

# Human Milk Microbiota and Mycobiota in the CHILD Cohort Study

by  
Shirin Moossavi

A Thesis submitted to the Faculty of Graduate Studies of The University of  
Manitoba  
in partial fulfillment of the requirements of the degree of

DOCTOR OF PHILOSOPHY

Department of Medical Microbiology and Infectious Diseases  
Max Rady Faculty of Health Sciences  
University of Manitoba  
Winnipeg

Copyright © 2020 by Shirin Moossavi

## **Abstract**

Introduction: Previous studies have confirmed the existence of a highly diverse bacterial community in human milk and suggested that its composition might be affected by mode of delivery and maternal health. However, these findings have not yet been reproduced in large-scale studies. Additionally, the compositions of milk microbiota such as milk fungi are less frequently studied.

Methods: Milk microbiota (bacteria) and mycobiota (fungi) were profiled in 393 and 271 mother-infant dyads from the CHILD Cohort Study using amplicon sequencing. Additionally, the milk microbiota was examined in a pilot study using culture-enriched molecular profiling. Finally, the impact of the bioinformatics processing on the results of downstream statistical analysis was compared.

Results: Using multiple analytic techniques including causal modelling, evidence is provided that mode of breastfeeding (directly at the breast vs. pumped and bottle-fed) has a significant association with the composition of milk bacteria potentially highlighting the importance of exogenously-derived bacteria as sources of milk bacteria. The composition of other major milk components were associated with the overall composition of the milk microbiota while controlling for relevant confounding factors.

Defining fungal presence at minimum threshold of 1000 reads per sample, the majority of milk samples did not contain detectable fungi and that presence of fungi was associated with residential characteristics and milk microbiota composition. Fungi diversity was associated with mode of delivery, maternal atopy, and home environment.

We were able to isolate major milk bacteria using culture-enriched methods. Short-term freezing considerably impacted the composition of milk bacteria while we were not able to grow any bacterial isolates from the long-term frozen samples.



Finally, milk microbiota richness was strongly impacted by the choice of the bioinformatics approach. However, there was an acceptable agreement and consistency in the relative abundances of the dominant milk bacteria and milk diversity. Importantly, the main conclusions remained robust to the choice of data processing.

Conclusion: This is among the largest studies of human milk microbiota and mycobiota performed to date. Moreover, this is among the first studies of milk microbiota culture-enriched molecular profiling. Together, these results considerably expand upon existing knowledge about milk microbiota and mycobiota.

## **Acknowledgements**

I would like to thank Dr Meghan Azad for her unwavering support and being a role model in professionalism, collegiality, and strategic planning. I would also like to thank Dr Denice Bay for her invaluable support. I am grateful to Dr Padmaja Subbarao, Director of the CHILD Cohort Study and site leaders (Theo Moraes, Allan Becker, Stuart Turvey, and Piushkumar Mandhane) for their support and enthusiastic encouragement throughout my studies. I would also like to thank Dr Qingling Duan (Queen's University) and Dr Hind Sbihi (University of British Columbia) for providing unpublished genetics and environmental data.

The experiments were performed in Gut Microbiome Laboratories, University of Manitoba. I would like to thank Prof Martin Nyachoti, Dr Ehsan Khafipour, and Dr Shadi Sepehri and the lab members for their support.

I would like to thank members of Dr Azad's group specifically Deborah Chan and Kozeta Miliku. Additionally, I am thankful to Dr Hooman Derakhshani, Kelsey Fehr, Faisal Atakora, and Dr John Schellenberg for helping with experimental work, critical review of the results, graphics, and/or data management/analysis. I would like to thank Angela Nelson and other Med Micro administrative staff for their support.

Milk microbiota project was funded by the Heart and Stroke Foundation and Canadian Lung Association, in partnership with the Canadian Respiratory Research Network and AllerGen NCE. Milk mycobiota project was funded by Children's Hospital Research Institute of Manitoba. I would also like to acknowledge the Research Manitoba Doctoral Studentship and Department of Medical Microbiology Allan Roland Scholarship.

The culture-enriched molecular profiling was funded by a generous travel award from the International Society of Research in Human Milk and Lactation and the Family Larsson Rosenquist Foundation Trainee Expansion Program. I am grateful to Prof Michael Surette for hosting me during this work. Prof Stephanie Atkinson and Dr Gerhard Fusch were

instrumental in obtaining the milk samples and the ethics exemption. Michelle Shah performed the DNA extractions and Laura Rossi prepared the sequencing libraries.

