+ or Se- mothers.

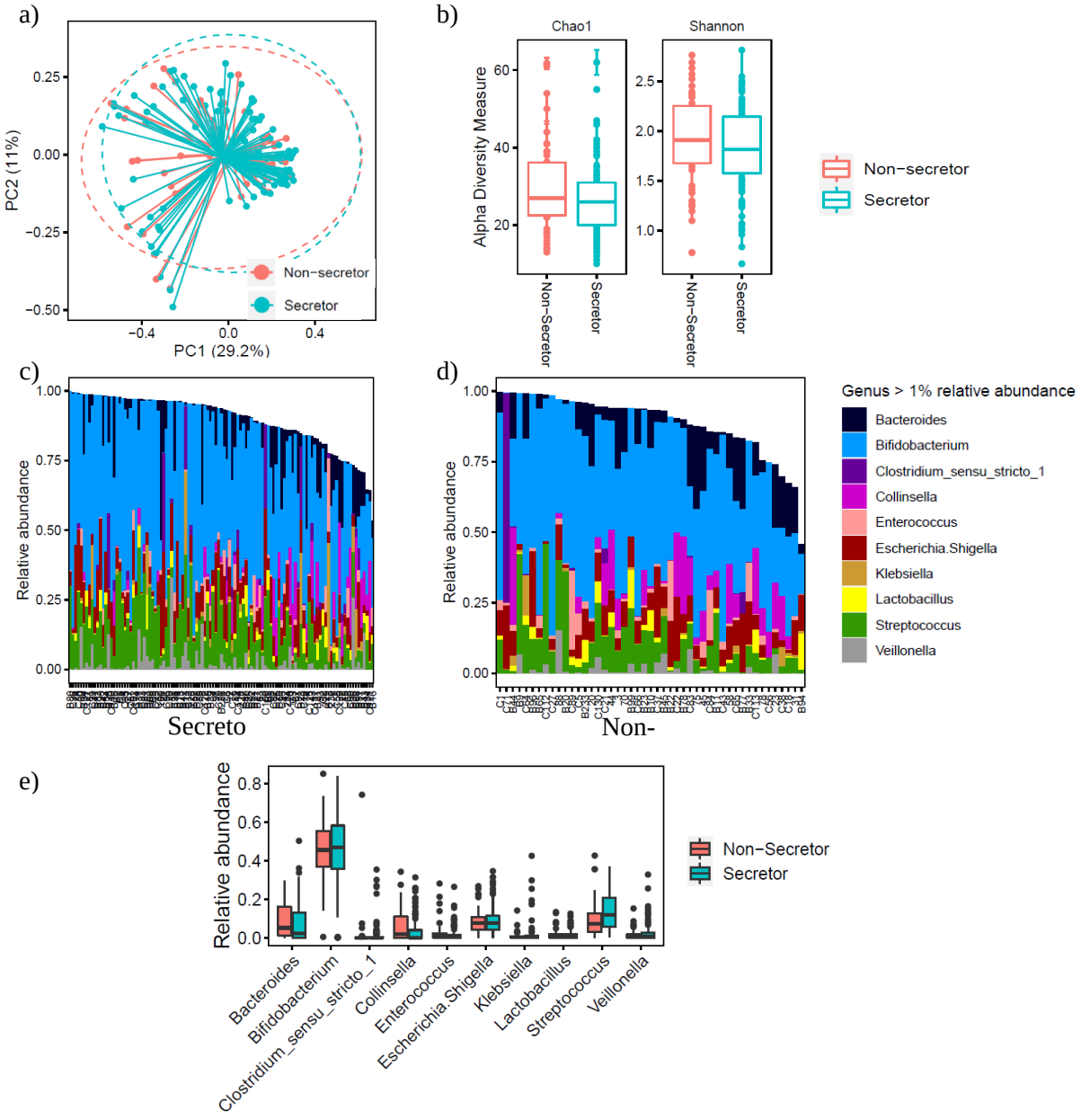


Figure 3.2. Infant fecal microbiome data in a) Principal Coordinates Analysis (PCoA), b) alpha diversity analysis, and c), d) relative abundance of genus over 1%. PCoA was based on bray distance matrix data. The centroids of each cluster (centroid of mass) were calculated as the average PC1 and PC2 of all samples for each group. The ellipses were constructed based on multivariate normal distribution at 95% confidence level.

Impact of milk secretor phenotype on the infant fecal metabolome. Fecal metabolites were examined to evaluate the impact of secretor phenotypes on intestinal microbial fermentation

products. PCoA of fecal metabolites showed no clear separation between the Se+ and Se- group (Figure 3.3 a). To determine if subtle differences in specific metabolites could be detected, fecal metabolites were compared using a combination of effect size and the Mann-Whitney U test (Figure 3.3 b). 2'FL, LDFT, and 1,2-propanediol were significantly higher in infant feces from the Se+ group compared to the Se- group ($p < 0.05$, $-0.32 < \delta < -0.16$). Another HMO related structure, N-acetylglucosamine (GlcNAc) was significantly lower in the Se+ group ($p = 0.05$, $\delta = 0.20$). Amino acids and related metabolites (glycine, pyroglutamate, 4-hydroxyphenylacetate, and 2-hydroxyphenylacetate) were also significantly lower in the Se+ group ($p < 0.05$, $0.20 < \delta < 0.24$ for all). Succinate and 2'-deoxyinosine were lower in the Se+ group as compared to the Se- group ($p < 0.05$, $0.25 < \delta$ 0.29).

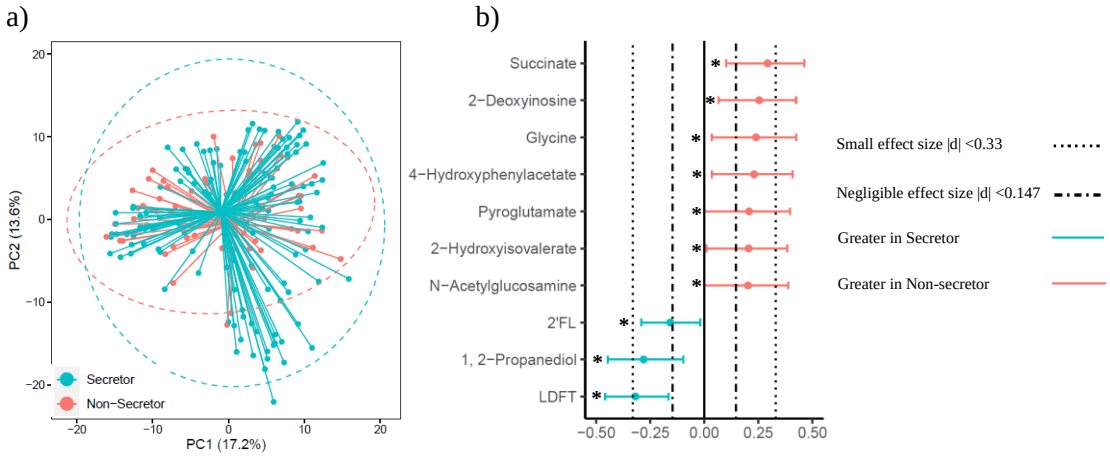


Figure 3.3. Infant fecal metabolome data in a) Principle Coordinate Analysis (PCoA) and b) Cliff's delta effect size analysis and Mann-Whiteney U test. * $p < 0.05$ in Mann-Whitney test with FDR correction

Association between HMO, microbiome and metabolome. To further investigate the association between milk HMO, fecal microbiota and the metabolome, Spearman correlations were conducted between fecal metabolites and milk HMOs, and between fecal metabolites and

microbiome relative abundance data. Fecal 1,2-propanediol concentration, a fermentation byproduct shown to be produced by multiple *Bifidobacterium* species (Bunesova, Lacroix, and Schwab 2016), was positively correlated with the concentration of total milk fucosylated HMOs in the Se- group ($P < 0.05$) but not in the Se+ group (Figure 3.4 a). Interestingly, no correlation between milk fucosylated HMOs and *Bifidobacterium* relative abundance was observed (data not shown).

A positive correlation was observed between the relative abundance of *Bifidobacterium* and *Streptococcus* ($p < 0.001$) in infants in the Se+ group but not in the Se- group (Figure 3.4 b). Interestingly, *Bifidobacterium* and *Streptococcus* had similar correlations with fecal metabolites in the Se+ group but not in Se- group (Figure 3.4 c). In the Se+ group, the relative abundance of *Bifidobacterium* and *Streptococcus* were negatively correlated with several short chain fatty acids and organic acids including butyrate ($p < 0.001$), propionate ($p < 0.001$), valerate ($p < 0.001$), and succinate ($p < 0.001$), as well as the amino acid breakdown metabolite 2-hydroxyisovalerate ($p < 0.01$ for *Bifidobacterium* and $p < 0.001$ for *Streptococcus*) and a purine degradation metabolite, hypoxanthine ($p < 0.01$ for *Bifidobacterium* and $p < 0.001$ for *Streptococcus*). *Bifidobacterium* and *Streptococcus* relative abundances were both positively correlated with acetoin ($p < 0.001$), indole-3-lactate ($p < 0.001$ for *Bifidobacterium* and $p < 0.05$ for *Streptococcus*), lactate ($p < 0.001$), 1,2-propaendiol ($p < 0.001$ for *Bifidobacterium* and $p < 0.01$ for *Streptococcus*) and pyruvate ($p < 0.001$) in the Se+ group. *Bifidobacterium* relative abundance was also positively correlated with 4-hydroxyphenyllactate ($p < 0.001$), 2-hydroxyglutarate ($p < 0.001$), acetate ($p < 0.01$), and aspartate ($p < 0.001$) in the Se+ group. In the Se- group, *Bifidobacterium* relative abundance was positively correlated with indole-3-lactate ($p < 0.05$), whereas *Streptococcus* relative abundance was positively correlated with lactate ($p < 0.01$), 1, 2-propanediol ($p < 0.05$), pyruvate ($p < 0.01$),

and negatively correlated with butyrate ($p < 0.001$), hypoxanthine ($p < 0.01$) propionate ($p < 0.01$), succinate ($p < 0.05$), and valerate ($p < 0.01$).

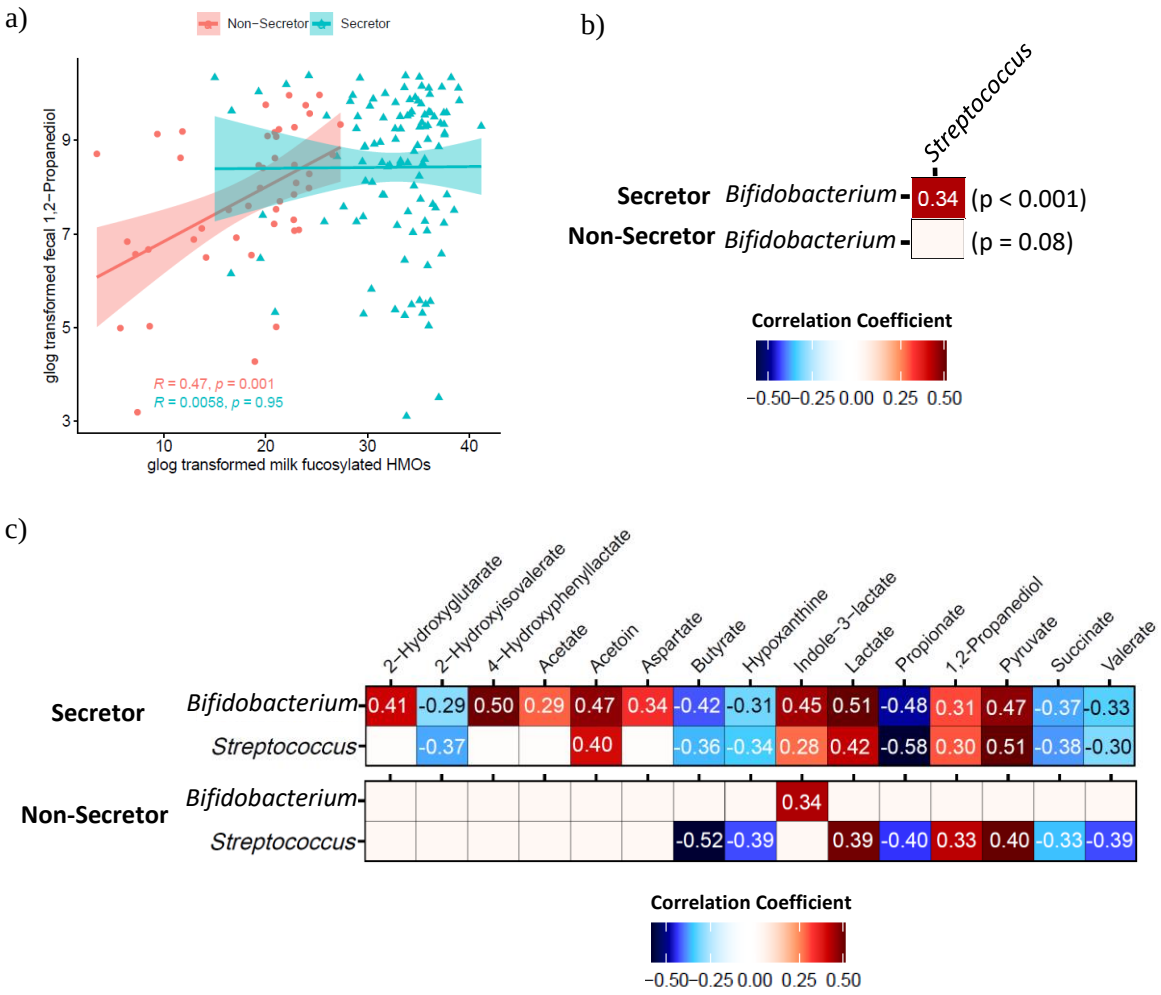


Figure 3.4. Association between a) milk fucosylated HMOs and infant fecal 1, 2-propanediol, b) gut *Bifidobacterium* and *Streptococcus* and c) *Bifidobacterium* or *Streptococcus* and fecal metabolites.

Discussion

HMOs have been reported to have beneficial effects on infants through shaping infant gut microbial composition. Receiving Se+ milk has been reported to be positively associated with gut *Bifidobacterium* in some US (Lewis et al. 2015) and Australian infants (Smith-Brown et al. 2016),

not correlated with gut microbiota in some Finnish vaginally born infants (Korpela et al. 2018), and negatively associated with *Bifidobacterium* in some EBF US infants (Wang et al. 2015). One thing in common with all of these studies, is that reports of exclusive feeding were solely based on recall of the mothers. Moreover, these studies all took place in resource-rich countries. Our study on EBF infants was performed in a rural setting in Indonesia, and exclusive breastfeeding was validated through the DTM method. Here, no significant difference in gut microbiota between infants receiving Se+ and Se- milk based on results from PCoA (Figure 3.2 a), alpha diversity measurement (Figure 3.2 b), and comparison of microbial genera (Figure 3.2 e) was observed.

We speculate that geographic and/or socioeconomic factors may play a role in how maternal secretor status shapes infant gut microbial composition and accommodation to the available HMOs entering the gut. Indeed, a difference in infant gut microbiome between WEIRD (Western, educated, industrialized, rich and democratic) societies and other parts of the world has recently drawn attention. Specifically, a loss of highly specialized *Bifidobacterium* species in the infant gut and an increase of fecal pH has been recognized in resource-rich countries over the past century (Henrick et al. 2018). *Bifidobacterium longum* subsp. *infantis*, which has a complete capacity for HMO utilization, was found to be rare in US infants (Casaburi et al. 2021). Though without direct evidence, interventions such as provision of antibiotics and caesarean section were speculated to be part of the reason causing the loss of this bacterial taxon (Duranti et al. 2017; Betrán et al. 2016). All infants in the present study were from a rural area in Indonesia, and the sample size reported here is much larger (160 infants) compared to previously published studies (22 to 76 infants, Lewis et al. 2015; Smith-Brown et al. 2016; Korpela et al. 2018; Wang et al. 2015). The current study adds insight into how maternal secretor status influences infant gut microbiota without confounding factors such as antibiotic use, caesarean section, and formula use.

Although no difference at the genus level from the 16s rRNA gene sequencing results was observed, there could be a shift of *Bifidobacterium* at the species level due to the difference in HMO profile (Figure 3.1 a-g). Different species of *Bifidobacterium* have been reported to have different capability and preference for utilizing HMOs (Underwood et al. 2014). For instance, several *Bifidobacterium* subspecies (including *Bifidobacterium longum* subsp. *infantis*, *Bifidobacterium longum* subsp. *suis* BSM11-5, and *Bifidobacterium kashiwanohense*) have been reported to grow in the presence of 2'FL and 3FL (Bunesova, Lacroix, and Schwab 2016), whereas *Bifidobacterium breve* KA179 was reported to grow on 2'FL but not on 3FL (Ruiz-Moyano et al. 2013). The difference in concentration of 1,2-propanediol in infant feces (Figure 3.3 b) between infants consuming milk from Se+ or Se- mothers suggests that there are differences at the species and strain levels of *Bifidobacterium*. 1,2-propanediol has been reported to be produced by *Bifidobacterium longum* subsp. *infantis* and *Bifidobacterium longum* subsp. *suis* BSM11-5 but not by *Bifidobacterium kashiwanohense* DSM 21854 when grown on 2'FL and 3FL (Bunesova, Lacroix, and Schwab 2016; Zabel et al. 2019; Dedon et al. 2020). The higher level of 1,2-propanediol in the feces of infant fed Se+ milk is likely a result of colonization of *Bifidobacterium* species that are able to produce 1,2-propanediol from 2'FL. Importantly, *Bifidobacterium* was positively correlated with 1,2-propanediol in infants fed milk from Se+ mothers but not Se- mothers (Figure 3.4 c).

It is worth noting that 1,2-propanediol was still present in the feces of infants fed milk from Se- mothers. This could be attributed to species such as *B. longum* subsp. *suis* BSM 11–5 which synthesize 1,2-propanediol mainly from 3FL when given both 2'FL and 3FL (Bunesova, Lacroix, and Schwab 2016). Furthermore, 1,2-propanediol showed a positive correlation with fucosylated HMOs in the Se- group but not in the Se+ group (Figure 3.4 a). This suggests that production of

1,2-PD arises from fermentation of fucosylated HMOs, and that production may be slower when 2'FL is not present. Once a certain amount of 1,2-propanediol is produced, it could be that a "steady-state" is reached. Whether this "steady state" is due to increased absorption, or utilization by other microbes (Cheng et al. 2020) remains to be studied.

A lower level of succinate as well as amino acids and related metabolites (glycine, pyroglutamate, 4-hydroxyphenylacetate, and 2-hydroxyisovalerate) (Figure 3.3 b) in the feces of infants fed Se+ milk may suggest a shift in carbohydrate and amino acid fermentation. There were no significant differences in fecal organic acids associated with colonic pH (other than succinate), so it is unlikely that there are differences in stool pH between the groups (although this was not specifically measured). Lower overall levels of fucosylated HMOs in the milk could lead to slight changes in the fermentative capacity of the microbiome, and thus slight changes in abundance of certain bacterial taxa resulting in the differences observed here. Higher fecal succinate has been associated with inflammatory bowel disease (IBD) (Ariake et al. 2000), and succinate accumulation has been attributed to changes in the abundance of succinate-consuming gut microbes (Morgan et al. 2012). One study suggested that higher succinate in the neonate may favor the colonization of strict anaerobes (Kim et al. 2017). It is clear that the potential impact of succinate in the neonatal gut on infant health needs to be further studied.

Besides shaping infant gut microbiota, maternal secretor status has also been reported to protect against diarrhea. The pathogen responsible for the most common cause of infant bacterial diarrhea, *Campylobacter jejuni*, targets α-1,2-fucosylated glycans on epithelial cells (Ruiz-Palacios et al. 2003), and 2'FL has been reported to protect breastfed infants against *Campylobacter* diarrhea by competing for binding (Morrow et al. 2005). In this study, a slightly lower prevalence of diarrhea was observed in Se+ group (8.0%) compared to Se- group (17.0%) (Table 3.1), although the cause

of the diarrhea was not determined. Breastfeeding has also been reported to decrease respiratory infection in both infancy (Raheem, Binns, and Chih 2017) and childhood (Tromp et al. 2017), with the benefit attributed to the antiadhesive function of HMOs (Stepans et al. 2006). 2'FL and lacto-N-neotetraose (LNnT) supplemented infant formula was shown to decrease the risk of lower respiratory tract infection in infants (Puccio et al. 2017), and 2'FL was shown to be the most effective HMO in decreasing respiratory syncytial virus load and cytokines in epithelial cells when compared to 3FL, 6'-sialyllactose (6'SL) and other oligosaccharides (Duska-McEwen et al. 2014). In this study, cough was also more common in infants fed Se- milk (34.0%) compared to infants fed Se+ milk (14.2%) (Table 3.1).

A positive association between *Bifidobacterium* and *Streptococcus* (Figure 3.4 b), and a similar trend of correlation with fecal metabolites between the two microbes (Figure 3.4 c) in Se+ group indicate that there could be cross-talk between these two microbial taxa in infants fed Se+ milk. A previous study reported a synergistic effect between *Bifidobacterium lactis* and *Streptococcus thermophiles* when cultured in skim milk, and such positive mutual interaction was further improved by inulin due to its prebiotic and bifidogenic function (Oliveira et al. 2012). Specifically, better growth of both organisms was observed in co-culture compared to pure cultures, and the increase in the biomass of the two microbes was larger in the presence of inulin (Oliveira et al. 2012). In the present study, Se+ associated HMOs may act as prebiotics favoring the growth of certain species of *Bifidobacterium* and *Streptococcus* that have a synergistic relationship.

The positive correlation between *Bifidobacterium* and 4-hydroxyphenyllactate in the Se+ group but not in the Se- group further delineates differences at the species and strain level of *Bifidobacterium* between the two groups. 4-Hydroxyphenyllactate, a tyrosine metabolite, has been shown to be produced by *Bifidobacterium in vitro* (Beloborodova et al. 2012; Windey, De Preter,

and Verbeke 2012). A recent study showed that 4-hydroxyphenyllactate was primarily produced by *Bifidobacterium* species that utilize HMOs including *Bifidobacterium bifidum*, *Bifidobacterium breve*, *Bifidobacterium longum* subsp. *longum*, *Bifidobacterium longum* subsp. *infantis* and *Bifidobacterium sardovii* via the function of their aromatic lactate dehydrogenase (Laursen et al. 2020). Other *Bifidobacterium* species such as *Bifidobacterium adolescentis* and *Bifidobacterium pseudocatenulatum* have not been observed to produce 4-hydroxyphenyllactate (Laursen et al. 2020). Interestingly, the fecal concentration of 4-hydroxyphenyllactate was not significantly different between infants from Se+ and Se- mothers, which suggests that other bacterial species may be creating this compound. Indeed, 4-hydroxyphenyllactate has been shown to be produced by other microbes including *Lactobacillus fermentum* and *Eubacterium lentum* (Beloborodova et al. 2012).

Bifidobacteria have been reported to be capable of producing lactate (Peirotén et al. 2019; Oliveira et al. 2012), and their ability to produce lactate could be affected by the availability of different HMOs. Specifically, *Bifidobacterium longum* JCM 1260, JCM7011, and JCM 7009 produce lactate in the presence of 2'FL but not 3FL (Yu, Chen, and Newburg 2013). This could shed light on the presence of the positive correlation between *Bifidobacterium* and lactate in the Se+ group but not in the Se- group. The utilization of lactate by microbes such as *Veillonella* in the infant gut could also be a contributing factor. *Streptococcus thermophilus* ATCC19258, a common human gut microbe, was reported to utilize either 2'FL or 3FL to produce lactate (Yu, Chen, and Newburg 2013). This could explain the positive association between *Streptococcus* and lactate in both the Se+ and Se- groups.