Completing a PhD in essence is a philosophical exercise, a meditative contemplation of the curious mind, a deep dive into the meaning of life and purpose, a rebellious disbelief in the common knowledge and the common practice. My rebellious anarchist nature is thanks to brilliantly unique mother. Peace of mind is the product of one's internal self and the people around. My father has a prominent role in providing an unwavering support throughout my studies. Ching Wu Club, Sigung David Cliffe, and my Kung Fu brothers specially David Anema, David Lane, Randall Flower, and John McAmmond have been my family in Winnipeg. Music is life and life is music and my music teacher, Masoud Arzanlou is the master of the art of conquering the souls. Maureen Brannigan and Tringë Lila Vejsa have been good friends.

## List of papers

This thesis is based on the following manuscripts:

### Original papers

- I. Moossavi S, Sepehri S, Robertson B, Bode L, Goruk S, Field CJ, Lix LM, de Souza, RJ, Becker AB, Mandhane PJ, Turvey SE, Subbarao P, Moraes TJ, Lefebvre DL, Sears MR, Khafipour E, Azad MB. Composition and variation of the human milk microbiome is influenced by maternal and early life factors. *Cell Host Microbe* 2019; 25: 324-335
- II. Moossavi S, Atakora F, Miliku K, Sepehri S, Robertson B, Duan QL, Becker AB, Mandhane PJ, Turvey SE, Moraes TJ, Lefebvre DL, Sears, MR, Subbarao P, Field CJ, Bode L, Khafipour EK, Azad MB. Integrated analysis of human milk microbiota with oligosaccharides and fatty acids in the CHILD cohort. *Front Nutr* 2019; 6: 58
- III. Moossavi S, Fehr K, Derakhshani H, Sbihi H, Robertson B, Bode L, Brook J, Turvey SE, Moraes TJ, Becker AB, Mandhane PJ, Sears MR, Khafipour E, Subbarao P, Azad MB. Human milk fungi: environmental determinants and inter-kingdom associations with milk bacteria (Submitted)
- IV. Moossavi S, Atakora F, Fehr K, Khafipour E, Azad MB. Biological observations in microbiota analysis are robust to the choice of 16S rRNA gene sequencing processing algorithm: case study on human milk microbiota. (Submitted)
- V. Moossavi S, Shah M, Rossi L, Fusch G, Surette M, Azad MB. Capturing the diversity of the human milk microbiota through culture-enriched molecular profiling reveals the impact of storage on milk microbiota composition (In preparation)

### Review papers

- VI. Moossavi S, Miliku K, Sepehri S, Khafipour E, Azad MB. The prebiotic and probiotic properties of human milk: implications for infant immune development and pediatric asthma. *Front Pediatr* 2018; 6: 197
- VII. Moossavi S, Azad MB. Origins of human milk microbiota: new evidence and arising questions. *Gut Microbes* 2019; doi: 10.1080/19490976.2019.1667722

As the author of paper I, I have the rights to include the material in my thesis while the original publication is acknowledged and cited. Papers II, VI, and VII are reproduced under Creative Commons Licence.

## Table of Contents

|                                                                                                                 |       |
|-----------------------------------------------------------------------------------------------------------------|-------|
| Abstract.....                                                                                                   | ii    |
| Acknowledgements .....                                                                                          | iv    |
| List of papers .....                                                                                            | vi    |
| Table of Contents .....                                                                                         | viii  |
| List of Tables .....                                                                                            | xii   |
| List of Figures.....                                                                                            | xiii  |
| List of Boxes.....                                                                                              | xviii |
| List of Copyright Permissions .....                                                                             | xix   |
| List of Abbreviations .....                                                                                     | xx    |
| 1. INTRODUCTION .....                                                                                           | 1     |
| 1.1. Milk Microbiota.....                                                                                       | 2     |
| 1.2. Culture-dependent analysis of milk microbiota .....                                                        | 6     |
| 1.3. Bacterial profile of different milk fractions.....                                                         | 7     |
| 1.4. Milk Mycobiota .....                                                                                       | 8     |
| 1.5. Origins of milk microbes.....                                                                              | 8     |
| 1.6. Determinants of milk microbiota.....                                                                       | 10    |
| 1.7. Milk environment as a microbial niche .....                                                                | 14    |
| 1.8. Influence of the choice of 16S rRNA gene sequencing processing algorithm on milk microbiota analysis ..... | 16    |
| 1.9. Gaps in knowledge .....                                                                                    | 18    |
| 1.10. Hypotheses and objectives .....                                                                           | 19    |
| 2. MATERIALS AND METHODS .....                                                                                  | 21    |
| 2.1. Culture-independent study.....                                                                             | 21    |
| 2.1.1. Study Design .....                                                                                       | 21    |
| 2.1.2. Data accessed from the CHILD database.....                                                               | 23    |
| 2.1.3. Milk sample collection.....                                                                              | 28    |
| 2.1.4. DNA extraction and microbiota analysis .....                                                             | 28    |
| 2.1.5. 16S rRNA and ITS gene sequencing Analysis .....                                                          | 30    |
| 2.1.6. Contamination removal from the sequencing output .....                                                   | 30    |
| 2.1.7. Sequencing output processing following contaminant removal .....                                         | 30    |
| 2.2. Analysis of other milk components.....                                                                     | 31    |
| 2.2.1. HMO Analysis .....                                                                                       | 31    |
| 2.2.2. Fatty acids analysis .....                                                                               | 32    |
| 2.2.3. Metabolic hormones.....                                                                                  | 33    |
| 2.3. Statistical analysis .....                                                                                 | 33    |
| 2.3.1. Handling Missing Data .....                                                                              | 33    |
| 2.3.2. Exclusion of data.....                                                                                   | 33    |
| 2.3.3. Linear association of alpha diversity with maternal, infant, and milk factors ....                       | 34    |
| 2.3.4. Linear relationships between taxa abundance and maternal, infant, and milk factors.....                  | 34    |
| 2.3.5. Redundancy analysis.....                                                                                 | 34    |
| 2.3.6. Hierarchical clustering .....                                                                            | 34    |
| 2.3.7. Structural equation modeling .....                                                                       | 35    |

|                                                                                                                                         |    |
|-----------------------------------------------------------------------------------------------------------------------------------------|----|
| 2.3.8. Linear discriminant analysis .....                                                                                               | 35 |
| 2.3.9. Presence/absence analysis .....                                                                                                  | 35 |
| 2.3.10. Inter-kingdom bacteria fungi correlation network analysis .....                                                                 | 36 |
| 2.4. Comparing sequencing processing approaches .....                                                                                   | 36 |
| 2.4.1. Milk microbiota sequencing data re-processing using UCLUST .....                                                                 | 36 |
| 2.4.2. Comparison of clustering-based and denoising algorithms .....                                                                    | 38 |
| 2.4.3. Interclass correlation analysis .....                                                                                            | 38 |
| 2.5. Culture-enriched molecular profiling .....                                                                                         | 38 |
| 2.5.1. Sample collection .....                                                                                                          | 38 |
| 2.5.2. Culturing of milk samples .....                                                                                                  | 38 |
| 2.5.3. DNA extraction and 16S rRNA gene sequencing .....                                                                                | 40 |
| 2.5.4. Microbial data pre-processing and reagent contaminant removal .....                                                              | 41 |
| 2.5.5. Sequencing output processing following contaminant removal .....                                                                 | 41 |
| 2.5.6. Statistical analysis .....                                                                                                       | 42 |
| 3. MILK MICROBIOTA AND MYCOBIOTA PROFILE IN THE CHILD COHORT STUDY .....                                                                | 44 |
| 3.1. Introduction .....                                                                                                                 | 44 |
| 3.2. Milk microbiota profile in human milk from CHILD Cohort Study .....                                                                | 44 |
| 3.2.1. Batch effect and sequencing inaccuracy were ruled out by quality control assessment .....                                        | 44 |
| 3.2.2. Milk microbiota was dominated by inversely correlated Proteobacteria and Firmicutes with high inter-individual variability ..... | 47 |
| 3.2.3. Core milk microbiota .....                                                                                                       | 49 |
| 3.2.4. Milk microbiota profiles exhibited discrete compositional patterns .....                                                         | 49 |
| 3.3. Milk mycobiota profile in human milk from CHILD Cohort Study .....                                                                 | 52 |
| 3.3.1. Quality control .....                                                                                                            | 52 |
| 3.3.2. Detectable fungi were absent in the majority of milk samples .....                                                               | 53 |
| 3.3.3. Taxonomic composition of milk mycobiota .....                                                                                    | 54 |
| 3.4. Inter-kingdom association of milk bacteria and fungi .....                                                                         | 55 |
| 3.5. Conclusion .....                                                                                                                   | 58 |
| 4. CAPTURING THE DIVERSITY OF THE HUMAN MILK MICROBIOTA THROUGH CULTURE-ENRICHED MOLECULAR PROFILING .....                              | 60 |
| 4.1. Introduction .....                                                                                                                 | 60 |
| 4.2. Bacteria were not isolated from long-term frozen milk samples .....                                                                | 60 |
| 4.3. Milk microbiota richness is impacted by short-term storage .....                                                                   | 61 |
| 4.4. Milk microbiota prevalence and relative abundance are influenced by short-term storage and milk fractions .....                    | 63 |
| 4.5. Milk microbiota composition exhibit inter-individual variability .....                                                             | 65 |
| 4.6. Conclusion .....                                                                                                                   | 66 |
| 5. DETERMINANT OF HUMAN MILK MICROBIOTA .....                                                                                           | 69 |
| 5.1. Introduction .....                                                                                                                 | 69 |
| 5.2. Milk microbiota diversity is associated with mode of breastfeeding, method of milk expression and other maternal factors .....     | 70 |
| 5.3. Associations of maternal and infant factors with milk microbiota alpha diversity are phylum-specific .....                         | 73 |
| 5.4. Milk microbiota taxonomic composition is associated with mode of breastfeeding .....                                               | 74 |
| 5.5. Milk microbiota clusters are associated with mode of breastfeeding, method of milk expression and other maternal factors .....     | 76 |