The tryptophan metabolite indole-3-lactate has been shown to be produced by *Bifodobacterium longum subsp. infantis* in the presence HMOs (Ehrlich et al. 2020). However, the production of

indole-3-lacate is not necessarily dependent on the presence of HMOs, as multiple *Bifidobacterium* species (including *Bifidobacterium longum* subsp. *longum*, *Bifidobacterium longum* subsp. *infantis*, *Bifidobacterium breve*, and *Bifidobacterium bifidum*) are also able to produce indole-3-lactate when cultured in DeMan, Rogosa and Sharpe (MRS) medium (Sakurai, Odamaki, and Xiao 2019). This could explain the presence of the positive association between *Bifidobacterium* in general (Figure 3.4 c) and indole-3-lactate in both Se+ and Se- groups, and a similar concentration of fecal indole-3-lactate between the two groups. Indole-3-lactate has been reported to have an anti-inflammatory effect (Ehrlich et al. 2020).

Conclusion

The results obtained in the present study show the impact of maternal secretor status on infant gut microbial composition and their fermentation capability. Distinct HMO profiles between breast milk from Se+ and Se- women were confirmed. Though previous studies have reported differences in infant microbial composition when consuming Se+ milk and Se- milk (Lewis et al. 2015; Smith-Brown et al. 2016), no significant difference in the infant gut microbiome based on 16s sequencing was observed in this study, indicating that there could be differences at the species and strain level, and further that there may be other factors playing a role in shaping the gut microbiome and its function.

One limitation of the current study is that the secretor status of the infants were not measured. Studies have reported that being a Se- infant was associated with resistance against diarrhea prevalence (Muthumuni et al. 2021) and norovirus (Thorven et al. 2005). Therefore, information on the secretor status of both the mothers and the infants could provide better understanding of how HMOs function to protect infants. Future studies should also investigate the microbiome with

higher resolution to evaluate specific species and strains of *Bifidobacterium*, and study the interaction of *Bifidobacterium* and *Streptococcus* at the species or strain level.

Abbreviations

LNFP I: lacto-N-fucopentose I, LNFP II: lacto-N-fucopentose II, LNFP III: lacto-N-fucopentose III, LDFT: lactodifucotetraose, 2'FL: 2'fucosyllactose, 3FL: 3-fucosyllactose, LNnT: lacto-N-neotetraose, GlcNAc: N-Acetylglucosamine, MFGM: milk fat globule membrane, SCFAs: short chain fatty acids, Se+: secretor, Se-: non-secretor, PCoA: Principal coordinate analysis, ANCOM: Analysis of composition of microbiomes, BMI: body mass index

References

Azad, Meghan B, Bianca Robertson, Faisal Atakora, Allan B Becker, Padmaja Subbarao, Theo J Moraes, Piushkumar J Mandhane, et al. 2018. "Human Milk Oligosaccharide Concentrations Are Associated with Multiple Fixed and Modifiable Maternal Characteristics, Environmental Factors, and Feeding Practices." *The Journal of Nutrition* 148 (11): 1733–42. doi: 10.1093/jn/nxy175.

Beloborodova, Natalia, Iskander Bairamov, Andrei Olenin, Victoria Shubina, Vera Teplova, and Nadezhda Fedotcheva. 2012. "Effect of Phenolic Acids of Microbial Origin on Production of Reactive Oxygen Species in Mitochondria and Neutrophils." *Journal of Biomedical Science* 19 (1): 89. doi: 10.1186/1423-0127-19-89.

Betrán, Ana Pilar, Jianfeng Ye, Anne-Beth Moller, Jun Zhang, A. Metin Gülmezoglu, and Maria Regina Torloni. 2016. "The Increasing Trend in Caesarean Section Rates: Global, Regional and National Estimates: 1990-2014." *PloS One* 11 (2): e0148343. doi: 10.1371/journal.pone.0148343.

Bunesova, Vera, Christophe Lacroix, and Clarissa Schwab. 2016. "Fucosyllactose and L-Fucose Utilization of Infant Bifidobacterium Longum and Bifidobacterium Kashiwanohense." *BMC Microbiology* 16 (1): 248. doi: 10.1186/s12866-016-0867-4.

Callahan, Benjamin J, Paul J McMurdie, Michael J Rosen, Andrew W Han, Amy Jo A Johnson, and Susan P Holmes. 2016. "DADA2: High-Resolution Sample Inference from Illumina Amplicon Data." *Nature Methods* 13 (7): 581–83. doi: 10.1038/nmeth.3869.

Casaburi, Giorgio, Rebbeca M. Duar, Heather Brown, Ryan D. Mitchell, Sufyan Kazi, Stephanie Chew, Orla Cagney, et al. 2021. "Metagenomic Insights of the Infant Microbiome Community Structure and Function across Multiple Sites in the United States." *Scientific Reports* 11 (1): 1472. doi: 10.1038/s41598-020-80583-9.

Cheng, Christopher C., Rebbeca M. Duar, Xiaoxi Lin, Maria Elisa Perez-Munoz, Stephanie Tollenaar, Jee-Hwan Oh, Jan-Peter van Pijkeren, et al. 2020. "Ecological Importance of Cross-Feeding of the Intermediate Metabolite 1,2-Propanediol between Bacterial Gut Symbionts." *Applied and Environmental Microbiology*.

Duranti, Sabrina, Gabriele Andrea Lugli, Leonardo Mancabelli, Francesca Turroni, Christian Milani, Marta Mangifesta, Chiara Ferrario, et al. 2017. "Prevalence of Antibiotic Resistance Genes among Human Gut-Derived Bifidobacteria." *Applied and Environmental Microbiology* 83 (3). doi: 10.1128/AEM.02894-16.

Duska-McEwen, Geralyn, Albert P. Senft, Teah L. Ruetschilling, Edward G. Barrett, and Rachael H. Buck. 2014. "Human Milk Oligosaccharides Enhance Innate Immunity to Respiratory Syncytial Virus and Influenza" *Food and Nutrition Sciences*.

Ehrlich, Amy M., Alline R. Pacheco, Bethany M. Henrick, Diana Taft, Gege Xu, M. Nazmul Huda, Darya Mishchuk, et al. 2020. "Indole-3-Lactic Acid Associated with Bifidobacterium-Dominated Microbiota Significantly Decreases Inflammation in Intestinal Epithelial Cells." *BMC Microbiology* 20 (1): 357. doi: 10.1186/s12866-020-02023-y.

He, Xuan, Mariana Parenti, Tove Grip, Magnus Domellöf, Bo Lönnerdal, Olle Hernell, Niklas Timby, and Carolyn M. Slupsky. 2019b. "Metabolic Phenotype of Breast-Fed Infants, and Infants Fed Standard Formula or Bovine MFGM Supplemented Formula: A Randomized Controlled Trial." *Scientific Reports* 9 (1): 339. doi: 10.1038/s41598-018-36292-5.

He, Xuan, Mariana Parenti, Tove Grip, Bo Lönnerdal, Niklas Timby, Magnus Domellöf, Olle Hernell, and Carolyn M. Slupsky. 2019a. "Fecal Microbiome and Metabolome of Infants Fed Bovine MFGM Supplemented Formula or Standard Formula with Breast-Fed Infants as Reference: A Randomized Controlled Trial." *Scientific Reports* 9 (1): 11589. doi: 10.1038/s41598-019-47953-4.

Henrick, Bethany M., Andra A. Hutton, Michelle C. Palumbo, Giorgio Casaburi, Ryan D. Mitchell, Mark A. Underwood, Jennifer T. Smilowitz, and Steven A. Frese. 2018. "Elevated Fecal PH Indicates a Profound Change in the Breastfed Infant Gut Microbiome Due to Reduction of *Bifidobacterium* over the Past Century." Edited by Julia Oh. *MSphere* 3 (2): e00041-18. doi: 10.1128/mSphere.00041-18.

Korpela, Katri, Anne Salonen, Brandon Hickman, Clemens Kunz, Norbert Sprenger, Kaarina Kukkonen, Erkki Savilahti, Mikael Kuitunen, and Willem M. de Vos. 2018. “Fucosylated Oligosaccharides in Mother’s Milk Alleviate the Effects of Caesarean Birth on Infant Gut Microbiota.” *Scientific Reports* 8 (1): 13757. doi: 10.1038/s41598-018-32037-6.

Kunz, Clemens, Christina Meyer, Maria Carmen Collado, Lena Geiger, Izaskun García-Mantrana, Bibiana Bertua-Ríos, Cecilia Martínez-Costa, Christian Borsch, and Silvia Rudloff. 2017. “Influence of Gestational Age, Secretor, and Lewis Blood Group Status on the Oligosaccharide Content of Human Milk.” *Journal of Pediatric Gastroenterology and Nutrition* 64 (5): 789–98.

Laursen, Martin F., Mikiyasu Sakanaka, Nicole von Burg, Urs Mörbe, Daniel Andersen, Janne Marie Moll, Ceyda T. Pekmez, et al. 2020. “Breastmilk-Promoted Bifidobacteria Produce Aromatic Amino Acids in the Infant Gut.” *BioRxiv*, January, 2020.01.22.914994. doi: 10.1101/2020.01.22.914994.

Leong, Claudia, Rosalind S Gibson, Aly Diana, Jillian J Haszard, Sofa Rahmannia, Mohammad Brachim Ansari, Lina Sofiatul Inayah, Afini Dwi Purnamasari, and Lisa A Houghton. 2021. “Differences in Micronutrient Intakes of Exclusive and Partially Breastfed Indonesian Infants from Resource-Poor Households Are Not Accompanied by Differences in Micronutrient Status, Morbidity, or Growth.” *The Journal of Nutrition* 151 (3): 705–15. doi: 10.1093/jn/nxaa381.

Lewis, Zachery T., Sarah M. Totten, Jennifer T. Smilowitz, Mina Popovic, Evan Parker, Danielle G. Lemay, Maxwell L. Van Tassell, Michael J. Miller, Yong-Su Jin, and J. Bruce German. 2015. “Maternal Fucosyltransferase 2 Status Affects the Gut Bifidobacterial Communities of Breastfed Infants.” *Microbiome* 3 (1): 13.

Liu, Zheng, Aly Diana, Christine Slater, Thomas Preston, Rosalind S. Gibson, Lisa Houghton, and Stephen B. Duffull. 2019. “Development of a Nonlinear Hierarchical Model to Describe the Disposition of Deuterium in Mother-Infant Pairs to Assess Exclusive Breastfeeding Practice.” *Journal of Pharmacokinetics and Pharmacodynamics* 46 (1): 1–13. doi: 10.1007/s10928-018-9613-x.

Mazariegos, Monica, Christine Slater, and Manuel Ramirez-Zea. 2016. “Validity of Guatemalan Mother’s Self-Reported Breast-Feeding Practices of 3-Month-Old Infants.” *Food and Nutrition Bulletin* 37 (4): 494–503. doi: 10.1177/0379572116654644.

Medoua, Gabriel Nama, Estelle C. Sajo Nana, Anne Christine A. Ndzana, Caroline S. Makamto, Lucien S. Etame, Honorine A. Rikong, and Jean Louis E. Oyono. 2012. “Breastfeeding Practices of Cameroonian Mothers Determined by Dietary Recall since Birth and the Dose-to-the-Mother Deuterium-Oxide Turnover Technique.” *Maternal & Child Nutrition* 8 (3): 330–39. doi: 10.1111/j.1740-8709.2011.00293.x.

Morrow, Ardythe L., Guillermo M. Ruiz-Palacios, Xi Jiang, and David S. Newburg. 2005. “Human-Milk Glycans That Inhibit Pathogen Binding Protect Breast-Feeding Infants against Infectious Diarrhea.” *The Journal of Nutrition* 135 (5): 1304–7. doi: 10.1093/jn/135.5.1304.

Muthumuni, Dhasni, Kozeta Miliku, Kaitlin H. Wade, Nicholas J. Timpson, and Meghan B. Azad. 2021. “Enhanced Protection Against Diarrhea Among Breastfed Infants of Nonsecretor Mothers.” *The Pediatric Infectious Disease Journal* 40 (3): 260–63. doi: 10.1097/INF.0000000000003014.

Newburg, David S., Guillermo M. Ruiz-Palacios, and Ardythe L. Morrow. 2005. “Human Milk Glycans Protect Infants Against Enteric Pathogens.” *Annual Review of Nutrition* 25 (1): 37–58. doi: 10.1146/annurev.nutr.25.050304.092553.

Oliveira, Ricardo Pinheiro de Souza, Patrizia Perego, Maricê Nogueira de Oliveira, and Attilio Converti. 2012. “Growth, Organic Acids Profile and Sugar Metabolism of Bifidobacterium Lactis in Co-Culture with Streptococcus Thermophilus: The Inulin Effect.” *Food Research International* 48 (1): 21–27. doi: 10.1016/j.foodres.2012.02.012.

Peirotén, Angela, Pilar Gaya, Juan Luis Arqués, Margarita Medina, and Eva Rodríguez. 2019. “Technological Properties of Bifidobacterial Strains Shared by Mother and Child.” *BioMed Research International* 2019 (January): 9814623. doi: 10.1155/2019/9814623.

Puccio, Giuseppe, Philippe Alliet, Cinzia Cajozzo, Elke Janssens, Giovanni Corsello, Norbert Sprenger, Susan Wernimont, Delphine Egli, Laura Gosoniu, and Philippe Steenhout. 2017. “Effects of Infant Formula with Human Milk Oligosaccharides on Growth and Morbidity: A Randomized

Multicenter Trial." *Journal of Pediatric Gastroenterology and Nutrition* 64 (4): 624–31. doi: 10.1097/MPG.0000000000001520.

Quast, Christian, Elmar Pruesse, Pelin Yilmaz, Jan Gerken, Timmy Schweer, Pablo Yarza, Jörg Peplies, and Frank Oliver Glöckner. 2013. "The SILVA Ribosomal RNA Gene Database Project: Improved Data Processing and Web-Based Tools." *Nucleic Acids Research* 41 (Database issue): D590–96. doi: 10.1093/nar/gks1219.

Raheem, Raheema Abdul, Colin W Binns, and Hui J Chih. 2017. "Protective Effects of Breastfeeding against Acute Respiratory Tract Infections and Diarrhoea: Findings of a Cohort Study." *Journal of Paediatrics and Child Health* 53 (3): 271–76. doi: 10.1111/jpc.13480.

Ruiz-Moyano, Santiago, Sarah M. Totten, Daniel A. Garrido, Jennifer T. Smilowitz, J. Bruce German, Carlito B. Lebrilla, and David A. Mills. 2013. "Variation in Consumption of Human Milk Oligosaccharides by Infant Gut-Associated Strains of Bifidobacterium Breve." *Applied and Environmental Microbiology* 79 (19): 6040–49. doi: 10.1128/AEM.01843-13.

Ruiz-Palacios, Guillermo M., Luz Elena Cervantes, Pilar Ramos, Bibiana Chavez-Munguia, and David S. Newburg. 2003. "Campylobacter Jejuni Binds Intestinal H(O) Antigen (Fuc Alpha 1, 2Gal Beta 1, 4GlcNAc), and Fucosyloligosaccharides of Human Milk Inhibit Its Binding and Infection." *The Journal of Biological Chemistry* 278 (16): 14112–20. doi: 10.1074/jbc.M207744200.

Sakurai, Takuma, Toshitaka Odamaki, and Jin-Zhong Xiao. 2019. "Production of Indole-3-Lactic Acid by Bifidobacterium Strains Isolated FromHuman Infants." *Microorganisms* 7 (9). doi: 10.3390/microorganisms7090340.

Smilowitz, Jennifer T., Aifric O'sullivan, Daniela Barile, J. Bruce German, Bo Lönnerdal, and Carolyn M. Slupsky. 2013. "The Human Milk Metabolome Reveals Diverse Oligosaccharide Profiles." *The Journal of Nutrition* 143 (11): 1709–18. doi: 10.3945/jn.113.178772.

Smith-Brown, Paula, Mark Morrison, Lutz Krause, and Peter S. W. Davies. 2016. "Mothers Secretor Status Affects Development of Childrens Microbiota Composition and Function: A Pilot Study." *PLOS ONE* 11 (9): e0161211. doi: 10.1371/journal.pone.0161211.

Spevacek, Ann R, Jennifer T Smilowitz, Elizabeth L Chin, Mark A Underwood, J Bruce German, and Carolyn M Slupsky. 2015. "Infant Maturity at Birth Reveals Minor Differences in the Maternal Milk Metabolome in the First Month of Lactation." *The Journal of Nutrition* 145 (8): 1698–1708. doi: 10.3945/jn.115.210252.

Stepans, Mary Beth Flanders, Susan L. Wilhelm, Melody Hertzog, T. Kim Callahan Rodehorst, Susan Blaney, Beth Clemens, Josef J. Polak, and David S. Newburg. 2006. "Early Consumption of Human Milk Oligosaccharides Is Inversely Related to Subsequent Risk of Respiratory and Enteric Disease in Infants." *Breastfeeding Medicine : The Official Journal of the Academy of Breastfeeding Medicine* 1 (4): 207–15. doi: 10.1089/bfm.2006.1.207.

Thorven, Maria, Ammi Grahn, Kjell-Olof Hedlund, Hugo Johansson, Christer Wahlfrid, Göran Larson, and Lennart Svensson. 2005. "A Homozygous Nonsense Mutation (428G→A) in the Human Secretor (*FUT2*) Gene Provides Resistance to Symptomatic Norovirus (GGII) Infections." *Journal of Virology* 79 (24): 15351. doi: 10.1128/JVI.79.24.15351-15355.2005.

Thurl, Stephan, Manfred Munzert, Jobst Henker, Günther Boehm, Beate Müller-Werner, Jürgen Jelinek, and Bernd Stahl. 2010. "Variation of Human Milk Oligosaccharides in Relation to Milk Groups and Lactational Periods." *British Journal of Nutrition* 104 (9): 1261–1271.

Totten, Sarah M., Angela M. Zivkovic, Shuai Wu, UyenThao Ngyuen, Samara L. Freeman, L. Renee Ruhaak, Momodou K. Darboe, J. Bruce German, Andrew M. Prentice, and Carlito B. Lebrilla. 2012. "Comprehensive Profiles of Human Milk Oligosaccharides Yield Highly Sensitive and Specific Markers for Determining Secretor Status in Lactating Mothers." *Journal of Proteome Research* 11 (12): 6124–33. doi: 10.1021/pr300769g.

Tromp, Ilse, Jessica Kiefte-de Jong, Hein Raat, Vincent Jaddoe, Oscar Franco, Albert Hofman, Johan de Jongste, and Henriëtte Moll. 2017. "Breastfeeding and the Risk of Respiratory Tract Infections after Infancy: The Generation R Study." *PLOS ONE* 12 (2): e0172763. doi: 10.1371/journal.pone.0172763.

Underwood, Mark A., Stephanie Gaerlan, Maria Lorna A. De Leoz, Lauren Dimapasoc, Karen M. Kalanetra, Danielle G. Lemay, J. Bruce German, David A. Mills, and Carlito B. Lebrilla. 2015. "Human Milk Oligosaccharides in Premature Infants: Absorption, Excretion, and Influence on the Intestinal Microbiota." *Pediatric Research* 78 (August): 670.

Underwood, Mark A., J. Bruce German, Carlito B. Lebrilla, and David A. Mills. 2014. "*Bifidobacterium Longum* Subspecies *Infantis*: Champion Colonizer of the Infant Gut." *Pediatric Research* 77 (October): 229.

Wang, Aidong, Petya Koleva, Elloise du Toit, Donna T Geddes, Daniel Munblit, Susan L Prescott, Merete Eggesbø, et al. 2021. "The Milk Metabolome of Non-Secretor and Lewis Negative Mothers." *Frontiers in Nutrition* 7 (February): 576966–576966. doi: 10.3389/fnut.2020.576966.

Wang, Mei, Min Li, Shuai Wu, Carlito B Lebrilla, Robert S Chapkin, Ivan Ivanov, and Sharon M Donovan. 2015. "Fecal Microbiota Composition of Breast-Fed Infants Is Correlated with Human Milk Oligosaccharides Consumed." *Journal of Pediatric Gastroenterology and Nutrition* 60 (6): 825–33. doi: 10.1097/MPG.0000000000000752.

Weljie, Aalim M., Jack Newton, Pascal Mercier, Erin Carlson, and Carolyn M. Slupsky. 2006. "Targeted Profiling: Quantitative Analysis of 1H NMR Metabolomics Data." *Analytical Chemistry* 78 (13): 4430–42. doi: 10.1021/ac060209g.

Wesolowska-Andersen, Agata, Martin Iain Bahl, Vera Carvalho, Karsten Kristiansen, Thomas Sicheritz-Pontén, Ramneek Gupta, and Tine Rask Licht. 2014. "Choice of Bacterial DNA Extraction Method from Fecal Material Influences Community Structure as Evaluated by Metagenomic Analysis." *Microbiome* 2 (1): 19. doi: 10.1186/2049-2618-2-19.

Yilmaz, Pelin, Laura Wegener Parfrey, Pablo Yarza, Jan Gerken, Elmar Pruesse, Christian Quast, Timmy Schweer, Jörg Peplies, Wolfgang Ludwig, and Frank Oliver Glöckner. 2014. "The SILVA and 'All-Species Living Tree Project (LTP)' Taxonomic Frameworks." *Nucleic Acids Research* 42 (Database issue): D643–48. doi: 10.1093/nar/gkt1209.

Yu, Zhuo-Teng, Ceng Chen, and David S. Newburg. 2013. "Utilization of Major Fucosylated and Sialylated Human Milk Oligosaccharides by Isolated Human Gut Microbes." *Glycobiology* 23 (11): 1281–92. doi: 10.1093/glycob/cwt065.

Romano, J., Kromrey, J., Coraggio, J. & Skowronek, J. (2006). Appropriate statistics for ordinal level data: Should we really be using t-test and Cohen'sd for evaluating group differences on the NSSE and other surveys? In annual meeting of the Florida Association of Institutional Research (pp. 1--3).

Chapter 4 Infant fecal metabolites as potential biomarkers for exclusive breastfeeding

Abstract

Diet is known to impact the colonic microbiota, and thus the fecal metabolome may provide an important opportunity to find biomarkers to distinguish between exclusively breastfed (EBF) and non-exclusively breastfed (non-EBF) infants. The present study aimed to characterize differences in the metabolic profile between EBF and non-EBF infants, and to identify metabolic biomarkers to distinguish between the feeding practices. Utilizing a combination of the dose-to-mother (DTM) technique combined with food records, EBF infants were identified and compared with non-EBF infants who consumed either food (non-EBF_food), or only water (non-EBF_water). Metabolic and microbial differences were observed between feeding practices at both 2 and 5 months. A set of metabolic biomarkers including human milk oligosaccharides, short chain fatty acids, amino acids and organic acids with excellent predictive performance (area under the receiver operating curve (AUC) of 0.79 – 0.84) was developed utilizing the random forest and backwards selection models. Our findings indicate that the introduction of complementary food or water to an exclusively breastfed infant has an impact on the infant gut microbiome and metabolome, and that the infant fecal metabolic profile is a promising novel approach to distinguish between infant feeding practices.

Introduction

Breastfeeding is the gold standard in infant nutrition as it provides both nutrients and bioactive components that protect infants in their early development against many health risks (Smith et al. 2017; Lee and Binns 2020; Raheem, Binns, and Chih 2017). The World Health Organization (WHO) recommends infants to be exclusively breastfed (EBF) for the first 6 months of age without any other liquids or solids, including water (WHO, 2019). An accurate measure of the true prevalence of exclusive breastfeeding (EBF) practice is vital for researchers to further reveal the mechanisms of the benefits of breastfeeding and to test the effectiveness of breastfeeding education. Most studies determine infant feeding status through parental recall methodologies using 24 h recalls, recall since birth, and reports of current feeding practices (Mazariegos, Slater, and Ramirez-Zea 2016). However, evidence using a stable isotope deuterium dose-to-mother (DTM) technique to estimate the intake of breast milk and non-breastmilk water demonstrated that self-reported methods substantially overestimate the true prevalence of EBF practice. One study showed that 75% of the self-reported EBF infants were consuming other foods based on the DTM method (Medoua et al. 2012). In another study, 60% of parents self-reported EBF practice through 24 h recall, whereas only 36% of infants were determined to be EBF as calculated by the DTM technique (Mazariegos, Slater, and Ramirez-Zea 2016). Although the DTM method is accurate, there are limitations to use it, such as the high workload burden (collection of saliva samples from both the mothers and infants over the study period) and the fact that the determined EBF status is only accurate within the study period. Therefore, a less time and resource intensive method with the ability to objectively assess EBF practice needs to be developed.

Infant fecal metabolic and/or microbial profiles are a promising approach to determining EBF status as many studies have reported the impact of diet on the infant fecal metabolome and

microbiome (Dotz et al. 2016; Chow et al. 2014; Tao et al. 2011; Bridgman et al. 2017; Martin et al. 2014; He, et al. 2019). One study reported that the fecal metabolome of EBF infants is higher in human milk oligosaccharides (HMOs) including 2'-fucosyllactose (2'FL), lacto-*N*-fucopentaose (LNFP), fucose (Chow et al. 2014), and HMO metabolites than the fecal metabolome of non-EBF infants (Dotz et al. 2016). Human milk is unique as it contains higher amounts and more complex structures of soluble oligosaccharides than any other mammalian milk (Tao et al. 2011). While the presence of HMOs in infant feces can contribute to the classification of EBF and non-exclusively breastfed (non-EBF) infants, this measurement alone is problematic as the fecal HMO profile is also closely related to the species of gut microbiota present, and in the case where *Bifidobacterium longum* subsp *infantis (B. infantis)* is present in the infant gut, some HMOs could be fully utilized (Frese et al. 2017). SCFAs and organic acids have also been reported to differ between EBF and non-EBF infants. Acetate, butyrate, propionate, valerate, isobutyrate and isovalerate were reported to be lower while lactate was higher in EBF infants (Bridgman et al. 2017). A higher level of fecal propionate, acetate and butyrate in formula-fed infants has also been reported (Martin et al. 2014). Free AAs, such as phenylalanine, tyrosine, leucine and isoleucine (Hellmuth et al. 2016), as well as AAs catabolites, such as phenyllactate, 4-hydroxyphenylacetate, and 5-aminovalerate (Martin et al. 2014; He, et al. 2019), have been demonstrated to be higher in formula-fed infants relative to breastfed infants. The metabolic signature was reported to remain indicative of the feeding practice throughout first year of life (Martin et al. 2014). These results suggest that fecal metabolic profiling is a promising approach to differentiate EBF infants from non-EBF infants.

Although studies have characterized the impact of food intake on the infant fecal metabolome and microbiome of breastfed infants, no study has investigated if water intake influence the fecal

microbiome and metabolome. Water or other diluted liquid intake has been reported to cause water intoxication in infants, and thus is not recommended for infants under 6 months (Bruce and Kliegman 1997; Boussemart et al. 2006). It was also shown that water supplementation is not needed for EBF infants even in tropical area (Sachdev et al. 1991). However, evidence shows that ~25% of infants under 5 months are provided water (Demmer et al. 2018). Therefore, it is important to understand how water consumption impacts infants.

The present study targeted EBF and non-EBF infants in Indonesia and compared the fecal metabolome and microbiome between the feeding practices to identify metabolic biomarkers with predictive capability. Using a combination of the DTM technique and the weighed food record to determine exclusivity of breastfeeding, this study has the unique opportunity to distinguish non-EBF infants consuming water (non-EBF_water) and non-EBF infants consuming food (non-EBF_food) from EBF infants. This study provides additional insights into the potential impact of water intake on the infant gut microbiome.

Materials and methods

Mother–infant pairs were recruited from the Bandung municipality and Sumedang district area in West Java, Indonesia at post-natal age of 2 months and followed up at 5 months. Written informed consent was obtained from all participating mothers. The inclusion and exclusion criteria were described in detail previously (Leong et al. 2021). Ethical approval was obtained from the Health Research Ethics Committee, Faculty of Medicine, Universitas Padjadjaran, Bandung, Indonesia (05/UN6.C1.3.2/KEPK/PN/2017).

Breastfeeding status and maternal secretor status determination

The feeding practice (EBF or non-EBF) of infants was determined using the DTM method as previously described (Liu et al. 2019). Briefly, the recruited mothers were given oral deuterium oxide and the deuterium transition from mothers to their infants was measured. The water intake of the infants from non-breastmilk sources was estimated and compared to a cut off value of 86.6 g/d (Liu et al. 2019a) to decide if the infant was EBF or non-EBF. To assess the complementary food intake of the non-EBF infants, in-home weighed food records on three consecutive days at both 2 and 5 months were taken as previously described (Daniels et al. 2019). Daily percent energy intake from breast milk of the infants was calculated based on these food records. Non-EBF infants with 100% of daily energy intake from breast milk were assumed taking only water besides breast milk thus assigned as non-EBF_water, and non-EBF infants with less than 100% daily energy intake from breast milk were assumed consuming other food sources so they were classified as non-EBF_food.

Maternal secretor phenotype was determined based on the presence or near absence of 2'FL as described in Chapter 3. Briefly, secretor (Se+) status was assigned when milk 2'FL concentration was greater than 200 μM while non-secretor (Se-) status was assigned when milk 2'FL concentration was below 200 μM.

Fecal metabolite measurement

Fecal samples were collected from the diaper into Eppendorf tubes, which were placed in Bio-freeze bottle transport containers, transported to the laboratory and stored at -80 °C after collection. Samples were then shipped to UC Davis on dry ice and kept at -80 °C until metabolite and DNA extraction. Prior to metabolite extraction, fecal samples were thawed on ice and homogenized

manually with a sterile microspatula. Ice-cold Dulbecco's phosphate buffered saline (DPBS, 1X, pH 7.4) was used to extract polar metabolites as described previously (He et al. 2019), keeping the fecal pellet at 4 °C for DNA extraction after metabolite extraction. 207 μL of the extracted sample and 23 μL of internal standard (5 mM 3-(trimethylsilyl)-1-propanesulfonic acid-d6 (DSS-d6) in 99.8% D_2O (to serve as a lock) and 0.2% NaN_3) was mixed together and the pH was adjusted to reach 6.78 - 6.92 before data acquisition on NMR. ^{1}H NMR spectra were acquired at 298K using a NOESY experiment ('noesypr1d') on a Bruker Avance 600 MHz NMR spectrometer (Bruker BioSpin, Germany) equipped with a SampleJet autosampler (Bruker BioSpin, Germany) as previously described (He et al. 2019b). Metabolites was quantified via manual profiling in Chenomx Profiler as described previously (Weljie et al. 2006) after processing with Chenomx Processor (Chenomx NMR Suite v8.3, Chenomx Inc, Edmonton, Alberta, Canada).

To normalize the metabolite concentrations based on the water content of the fecal samples, water estimate of each sample was determined as described in chapter 3. Briefly ~75 mg of feces was analytically weighed, lyophilized (Labconco FreeZone 4.5 L Freeze Dry System, Labconco, Kansas City, MO), and the difference between the wet and dry weight was used to determine the amount of fecal water:

$$\text{Fecal water estimate} = \left(1 - \frac{Dry\ weight}{Wet\ weight}\right) \times Weight\ of\ extracted\ sample \quad (1)$$

Metabolite concentrations were corrected using the water estimate as described previously (He et al. 2019a).

Fecal microbial DNA extraction, library preparation and 16S rRNA gene sequencing

The fecal pellet saved from above step was used for DNA extraction based on the Human Microbiome Project (HMP) protocol (Wesolowska-Andersen et al. 2014) and our previous study (He et al. 2019a). The 16S rRNA gene V4 hypervariable region was targeted using a two-step PCR protocol as described in chapter 3. The first step was to amplify the V4 region using F515/R806 primers with an Illumina overhang sequence. The second step aimed to add sample-specific barcodes to multiplex the samples. The PCR reaction procedure for both steps included an initial denaturation at 95 °C for 5 min, a 10-cycle program of amplifying (a denaturation at 98 °C for 20 s, an annealing at 55 °C for 15 s, and a primer extension at 72 °C for 60 s), and a final extension step to allow DNA reannealing at 72 °C for 10 min.

Gel electrophoresis was conducted to quantify the PCR products by comparing the band intensity with a molecular ladder with known concentration (BioRad EZ ladder 1 kb) through ImageLab software (v5.2.1, BioRad, Hercules, CA). An equal molar amount of all the amplicons were pooled together and purified before quality checking via Bioanalyzer. The library was submitted to the UC Davis Genome Center DNA Technologies Core for 300 bp paired-end sequencing on the Illumina MiSeq platform.

Successfully demultiplexed sequencing data was then fed into the pipeline to trim adapters (*cutadapt* function) and check quality (*fastqc* and *multiqc* function). Filtering, trimming, and merging was conducted in DADA2 (version 1.12.1) with a maxEE of 2. The SILVA 16S rRNA database version 138 (Callahan et al. 2016; Quast et al. 2013; Yilmaz et al. 2014) was used to assign taxonomy and generate ASV table.

Statistical analysis

Statistical analysis was performed using R (4.0.3). The alpha-diversity analysis was conducted using the *phyloseq* package and the comparisons of microbes was performed using the *ANCOM-BC* package (Lin and Peddada 2020) using the ASV table generated from the DADA2 pipeline. Metabolomics data was generalized log transformed (defined as $log(y + 1)$) prior to analysis to ensure normality. All plots were generated via the *ggplot2* package.

The difference in microbial composition was measured using the analysis of composition of microbiome compositions with bias correction (ANCOMBC) and the significance level was set at $p < 0.05$ after the Benjamini-Hochberg false discovery rate (FDR) control procedure. The significance of metabolomics data between feeding practices was evaluated using the Mann-Whitney Test (*wilcox.test* function) with p-value adjustment (*p.adjust(, method = 'fdr')*). The overall level of significance was set at $p < 0.05$. Differences in metabolome data were also assessed using Cliff's delta (δ) effect size (*cliff.delta* function) with a 95% confidence interval. According to Romano *et al* (Romano et al. 2006), $|\delta|<0.147$ corresponds to negligible, $0.147<|\delta|<0.33$ corresponds to small, and $|\delta|>0.474$ large effect size.

Principal coordinate analysis (PCoA) was computed based on the Bray distance (for microbiome data) and Euclidian distance (for metabolome data) using the *pcoa* package. Centroids were calculated by averaging PC1 and PC2 for each group. The 95% confidence level constructed based on multivariate normal distribution was also calculated and plotted. The difference in beta-dispersion was tested using the *betadisper* function followed by TukeyHSD post-hoc analysis.

To generate the biomarker set, a Random Forest (RF) algorithm (*randomForest*) in R was used. The fecal metabolome data from infant subjects were randomly divided into two sets: a training

set, which contained 70% of the subjects, and a test set with the remaining 30% of the subjects. A similar ratio between feeding practices in each dataset was confirmed. The number of trees (ntree) was set to 4000, and the optimal number of metabolites at each RF split was calculated using *tuneRF* function. To select differentiating metabolites, backwards elimination via the *varSelRF* function, using discarding rate of 20%, was conducted. The selected metabolites were fit with the RF model to evaluate performance and calculate area under receiver operating curve (AUC). Model training and testing were computed in three different ways: EBF vs. non-EBF (combining non-EBF_water and non-EBF_food), EBF vs. non-EBF_water, and EBF vs. non-EBF_food.

Results

The present study aimed to compare the fecal metabolome and gut microbiome between EBF and non-EBF infants in an effort to identify metabolic markers to distinguish between feeding practices. EBF status was determined using the DTM method, and results from this study have been published previously for this study (Leong et al. 2021). Based on energy intake, non-EBF infants were divided into two categories: non-EBF with food intake (non-EBF_food, non-EBF with < 100% daily energy intake from breast milk) and non-EBF with water intake (non-EBF_water, non-EBF with 100% daily energy intake from breast milk). We compared infants in these two categories with EBF infants separately to characterize the potential change in infant fecal metabolome and gut microbiome caused by introducing water or complementary food. After data generation and cleaning, data from a total of 152 and 133 mother-infant pairs were included at 2 and 5 months of age, respectively.

Characteristics of the study population. As shown in Table 4.1, at 2 months, most infants (n = 128) were EBF, with a small number of infants non-EBF_water (n = 8) and non-EBF_food (n =

16). Some EBF infants switched to non-EBF at 5 months, making the number of infants in the EBF, non-EBF_water, and non-EBF_food groups 57, 46, and 30, respectively at this time point. Maternal age, body mass index (BMI) and secretor status were comparable among groups. Infant length, weight, and *z*-score were comparable among feeding groups at 2 months and 5 months. Breast milk intake at 2 months was significantly higher in EBF infants ($p < 0.05$) relative to infants in the two non-EBF groups. At 5 months, EBF infants had the highest breast milk intake, followed by non-EBF_water infants, and then non-EBF_food, with the difference between any two groups statistically significant ($p < 0.05$). At both 2 and 5 months, the introduction of water or food increased fever prevalence (at 2 months: 25.0% with water intake, 31.3% with food intake; at 5 months: 26.1% with water intake, 60.0% with food intake) when compared to EBF infants (16.4% at 2 months, and 22.8% at 5 months). At 2 months, compared to EBF infants (7.8%), diarrhea prevalence was slightly higher with the intake of water (12.5%) or food (12.5%). At 5 months, the diarrhea rate in EBF infants (7.0%) and non-EBF_water (8.7%) was comparable while the diarrhea rate in non-EBF_water infants was slightly higher (13.3%). Cough was more commonly seen in non-EBF_food infants (60%) when compared to EBF infants (35.1%) and non-EBF_water infants (26.1%) at 5 months.

Impact of complementary food introduction on the infant gut microbiome. To investigate the impact of feeding practices on the infant gut microbiome, alpha diversity, beta-diversity, and microbial relative abundance were compared among feeding groups at each time point. Overall, the infant gut microbiome had significantly higher Chao1 and Shannon diversity at 5 months when compared to 2 months. At 2 months, infants who were non-EBF_water had a significantly increased Chao1 diversity when compared to infants who were EBF ($p < 0.05$). Comparison of the Chao1 diversity between non-EBF_food and EBF infants was not significantly different at 2

months ($p = 0.7$) (Figure 4.1 a). In spite of this, the PCoA revealed a slight difference of the non-EBF_food group from the EBF group at 2 months (Figure 4.1 b). At 5 months, no separation among groups was observed through PCoA (Figure 4.1 c). Comparison of the relative abundance of various taxa revealed that *Collinsella* was significantly higher in non-EBF_water infants compared to EBF infants ($p < 0.05$), and *Staphylococcus* was significantly higher in EBF infants compared to non-EBF_food infants ($p < 0.05$) at 2 months (Figure 4.1 d). At 5 months, the relative abundance of *Akkermansia* was significantly higher in non-EBF_food infants compared to EBF infants ($p < 0.05$) (Figure 4.1 e).

Table 4.1. Characteristics of mother and infant subjects.

Characteristics	Infant age (months) and feeding practice					
	2			5		
	EBF	non-EBF_water	non-EBF_food	EBF[1]	non-EBF_water[2]	non-EBF_food[3]
Sample size, n	128	8	16	57	46	30
Mother						
Age, year	26.9 (6.3)	19.1(6.5)	31 (6.6)	27.4 (6.3)	26.3 (6.3)	27.8 (6.0)
BMI, kg/m^2	24.6 (3.6)	22.8 (3.2)	26.1 (4.3)	23.5 (3.4)	24.2 (3.9)	24.1 (4.2)
Secretor, %	83 (64.8%)	5 (62.5)%	12 (75.0%)	40 (70.2%)	31 (67.4%)	18 (60.0%)
Infant						
%energy from breast milk	100	100	59 (26.4)	100	100	64.8 (30.7)
Breast milk intake, mL/day	785 (170) [a]	530 (117)	356 (207)	809 (115) [b]	656 (136) [b]	486 (259) [b]
Sex, %female	60 (46.9%）	2 (25.0%)	7 (43.8%)	34 (59.6%)	15 (32.6%)	16 (53.3%)
Length, cm	55.9 (2.1)	57.8 (1.0)	56.3 (2.6)	62.8 (2.1)	63.0 (2.3)	62.9 (2.3)
Weight, kg	5.3 (0.6)	5.7 (0.3)	5.0 (0.7)	6.9 (0.8)	7.0 (0.8)	6.7 (0.8)
z score	-0.2 (1.0)	0.1 (0.6)	-0.9 (0.6)	0 (1.0)	0 (1.0)	-0.4 (1.2)
Vomit, n (%)[#]	4 (3.1%)	0	1 (6.3%)	1 (1.8%)	3 (6.5%)	0
Fever, n (%)[#]	21 (16.4%)	2 (25.0%)	5 (31.3%）	13 (22.8%)	12 (26.1%)	18 (60.0%)
Diarrhea, n (%)[#]	10 (7.8%)	1 (12.5%)	2 (12.5%)	4 (7.0%)	4 (8.7%)	4 (13.3%)
Cough, n (%)[#]	29 (22.7%)	2 (25.0%)	5 (31.25%)	20 (35.1%)	12 (26.1%)	18 (60%)
Tachypnea, n (%)[#]	5 (3.9%)	1 (12.5%)	0	4 (7.0%)	1 (2.2%)	3 (10.0%)

Mother age, BMI (body mass index), breast milk intake, infant length, infant weight, and infant *z*-score are shown as mean (standard deviation).
[#]Morbidity data reflects the prevalence of disease over a two week time period before sample collection.
[1]All EBF infants at 5 months were EBF at 2 months. [2] In non-EBF_water group at 5 months, 41 infants were in EBF, and 5 infants were in non-EBF_water groups at 2 months. [3]In non-EBF_food group at 5 months, 19 infants were in EBF, 1 infant was in non-EBF_water, and 10 infants were in non-EBF_food groups at 2 months. [a]$p < 0.05$ between EBF group and non-EBF_water, and between EBF and non-EBF_water groups assessed by Dunn's test, [b]$p < 0.05$ between any two groups.

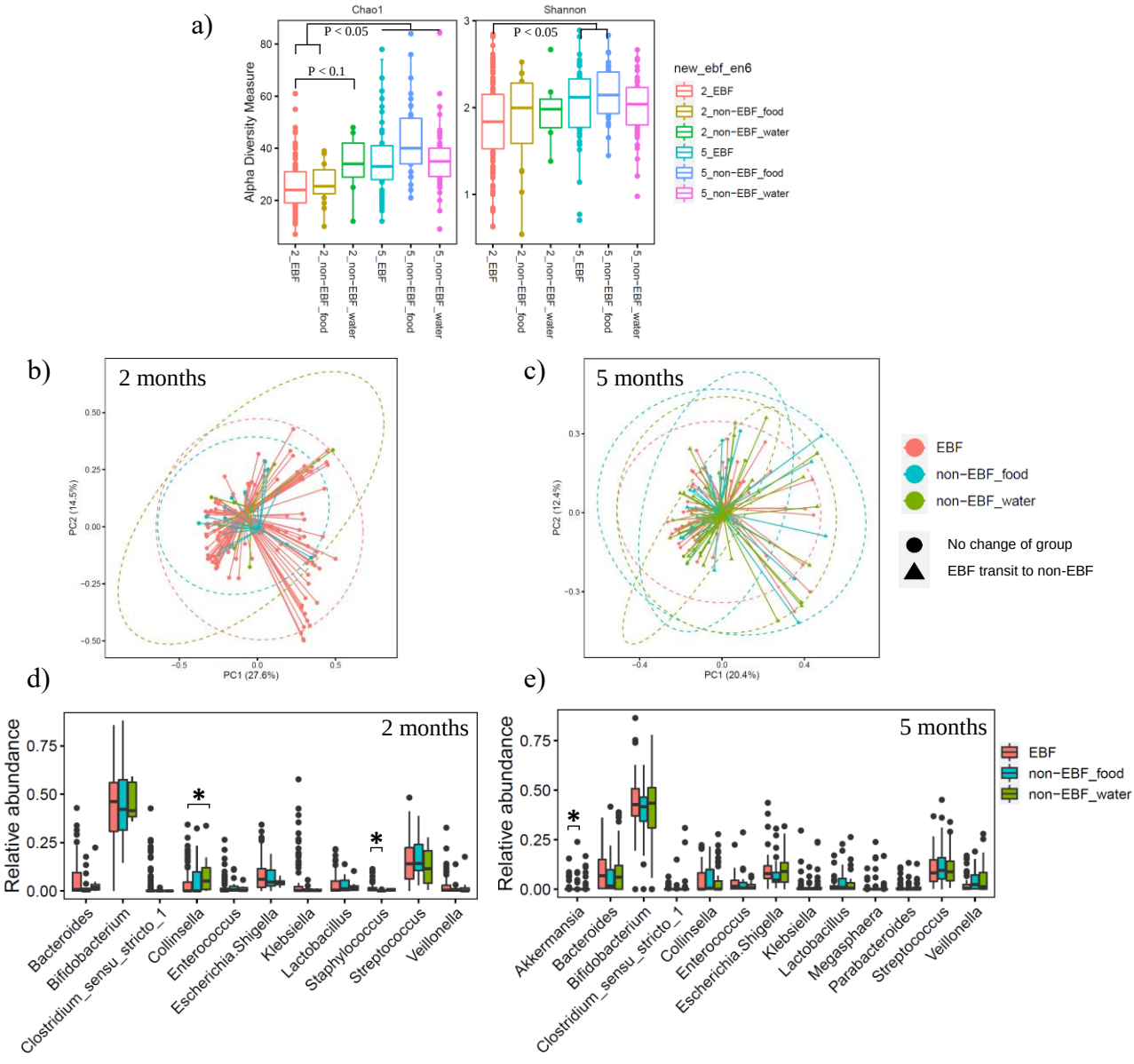


Figure 4.1. Impact of infant feeding mode at 2 and 5 months on the fecal microbiome. a) alpha diversity analysis (Chao1 and Shannon diversity), b) PCoA (principal coordinate analysis) at 2 months, c) PCoA at 5 months, d) comparison of the average relative abundance for genera over 1% prevalence at 2 months, e) comparison of the average relative abundance for genera over 1% prevalence at 5 months. * $p < 0.05$ in ANCOM (analysis of composition of microbiomes).