|      |                                                                                                                                                                               |     |
|------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| 5.6. | Overall milk microbiota composition is associated with breastfeeding practices, multiparity, and other maternal factors in a sex-specific manner.....                         | 78  |
| 5.7. | Structural equation modeling identifies mode of breastfeeding as a key determinant of milk microbiota, and defines other causal pathways influencing milk composition .....   | 79  |
| 5.8. | Indirect breastfeeding is associated with enrichment of potential pathogens and depletion of <i>Bifidobacteria</i> in milk microbiota .....                                   | 83  |
| 5.9. | Conclusion.....                                                                                                                                                               | 84  |
| 6.   | INTEGRATED ANALYSIS OF MILK MICROBIOTA WITH HUMAN MILK OLIGOSACCHARIDES AND FATTY ACIDS .....                                                                                 | 86  |
| 6.1. | Introduction .....                                                                                                                                                            | 86  |
| 6.2. | Overview of milk microbiota, fatty acids, and HMOs profiles in human milk in the CHILD Cohort Study .....                                                                     | 86  |
| 6.3. | Correlations are stronger and more frequent within than between different types of milk components .....                                                                      | 88  |
| 6.4. | Some HMOs and fatty acids are associated with milk microbiota alpha diversity and overall composition .....                                                                   | 90  |
| 6.5. | Some HMOs and fatty acids are weakly associated with specific milk bacteria relative abundances .....                                                                         | 93  |
| 6.6. | Maternal and infant secretor status are not strongly associated with overall milk microbiota composition.....                                                                 | 96  |
| 6.7. | HMOs are associated with milk <i>Bifidobacterium</i> prevalence and abundance .....                                                                                           | 97  |
| 6.8. | Fatty acids profile is associated with milk microbiota composition when controlling for relevant confounding factors .....                                                    | 99  |
| 6.9. | Conclusion.....                                                                                                                                                               | 100 |
| 7.   | DETERMINANTS OF HUMAN MILK MYCOBIOTA .....                                                                                                                                    | 102 |
| 7.1. | Introduction .....                                                                                                                                                            | 102 |
| 7.2. | Participant characteristics.....                                                                                                                                              | 102 |
| 7.3. | Prevalence of milk fungi was significantly and independently associated with environmental characteristics, human milk oligosaccharides, and milk bacterial composition ..... | 102 |
| 7.4. | Milk mycobiota profile was not clearly associated with any available metadata characteristics .....                                                                           | 109 |
| 7.5. | Conclusion.....                                                                                                                                                               | 110 |
| 8.   | COMPARISON OF SEQUENCING PROCESSING APPROACHES ON DOWNSTREAM STATISTICAL RESULTS.....                                                                                         | 112 |
| 8.1. | Introduction .....                                                                                                                                                            | 112 |
| 8.2. | Mock community composition is more accurately resolved using DADA2.....                                                                                                       | 112 |
| 8.3. | Observed OTUs/ASVs (richness) is influenced by the choice of sequencing processing method.....                                                                                | 114 |
| 8.4. | Milk microbiota structure and relative abundances are not strongly impacted by the choice of sequencing processing method .....                                               | 117 |
| 8.5. | Downstream statistical analyses are robust to the choice of data processing approach.. .....                                                                                  | 120 |
| 8.6. | Conclusion.....                                                                                                                                                               | 123 |
| 9.   | DISCUSSION.....                                                                                                                                                               | 124 |
| 9.1. | Summary.....                                                                                                                                                                  | 124 |
| 9.2. | Origins of milk bacteria.....                                                                                                                                                 | 126 |
| 9.3. | Are milk bacteria viable and/or metabolically active? .....                                                                                                                   | 127 |