Impact of complementary food introduction on the infant fecal metabolome. To characterize the impact of complementary food introduction on infant gut microbial function, the fecal metabolome was analyzed. At 2 months, PCoA revealed that the non-EBF_water group significantly differed from the EBF group ($p < 0.05$) while the difference between the non-EBF_food group and EBF group trended significant ($p = 0.08$) (Figure 4.2 a). For infants provided other food, the infant fecal metabolome revealed higher levels of short chain fatty acids (SCFAs) (valerate ($p < 0.05$, $\delta = -0.65$), acetate ($p < 0.05$, $\delta = -0.38$), and propionate ($p < 0.05$, $\delta = -0.36$)), the amino acid fermentation product 4-hydroxyphenylacetate ($p < 0.05$, $\delta = -0.46$), nucleotides (2'-deoxyuridine ($p < 0.05$, $\delta = -0.45$), thymine ($p < 0.05$, $\delta = -0.40$), and uracil ($p < 0.05$, $\delta = -0.38$)), and a lower level of human milk oligosaccharide (HMO) related structures including lactodifucotetraose (LDFT, $p < 0.05$, $\delta = 0.40$), 2'FL ($p < 0.05$, $\delta = 0.42$), and 3-fucosyllactose (3-FL, $p < 0.05$, $\delta = 0.48$) (Figure 4.2 c). In fecal samples of infants who consumed only water in addition to breast milk, greater concentrations of SCFAs such as valerate ($\delta = -0.37$) and acetate ($\delta = -0.35$), nucleotides, including thymine ($\delta = -0.43$) and uracil ($\delta = -0.55$), and a lower level of 2'FL ($\delta = 0.43$), 3-FL ($\delta = 0.52$), lacto-N-fucopentose I (LNFP I, $\delta = 0.39$) and 6'-sialyllactose (6'SL, $\delta = 0.48$) were also observed when compared to EBF infants based on Cliff's delta effect size. Additionally, a broad range of amino acids and their derivatives (including tryptophan ($\delta = -0.40$), pyroglutamate ($\delta = -0.40$), proline ($\delta = -0.57$), phenylacetate ($\delta = -0.36$), lysine ($\delta = -0.34$), isoleucine ($\delta = -0.41$), glycine ($\delta = -0.41$), glutamine ($\delta = -0.33$), alanine ($\delta = -0.50$), 4-aminobutyrate ($\delta = -0.35$), phenyllactate ($\delta = -0.41$), 2-oxoglutarate ($\delta = -0.49$) and 2-hydroxybutyrate ($\delta = -0.40$)) were higher in fecal samples from non-EBF_water infants when compared to EBF infants with a medium to large effect size. However, these metabolites were not significant using the Mann-Whitney test after FDR correction. At 5 months, the fecal metabolome

of infants fed other foods in addition to breast milk had a higher level of propionate ($p < 0.05$, $\delta = -0.34$) and lower levels of malate ($p < 0.05$, $\delta = 0.33$), LDFT ($p < 0.05$, $\delta = 0.34$), 2-hydroxyisobutyrate ($p < 0.05$, $\delta = 0.39$), fucose ($p < 0.05$, $\delta = 0.43$), and citrate ($p < 0.05$, $\delta = 0.45$). Infants in the non-EBF_water group did not have a different infant fecal metabolome from the EBF group at 5 months (Figure 4.2 d). No impact of maternal secretor phenotype on the infant fecal metabolomes were observed at either timepoint (Supplementary Figure 4.1).

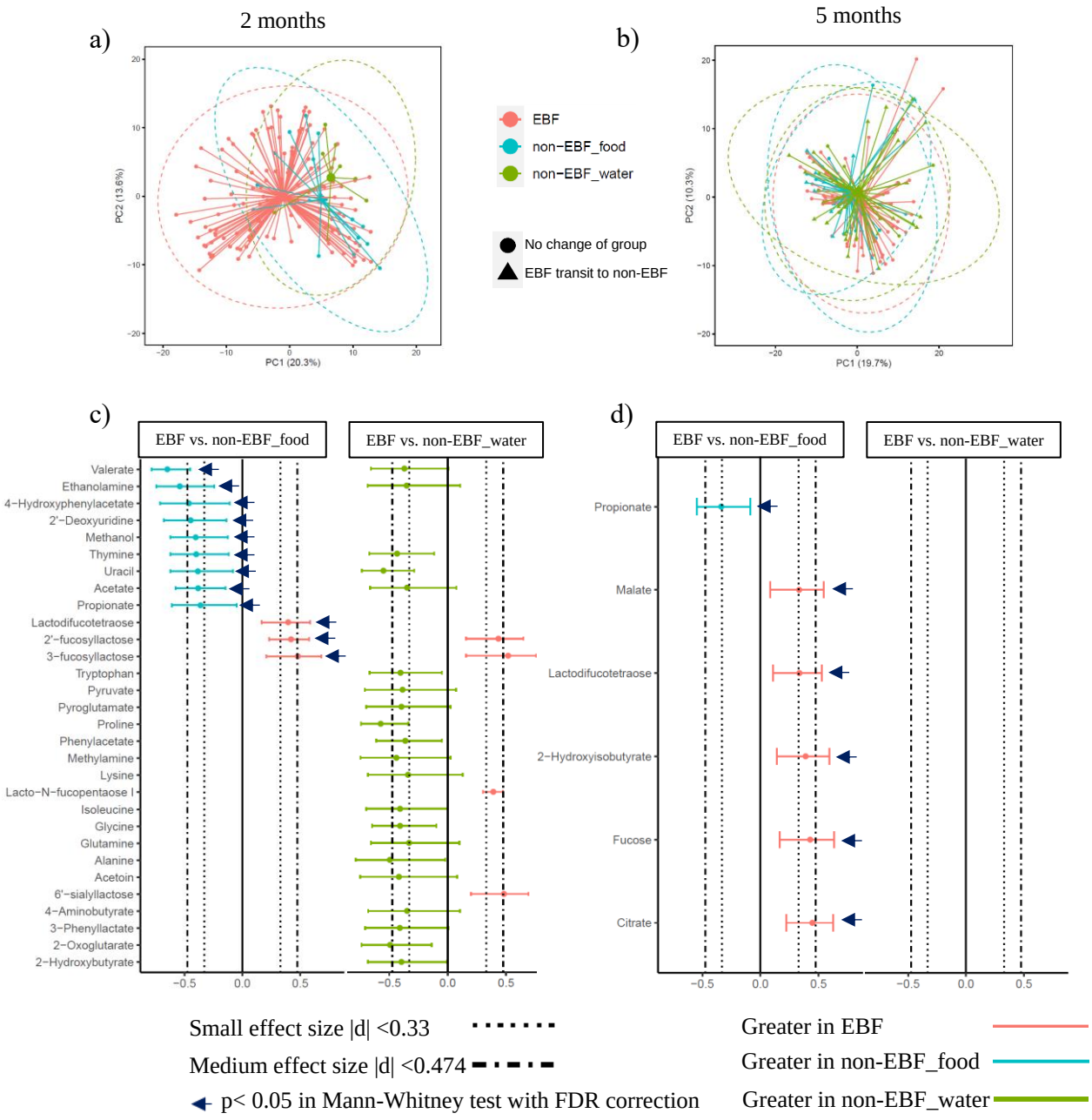


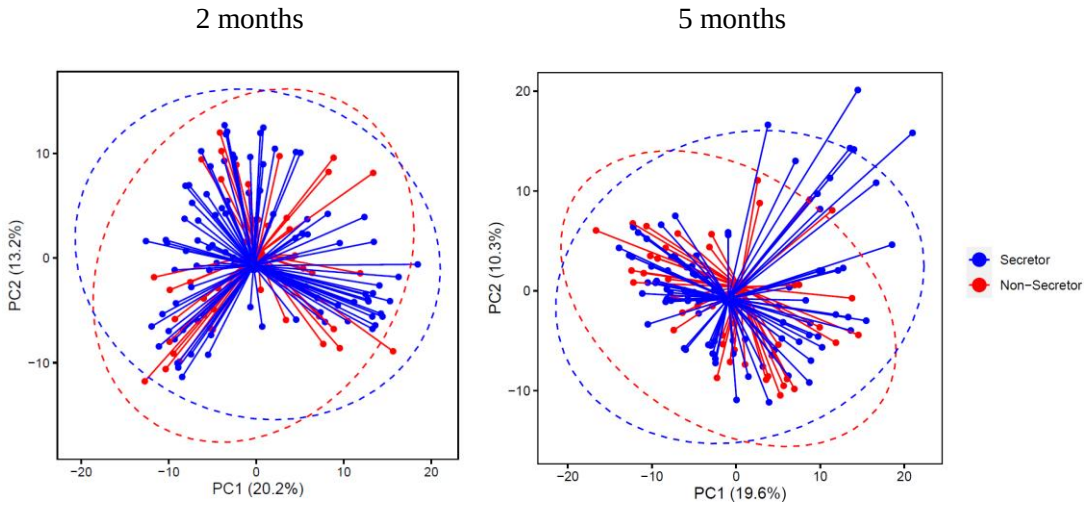
Figure 4.2. Impact of infant feeding mode at 2 and 5 months on the fecal metabolome. a) PCoA at 2 months b) PCoA at 5 months c) Cliff's delta effect sizes at 2 months, d) Cliff's delta effect at 5 months. Arrows indicate $p < 0.05$ assess with the Mann Whitney U test.

Prediction of feeding practice based on the infant fecal metabolome. To investigate if the infant fecal metabolome can predict feeding practice, infant fecal metabolome data were used as input to develop random forest (RF) models in three different ways: EBF vs. non-EBF (combining non-EBF_water and non-EBF_food), EBF vs. non-EBF_water, and EBF vs. non-EBF_food (Table 4.2). Under each classification scenario, the differentiating metabolites were selected via the backwards selection algorithm using the training data set. Area under the receiver operating curve (AUC) analysis was used to evaluate the model performance using the test set. Prediction between EBF and non-EBF_food achieved the highest AUC (0.84), while the prediction between EBF and non-EBF_water had with the lowest AUC (0.79). Among the 97 metabolites measured, the major differentiating metabolites included HMO related structures (6'SL, 3'sialylactose, 2'FL, fucose, sialic acid), SCFAs (valerate, propionate, acetate), organic acids (malate, citrate), and amino acid fermentation products (2-hydroxyisobutyrate, 3-phenyllactate).

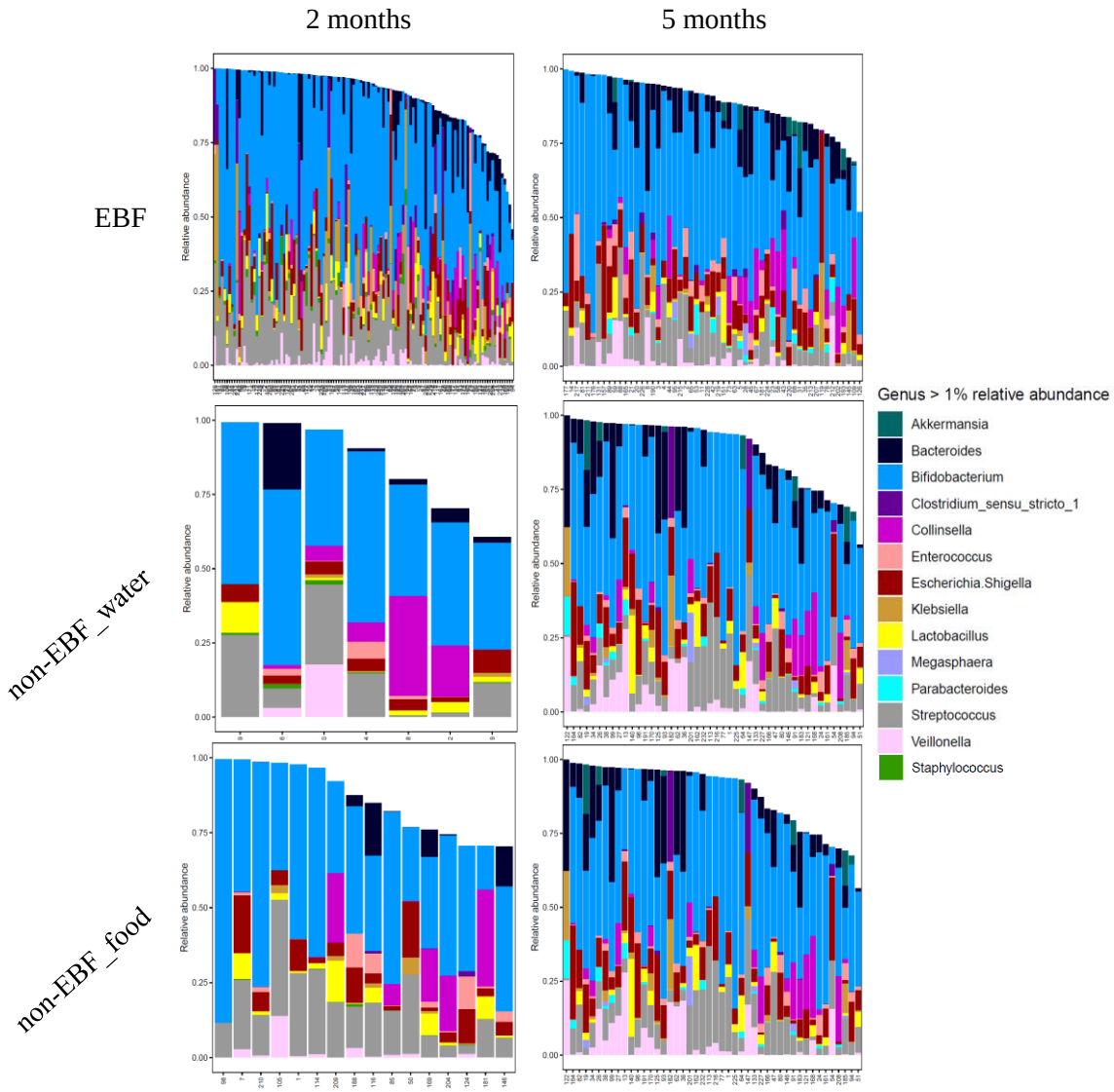
Table 4.2. Predictive performance of selected metabolic biomarkers.

Classification	AUC	Sensitivity	Specificity	Features selected
EBF vs. **non-EBF**	0.80 (0.04)	0.81 (0.05)	0.60 (0.1)	Infant age, 6'SL, 3'SL, 2'FL, Valerate, Propionate, Malate, Citrate, Methanol
EBF vs. **non-EBF_food**	0.84 (0.04)	0.84 (0.05)	0.63 (11)	Valerate, 3'SL, Malate, Propionate, Methanol, Acetate, 2-Hydroxyisobutyrate
EBF vs. **non-EBF_water**	0.79 (0.05)	0.78 (0.07)	0.61 (0.13)	Infant age, Fucose, Taurine, Sialic acid, Methionine, Methylsuccinate, Glucose, Methanol, 3-Phenyllactate

AUC, area under receiver operator characteristics curve; EBF, exclusively breastfed; non-EBF_food, non-exclusively breastfed with daily percent energy intake from breast milk < 100%; non-EBF_water, non-exclusively breastfed with water intake and daily percent energy intake from breast milk = 100%; AUC is shown as mean (standard deviation) of 20 runs.



Supplementary Figure 4.1. Impact of maternal secretor status on the infant fecal metabolome at 2 and 5 months.



Supplementary Figure 4.2. Comparison of the relative abundance of genera measured in feces of EBF infants, non-EBF_water infants, and non-EBF_food infants. Shown are microbes with over 1% relative abundance.

Discussion

Several studies have reported that the introduction of complementary food alters the infant gut microbiome and fecal metabolome (He, et al. 2019; Jiménez et al. 2008; Ma et al. 2020). Unfortunately, these studies utilize parental recall methods to determine infant feeding practice. It has been shown in many studies that there is a discrepancy between parental recall and exclusive breastfeeding (Medoua et al. 2012; Mazariegos, Slater, and Ramirez-Zea 2016). Moreover, provision of water to infants is seldom recorded in such studies. Here, a large group of mother-infant pairs whose breastfeeding practice (EBF or non-EBF) was objectively determined via the DTM technique were recruited to investigate the impact of complementary food (non-EBF_food) or water (non-EBF_water) on the infant gut microbiome and fecal metabolome, and to potentially develop metabolic biomarkers to discriminate between infant feeding practices. To our knowledge, this is the first study to show the impact of providing water on the gut microbiome and metabolome of exclusively breastfed infants.

In agreement with previous studies (Koenig et al. 2011; Ho et al. 2018; Hill et al. 2017), we observed an overall increase in alpha-diversity from 2 months to 5 months (Figure 4.1 a). For infants at 2 months of age, the increased Chao1 richness in the feces of the non-EBF_water infants compared to EBF infants (Figure 4.1 a) suggests that water intake introduced microbes to the infants thereby increasing species richness. Water intake is not recommended for infants younger than 6 months, as impure water could contain bacterial pathogens (Balbus and Lang 2001) and the risk of water intoxication (Boussemart et al. 2006). *Collinsella,* reported to be higher in formula-fed infants when compared to breastfed infants in our previous study (He, et al. 2019), was significantly higher in non-EBF_water infants when compared to EBF infants at 2 months (Figure 4.1 c) in the present study. *Collinsella* could be a contaminant in drinking water (Cabral 2010).

Evidence has also shown that this microbe is elevated in the gut microbiome of infants exposed to lead (Pb) (Sitarik et al. 2020), a possible contaminant in drinking water from water pipes containing lead (Balbus and Lang 2001). No data were collected in this study regarding the presence of lead in the drinking water, so further studies investigating the association between elevated intestinal *Collinsella* and lead contamination in water in this region of Peru may be warranted. The decreased relative abundance of *Staphylococcus* at 2 months in non-EBF_food infants compared to EBF infants may be attributed to the potential reduction of breast milk intake (Table 4.1), as some *Staphylococcus* species in the infant gut have been hypothesized to come from breast milk (Martín et al. 2007). A lower relative abundance of *Staphylococcus* in formula-fed infants when compared to breastfed infants has also been reported in previous studies (He, et al. 2019; Jiménez et al. 2008). Although some genera showed differences in relative abundance between feeding practices, no genera were identified by a random forest (RF) model as important features to discriminate between feeding practices, even if both metabolite and relative abundance data were used together as one large data set (data not shown). We speculate that this is because the difference in the microbiome between feeding practices was not as profound as in the metabolome, since all the infants were consuming at least 59% of their calories as breast milk (Table 4.1), and breast milk will dominate bacterial selection for colonization of the gastrointestinal tract.

Using 70% of the infant fecal metabolome data, metabolites differentiating EBF from non-EBF were identified (using a combination of 2 and 5 months data) via the varSelRF package, and the validity of the model was subsequently tested using the remaining 30% of data to obtain the AUC (Table 4.2). The profile of metabolites with discriminate capability included HMOs, SCFAs, AAs and their derivatives (Table 4.2).

Higher concentrations of some of the HMOs (2'FL, 3-FL, LDTF, LNFP I, and 6'SL, Figure 4.2 c, d) in the EBF group at both 2 and 5 months can be attributed to a higher breast milk intake in EBF infants when compared to the non-EBF groups at both ages (Table 4.1). This difference in HMO concentrations made them good predictors for infant feeding practices. Specifically, 6'SL, 3'SL, and 2'FL were selected when discriminating between EBF and all non_EBF infants, or between EBF and non_EBF_food infants (Table 4.2). The impact of consuming food or water on infants in this study was to essentially reduce the amount of breast milk consumed (Table 4.1). It stands to reason, then, that reduced breast milk intake results in a lower HMO intake and thus a lower concentration of HMOs in the feces of the non-EBF groups. However, the concentration of HMOs in the feces could be altered by the presence of the major HMO-utilizer, *Bifidobacterium* (Underwood et al. 2014). Here, *Bifidobacterium* relative abundance was comparable across feeding practices at both 2 and 5 months (Figure 4.1 d, e); however, the relative abundance of specific HMO-utilizing strains was not measured.

SCFAs in feces were previously reported to be different between breastfed and formula-fed infants due to colonic microbial fermentation preferences between carbohydrates and protein ((Bridgman et al. 2017; He, et al. 2019; Chow et al. 2014). In the present study, valerate, propionate, and acetate were all shown to have predictive capability between EBF and non-EBF, or between EBF and non-EBF_food infants (Table 4.2), indicating a shift in fermentation of the gut microbes when complementary food was introduced. Valerate was significantly higher in infants from the non-EBF_food group when compared to the EBF group at 2 months (Figure 4.2 c). Infant formulas are known to contain higher protein than breast milk (Macé et al. 2006), and the excess protein that escapes absorption could be fermented by gut microbes to produce valerate. Valerate was previously reported to be higher in the gut of partially breastfed and formula-fed infants compared

to EBF infants (Bridgman et al. 2017; He, et al. 2019). Interestingly, a higher level of valerate was also observed in non-EBF_water infants (Figure 4.2 c) at 2 months. We speculate that this observation may be because water intake might interfere with the digestion of proteins from breast milk due to dilution of both proteins and digestive enzymes in the gastrointestinal tract. This could result more protein escaping digestion and entering into the colon to be fermented by the microbes. Propionate was significantly higher in non-EBF_food infants when compared to EBF infants at both 2 and 5 months (Figure 4.2 c, d). Microbial pathways that produce propionate from AAs, organic acids and carbohydrates have been reviewed previously (Louis and Flint 2017). For infants in the non-EBF_food group, propionate could be derived from different precursors, and thus explain the higher level of propionate (Wang et al. 2019). Acetate, higher in both non-EBF groups at 2 months (Figure 4.2 c), could be produced by different microbes from acetogenic fibers, residual peptides, and fats (Koh et al. 2016; Wong et al. 2006). Acetate has been shown to be higher in formula-fed or partially breastfed infants when compared to EBF infants (Bridgman et al. 2017), and we observed this in the present study. It has been proposed that the fecal acetate could be an indicator of absorption of acetate by the host instead of the microbial production of acetate since a negative association was observed between fecal acetate and acetate infusion in the colon (Vogt and Wolever 2003). Indeed, higher circulating acetate has been observed in breastfed infants compared to formula-fed infants in our previous studies (He, et al. 2019; Slupsky et al. 2017).

In the current study, AA derivatives were another set of metabolites selected by the RF model as biomarkers to discriminate between feeding practices. 2-Hydroxyisobutyrate and 3-phenyllactate were identified as biomarkers distinguishing between EBF and non-EBF_food infants, and between EBF and non-EBF_water infants, respectively (Table 4.2). Higher levels of AA

derivatives were observed in formula-fed infants when compared to breastfed infants in previous studies and were attributed to an increased breakdown of AAs (Kirchberg et al. 2015; He, et al. 2019). 2-Hydroxyisobutyrate is a valine degradation product by gut microbes and was found to be higher in breastfed infants in both the present study (Figure 4.2 d) and our previous study (H. Lee et al. 2021). 3-Phenyllactate and tryptophan were both higher in non-EBF_water infants relative to EBF infants at 2 months (Figure 4.2 c). A previous study reported that 3-phenyllactate could be produced by *Bifidobacterium* from aromatic AAs (Laursen et al. 2020). Thus we hypothesize that the higher concentration of tryptophan in non-EBF_water infants at 2 months could be utilized to produce 3-phenyllactate.

Infant age was identified by the RF model as an important feature when predicting feeding practices only in the situation when non-EBF_water infants were included (Table 4.2, EBF vs. non-EBF, EBF vs. non-EBF_water). This could be because water intake altered the infant fecal metabolome at 2 months but not at 5 months (Table 4.2 c, d), and thus at different ages, the metabolic markers cooperate with each other differently to discriminate between feeding practices. Infant intestinal microbial community and metabolic function gradually matures during the first year (Yatsunenko et al. 2012). Previous studies have shown that the major resident, *Bifidobacterium*, does not dominate the microbiome until approximately 1 to 2 months of age in Armenian and American infants (Lewis et al. 2017) (Baumann-Dudenhoeffer et al. 2018). A separate study in Japan reported that 21% of infants had *Bifidobacteria* measurable in their feces soon after birth, with the number of infants having *Bifidobacteria* in their feces increasing to 97% at the age of 6 months (Nagpal et al. 2017). We observed that at 2 months, fewer infants had significant levels of *Bifidobacterium* in their gut than at 5 months (Supplementary Figure 4.2), suggesting that the infant gut could be more vulnerable at this age.

Conclusion

The results obtained in this study revealed the impact of non-exclusive breastfeeding through incorporation of water or foods other than breast milk on the infant fecal metabolome and microbiome at 2 and 5 months. In agreement with previous studies (Dotz et al. 2016; Chow et al. 2014; Tao et al. 2011; Bridgman et al. 2017; Martin et al. 2014; He, et al. 2019), the introduction of food other than breast milk altered the infant fecal metabolome and microbiome. The impact of water intake on infant metabolism has not been reported previously. The present study revealed that water intake had a more profound impact on the metabolome at 2 months compared to 5 months. A profile of metabolic markers consisting of HMOs, SCFAs, and AAs was identified with excellent predictive capability to discriminate between EBF and non-EBF feeding practices. One limitation of this study is the imbalanced sample size among the feeding practice groups, in particular, the sample size of the non-EBF groups at 2 months was small. Further studies with larger sample sizes should be conducted to validate the metabolic markers identified in the present study.

References

WHO. Exclusive Breastfeeding for Optimal Growth, Development and Health of Infants. Geneva: World Health Organization; 2019.

Balbus, John M., and Michael E. Lang. 2001. “Is The Water Safe For My Baby?” *Pediatric Clinics of North America* 48 (5): 1129–52. doi: 10.1016/S0031-3955(05)70365-5.

Baumann-Dudenhoeffer, Aimee M., Alaric W. D’Souza, Phillip I. Tarr, Barbara B. Warner, and Gautam Dantas. 2018. “Infant Diet and Maternal Gestational Weight Gain Predict Early Metabolic Maturation of Gut Microbiomes.” *Nature Medicine* 24 (12): 1822–29. doi: 10.1038/s41591-018-0216-2.

Boussemart, T, V Flurin, M-C Chevalier, A Godoy-Chaabane, and J-M Bonardi. 2006. “Dietary water intoxication: a preventable pathology.” *Arch Pediatr* 13 (7): 1050–52. doi: 10.1016/j.arcped.2006.02.011.

Bridgman, Sarah L., Meghan B. Azad, Catherine J. Field, Andrea M. Haqq, Allan B. Becker, Piushkumar J. Mandhane, Padmaja Subbarao, et al. 2017. “Fecal Short-Chain Fatty Acid Variations by Breastfeeding Status in Infants at 4 Months: Differences in Relative versus Absolute Concentrations.” *Frontiers in Nutrition* 4: 11. doi: 10.3389/fnut.2017.00011.

Bruce, Robert C., and Robert M. Kliegman. 1997. “Hyponatremic Seizures Secondary to Oral Water Intoxication in Infancy: Association with Commercial Bottled Drinking Water.” *Pediatrics* 100 (6): e4. doi: 10.1542/peds.100.6.e4.

Cabral, João P S. 2010. “Water Microbiology. Bacterial Pathogens and Water.” *International Journal of Environmental Research and Public Health* 7 (10): 3657–3703. doi: 10.3390/ijerph7103657.

Callahan, Benjamin J, Paul J McMurdie, Michael J Rosen, Andrew W Han, Amy Jo A Johnson, and Susan P Holmes. 2016. “DADA2: High-Resolution Sample Inference from Illumina Amplicon Data.” *Nature Methods* 13 (7): 581–83. doi: 10.1038/nmeth.3869.

Chow, JoMay, Matthew R. Panasevich, Danny Alexander, Brittany M. Vester Boler, Mariana C. Rossoni Serao, Trevor A. Faber, Laura L. Bauer, and George C. Fahey. 2014. “Fecal Metabolomics of Healthy Breast-Fed versus Formula-Fed Infants before and during In Vitro Batch Culture Fermentation.” *Journal of Proteome Research* 13 (5): 2534–42. doi: 10.1021/pr500011w.

Daniels, Lisa, Rosalind S. Gibson, Aly Diana, Jillian J. Haszard, Sofa Rahmannia, Dimas E. Luftimas, Daniela Hampel, et al. 2019. “Micronutrient Intakes of Lactating Mothers and Their Association with Breast Milk Concentrations and Micronutrient Adequacy of Exclusively Breastfed Indonesian Infants.” *The American Journal of Clinical Nutrition* 110 (2): 391–400. doi: 10.1093/ajcn/nqz047.

Demmer, Elieke, Christopher J. Cifelli, Jenny A. Houchins, and Victor L. Fulgoni. 2018. “Ethnic Disparities of Beverage Consumption in Infants and Children 0–5 Years of Age; National Health and Nutrition Examination Survey 2011 to 2014.” *Nutrition Journal* 17 (1): 78. doi: 10.1186/s12937-018-0388-0.

Dotz, Viktoria, Rüdiger Adam, Günter Lochnit, Horst Schroten, and Clemens Kunz. 2016. “Neutral Oligosaccharides in Feces of Breastfed and Formula-Fed Infants at Different Ages.” *Glycobiology* 26 (12): 1308–16. doi: 10.1093/glycob/cww087.

Frese, Steven A., Andra A. Hutton, Lindsey N. Contreras, Claire A. Shaw, Michelle C. Palumbo, Giorgio Casaburi, Gege Xu, et al. 2017. “Persistence of Supplemented Bifidobacterium Longum Subsp. Infantis EVC001 in Breastfed Infants.” *MSphere* 2 (6). doi: 10.1128/mSphere.00501-17.

He, Xuan, Mariana Parenti, Tove Grip, Magnus Domellöf, Bo Lönnerdal, Olle Hernell, Niklas Timby, and Carolyn M. Slupsky. 2019. “Metabolic Phenotype of Breast-Fed Infants, and Infants Fed Standard Formula or Bovine MFGM Supplemented Formula: A Randomized Controlled Trial.” *Scientific Reports* 9 (1): 339. doi: 10.1038/s41598-018-36292-5.

He, Xuan, Mariana Parenti, Tove Grip, Bo Lönnerdal, Niklas Timby, Magnus Domellöf, Olle Hernell, and Carolyn M. Slupsky. 2019. “Fecal Microbiome and Metabolome of Infants Fed Bovine MFGM Supplemented Formula or Standard Formula with Breast-Fed Infants as Reference: A Randomized Controlled Trial.” *Scientific Reports* 9 (1): 11589. doi: 10.1038/s41598-019-47953-4.

Hellmuth, Christian, Olaf Uhl, Franca F. Kirchberg, Veit Grote, Martina Weber, Peter Rzehak, Clotilde Carlier, et al. 2016. "Effects of Early Nutrition on the Infant Metabolome." *Nestle Nutrition Institute Workshop Series* 85: 89–100. doi: 10.1159/000439491.

Hill, Cian J., Denise B. Lynch, Kiera Murphy, Marynka Ulaszewska, Ian B. Jeffery, Carol Anne O'Shea, Claire Watkins, et al. 2017. "Evolution of Gut Microbiota Composition from Birth to 24 Weeks in the INFANTMET Cohort." *Microbiome* 5 (1): 4. doi: 10.1186/s40168-016-0213-y.

Ho, Nhan T., Fan Li, Kathleen A. Lee-Sarwar, Hein M. Tun, Bryan Brown, Pia S. Pannaraj, Jeffrey M. Bender, et al. 2018. "Effects of Exclusive Breastfeeding on Infant Gut Microbiota: A Meta-Analysis across Studies and Populations." *BioRxiv*, January, 292755. doi: 10.1101/292755.

Jiménez, Esther, Susana Delgado, Antonio Maldonado, Rebeca Arroyo, Mar Albújar, Natalia García, Manel Jariod, Leonides Fernández, Adolfo Gómez, and Juan M. Rodríguez. 2008. "Staphylococcus Epidermidis: A Differential Trait of the Fecal Microbiota of Breast-Fed Infants." *BMC Microbiology* 8 (1): 143. doi: 10.1186/1471-2180-8-143.

Kirchberg, Franca F., Ulrike Harder, Martina Weber, Veit Grote, Hans Demmelmair, Wolfgang Peissner, Peter Rzehak, et al. 2015. "Dietary Protein Intake Affects Amino Acid and Acylcarnitine Metabolism in Infants Aged 6 Months." *The Journal of Clinical Endocrinology and Metabolism* 100 (1): 149–58. doi: 10.1210/jc.2014-3157.

Koenig, Jeremy E., Aymé Spor, Nicholas Scalfone, Ashwana D. Fricker, Jesse Stombaugh, Rob Knight, Largus T. Angenent, and Ruth E. Ley. 2011. "Succession of Microbial Consortia in the Developing Infant Gut Microbiome." *Proceedings of the National Academy of Sciences* 108 (Supplement 1): 4578. doi: 10.1073/pnas.1000081107.

Koh, Ara, Filipe De Vadder, Petia Kovatcheva-Datchary, and Fredrik Bäckhed. 2016. "From Dietary Fiber to Host Physiology: Short-Chain Fatty Acids as Key Bacterial Metabolites." *Cell* 165 (6): 1332–45. doi: 10.1016/j.cell.2016.05.041.

Laursen, Martin F., Mikiyasu Sakanaka, Nicole von Burg, Urs Mörbe, Daniel Andersen, Janne Marie Moll, Ceyda T. Pekmez, et al. 2020. "Breastmilk-Promoted Bifidobacteria Produce Aromatic Amino Acids in the Infant Gut." *BioRxiv*, January, 2020.01.22.914994. doi: 10.1101/2020.01.22.914994.

Lee, Hanna, Carolyn M. Slupsky, Anne B. Heckmann, Britt Christensen, Yongmei Peng, Xiaonan Li, Olle Hernell, Bo Lönnerdal, and Zailing Li. 2021. "Milk Fat Globule Membrane as a Modulator of Infant Metabolism and Gut Microbiota: A Formula Supplement Narrowing the Metabolic Differences between Breastfed and Formula-Fed Infants." *Molecular Nutrition & Food Research* 65 (3): 2000603. doi: 10.1002/mnfr.202000603.

Lee, Mi K., and Colin Binns. 2020. "Breastfeeding and the Risk of Infant Illness in Asia: A Review." *International Journal of Environmental Research and Public Health* 17 (1). doi: 10.3390/ijerph17010186.

Leong, Claudia, Rosalind S Gibson, Aly Diana, Jillian J Haszard, Sofa Rahmannia, Mohammad Brachim Ansari, Lina Sofiatul Inayah, Afini Dwi Purnamasari, and Lisa A Houghton. 2021. "Differences in Micronutrient Intakes of Exclusive and Partially Breastfed Indonesian Infants from Resource-Poor Households Are Not Accompanied by Differences in Micronutrient Status, Morbidity, or Growth." *The Journal of Nutrition* 151 (3): 705–15. doi: 10.1093/jn/nxaa381.

Lewis, Zachery T, Ketevan Sidamonidze, Vardan Tsaturyan, David Tsereteli, Nika Khachidze, Astghik Pepoyan, Ekaterine Zhgenti, et al. 2017. "The Fecal Microbial Community of Breast-Fed Infants from Armenia and Georgia." *Scientific Reports* 7 (February): 40932–40932. doi: 10.1038/srep40932.

Liang, Xiaona, Hongjiao Han, Xue Zhao, Xueyan Cao, Mei Yang, Dongbing Tao, Rina Wu, and Xiqing Yue. 2018. "Quantitative Analysis of Amino Acids in Human and Bovine Colostrum Milk Samples through ITRAQ Labeling." *Journal of the Science of Food and Agriculture* 98 (13): 5157–63. doi: 10.1002/jsfa.9032.

Lin, Huang, and Shyamal Das Peddada. 2020. "Analysis of Compositions of Microbiomes with Bias Correction." *Nature Communications* 11 (1): 3514. doi: 10.1038/s41467-020-17041-7.

Liu, Zheng, Aly Diana, Christine Slater, Thomas Preston, Rosalind S. Gibson, Lisa Houghton, and Stephen B. Duffull. 2019a. “Development of a Nonlinear Hierarchical Model to Describe the Disposition of Deuterium in Mother-Infant Pairs to Assess Exclusive Breastfeeding Practice.” *Journal of Pharmacokinetics and Pharmacodynamics* 46 (1): 1–13. doi: 10.1007/s10928-018-9613-x.

Liu, Zheng, Aly Diana, Christine Slater, Thomas Preston, Rosalind S. Gibson, Lisa Houghton, and Stephen B. Duffull. 2019b. “Development of a Parsimonious Design for Optimal Classification of Exclusive Breastfeeding.” *CPT: Pharmacometrics & Systems Pharmacology* 8 (8): 596–605. doi: 10.1002/psp4.12428.

Louis, Petra, and Harry J. Flint. 2017. “Formation of Propionate and Butyrate by the Human Colonic Microbiota.” *Environmental Microbiology* 19 (1): 29–41. doi: 10.1111/1462-2920.13589.

Ma, Jingran, Zhenghong Li, Wenjuan Zhang, Chunli Zhang, Yuheng Zhang, Hua Mei, Na Zhuo, Hongyun Wang, Lin Wang, and Dan Wu. 2020. “Comparison of Gut Microbiota in Exclusively Breast-Fed and Formula-Fed Babies: A Study of 91 Term Infants.” *Scientific Reports* 10 (1): 15792. doi: 10.1038/s41598-020-72635-x.

Macé, Katherine, Philippe Steenhout, Petra Klassen, and Anne Donnet. 2006. “Protein Quality and Quantity in Cow’s Milk-Based Formula for Healthy Term Infants: Past, Present and Future.” *Nestle Nutrition Workshop Series. Paediatric Programme* 58: 189–203; discussion 203-205. doi: 10.1159/000095063.

Martin, François-Pierre J., Sofia Moco, Ivan Montoliu, Sebastiano Collino, Laeticia Da Silva, Serge Rezzi, Ruth Prieto, Martin Kussmann, Jaime Inostroza, and Philippe Steenhout. 2014. “Impact of Breast-Feeding and High- and Low-Protein Formula on the Metabolism and Growth of Infants from Overweight and Obese Mothers.” *Pediatric Research* 75 (4): 535–43. doi: 10.1038/pr.2013.250.

Martín, Rocío, Hans G.H.J. Heilig, Erwin. G. Zoetendal, Esther Jiménez, Leónides Fernández, Hauke Smidt, and Juan M. Rodríguez. 2007. “Cultivation-Independent Assessment of the Bacterial Diversity of Breast Milk among Healthy Women.” *Research in Microbiology* 158 (1): 31–37. doi: 10.1016/j.resmic.2006.11.004.

Mazariegos, Monica, Christine Slater, and Manuel Ramirez-Zea. 2016. “Validity of Guatemalan Mother’s Self-Reported Breast-Feeding Practices Of.” *Food and Nutrition Bulletin* 37 (4): 494–503. doi: 10.1177/0379572116654644.

Medoua, Gabriel Nama, Estelle C. Sajo Nana, Anne Christine A. Ndzana, Caroline S. Makamto, Lucien S. Etame, Honorine A. Rikong, and Jean Louis E. Oyono. 2012. “Breastfeeding Practices of Cameroonian Mothers Determined by Dietary Recall since Birth and the Dose-to-the-Mother Deuterium-Oxide Turnover Technique.” *Maternal & Child Nutrition* 8 (3): 330–39. doi: 10.1111/j.1740-8709.2011.00293.x.

Nagpal, Ravinder, Takashi Kurakawa, Hirokazu Tsuji, Takuya Takahashi, Kazunari Kawashima, Satoru Nagata, Koji Nomoto, and Yuichiro Yamashiro. 2017. “Evolution of Gut Bifidobacterium Population in Healthy Japanese Infants over the First Three Years of Life: A Quantitative Assessment.” *Scientific Reports* 7 (1): 10097. doi: 10.1038/s41598-017-10711-5.

Quast, Christian, Elmar Pruesse, Pelin Yilmaz, Jan Gerken, Timmy Schweer, Pablo Yarza, Jörg Peplies, and Frank Oliver Glöckner. 2013. “The SILVA Ribosomal RNA Gene Database Project: Improved Data Processing and Web-Based Tools.” *Nucleic Acids Research* 41 (Database issue): D590–96. doi: 10.1093/nar/gks1219.

Raheem, Raheema Abdul, Colin W Binns, and Hui J Chih. 2017. “Protective Effects of Breastfeeding against Acute Respiratory Tract Infections and Diarrhoea: Findings of a Cohort Study.” *Journal of Paediatrics and Child Health* 53 (3): 271–76. doi: 10.1111/jpc.13480.

Sachdev, H.P.S., J. Krishna, R.K. Puri, L. Satyanarayana, and S. Kumar. 1991. “Water Supplementation in Exclusively Breastfed Infants during summer in the Tropics.” *Originally Published as Volume 1, Issue 8747* 337 (8747): 929–33. doi: 10.1016/0140-6736(91)91568-F.

Sitarik, Alexandra R., Manish Arora, Christine Austin, Lawrence F. Bielak, Shoshannah Eggers, Christine C. Johnson, Susan V. Lynch, et al. 2020. “Fetal and Early Postnatal Lead Exposure Measured in

Teeth Associates with Infant Gut Microbiota." *Environment International* 144 (November): 106062. doi: 10.1016/j.envint.2020.106062.

Slupsky, Carolyn M., Xuan He, Olle Hernell, Yvonne Andersson, Colin Rudolph, Bo Lönnerdal, and Christina E. West. 2017. "Postprandial Metabolic Response of Breast-Fed Infants and Infants Fed Lactose-Free vs Regular Infant Formula: A Randomized Controlled Trial." *Scientific Reports* 7 (1): 3640. doi: 10.1038/s41598-017-03975-4.

Smith, Emily R., Lindsey M. Locks, Karim P. Manji, Christine M. McDonald, Roland Kupka, Rodrick Kisenge, Said Aboud, Wafaie W. Fawzi, and Christopher P. Duggan. 2017. "Delayed Breastfeeding Initiation Is Associated with Infant Morbidity." *The Journal of Pediatrics* 191 (December): 57-62.e2. doi: 10.1016/j.jpeds.2017.08.069.

Tao, Nannan, Shuai Wu, Jaehan Kim, Hyun Joo An, Katie Hinde, Michael L. Power, Pascal Gagneux, J. Bruce German, and Carlito B. Lebrilla. 2011. "Evolutionary Glycomics: Characterization of Milk Oligosaccharides in Primates." *Journal of Proteome Research* 10 (4): 1548–1557.

Underwood, Mark A., J. Bruce German, Carlito B. Lebrilla, and David A. Mills. 2014. "Bifidobacterium Longum Subspecies Infantis: Champion Colonizer of the Infant Gut." *Pediatric Research* 77 (October): 229.

Vogt, Janet A., and Thomas M. S. Wolever. 2003. "Fecal Acetate Is Inversely Related to Acetate Absorption from the Human Rectum and Distal Colon." *The Journal of Nutrition* 133 (10): 3145–48. doi: 10.1093/jn/133.10.3145.

Wang, Lu, Amy van Grieken, Laura A. van der Velde, Eline Vlasblom, Maaike Beltman, Monique P. L'Hoir, Magda M. Boere-Boonekamp, and Hein Raat. 2019. "Factors Associated with Early Introduction of Complementary Feeding and Consumption of Non-Recommended Foods among Dutch Infants: The BeeBOFT Study." *BMC Public Health* 19 (1): 388. doi: 10.1186/s12889-019-6722-4.

Weljie, Aalim M., Jack Newton, Pascal Mercier, Erin Carlson, and Carolyn M. Slupsky. 2006. "Targeted Profiling: Quantitative Analysis of 1H NMR Metabolomics Data." *Analytical Chemistry* 78 (13): 4430–42. doi: 10.1021/ac060209g.

Wesolowska-Andersen, Agata, Martin Iain Bahl, Vera Carvalho, Karsten Kristiansen, Thomas Sicheritz-Pontén, Ramneek Gupta, and Tine Rask Licht. 2014. "Choice of Bacterial DNA Extraction Method from Fecal Material Influences Community Structure as Evaluated by Metagenomic Analysis." *Microbiome* 2 (1): 19. doi: 10.1186/2049-2618-2-19.

Wong, Julia M. W., Russell de Souza, Cyril W. C. Kendall, Azadeh Emam, and David J. A. Jenkins. 2006. "Colonic Health: Fermentation and Short Chain Fatty Acids." *Journal of Clinical Gastroenterology* 40 (3): 235–43. doi: 10.1097/00004836-200603000-00015.

Yatsunenko, Tanya, Federico E. Rey, Mark J. Manary, Indi Trehan, Maria Gloria Dominguez-Bello, Monica Contreras, Magda Magris, et al. 2012. "Human Gut Microbiome Viewed across Age and Geography." *Nature* 486 (7402): 222–27. doi: 10.1038/nature11053.

Yilmaz, Pelin, Laura Wegener Parfrey, Pablo Yarza, Jan Gerken, Elmar Pruesse, Christian Quast, Timmy Schweer, Jörg Peplies, Wolfgang Ludwig, and Frank Oliver Glöckner. 2014. "The SILVA and 'All-Species Living Tree Project (LTP)' Taxonomic Frameworks." *Nucleic Acids Research* 42 (Database issue): D643–48. doi: 10.1093/nar/gkt1209.

Chapter 5 Infant urine metabolic profile as a source of potential biomarkers for exclusive breastfeeding

Abstract

Early postnatal feeding mode (breastfed versus formula-fed) has been shown to dramatically impact the urine metabolic profile. Here, we aimed to characterize differences in the infant urine metabolome among exclusively breastfed (EBF) infants, partially breastfed (non-EBF_food) infants, and infants provided water in addition to breast milk (non-EBF_water) in order to develop urine metabolic biomarkers with the capability to distinguish between these different feeding modes. Infant feeding practices were determined using the dose-to-mother (DTM) technique combined with weighed food records. Urinary metabolic differences were observed between EBF and both groups of non-EBF infants at both 2 and 5 months. Utilizing a Random Forest with backwards selection model, a group of urine metabolites was identified with excellent predictive capability (area under the receiver operating curve (AUC) of 0.79 – 0.87). The selected biomarkers mainly included amino acid derivatives, nicotinamide-adenine dinucleotide degradation products, as well as metabolites from dietary sources and/or host-microbial co-metabolism. This study demonstrates that the inclusion of food or water in the infant diet prior to 5 months impacts the infant urine metabolome, and further that the urine metabolome can serve as a powerful tool to determine infant feeding practices.

Introduction

Breastfeeding plays a vital role in ensuring optimal development of infants in their first 6-months of life. The World Health Organization (WHO) has set the target to increase exclusive breastfeeding (EBF) rate during the first 6 months of infants to 50% by 2025 (WHO 2018). Accurate measurement of the infant feeding modality is needed to monitor progress. Moreover, accurate assessment of infant feeding practice (EBF versus nonexclusive breastfeeding (non-EBF)) is important to understand how infant feeding practices are related to health outcomes, and understand societal barriers (Greiner 2014; Kotowski et al. 2020). Several parental recall methods, such as recall since birth, recall over the past 24 h, and report of current feeding practice, have been routinely used to determine infant feeding practice in field studies (Aarts et al. 2000; Mazariegos, Slater, and Ramirez-Zea 2016). However, these methods tend to overestimate the rate of exclusive breastfeeding when compared to the deuterium oxide dose-to-mother (DTM) method (Mazariegos, Slater, and Ramirez-Zea 2016; Noel-Weiss, Boersma, and Kujawa-Myles 2012). This technique works by providing a dose of deuterium oxide (D_2O) to a mother, and measuring the amount of deuterium in saliva of the mother and baby over the period of 2 weeks. While the DTM technique provides an accurate measurement of infant feeding practices over the period of sample collection, its utility in the field is limited due to the high workload of sample collection. Moreover, questions as to its validity outside the sample collection window limit its use generally. There is a need to develop novel approaches that utilize easily collectable samples from infants that allow accurate estimates of infant feeding practices.