|         |                                                                                                    |     |
|---------|----------------------------------------------------------------------------------------------------|-----|
| 9.4.    | Determinants of milk microbiota composition.....                                                   | 129 |
| 9.5.    | Sex-specific effect of factors on milk microbiota .....                                            | 130 |
| 9.6.    | Pumping, milk microbiota and associated controversies .....                                        | 133 |
| 9.7.    | Culture-dependent milk microbiota analysis.....                                                    | 134 |
| 9.8.    | Milk microbiota analyses are robust to the choice of data processing and contaminant removal ..... | 136 |
| 9.9.    | Are fungi universally present in milk? .....                                                       | 137 |
| 9.10.   | Origins and determinants of milk mycobiota.....                                                    | 138 |
| 9.11.   | Regional and seasonal differences in the prevalence of fungi .....                                 | 139 |
| 9.12.   | Bacteria-fungi interaction in the milk: direct or indirect interaction?.....                       | 140 |
| 9.13.   | Milk microbiota and other milk components .....                                                    | 141 |
| 9.14.   | HMOs and milk Bifidobacterium.....                                                                 | 143 |
| 9.15.   | Milk fatty acids and milk microbiota .....                                                         | 143 |
| 9.16.   | Functional importance of milk microbiota.....                                                      | 145 |
| 10.     | OVERALL CONCLUSION AND FUTURE DIRECTION .....                                                      | 150 |
| 10.1.   | Conclusion.....                                                                                    | 150 |
| 10.2.   | Strengths.....                                                                                     | 153 |
| 10.3.   | Limitations .....                                                                                  | 153 |
| 10.4.   | Future direction .....                                                                             | 154 |
| 10.4.1. | Scientific considerations .....                                                                    | 154 |
| 10.4.2. | Technical considerations.....                                                                      | 156 |
| 10.4.3. | Practical and translational considerations .....                                                   | 157 |
| 10.5.   | Closing remark .....                                                                               | 158 |
| 11.     | REFERENCES .....                                                                                   | 160 |

## List of Tables

|                                                                                                                                                                                                                                                                                                           |     |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Table 1. Characteristics of mother-infant dyads from the CHILD Cohort Study included in the microbiota analysis (n=393) in comparison with all eligible dyads (n=2536) .....                                                                                                                              | 24  |
| Table 2. List and definition of variables used in this study .....                                                                                                                                                                                                                                        | 27  |
| Table 3. Human milk oligosaccharides (HMOs) and fatty acids structural categories.. .....                                                                                                                                                                                                                 | 32  |
| Table 4. Key resource table for the culture-independent study of milk microbiota in the CHILD Cohort Study.. .....                                                                                                                                                                                        | 37  |
| Table 5. Culture media used in the culture-enriched molecular profiling of human milk samples. ....                                                                                                                                                                                                       | 39  |
| Table 6. Key resource table for the culture-enriched molecular profiling of milk microbiota. ....                                                                                                                                                                                                         | 42  |
| Table 7. Most abundant bacterial genera (>1% mean relative abundance) in the human milk microbiota among 393 mothers in the CHILD Cohort Study related to Figure 11. ....                                                                                                                                 | 48  |
| Table 8. Most frequently isolated bacterial families in the human milk microbiota in 5 human milk samples from McMaster University assessed by culture-enriched molecular profiling related to Figure 22.....                                                                                             | 64  |
| Table 9. Comparison of milk samples assessed in the CHILD Cohort Study to the milk samples collected at McMaster and assessed by culture-enriched molecular profiling.....                                                                                                                                | 67  |
| Table 10. Factors associated with milk microbiota alpha diversity in 393 human milk samples from the CHILD Cohort Study.....                                                                                                                                                                              | 72  |
| Table 11. Comparison of the equivalent causal models of association of maternal and infant factors influencing the milk microbiota to the final model (see Figure 31). ....                                                                                                                               | 83  |
| Table 12. Association of milk microbiota overall composition with human milk oligosaccharides (HMOs) and fatty acids in 393 human milk samples from the CHILD Cohort Study.....                                                                                                                           | 93  |
| Table 13. Characteristics of mother-infant dyads from the CHILD Cohort Study included in this analysis in comparison with previous microbiota analysis.....                                                                                                                                               | 103 |
| Table 14. Milk characteristics of mother-infant dyads in milk mycobiota analysis in the CHILD Cohort Study. ....                                                                                                                                                                                          | 104 |
| Table 15. Comparison of the number of OTUs/ASVs in the mock community, negative controls, and milk microbiota datasets processed by clustering-based OTU method (UCLUST) and a denoising algorithm (DADA2) with or without contaminant removal in 393 human milk samples from the CHILD Cohort Study..... | 114 |