Developing infant urine metabolic biomarkers is a promising way to fulfill this purpose as studies have shown that the urine metabolome is impacted by infant feeding practice (Martin et al. 2014; O'Sullivan et al. 2013; He et al. 2020; Cesare Marincola et al. 2016; Hellmuth et al. 2016; Dessì

et al. 2016; Shoji et al. 2017). Exclusive formula feeding has been associated with higher levels of urinary amino acids (AAs) and their derivatives in both human (F.-P. J. Martin et al. 2014) and rhesus monkey (O'Sullivan et al. 2013; He et al. 2020) infants. It is thought that this may be due to a higher protein level in infant formula relative to breast milk, resulting in an increase in AAs and AA derivatives that are directly absorbed or produced by gut microbes. Urine metabolites involved in the tricarboxylic acid (TCA) cycle have also been shown to be altered by infant feeding practice (Cesare Marincola et al. 2016). Specifically, citrate and formate (a precursor of malate) were observed higher in breastfed infants while *cis*-aconitate and pantothenate (a precursor of coenzyme A) were higher in formula-fed infants (Cesare Marincola et al. 2016). Two end products of nicotinamide-adenine dinucleotide (NAD) degradation, N-methyl-2-pyridone-5-carboxamide (2PY) and N-methyl-4-pyridone-5-carboxamide (4PY), were observed to be higher in urine from formula-fed infants when compared to urine from breastfed infants (Martin et al. 2014). Other metabolites observed to be higher in infants consuming formula include choline, taurine, carnitine, and creatine (Cesare Marincola et al. 2016; Martin et al. 2014).

To date, no studies have reported on the impact of partial breastfeeding or the impact of providing water to infants on the urine metabolome. Here, urine samples from EBF and non-EBF infants (provided water or food in addition to breast milk) were analyzed and quantitative urine metabolome data were used to develop and validate biomarkers with the capability to discriminate between feeding practices.

Materials and methods

This current study took place in Bandung municipality and Sumedang district area in West Java, Indonesia. After obtaining parental consent, breast milk samples from mothers, and urine samples

from their infants were taken at postnatal ages of 2 months and 5 months. Information on the inclusion and exclusion criteria was reported in a previous study (Leong et al. 2021). Ethical approval was obtained from the Health Research Ethics Committee, Faculty of Medicine, Universitas Padjadjaran, Bandung, Indonesia (05/UN6.C1.3.2/KEPK/PN/2017).

Feeding practice, breast milk intake, and maternal secretor status determination

The DTM method was performed to determine feeding practice (EBF or non-EBF) of the infants as described previously (Liu et al. 2019a; Liu et al. 2019b). In brief, mothers were provided with a dose of oral D_2O. Saliva samples from both the mother and infant were collected, and the infant's water intake from sources other than breast milk was calculated as described (Liu et al. 2019a; Liu et al. 2019b). A cut off value of 86.6 g/d was used to assign the infant as EBF or non-EBF (Liu et al. 2019a; Liu et al. 2019b). Infant breast milk intake was calculated based on the equations described by the International Atomic Energy Agency (IAEA) (IAEA 2010). Based on the weighed food record of the infant, the daily percent energy intake from breast milk was also calculated and used to divide non-EBF infants into two categories: non-EBF_water (infants with 100% daily energy intake from breast milk and consuming water in addition to breast milk), and non-EBF_food (infants with less than 100% daily energy from breast milk and consuming food in addition to breast milk).

As described in Chapter 4, milk 2'FL was used as the phenotypic marker for maternal secretor status. Milk with greater than 200 μM 2'FL was assigned secretor (Se+), and milk with less than 200 μM 2'FL was assigned non-secretor (Se-).

Urine metabolite measurement

Urine samples were collected from infants using a urine pot and transferred into 1 mL microcentrifuge tubes for storage at -80 °C. To prepare for analysis, urine samples were thawed on ice and centrifuged to remove potential particulate matter (14k rcf, 4 °C, 5 min). 350 μL of the supernatant was transferred into pre-washed 3 kDa Amicon filters (Amicon ultracentrifugal filter, Millipore, Billerica, MA) and centrifuged at 14k rcf at 4 °C for 45 min. 207 μL of filtrate was mixed with 23 μL internal standard (5 mM 3-(trimethylsilyl)-1-propanesulfonic acid-d6 (DSS-d6) in 99.8% D_2O (to serve as a lock) and 0.2% NaN_3) and the pH was adjusted to between 6.78 - 6.92 before data acquisition on NMR.

^1H NMR spectra were acquired at 298K using a NOESY ^1H presaturation experiment ('noesypr1d') on a Bruker Avance 600 MHz NMR spectrometer (Bruker BioSpin, Germany) equipped with a SampleJet autosampler (Bruker BioSpin, Germany) as previously described (He et al. 2019). NMR spectra were Fourier transformed using 0.5 Hz line broadening, and then phase and baseline adjusted using Chenomx Processor. Urine metabolites were quantified using Chenomx Profiler (Chenomx NMR Suite v8.3, Chenomx Inc, Edmonton, Alberta, Canada) as described previously (Weljie et al. 2006).

Statistical analysis

Statistical analysis was performed using R (4.0.3). Urine metabolite concentration data were subjected to a generalized log transform (defined as $log(y + 1)$) before analysis. All plots shown were generated using the *ggplot2* package.

Significant differences in the urine metabolome between EBF and non-EBF infants were evaluated using the Mann-Whitney Test (*wilcox.test* function) and the p-value was adjusted (*p.adjust(,*

method = 'fdr')) with an overall significance of $p < 0.05$. Effect sizes were calculated using Cliff's delta (δ) (*cliff.delta* function) with a 95% confidence interval. Effect size thresholds were set according to Romano *et al* (Romano et al. 2006): $|\delta|<0.147$ corresponds to negligible; $0.147<|\delta|<0.33$ corresponds to small; $0.33<|\delta|<0.474$ corresponds to medium; and $|\delta|>0.474$ corresponds to a large effect size.

The *randomForest* package in R was used to perform the random forest (RF) algorithm. Urine metabolome data were randomly separated into a training data set (70% of data) and a testing data set (30% of data). A comparable percentage of EBF rate in the two data sets was confirmed. When constructing the RF model, the number of trees (ntree) was set to 4000 and the tuneRF function was used to estimate the best value of mytry (number of metabolites at each split). Potential urine metabolic markers were selected by the backwards elimination algorithm (*varSelRF*) and 20% of the metabolites were discarded each time. The selected marker metabolites were used to construct new RF models to predict the feeding practices in the testing data set and the model performance was calculated using area under receiver operating curve (ACU). The prediction was conducted in three ways: EBF vs. non-EBF (combining non-EBF_water and non-EBF_food), EBF vs. non-EBF_water, and EBF vs. non-EBF_food.

Results

This study investigated the impact of consumption of complementary food or water on the infant urine metabolome. The DTM technique used to determine infant feeding practice as EBF or non-EBF, and the results were published previously (Liu et al. 2019a; 2019b). Weighed food records of the infants were used to identify non-EBF infants with only water intake (non-EBF_water) from

non-EBF infants who were consuming food (non-EBF_food). Data from a total of 190 infants at 2 months and 169 infants at 5 months were included after data screening.

Characteristics of study population. As shown in Table 5.1, most of the infants (161 out of 190) at 2 months were EBF. Of the non-EBF infants, 9 were consuming only water in addition to breast milk, and 20 were consuming food in addition to breast milk. At 5 months, 76 infants remained EBF, while the number of infants in non-EBF_water and non-EBF_food group increased to 55 and 38, respectively. Maternal age and BMI (body mass index) were comparable among feeding groups. EBF infants at 2 months consumed a significantly larger daily volume of breast milk than non-EBF_water and non-EBF_food infants ($p < 0.05$) (Table 5.1). At 5 months, EBF infants had significantly higher daily breast milk intake than all non-EBF infants while between the two non-EBF groups, non-EBF_water infants consumed a significantly larger volume of breast milk than the non-EBF_food infants ($p < 0.05$). Infant length, weight, and z-scores were comparable among feeding groups at each time point.

In terms of morbidity, the prevalence of vomit and tachypnea were comparable among groups. Non-EBF_food infants had a slightly higher prevalence of fever at both 2 (30.0%) and 5 (50%) months when compared to other feeding types (at 2 months: 15.5% in EBF, 22.2% in non-EBF_water; at 5 months: 22.4% in EBF, 27.3% in non-EBF_water). Diarrhea was more commonly seen in non-EBF_water infants at 2 months (22.2%) compared to other feeding groups (2.6% - 10.0%). Prevalence of cough trended higher in non-EBF_food infants at 5 months (52.6%) when compared to other feeding groups (20.0% - 33.3%).

Table 5.1. Characteristics of mother and infant subjects.

Characteristics	Infant age (months) and feeding practice					
	2			5		
	EBF	non-EBF_water	non-EBF_food	EBF[1]	non-EBF_water[2]	non-EBF_food[3]
Sample size, n	161	9	20	76	55	38
Mother						
Age, year	27.2 (6.2)	27.0 (7.5)	31.1 (6.0)	27.2 (5.8)	26.7 (6.7)	29.1 (6.6)
BMI, kg/m^2	24.5 (3.4)	23.1 (3.3)	25.8 (4.0)	23.6 (3.4)	24.6 (4.1)	24.0 (3.5)
Secretor, %	106 (65.8%)	1 (11.1)%	15 (66.7%)	53 (69.7%)	37 (67.3%)	23 (60.5%)
Infant						
Breast milk take, mL/day	759 (117) [a]	540 (128)	367 (231)	803 (117)[b]	666 (149)[b]	498 (251)[b]
%energy from breast milk	100	100	58.5 (31.0)	100	100	66.0 (30.1)
Sex, %female	87 (54.0%）	7 (31.9%)	8 (40.0%)	47 (61.8%)	21 (38.1%)	22 (57.9%)
Length, cm	55.8 (2.0)	57.8 (1.4)	56.4 (2.4%)	62.5 (1.9)	63.1 (2.4)	62.7 (2.2)
Weight, kg	5.3 (0.6)	5.8 (0.6)	5.1 (0.7)	6.8 (0.8)	7.0 (0.9)	6.7 (0.8)
z score	-0.2 (1.0)	0.1 (0.8)	-0.8 (0.6)	0 (1.0)	0 (1.0)	-0.3 (1.1)
Vomit, n (%) #	5 (5.2%)	0	1 (5.0%)	2 (2.6%)	2 (3.6%)	0
Fever, n (%) #	25 (15.5%)	2 (22.2%)	6 (30.0%)	17 (22.4%)	15 (27.3%)	19 (50%)
Diarrhea, n (%) #	12 (7.5%)	2 (22.2%)	2 (10.0%)	2 (2.6%)	04 (7.3%)	2 (5.3%)
Cough, n (%) #	34 (21.1%)	3 (33.3%)	6 (30.0%)	24 (31.6%)	11 (20%)	20 (52.6%)
Tachypnea, n (%) #	6 (3.7%)	1 (0.4%)	0	5 (6.6%)	1 (1.8%)	3 (7.9%)

Mother age, BMI (body mass index), breast milk intake, infant length, infant weight, and infant *z*-score are shown as mean (standard deviation).
[#]Morbidity data reflects the prevalence of disease over a two week time period before sample collection.
[1]All EBF infants at 5 months were EBF at 2 months. [2] In non-EBF_water group at 5 months, 41 infants were in EBF, and 5 infants were in non-EBF_water groups at 2 months. [3]In non-EBF_food group at 5 months, 19 infants were in EBF, 1 infant was in non-EBF_water, and 10 infants were in non-EBF_food groups at 2 months. [a]$p < 0.05$ between EBF group and non-EBF_water, and between EBF and non-EBF_water groups assessed by Dunn's test, [b]$p < 0.05$ between any two groups.

Impact of food on the infant urine metabolome. To investigate the impact of partial breastfeeding on the infant urine metabolome, each metabolite was compared between the non-EBF_food and EBF groups, and between the non-EBF_water and EBF groups (Figure 5.1 a, b). At 2 months (Figure 5.1 a), a broad profile of amino acids (AAs) and their derivatives including threonine ($p < 0.01$, $\delta = -0.49$), serine ($p < 0.01$, $\delta = -0.48$), isoleucine ($p < 0.01$, $\delta = -0.46$), carnitine ($p < 0.01$, $\delta = -0.46$), tryptophan ($p < 0.01$, $\delta = -0.45$), valine ($p < 0.01$, $\delta = -0.45$), lysine ($p < 0.01$, $\delta = -0.41$), leucine ($p < 0.01$, $\delta = -0.41$), alanine ($p < 0.01$, $\delta = -0.41$), proline ($p < 0.01$, $\delta = -0.40$), tyrosine ($p < 0.01$, $\delta = -0.39$), taurine ($p < 0.01$, $\delta = -0.50$), phenylacetyl-L-glutamine ($p < 0.01$, $\delta = -0.41$), 2-aminobutyrate ($p < 0.01$, $\delta = -0.40$), 4-hydroxyphenylacetate ($p < 0.01$, $\delta = -0.40$), urea ($p < 0.01$, $\delta = -0.38$), 3-hydroxy-3-methylglutarate ($p < 0.05$, $\delta = -0.34$) and guanidoacetate ($p < 0.05$, $\delta = -0.35$) were significantly higher in urine from non-EBF_food infants when compared to EBF infants. Organic acids, including succinate ($p < 0.01$, $\delta = -0.45$) and lactate ($p < 0.01$, $\delta = -0.37$), were significantly higher in urine from non-EBF_food infants. Pantothenate ($p < 0.01$, $\delta = -0.69$), dimethyl sulfone ($p < 0.01$, $\delta = -0.68$), hippurate ($p < 0.01$, $\delta = -0.52$) and N-methyl-2-pyridone-5-carboxamide (2PY) ($p < 0.01$, $\delta = -0.49$) were also higher in non-EBF_food infants relative to EBF infants (Figure 5.1 a). When compared to EBF infants at 2 months, non-EBF_water infants had a lower level of several AAs including threonine ($\delta = 0.41$), serine ($\delta = 0.33$), valine ($\delta = 0.37$), tyrosine ($\delta = 0.34$), and glutamine ($\delta = 0.40$) as well as two other metabolites: galactose ($\delta = 0.34$) and formate ($\delta = 0.34$). However, these metabolites were not significantly different when assessed via the Mann-Whitney test after FDR correction (Figure 5.1 a). At 5 months (Figure 5.1 b), fewer metabolites were different between feeding practices. When compared to EBF infants, non-EBF_food infants had significantly higher levels of pantothenate ($p < 0.01$, $\delta = -0.38$), 2PY ($p < 0.01$, $\delta = -0.37$), dimethyl sulfone ($p < 0.01$, $\delta = -0.35$), succinate ($p < 0.01$, $\delta = -0.34$),

and phenylacetyl-L-glutamine ($p < 0.01$, $\delta = -0.33$). Non-EBF_water infants had a higher level of phenylacetyl-L-glutamine ($p < 0.01$, $\delta = -0.41$), urea ($p < 0.01$, $\delta = -0.33$), 4-hydroxyphenyllactate ($p < 0.01$, $\delta = -0.35$), 4-hydroxyphenylacetate ($p < 0.01$, $\delta = -0.39$) and 3-indoxysulfate ($p < 0.01$, $\delta = -0.36$) relative to EBF infants. Maternal secretor status did not impact the infant urine metabolome (Supplementary Figure 5.1).

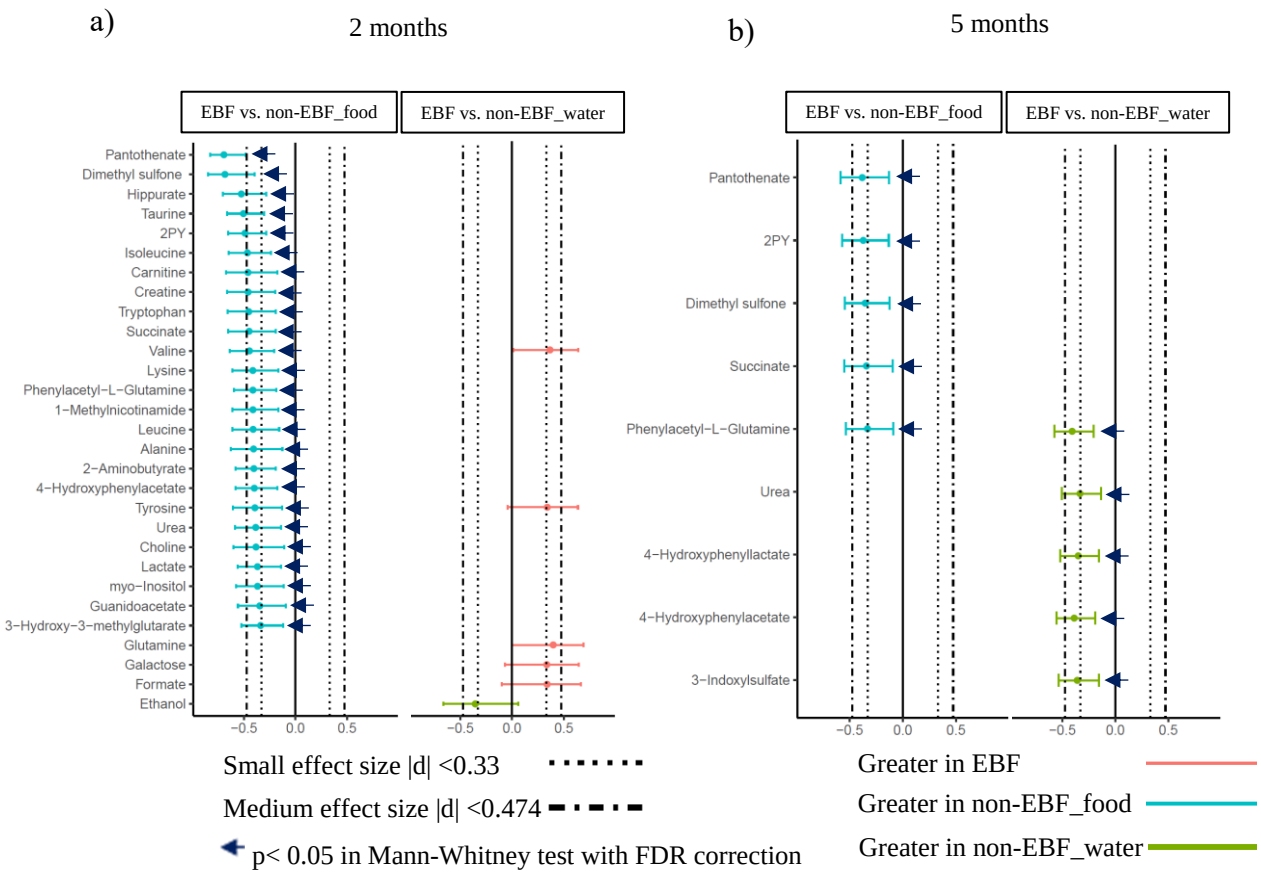


Figure 5.1. Impact of infant feeding mode at 2 and 5 months on the urine metabolome. a) Cliff's delta effect sizes at 2 months, b) Cliff's delta effect at 5 months. Arrows indicate $p < 0.05$ assess with the Mann Whitney U test.

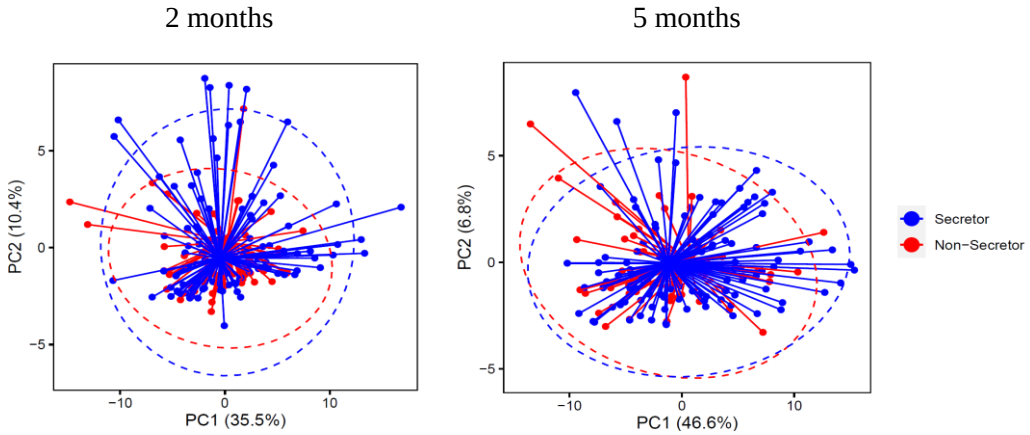
Prediction of feeding practice based on the infant urine metabolome. To investigate if the urine metabolome can predict infant feeding practices, 70% of the urine metabolome data were used to

generate random forest (RF) models and to identify metabolic biomarkers using a backwards selection algorithm (*varSelRF* function). The remaining 30 % of the urine metabolome data were used to measure the predictive performance of the selected metabolic biomarkers by calculating the area under the receiver operating curve (AUC). As shown in Table 5.2, among the three ways of prediction (EBF vs. non-EBF, EBF vs. non-EBF_water, and EBF vs. non-EBF_food), differentiating between EBF and non-EBF_food achieved the highest AUC of 0.87, and discriminating between EBF and non-EBF_water infants had the lowest AUC of 0.79. Among the 69 metabolites measured, the metabolites with a predictive capability selected by the RF model included dimethyl-sulfone, pantothenate, 2PY, 4PY, phenylacetyl-L-glutamine, hippurate, dimethylglycine, and myo-inositol.

Table 5.2. Predictive performance of selected metabolic biomarkers.

Classification	AUC	Features selected
EBF vs. **non-EBF**	0.81 (0.03)	Infant age, Dimethyl-sulfone, Pantothenate, 2PY, Phenylacetyl-L-Glutamine
EBF vs. **non-EBF_food**	0.87 (0.04)	Dimethyl sulfone, Pantothenate, Hippurate, 2PY, Dimethylglycine, myo-Inositol, Phenylacetyl-L-Glutamine, 4PY
EBF vs. **non-EBF_water**	0.79 (0.04)	Infant age, Acetone, Phenylacetyl-L-Glutamine, Adipate

AUC, area under receiver operator characteristics curve; EBF, exclusively breastfed; non-EBF_food, non-exclusively breastfed with daily percent energy intake from breast milk < 100%; non-EBF_water, non-exclusively breastfed with water intake and daily percent energy intake from breast milk = 100%; AUC is shown as mean (standard deviation) of 20 runs.



Supplementary Figure 5.1. Maternal secretor status did not impact the infant fecal metabolome at 2 or 5 months.

Discussion

In this study, we investigated the impact of partial breastfeeding or inclusion of water on the urine metabolome of infants. Although previous studies (Martin et al. 2014; Hellmuth et al. 2016; Cesare Marincola et al. 2016; Dessì et al. 2016; Shoji et al. 2017) have looked at the impact of feeding modality on the urine metabolome, they compared exclusive or partial breastfeeding to formula feeding. No one has compared exclusively breastfed infants (validated through the DTM method) with infants provided water or infants that are mostly breastfed, but provided less than 41% of their total daily energy intake of food. In addition to looking at subtle differences between exclusively breastfed and partially breastfed infants, it is important to understand whether water intake influences infant metabolism. Water intake in infants under 6 months of age has been reported although it is not recommended (Bruce and Kliegman 1997; Boussemart et al. 2006). The present study utilized the DTM technique to measure infant feeding practice (EBF or non-EBF) and used weighed food records to identify non-EBF infants with water (non-EBF_water) or food

(non-EBF_food) intake. Potential urine metabolic biomarkers with the capability to discriminate between the infant feeding practices were developed and validated using an RF algorithm.

Non-EBF_food infants in the present study were mainly given infant formula, infant cereals, vitamin supplements, other cereals, and fruits and vegetables (ranked by their percentage of the total food fed to the infants at 2 and 5 months) (Leong et al. 2021). Consuming infant formula or other foods by breastfed infants has been reported to alter the urine metabolome, resulting in higher concentration of AAs and their derivatives relative to EBF infants (Chiu et al. 2016). It has been reported that infant formula and other food introduced to infants significantly impact the infants' gut microbiome (Bäckhed et al. 2015; Bezirtzoglou, Tsiotsias, and Welling 2011; O'Sullivan, Farver, and Smilowitz 2015; He, et al. 2019). Overall, EBF infants have been reported to have a lower alpha diversity, and a higher relative abundance of taxa from the protective bacterial class *Actinobacteria* than formula-fed infants (Bäckhed et al. 2015; Bezirtzoglou, Tsiotsias, and Welling 2011; Chapter 4). HMOs in breast milk have been shown to selectively promote the growth of *Bifidobacterium* species (Ward et al. 2007). Higher levels of the proinflammatory bacterial class *γ-Proteobacteria* have been reported in formula-fed infants (Bäckhed et al. 2015; Bezirtzoglou, Tsiotsias, and Welling 2011). Formula and other food consumed in addition to breast milk was reported to vastly alter the infant gut microbiome and make it resemble that of exclusively formula-fed infants (O'Sullivan, Farver, and Smilowitz 2015).

Multiple metabolites reflective of gut microbial function were observed to be significantly altered by food or water introduction (Figure 5.1) and some of them were identified by RF model as biomarkers to discriminate between feeding practices (Table 5.2).

Consistent with previous studies (Martin et al. 2014; O'Sullivan et al. 2013; He et al. 2020), urine AAs and their derivatives were the major metabolites that showed significant differences between

feeding practices, and three metabolites (phenylacetyl-L-glutamine, hippurate, and dimethylglycine) were identified by the RF model as predictive biomarkers (Table 5.2). When compared to EBF infants, non-EFB_food infants had higher urinary concentrations of AAs and their derivatives, while non-EBF_water infants had lower levels of several AAs at 2 months (Figure 5.1). The most common food consumed by non-EBF_food infants was infant formula (Leong et al. 2021), and thus it makes sense that in comparison to EBF infants, these metabolites are different. The lower urinary AAs in non-EBF_water infants could be because of their lower intake of breast milk (Table 5.1). Phenylacetyl-L-glutamine is an amino acid acetylation product of phenylacetate and serves as an nitrogen clearance vehicle in the human body besides urea, and it is also a microbial metabolite produced by *Christensenellaceae*, *Lachnospiraceae* and *Ruminococcaceae* (Brusilow 1991). For healthy infants, phenylacetate could be produced by gut microbes by fermenting phenylalanine (Kumps, Duez, and Mardens 2002). Hippurate, another product of the nitrogen excretion pathway in which benzoate conjugates with glycine (van Straten et al. 2017), can also be formed through gut bacterial metabolism of dietary components, primarily polyphenols in fruit/vegetables (Gonthier et al. 2003; Walsh et al. 2007; Silva et al. 2000). The higher phenylacetyl-L-glutamine and hippurate in the urine of non-EBF_food infants could be attributed to altered microbiota and colon fermentative capacity. Interestingly, non-EBF_water infants at 5 months also had higher phenylacetyl-L-glutamine relative to EBF infants. This could be because water intake in non-EBF_water infants reduced the total amount of HMOs consumed, thus changing the fermentative capability of the some microbes.

Dimethyl sulfone is another metabolite with predictive capability selected by the RF model (Table 5.2). Its microbial and host co-metabolism pathway was reviewed previously (He and Slupsky 2014). Methionine can be metabolized by several gut microbes (*Proteus vulgaris, P. mirabilis, P.*

rettgeri, and Morganella morganii) to methanethiol and the host can enzymatically convert methanethiol to dimethyl sulfone (Hayward et al. 1977). Methionine, an essential AA, was found to be higher in some bovine milk and soy-based infant formulas (Agostoni et al. 2000). Formula-fed infants were reported to have higher fecal methanethiol when compared to EBF infants (Jiang et al. 2001). It can be hypothesized that the higher level of dimethyl sulfone in non-EBF_food infants relative to EBF infants could be the result of changes in the function of the gut microbiome. Higher dimethyl sulfone in the urine of non-EBF_food but not in non-EBF_water infants relative to EBF infants at both ages also suggests a dietary origin of the metabolite or its precursor, methionine. Interestingly, grain has been associated with production of dimethyl sulfone (Perkowski et al. 2012; Buśko et al. 2010), and thus it could be that infants consuming infant cereal had higher levels of this metabolite (Leong et al. 2021).

2PY and 4PY were also identified by the RF model as biomarkers of infant food intake (Table 5.2). Nicotinamide, a derivative of nicotinate (vitamin B3), can be metabolized in the liver to N-methyl-nicotinamide (MNA) by the enzyme nicotinamide N-methyltransferase (NNMT), and the latter can be further converted to 2PY and 4PY by aldehyde oxidase in mammals (Pissios 2017). Significantly higher urine 2PY induced by consuming non-breastmilk food at both ages (Figure 5.1) in this study is consistent with previous observations (Martin et al. 2014). A higher level of 2PY might indicate differences in kidney function, possibly in the activity of aldehyde oxidase, between EBF and non-EBF_food infants. Increased kidney size and renal workload (as shown by the serum urea/creatinine ratios) were observed in exclusively or partially formula-fed infants when compared to EBF infants (Escribano et al. 2011; Schmidt et al. 2004). Increased renal workload was previously shown to be secondary to the increased protein intake and protein metabolite filtration (mainly urea) (W. F. Martin, Armstrong, and Rodriguez 2005).

A higher concentration of pantothenate in urine samples from non-EBF infants observed in the current study was also reported in a previous study (Cesare Marincola et al. 2016). This could be attributed to the consumption of infant formula by infants in the non-EBF_food group because pantothenate is commonly supplemented in infant formula (Newberry 1982).

Conclusion

The findings in this study demonstrate the influence of complementary water and food intake on the urine metabolome of infants at 2 and 5 months. Major metabolic changes observed in the present study between EBF and non-EBF_food infants were supported by previous literature (Martin et al. 2014; Hellmuth et al. 2016; Cesare Marincola et al. 2016). The difference in the urine metabolome caused by water intake was not reported previously. Utilizing an RF algorithm, this study identified urine metabolte markers with excellent predictive capability. These markers included AA derivatives, NAD metabolites and metabolites from dietary sources and/or host-microbial co-metabolism. The results of this study have paved the path to a deeper understanding of how early diet impact infant metabolism. The current study was limited by the small number of non-EBF infants at 2 months. Further studies with a larger sample size should be conducted to confirm the validity of the urine metabolic biomarkers.

References

Aarts, Clara, Elisabeth Kylberg, Agneta Hörnell, Yngve Hofvander, Mehari Gebre-Medhin, and Ted Greiner. 2000. "How Exclusive Is Exclusive Breastfeeding? A Comparison of Data since Birth with Current Status Data." *International Journal of Epidemiology* 29 (6): 1041–46. doi: 10.1093/ije/29.6.1041.

Agostoni, Carlo, Brunella Carratù, Concetta Boniglia, Enrica Riva, and Elisabetta Sanzini. 2000. "Free Amino Acid Content in Standard Infant Formulas: Comparison with Human Milk." *Journal of the American College of Nutrition* 19 (4): 434–38. doi: 10.1080/07315724.2000.10718943.

Bäckhed, Fredrik, Josefine Roswall, Yangqing Peng, Qiang Feng, Huijue Jia, Petia Kovatcheva-Datchary, Yin Li, et al. 2015. "Dynamics and Stabilization of the Human Gut Microbiome during the First Year of Life." *Cell Host & Microbe* 17 (5): 690–703. doi: 10.1016/j.chom.2015.04.004.

Bezirtzoglou, Eugenia, Arsenis Tsiotsias, and Gjalt W. Welling. 2011. "Microbiota Profile in Feces of Breast- and Formula-Fed Newborns by Using Fluorescence in Situ Hybridization (FISH)." *Anaerobe* 17 (6): 478–82. doi: 10.1016/j.anaerobe.2011.03.009.

Brusilow, Saul W. 1991. "Phenylacetylglutamine May Replace Urea as a Vehicle for Waste Nitrogen Excretion." *Pediatric Research* 29 (2): 147–50. doi: 10.1203/00006450-199102000-00009.

Buśko, Maciej, Henryk Jeleń, Tomasz Góral, Jarosław Chmielewski, Kinga Stuper, Lidia Szwajkowska-Michałek, Bożena Tyrakowska, and Juliusz Perkowski. 2010. "Volatile Metabolites in Various Cereal Grains." *Food Additives & Contaminants: Part A* 27 (11): 1574–81. doi: 10.1080/19440049.2010.506600.

Cesare Marincola, Flaminia, Sara Corbu, Milena Lussu, Antonio Noto, Angelica Dessì, Stefania Longo, Elisa Civardi, et al. 2016. "Impact of Early Postnatal Nutrition on the NMR Urinary Metabolic Profile of Infant." *Journal of Proteome Research* 15 (10): 3712–23. doi: 10.1021/acs.jproteome.6b00537.

Chiu, Chih-Yung, Kuo-Wei Yeh, Gigin Lin, Meng-Han Chiang, Shu-Chen Yang, Wei-Ju Chao, Tsung-Chieh Yao, et al. 2016. "Metabolomics Reveals Dynamic Metabolic Changes Associated with Age in Early Childhood." *PLOS ONE* 11 (2): e0149823. doi: 10.1371/journal.pone.0149823.

Dessì, Angelica, Antonio Murgia, Rocco Agostino, Maria Grazia Pattumelli, Andrea Schirru, Paola Scano, Vassilios Fanos, and Pierluigi Caboni. 2016. "Exploring the Role of Different Neonatal Nutrition Regimens during the First Week of Life by Urinary GC-MS Metabolomics." *International Journal of Molecular Sciences* 17 (2): 265. doi: 10.3390/ijms17020265.

Escribano, Joaquin, Veronica Luque, Natalia Ferre, Marta Zaragoza-Jordana, Veit Grote, Berthold Koletzko, Dariusz Gruszfeld, et al. 2011. "Increased Protein Intake Augments Kidney Volume and Function in Healthy Infants." *Kidney International* 79 (7): 783–90. doi: 10.1038/ki.2010.499.

Gonthier, Marie-Paule, Marie-Anne Verny, Catherine Besson, Christian Rémésy, and Augustin Scalbert. 2003. "Chlorogenic Acid Bioavailability Largely Depends on Its Metabolism by the Gut Microflora in Rats." *The Journal of Nutrition* 133 (6): 1853–59. doi: 10.1093/jn/133.6.1853.

Greiner, Ted. 2014. "Exclusive Breastfeeding: Measurement and Indicators." *International Breastfeeding Journal* 9 (1): 18. doi: 10.1186/1746-4358-9-18.

Hayward, N J, T H Jeavons, A J Nicholson, and A G Thornton. 1977. "Methyl Mercaptan and Dimethyl Disulfide Production from Methionine by Proteus Species Detected by Head-Space Gas-Liquid Chromatography." *Journal of Clinical Microbiology* 6 (3): 187–94. doi: 10.1128/jcm.6.3.187-194.1977.

He, Xuan, Mariana Parenti, Tove Grip, Magnus Domellöf, Bo Lönnerdal, Olle Hernell, Niklas Timby, and Carolyn M. Slupsky. 2019. "Metabolic Phenotype of Breast-Fed Infants, and Infants Fed Standard Formula or Bovine MFGM Supplemented Formula: A Randomized Controlled Trial." *Scientific Reports* 9 (1): 339. doi: 10.1038/s41598-018-36292-5.

He, Xuan, Mariana Parenti, Tove Grip, Bo Lönnerdal, Niklas Timby, Magnus Domellöf, Olle Hernell, and Carolyn M. Slupsky. 2019. "Fecal Microbiome and Metabolome of Infants Fed Bovine MFGM

Supplemented Formula or Standard Formula with Breast-Fed Infants as Reference: A Randomized Controlled Trial." *Scientific Reports* 9 (1): 11589. doi: 10.1038/s41598-019-47953-4.

He, Xuan, and Carolyn M. Slupsky. 2014. "Metabolic Fingerprint of Dimethyl Sulfone (DMSO2) in Microbial–Mammalian Co-Metabolism." *Journal of Proteome Research* 13 (12): 5281–92. doi: 10.1021/pr500629t.

He, Xuan, Jennie Sotelo-Orozco, Colin Rudolph, Bo Lönnerdal, and Carolyn M. Slupsky. 2020. "The Role of Protein and Free Amino Acids on Intake, Metabolism, and Gut Microbiome: A Comparison between Breast-Fed and Formula-Fed Rhesus Monkey Infants." *Frontiers in Pediatrics* 7: 563. doi: 10.3389/fped.2019.00563.

Hellmuth, Christian, Olaf Uhl, Franca F. Kirchberg, Veit Grote, Martina Weber, Peter Rzehak, Clotilde Carlier, et al. 2016. "Effects of Early Nutrition on the Infant Metabolome." *Nestle Nutrition Institute Workshop Series* 85: 89–100. doi: 10.1159/000439491.

Jiang, T., F. L. Suarez, M. D. Levitt, S. E. Nelson, and E. E. Ziegler. 2001. "Gas Production by Feces of Infants." *Journal of Pediatric Gastroenterology and Nutrition* 32 (5): 534–41. doi: 10.1097/00005176-200105000-00009.

Kotowski, Judith, Cathrine Fowler, Christina Hourigan, and Fiona Orr. 2020. "Bottle-Feeding an Infant Feeding Modality: An Integrative Literature Review." *Maternal & Child Nutrition* 16 (2): e12939. doi: 10.1111/mcn.12939.

Kumps, Alain, Pierre Duez, and Yves Mardens. 2002. "Metabolic, Nutritional, Iatrogenic, and Artifactual Sources of Urinary Organic Acids: A Comprehensive Table." *Clinical Chemistry* 48 (5): 708–17. doi: 10.1093/clinchem/48.5.708.

Leong, Claudia, Rosalind S Gibson, Aly Diana, Jillian J Haszard, Sofa Rahmannia, Mohammad Brachim Ansari, Lina Sofiatul Inayah, Afini Dwi Purnamasari, and Lisa A Houghton. 2021. "Differences in Micronutrient Intakes of Exclusive and Partially Breastfed Indonesian Infants from Resource-Poor Households Are Not Accompanied by Differences in Micronutrient Status, Morbidity, or Growth." *The Journal of Nutrition* 151 (3): 705–15. doi: 10.1093/jn/nxaa381.

Liu, Zheng, Aly Diana, Christine Slater, Thomas Preston, Rosalind S. Gibson, Lisa Houghton, and Stephen B. Duffull. 2019a. "Development of a Nonlinear Hierarchical Model to Describe the Disposition of Deuterium in Mother-Infant Pairs to Assess Exclusive Breastfeeding Practice." *Journal of Pharmacokinetics and Pharmacodynamics* 46 (1): 1–13. doi: 10.1007/s10928-018-9613-x.

Liu, Zheng, Aly Diana, Christine Slater, Thomas Preston, Rosalind S. Gibson, Lisa Houghton, and Stephen B. Duffull. 2019b. "Development of a Parsimonious Design for Optimal Classification of Exclusive Breastfeeding." *CPT: Pharmacometrics & Systems Pharmacology* 8 (8): 596–605. doi: 10.1002/psp4.12428.

Martin, François-Pierre J., Sofia Moco, Ivan Montoliu, Sebastiano Collino, Laeticia Da Silva, Serge Rezzi, Ruth Prieto, Martin Kussmann, Jaime Inostroza, and Philippe Steenhout. 2014. "Impact of Breast-Feeding and High- and Low-Protein Formula on the Metabolism and Growth of Infants from Overweight and Obese Mothers." *Pediatric Research* 75 (4): 535–43. doi: 10.1038/pr.2013.250.

Martin, William F, Lawrence E Armstrong, and Nancy R Rodriguez. 2005. "Dietary Protein Intake and Renal Function." *Nutrition & Metabolism* 2 (September): 25–25. doi: 10.1186/1743-7075-2-25.

Mazariegos, Monica, Christine Slater, and Manuel Ramirez-Zea. 2016. "Validity of Guatemalan Mother's Self-Reported Breast-Feeding Practices of 3-Month-Old Infants." *Food and Nutrition Bulletin* 37 (4): 494–503. doi: 10.1177/0379572116654644.

Newberry, R. E. 1982. "The Infant Formula Act of 1980." *Journal - Association of Official Analytical Chemists* 65 (6): 1472–73.

Noel-Weiss, Joy, Sonya Boersma, and Sonya Kujawa-Myles. 2012. "Questioning Current Definitions for Breastfeeding Research." *International Breastfeeding Journal* 7 (1): 9. doi: 10.1186/1746-4358-7-9.

O'Sullivan, Aifric, Marie Farver, and Jennifer T Smilowitz. 2015. "The Influence of Early Infant-Feeding Practices on the Intestinal Microbiome and Body Composition in Infants." *Nutrition and Metabolic Insights* 8 (Suppl 1): 1–9. doi: 10.4137/NMI.S29530.

O'Sullivan, Aifric, Xuan He, Elizabeth M. S. McNiven, Neill W. Haggarty, Bo Lonnerdal, and Carolyn M. Slupsky. 2013. "Early Diet Impacts Infant Rhesus Gut Microbiome, Immunity, and Metabolism." *Journal of Proteome Research* 12 (6): 2833–45. doi: 10.1021/pr4001702.

Perkowski, Juliusz, Kinga Stuper, Maciej Buśko, Tomasz Góral, Anna Kaczmarek, and Henryk Jeleń. 2012. "Differences in Metabolomic Profiles of the Naturally Contaminated Grain of Barley, Oats and Rye." *Journal of Cereal Science* 56 (3): 544–51. doi: 10.1016/j.jcs.2012.07.012.

Pissios, Pavlos. 2017. "Nicotinamide N-Methyltransferase: More Than a Vitamin B3 Clearance Enzyme." *Trends in Endocrinology and Metabolism: TEM* 28 (5): 340–53. doi: 10.1016/j.tem.2017.02.004.

Schmidt, Ida M., Ida N. Damgaard, Kirsten A. Boisen, Claudia Mau, Marla Chellakooty, Klaus Olgaard, and Katharina M. Main. 2004. "Increased Kidney Growth in Formula-Fed versus Breast-Fed Healthy Infants." *Pediatric Nephrology* 19 (10): 1137–44. doi: 10.1007/s00467-004-1567-0.

Shoji, Hiromichi, Hikari Taka, Naoko Kaga, Naho Ikeda, Tomohiro Kitamura, Yoshiki Miura, and Toshiaki Shimizu. 2017. "A Pilot Study of the Effect of Human Breast Milk on Urinary Metabolome Analysis in Infants." *Journal of Pediatric Endocrinology and Metabolism* 30 (9): 939–46. doi: 10.1515/jpem-2017-0179.

International Atomic Energy Agency, 2010. Stable Isotope Technique to Assess Intake of Human Milk in Breastfed Infants. *Human Health Series 7. Vienna*.

Straten, G. van, M. G. M. de Sain-van der Velden, I. M. van Geijlswijk, R. P. Favier, S. J. Mesu, N. E. Holwerda-Loof, M. van der Ham, et al. 2017. "Saline Is as Effective as Nitrogen Scavengers for Treatment of Hyperammonemia." *Scientific Reports* 7 (1): 13112. doi: 10.1038/s41598-017-12686-9.

World Health Organisation. 2018. "2025, To Improve Maternal, Infant and Young Child Nutrition." Targets, Gobal.

Walsh, Marianne C., Lorraine Brennan, Estelle Pujos-Guillot, Jean-Louis Sébédio, Augustin Scalbert, Ailís Fagan, Desmond G. Higgins, and Michael J. Gibney. 2007. "Influence of Acute Phytochemical Intake on Human Urinary Metabolomic Profiles." *The American Journal of Clinical Nutrition* 86 (6): 1687–93. doi: 10.1093/ajcn/86.5.1687.

Ward, Robert E., Milady Niñonuevo, David A. Mills, Carlito B. Lebrilla, and J. Bruce German. 2007. "In Vitro Fermentability of Human Milk Oligosaccharides by Several Strains of Bifidobacteria." *Molecular Nutrition & Food Research* 51 (11): 1398–1405. doi: 10.1002/mnfr.200700150.

Weljie, Aalim M., Jack Newton, Pascal Mercier, Erin Carlson, and Carolyn M. Slupsky. 2006. "Targeted Profiling: Quantitative Analysis of 1H NMR Metabolomics Data." *Analytical Chemistry* 78 (13): 4430–42. doi: 10.1021/ac060209g.

Appendix: Multi-micronutrient and choline supplementation during pregnancy impacted children urine metabolome at age of 2 to 4 years

Abstract

Multi-micronutrients (MVM) and choline supplementation during pregnancy may help to mitigate the adverse effect of prenatal alcohol exposure. Here, the impact of prenatal alcohol exposure with and without MVM supplementation (with or without choline) on the offspring urine metabolome was studied. Prenatal alcohol exposure was not observed to alter the urine metabolome of the children; however, microbial related metabolites including 1,2-propanediol, isopropanol, methanol, 3-hydroxyphenylacetate as well as several amino acids showed differences between the MVM and choline supplement groups. This study demonstrated that prenatal MVM and choline supplementation has an impact on the offspring urine metabolome at least until the age of 4 years.

Introduction

Fetal alcohol spectrum disorder (FASD) is an umbrella term that includes a series of alcohol-induced effects: fetal alcohol syndrome, partial fetal alcohol syndrome, and alcohol-related neurodevelopmental disorder (O'Malley 2007). It is a leading preventable cause of neurodevelopmental disabilities worldwide (Popova et al. 2017). According to a recent study, the estimated prevalence of FASD among first-graders in U.S. communities ranged from 1.1% to 5.0% (May et al. 2018). In addition to reducing alcohol exposure at various stages of pregnancy and lactation, preventing or mitigating adverse birth outcomes after exposure is of great importance. Evidence suggests that maternal nutritional status could be a potential modulator of the adverse neurodevelopmental and behavioral effects of FASD (Keen et al. 2010).

It has been shown that when women take prenatal supplements that include choline during pregnancy, their offspring tend to have improved learning and memory regardless of alcohol-exposure status (Kable et al. 2015). In another study of women who consumed alcohol during pregnancy, infants from women taking prenatal choline supplements were reported to have better visual recognition memory as well as greater weight and head circumference when compared to the placebo group (Jacobson et al. 2018). Infant mental development index (MDI) was shown to be negatively associated with peri-conceptual alcohol usage, and MVM supplements showed a protective effect against the adverse outcomes associated with alcohol exposure (Coles et al. 2015). The beneficial effects of prenatal choline supplementation against alcohol exposure have also been reported in rat models (Thomas, Abou, and Dominguez 2009; Thomas et al. 2010).

Despite the observed protective effects of choline and MVM supplementation on reducing alcohol-related neurocognitive deficits and growth impairments in infants, the underlying mechanisms on how this works is poorly understood. The present study aimed to characterize the impact of

prenatal alcohol exposure with and without choline and MVM suplementation on the offspring urine metabolome in early childhood.