## List of Figures

|                                                                                                                                                                                                                               |    |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Figure 1. Repertoire of milk bacteria at the phylum level based on systematic review of the literature.....                                                                                                                   | 4  |
| Figure 2. Proportion of bacterial species identified from breast and breast milk according to their first isolation source.....                                                                                               | 5  |
| Figure 3. Origins of milk microbiota.....                                                                                                                                                                                     | 12 |
| Figure 4. Schematic overview of this thesis.....                                                                                                                                                                              | 20 |
| Figure 5. Flowchart of the participant selection for microbiota and mycobiota analysis in CHILD Cohort Study. ....                                                                                                            | 22 |
| Figure 6. Milk culture-enriched molecular profiling of different milk fractions. ....                                                                                                                                         | 40 |
| Figure 7. Milk microbiota beta diversity by sample type, sequencing run, and PCR plates, prior to removal of potential contaminant ASVs using the decontam package in 393 human milk samples from the CHILD Cohort Study..... | 45 |
| Figure 8. Data quality control in human milk samples from the CHILD Cohort Study. ....                                                                                                                                        | 45 |
| Figure 9. Relative abundance of most abundant bacterial genera (>1% mean relative abundance) in the mock community in comparison with the theoretical composition. ....                                                       | 46 |
| Figure 10. Sequencing output processing to remove low abundance ASVs from milk microbiota data in the CHILD Cohort Study. ....                                                                                                | 47 |
| Figure 11. Milk microbiota profile at A) phylum and B) genus levels in 393 human milk samples from the CHILD Cohort Study.....                                                                                                | 48 |
| Figure 12. Prevalence of core ASVs in 393 human milk samples from the CHILD Cohort Study compared to negative controls .....                                                                                                  | 50 |
| Figure 13. Hierarchical clustering of core milk microbiota among 393 mothers in the CHILD Cohort Study.....                                                                                                                   | 51 |
| Figure 14. Defining fungi presence in milk according to the depth of sequencing informed by the positive PCR bands in 271 human milk samples from the CHILD Cohort Study. ....                                                | 53 |
| Figure 15. Milk fungi taxonomy composition at A-B) phylum and C) species levels in 271 human milk samples from the CHILD Cohort Study.....                                                                                    | 54 |
| Figure 16. Association of human milk bacteria and fungi in 271 human milk samples from the CHILD Cohort Study.....                                                                                                            | 56 |

|                                                                                                                                                                                                                                                                                                                                                                  |    |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Figure 17. Co-occurrence analysis of most abundant bacteria and fungi (>1% mean relative abundance) in 271 human milk samples from the CHILD Cohort Study based on Spearman rank correlation.....                                                                                                                                                                | 57 |
| Figure 18. Schematic overview of milk microbiota culture-enriched molecular profiling in 8 human milk samples from McMaster University.....                                                                                                                                                                                                                      | 61 |
| Figure 19. Comparison of the observed ASVs between A) Samples and B) Culture media in 5 human milk samples from McMaster University assessed by culture-enriched molecular profiling.....                                                                                                                                                                        | 62 |
| Figure 20. Comparison of the number of detected amplicon sequence variants (ASVs) in 5 human milk samples from McMaster University using culture-enriched molecular profiling.....                                                                                                                                                                               | 62 |
| Figure 21. Sequencing output processing to remove low abundance ASVs in 5 human milk samples from McMaster University.....                                                                                                                                                                                                                                       | 63 |
| Figure 22. Comparison of milk microbiota composition in 5 human milk samples from McMaster University using culture-enriched molecular profiling at the family level based on storage (Fresh vs. Short-term Frozen, N samples=5), Oxygen availability during incubation (Aerobic vs. Anaerobic, N samples=5), and in different milk fractions (N samples=2)..... | 65 |
| Figure 23. Milk microbiota composition at phylum and family levels in human milk samples from McMaster using culture-enriched molecular profiling in individual milk samples comparing A) Fresh vs. Short-term Frozen samples (N=5), and B) Between different milk fractions (N=2). .....                                                                        | 66 |
| Figure 24. Theoretical framework of how different factors are associated with milk microbiota composition.....                                                                                                                                                                                                                                                   | 69 |
| Figure 25. Factors associated with milk microbiota alpha diversity in 393 human milk samples from the CHILD Cohort Study.....                                                                                                                                                                                                                                    | 71 |
| Figure 26. Milk microbiota structure is not strongly associated with maternal, infant, and early life factors in 393 human milk samples from the CHILD Cohort Study.. .....                                                                                                                                                                                      | 75 |
| Figure 27. Milk microbiota is associated with mode of breastfeeding by linear discriminant analysis of taxa enrichment in 393 human milk samples from the CHILD Cohort Study.....                                                                                                                                                                                | 76 |
| Figure 28. Univariate associations of maternal and infant factors with milk microbiota composition and cluster membership in 393 human milk samples from the CHILD Cohort Study.....                                                                                                                                                                             | 77 |
| Figure 29. Redundancy analysis of associations of maternal, infant, early life, and milk factors with the overall milk microbiota composition in 393 human milk samples from the CHILD Cohort Study. ....                                                                                                                                                        | 79 |
| Figure 30. Causal modelling of maternal, infant, and milk factors influencing the milk microbiota composition in 393 human milk samples from the CHILD Cohort Study. ....                                                                                                                                                                                        | 81 |