Methods

Urine sample collection and metabolite extraction

Urine samples (n = 76) analyzed in this study were from a larger dataset of the follow up study of previous work (Sowell et al. 2018) where women with high or low/no alcohol exposure during pregnancy were recruited at ~19 weeks of gestation at two locations in the Ukraine. Pregnant mothers were provided a multivitamin and mineral (MVM) supplement, a MVM and choline (MVM+) supplement, or no supplement. Offspring of these mothers were assessed for FASD outcomes within the first year of life as previously described (Sowell et al. 2018).

Urine specimens were collected using urine containers from children between 2 and 4 years. Urine samples were frozen and stored at -80 ℃ until analyzed. To prepare for analysis, urine samples were thawed on ice, and 300 μL was centrifuged at 10k rcf, 4 ℃ for 10 minutes to spin down and to remove sediments. A total of 207 uL of the supernatant was transferred to a new tube and mixed with 23 μL of the internal standard (5 mM 3-(trimethylsilyl)-1-propanesulfonic acid-d6 (DSS-d6) in 99.8% D_2O (to serve as a lock) and 0.2% NaN_3) and the pH was adjusted to 6.78 - 6.92 before data acquisition on NMR.

Urine metabolite measurements

^1H NMR spectra were acquired at 298K using a NOESY ^1H presaturation experiment ('noesypr1d') on a Bruker Avance 600 MHz NMR spectrometer (Bruker BioSpin, Germany) equipped with a SampleJet autosampler (Bruker BioSpin, Germany) as previously described (He et al. 2019). Using Chenomx Processor and Profiler (Chenomx NMR Suite v8.3, Chenomx Inc,

Edmonton, Alberta, Canada), urine metabolites were processed and manually profiled for quantification as described previously (Weljie et al. 2006).

Statistical analysis

Statistical analysis was performed using R (4.0.3) and plots were generated using the *ggplot2* package. Probable quotient normalization (Rosen Vollmar et al. 2019; Dieterle et al. 2006) was performed on the urine metabolite concentration data before analysis to adjust for potential differences in hydration status. Specifically, a reference spectrum was constructed by calculating the median concentration of each metabolite among all subjects with urea and creatinine excluded, followed by dividing the urine metabolite concentrations for each subject by the reference spectrum to calculate the quotients. The median of the all the quotients of each subject was subsequently calculated, and was used to correct the urine metabolite concentrations of the same subject. The corrected urine metabolite concentrations were then log transformed (defined as $log(y + 1)$) before statistical analysis.

To assess significance, both p-values and effect sizes were calculated. P-values were calculated based on robust linear model (*lmRob* function) with an overall significance of $p < 0.05$. Age of the subjects and the urine sample collection locations were controlled in the robust model as confounding factors. Effect sizes were calculated using Cliff's delta (δ) effect size (*cliff.delta* function) with a 95% confidence interval. Thresholds of the effect size were set according to Romano *et al* (Romano et al. 2006): $|\delta|<0.147$ corresponds to negligible, $0.147<|\delta|<0.33$ corresponds to small, and $|\delta|>0.474$ corresponds to a large effect size. Euclidean distances were calculated to perform the principal coordinate analysis (PCoA) of urine metabolome data using the *pcoa* package. The centroid of each group indicated in the figure was calculated by averaging PC1 and PC2 for each group and the 95% confidence level was calculated based on multivariate normal

distribution. The difference in beta-dispersion was tested using the *betadisper* function followed by TukeyHSD post-hoc analysis.

Results

This study investigated the influence of FASD outcomes, maternal alcohol exposure and micronutrient supplementation (with or without choline) during pregnancy on the offspring urine metabolome at 2 to 4 years. Urine samples from 76 children were included in the analysis. The urine metabolome was similar between FASD and non-FASD children, and between the alcohol-exposed and low/no alcohol exposure group (Figure A1). Amongst the offspring of mothers taking the MVM, MVM+, or no supplements, the urine metabolome was not significantly different in multivariate analysis ($p > 0.05$) (Figure A2 a). Pairwise analysis between groups revealed that children born to women taking the MVM supplement had higher levels of pyruvate ($\delta = 0.49$, $p < 0.05$), threonine ($\delta = 0.35$, $p > 0.05$) and isopropanol ($\delta = 0.34$, $p > 0.05$) and lower levels of 3-hydroxyphenylactate ($\delta = -0.38$, $p > 0.05$) in their urine when compared to children who were born to women not taking a supplement during pregnancy (Figure A2 b). Children born to women prenatally supplemented with MVM and choline (MVM+) had higher 1,2-propanediol ($\delta = 0.39$, $p > 0.05$), pyruvate ($\delta = 0.39$, $p > 0.05$) and methanol ($\delta = 0.33$, $p > 0.05$), as well as a lower asparagine ($\delta = -0.34$, $p > 0.05$), tryptophan ($\delta = -0.41$, $p > 0.05$) and serine ($\delta = -0.53$, $p < 0.05$) relative to children from women not taking any supplements (Figure A2 c). Compared to children of women taking only the MVM supplement, children of women taking the MVM+ supplement had lower levels of tryptophan ($\delta = -0.34$, $p > 0.05$), serine ($\delta = -0.37$, $p < 0.05$), threonine ($\delta = -0.38$, $p > 0.05$), isopropanol ($\delta = -0.38$, $p > 0.05$) and alanine ($\delta = -0.47$, $p > 0.05$).

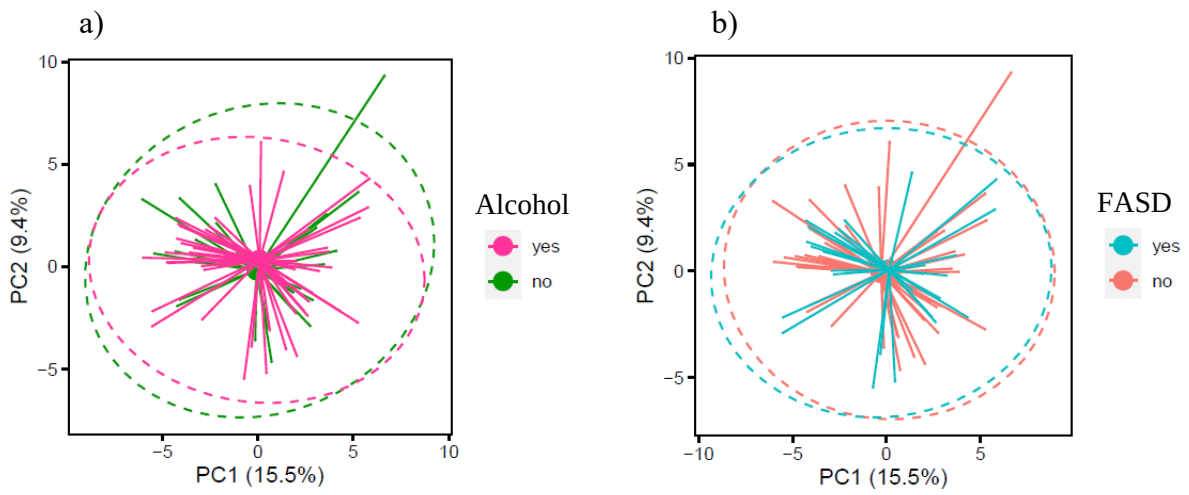


Figure A1. Principal coordinate analysis (PCoA) utilizing Euclidian distance of children's urine metabolome by a) prenatal alcohol exposure and b) fetal alcohol spectrum disorder (FASD) outcome.

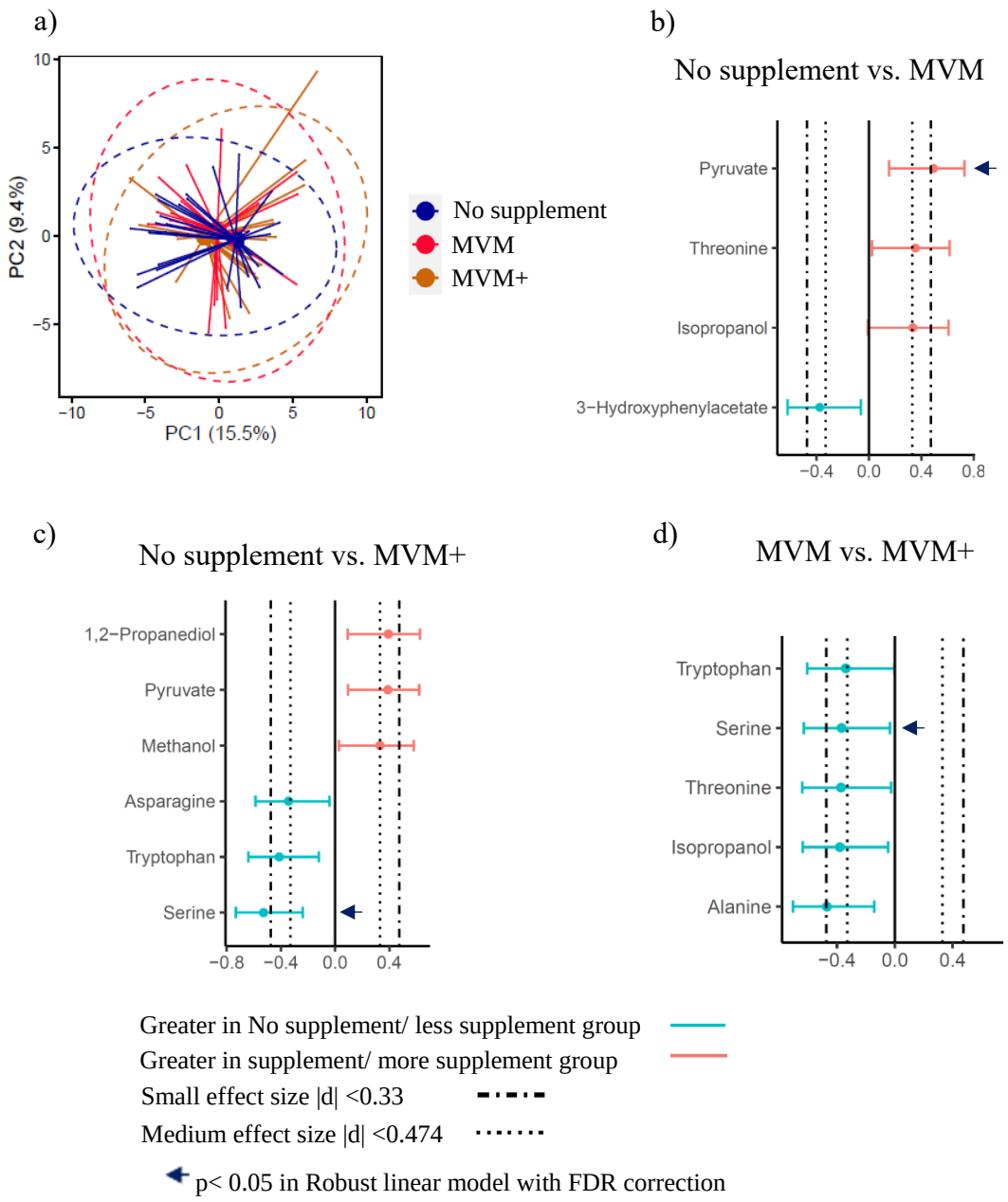


Figure A2. Principal coordinate analysis (PCoA) utilizing Euclidian distance of children's urine metabolome by micronutrient supplementation status a), and the metabolites showed difference between supplement groups based on Cliff's Delta effect size b)-d).

Discussion

In this study we investigated the impact of prenatal alcohol as well as MVM and choline supplementation, on the offspring urine metabolome at 2-4 years. While prenatal alcohol exposure did not have a measurable impact on the urine metabolome of offspring at 2 and 4 years of age, choline and MVM supplementation during pregnancy did have a measurable impact. These results are significant as choline and MVM supplementation has been shown to mitigate the adverse effects of prenatal alcohol exposure (Kable et al. 2015; Jacobson et al. 2018; Coles et al. 2015; Thomas, Abou, and Dominguez 2009; Thomas et al. 2010). In the present study, choline and MVM supplementation impacted several microbially-derived metabolites in the offspring urine (Figure A2).

The higher level of microbial-derived metabolites such as isopropanol, 1,2-propanediol and methanol in the children of women in the MVM or MVM+ group compared to the non-supplemented group indicate differences in microbial function in supplemented groups. Isopropanol was reported to be generated by *Clostridium* species under anaerobic growth and by *Propionibacteria* as a co-product when producing propionate (Walther and François 2016). 1,2-propanediol could be produced through fermenting sugar by multiple gut microbes including *Clostridium thermobutyricum*, *Clostridium sphenoides*, *Bacteroides ruminocola*, *Escherichia coli*, *Lactobacillus brevis*, *Lactobacillus buchneri* and *Bifidobacterium* species (Saxena et al. 2010; Bunesova, Lacroix, and Schwab 2016). 3-Hydroxypheneylacetate is a microbial fermentation product of tyrosine in the gut, and a higher level of 3-hydroxyphenylacetate in the non-supplemented group when compared to MVM group (Figure A2 b) may indicate higher activity of *Clostridium* (Xiong et al. 2016, 3) in the non-supplemented group. Higher pyruvate in both the MVM and MVM+ group could also be linked with the gut microbiome, as pyruvate can be

produced by gut microbes from carbohydrates and gluconeogenic amino acids (Alteri, Smith, and Mobley 2009; Oliphant and Allen-Vercoe 2019).

It has been shown that the gut microbiome is intimately connected with brain function and is capable of modulating neurodevelopment in early life (Cryan and Dinan 2012; Mayer et al. 2014). One study conducted in a large Finish cohort reported that the gut microbiota composition was associated with temperament in infancy in a sex-dependent manner (Aatsinki et al. 2019). Gut microbial disruption has been shown to predictive of anxiety in childhood (Callaghan et al. 2019). Neurodevelopmental disorders including autism spectrum disorder and attention deficit hyperactivity have been reported to be alleviated by prebiotic treatments (Grimaldi et al. 2018; Pärtty et al. 2015). Prenatal micronutrient supplementation such as vitamin D has been previously reported to alter the infant gut microbiome and reduce the colonization of *Clostridioides difficile* (Drall et al. 2020). However, few studies have reported the long term impact on children's urine metabolome of other prenatal micronutrient supplementation.

Conclusion

This present study investigated the impact of prenatal MVM and choline supplementation on the offspring urine metabolome at 2 to 4 years. With no alteration in children's urine metabolome caused by FASD outcomes or prenatal alcohol exposure observed in this study, the major metabolic differences induced by MVM and choline supplementation were microbial-related metabolites. This could suggest that the prenatal MVM and choline intake induced an altered gut microbiome in the children. The preliminary findings in the present study add new insights into the long-term effects of prenatal nutrition on infant health. The present study is limited by the small sample size and that the fecal microbiome was not measured. Further studies with

microbiome analysis are needed to investigate the impact of prenatal micronutrient supplementation on the offspring microbiome.

References

Aatsinki, Anna-Katariina, Leo Lahti, Henna-Maria Uusitupa, Eveliina Munukka, Anniina Keskitalo, Saara Nolvi, Siobhain O'Mahony, et al. 2019. "Gut Microbiota Composition Is Associated with Temperament Traits in Infants." *Brain, Behavior, and Immunity* 80 (August): 849–58. doi: 10.1016/j.bbi.2019.05.035.

Alteri, Christopher J., Sara N. Smith, and Harry L. T. Mobley. 2009. "Fitness of Escherichia Coli during Urinary Tract Infection Requires Gluconeogenesis and the TCA Cycle." *PLOS Pathogens* 5 (5): e1000448. doi: 10.1371/journal.ppat.1000448.

Bunesova, Vera, Christophe Lacroix, and Clarissa Schwab. 2016. "Fucosyllactose and L-Fucose Utilization of Infant Bifidobacterium Longum and Bifidobacterium Kashiwanohense." *BMC Microbiology* 16 (1): 248. doi: 10.1186/s12866-016-0867-4.

Callaghan, Bridget, Heidi Meyer, Maya Opendak, Michelle Van Tieghem, Chelsea Harmon, Anfei Li, Francis S. Lee, Regina M. Sullivan, and Nim Tottenham. 2019. "Using a Developmental Ecology Framework to Align Fear Neurobiology across Species." *Annual Review of Clinical Psychology* 15 (1): 345–69. doi: 10.1146/annurev-clinpsy-050718-095727.

Coles, Claire D., Julie A. Kable, Carl L. Keen, Kenneth Lyons Jones, Wladimir Wertelecki, Irina V. Granovska, Alla O. Pashtepa, and Christina D. Chambers. 2015. "Dose and Timing of Prenatal Alcohol Exposure and Maternal Nutritional Supplements: Developmental Effects on 6-Month-Old Infants." *Maternal and Child Health Journal* 19 (12): 2605–14. doi: 10.1007/s10995-015-1779-x.

Cryan, John F., and Timothy G. Dinan. 2012. "Mind-Altering Microorganisms: The Impact of the Gut Microbiota on Brain and Behaviour." *Nature Reviews Neuroscience* 13 (10): 701–12. doi: 10.1038/nrn3346.

Dieterle, Frank, Alfred Ross, Götz Schlotterbeck, and Hans Senn. 2006. "Probabilistic Quotient Normalization as Robust Method to Account for Dilution of Complex Biological Mixtures. Application in 1H NMR Metabonomics." *Analytical Chemistry* 78 (13): 4281–90. doi: 10.1021/ac051632c.

Drall, Kelsea M., Catherine J. Field, Andrea M. Haqq, Russell J. de Souza, Hein M. Tun, Nadia P. Morales-Lizcano, Theodore B. Konya, et al. 2020. "Vitamin D Supplementation in Pregnancy and Early Infancy in Relation to Gut Microbiota Composition and C. Difficile Colonization: Implications for Viral Respiratory Infections." *Gut Microbes* 12 (1): 1799734. doi: 10.1080/19490976.2020.1799734.

Grimaldi, Roberta, Glenn R. Gibson, Jelena Vulevic, Natasa Giallourou, Josué L. Castro-Mejía, Lars H. Hansen, E. Leigh Gibson, Dennis S. Nielsen, and Adele Costabile. 2018. "A Prebiotic Intervention Study in Children with Autism Spectrum Disorders (ASDs)." *Microbiome* 6 (1): 133. doi: 10.1186/s40168-018-0523-3.

Jacobson, Sandra W., R. Colin Carter, Christopher D. Molteno, Mark E. Stanton, Jane S. Herbert, Nadine M. Lindinger, Catherine E. Lewis, et al. 2018. "Efficacy of Maternal Choline Supplementation During Pregnancy in Mitigating Adverse Effects of Prenatal Alcohol Exposure on Growth and Cognitive Function: A Randomized, Double-Blind, Placebo-Controlled Clinical Trial." *Alcoholism, Clinical and Experimental Research* 42 (7): 1327–41. doi: 10.1111/acer.13769.

Kable, J.A., C.D. Coles, C.L. Keen, J.Y. Uriu-Adams, K.L. Jones, L. Yevtushok, Y. Kulikovsky, W. Wertelecki, T.L. Pedersen, and C.D. Chambers. 2015. "The Impact of Micronutrient Supplementation in Alcohol-Exposed Pregnancies on Information Processing Skills in Ukrainian Infants." *Alcohol* 49 (7): 647–56. doi: 10.1016/j.alcohol.2015.08.005.

Keen, Carl L., Janet Y. Uriu-Adams, Anatoly Skalny, Andrei Grabeklis, Sevil Grabeklis, Kerri Green, Lyubov Yevtushok, Wladimir W. Wertelecki, and Christina D. Chambers. 2010. "The Plausibility of Maternal Nutritional Status Being a Contributing Factor to the Risk for Fetal Alcohol Spectrum Disorders: The Potential Influence of Zinc Status as an Example." *BioFactors (Oxford, England)* 36 (2): 125–35. doi: 10.1002/biof.89.

May, Philip A., Christina D. Chambers, Wendy O. Kalberg, Jennifer Zellner, Haruna Feldman, David Buckley, David Kopald, et al. 2018. “Prevalence of Fetal Alcohol Spectrum Disorders in 4 US Communities.” *JAMA* 319 (5): 474–82. doi: 10.1001/jama.2017.21896.

Mayer, Emeran A., Rob Knight, Sarkis K. Mazmanian, John F. Cryan, and Kirsten Tillisch. 2014. “Gut Microbes and the Brain: Paradigm Shift in Neuroscience.” *Journal of Neuroscience* 34 (46): 15490–96.

Oliphant, Kaitlyn, and Emma Allen-Vercoe. 2019. “Macronutrient Metabolism by the Human Gut Microbiome: Major Fermentation by-Products and Their Impact on Host Health.” *Microbiome* 7 (1): 91. doi: 10.1186/s40168-019-0704-8.

O’Malley, Kieran D. 2007. *ADHD and Fetal Alcohol Spectrum Disorders (FASD)*. Nova Publishers.

Pärtty, Anna, Marko Kalliomäki, Pirjo Wacklin, Seppo Salminen, and Erika Isolauri. 2015. “A Possible Link between Early Probiotic Intervention and the Risk of Neuropsychiatric Disorders Later in Childhood: A Randomized Trial.” *Pediatric Research* 77 (6): 823–28. doi: 10.1038/pr.2015.51.

Popova, Svetlana, Shannon Lange, Charlotte Probst, Gerrit Gmel, and Jürgen Rehm. 2017. “Estimation of National, Regional, and Global Prevalence of Alcohol Use during Pregnancy and Fetal Alcohol Syndrome: A Systematic Review and Meta-Analysis.” *The Lancet. Global Health* 5 (3): e290–99. doi: 10.1016/S2214-109X(17)30021-9.

Rosen Vollmar, Ana K., Nicholas J. W. Rattray, Yuping Cai, Álvaro J. Santos-Neto, Nicole C. Deziel, Anne Marie Z. Jukic, and Caroline H. Johnson. 2019. “Normalizing Untargeted Periconceptional Urinary Metabolomics Data: A Comparison of Approaches.” *Metabolites* 9 (10). doi: 10.3390/metabo9100198.

Saxena, R. K., Pinki Anand, Saurabh Saran, Jasmine Isar, and Lata Agarwal. 2010. “Microbial Production and Applications of 1,2-Propanediol.” *Indian Journal of Microbiology* 50 (1): 2–11. doi: 10.1007/s12088-010-0017-x.

Thomas, Jennifer D., Elizabeth J. Abou, and Hector D. Dominguez. 2009. “Prenatal Choline Supplementation Mitigates the Adverse Effects of Prenatal Alcohol Exposure on Development in Rats.” *Neurotoxicology and Teratology* 31 (5): 303–11. doi: 10.1016/j.ntt.2009.07.002.

Thomas, Jennifer D., Nirelia M. Idrus, Bradley R. Monk, and Hector D. Dominguez. 2010. “Prenatal Choline Supplementation Mitigates Behavioral Alterations Associated with Prenatal Alcohol Exposure in Rats.” *Birth Defects Research Part A: Clinical and Molecular Teratology* 88 (10): 827–37. doi: 10.1002/bdra.20713.

Walther, Thomas, and Jean Marie François. 2016. “Microbial Production of Propanol.” *Biotechnology Advances* 34 (5): 984–96. doi: 10.1016/j.biotechadv.2016.05.011.

Xiong, Xiyue, Dan Liu, Yichao Wang, Ting Zeng, and Ying Peng. 2016. “Urinary 3-(3-Hydroxyphenyl)-3-Hydroxypropionic Acid, 3-Hydroxyphenylacetic Acid, and 3-Hydroxyhippuric Acid Are Elevated in Children with Autism Spectrum Disorders.” Edited by Jeroen Vermeulen. *BioMed Research International* 2016 (March): 9485412. doi: 10.1155/2016/9485412.