|                                                                                                                                                                                                                                                       |     |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Figure 31. Equivalent structural equation models assessing other potentially plausible models of associations of milk microbiota determinants in 393 milk samples from the CHILD Cohort Study. Reproduced from Moossavi et al. (139).....             | 82  |
| Figure 32. Milk microbiota is associated with mode of breastfeeding.....                                                                                                                                                                              | 84  |
| Figure 33. Milk microbiota, human milk oligosaccharides (HMOs), and fatty acids relative abundances in 393 human milk samples from the CHILD Cohort Study.....                                                                                        | 87  |
| Figure 34. Inter and intra-correlations of milk microbiota, human milk oligosaccharides (HMOs), and fatty acids by clustering heatmap based on Spearman rank correlation in 393 human milk samples from the CHILD Cohort Study.....                   | 88  |
| Figure 35. Inter and intra-correlations of milk microbiota, human milk oligosaccharides (HMOs), and fatty acids based on network visualisation based on partial Spearman rank correlations in 393 human milk samples from the CHILD Cohort Study..... | 89  |
| Figure 36. Association of A) human milk oligosaccharides (HMOs) and B) fatty acids with milk microbiota richness (observed ASVs) and diversity (inverse Simpson index) in 393 human milk samples from the CHILD Cohort Study.....                     | 91  |
| Figure 37. Association of milk fatty acids with milk microbiota richness (observed ASVs) and diversity (inverse Simpson index) in 393 human milk samples from the CHILD Cohort Study.....                                                             | 92  |
| Figure 38. Association of milk microbiota with human milk oligosaccharides stratified by maternal FUT2 secretor status.....                                                                                                                           | 94  |
| Figure 39. Correlation of core milk microbiota (>1% relative abundance) with milk components in 393 human milk samples from the CHILD Cohort Study. ....                                                                                              | 95  |
| Figure 40. Association of maternal and infant secretor status with milk microbiota alpha diversity in 393 human milk samples from the CHILD Cohort Study.....                                                                                         | 96  |
| Figure 41. Lack of association between maternal and infant FUT2 secretor status and milk microbiota overall composition in 393 human milk samples from the CHILD Cohort Study..                                                                       | 97  |
| Figure 42. Correlation of milk Bifidobacterium relative abundance (centre log-ratio transformed) with human milk oligosaccharides. ....                                                                                                               | 98  |
| Figure 43. Association of Bifidobacterium prevalence with human milk oligosaccharides in 393 human milk samples from the CHILD Cohort Study.....                                                                                                      | 99  |
| Figure 44. Structural equation modelling of the association of human milk oligosaccharides (HMOs) and fatty acids with microbiota in 393 human milk samples from the CHILD Cohort Study.....                                                          | 100 |

|                                                                                                                                                                                                                                                                                                                       |     |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Figure 45. Prevalence of fungi according to study site, population density, and season tested by chi-square test in 271 human milk samples from the CHILD Cohort Study. ....                                                                                                                                          | 105 |
| Figure 46. Unadjusted univariate associations of fungal prevalence with maternal, infant and early life factors, breastfeeding and selected HMOs, indoor and outdoor environmental factors, and milk microbiota composition tested by logistic regression in 271 human milk samples from the CHILD Cohort Study. .... | 106 |
| Figure 47. Relative abundances and absolute concentrations of Disialyllacto-N-tetraose (DSLNT) and Lacto-N-hexaose (LNH) based on milk fungi presence tested by ANOVA in 271 human milk samples from the CHILD Cohort Study. ....                                                                                     | 107 |
| Figure 48. Prediction accuracy of fungi presence based on unsupervised variable selection by LASSO in 271 human milk samples from the CHILD Cohort Study. ....                                                                                                                                                        | 108 |
| Figure 49. Adjusted multivariable associations of host and environmental factors with the prevalence of fungi in 271 human milk samples from the CHILD Cohort Study. ....                                                                                                                                             | 109 |
| Figure 50. Association of intrapartum antibiotics and study city with the prevalence of fungal genera tested by chi-squared test in 58 human milk samples from the CHILD Cohort Study. ....                                                                                                                           | 110 |
| Figure 51. Association of birth mode and city with fungal richness and maternal atopy and indoor moisture level with fungal diversity in 58 human milk samples from the CHILD Cohort Study. ....                                                                                                                      | 111 |
| Figure 52. Schematic of the study design to compare microbiota composition in a mock community and human milk samples using a clustering-based OTU method (UCLUST) and a denoising algorithm (DADA2) with and without contaminant removal in 393 human milk samples from the CHILD Cohort Study. ....                 | 113 |
| Figure 53. Composition of the mock community by UCLUST and DADA2 prior to contaminant removal. Each bar represents the average data from 8 replicates. ....                                                                                                                                                           | 115 |
| Figure 54. Comparison of the library size on datasets processed by clustering-based OTU method (UCLUST) and a denoising algorithm (DADA2) with or without contaminant removal in 393 human milk samples from the CHILD Cohort Study. ....                                                                             | 115 |
| Figure 55. Comparison of milk microbiota richness (observed OTUs/ASVs) and diversity (inverse Simpson index) between UCLUST and DADA2 with and without contaminant removal in 393 human milk samples from the CHILD Cohort Study. ....                                                                                | 116 |
| Figure 56. Comparison of the A) relative abundance and B) prevalence of bacterial genera on datasets processed by clustering-based OTU method (UCLUST) and a denoising algorithm (DADA2) without contaminant removal in 393 human milk samples from the CHILD Cohort Study. ....                                      | 117 |
| Figure 57. Comparison of the relative abundance and prevalence of abundant bacterial genera (>0.01% mean relative abundance) on datasets processed by clustering-based OTU method                                                                                                                                     |     |

|                                                                                                                                                                                                                                                                                                                              |     |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| (UCLUST) and a denoising algorithm (DADA2) without contaminant removal in 393 human milk samples from the CHILD Cohort Study.....                                                                                                                                                                                            | 118 |
| Figure 58. Comparison of the composition of abundant families (>1% mean relative abundance) between UCLUST and DADA2 with or without contaminant removal in 393 human milk samples from the CHILD Cohort Study.....                                                                                                          | 119 |
| Figure 59. Comparison of the relative abundances of four abundant families (>1% mean relative abundance) on datasets processed by clustering-based OTU method (UCLUST) and a denoising algorithm (DADA2) with or without contaminant removal in 393 human milk samples from the CHILD Cohort Study.....                      | 120 |
| Figure 60. A) Agreement and B) consistency between methods of sequencing processing by intraclass correlation for alpha diversity and 13 most abundant families in 393 human milk samples from the CHILD Cohort Study.....                                                                                                   | 121 |
| Figure 61. Impact of four sequence processing approaches on observed associations of milk microbiota richness (observed OTUs/ASVs), diversity (inverse Simpson index), and overall composition with maternal, infant, early life, breastfeeding, and milk factors in 393 human milk samples from the CHILD Cohort Study..... | 122 |
| Figure 62. Different dispersal routes of bacteria between maternal body sites, infant, and breastfeeding aids. ....                                                                                                                                                                                                          | 128 |
| Figure 63. Potential sources of milk mycobiota and factors influencing their prevalence and composition.. ....                                                                                                                                                                                                               | 140 |
| Figure 64. Association of breastfeeding with lung health and asthma.....                                                                                                                                                                                                                                                     | 149 |

**List of Boxes**

|                                                        |     |
|--------------------------------------------------------|-----|
| Box 1. Summary of Chapter 3 .....                      | 59  |
| Box 2. Summary of Chapter 4 .....                      | 68  |
| Box 3. Summary of Chapter 5 .....                      | 85  |
| Box 4. Summary of Chapter 6 .....                      | 101 |
| Box 5. Summary of Chapter 7 .....                      | 111 |
| Box 6. Summary of Chapter 8 .....                      | 123 |
| Box 7. Summary of the overall conclusions.....         | 152 |
| Box 8. Unanswered questions about milk microbiota..... | 155 |



## **List of Copyright Permissions**

Figures 1 and 2 are adapted from Togo et al. *Future Microbiology* 2019; 14: 623–641; Future Medicine Ltd.

Figures 3, 29, and 59 are reproduced from Moossavi & Azad. *Gut Microbes* 2019; doi: 10.1080/19490976.2019.1667722; Taylor & Francis Online under Creative Commons Licence.

Figure 6 is adapted from Lima et al. *PLoS One* 2018; 13: e0193671 under Creative Commons Licence.

Figures 7-10, 13, 24, 27, 30-32 and Tables 1, 4, 7, 10, and 11 are reproduced from Moossavi et al. *Cell Host Microbe* 2019; 25: 324-335; Elsevier Inc. As the author of this article, I retain the right to include it in my thesis, provided it is not published commercially. Permission is not required, and the journal is cited as the original source.

Figure 18 is adapted and modified from Lau et al. *Genome Medicine* 2016;8: 72; BioMed Central under Creative Commons Licence.

Figure 33-41 and Table 3, 12 are reproduced from Moossavi et al. *Frontiers in Nutrition* 2019; 6: 58 under Creative Commons Licence

Figure 61 is reproduced from Moossavi et al. *Frontiers in Pediatrics* 2018; 6: 197 under Creative Commons Licence.

## List of Abbreviations

|        |                                                  |
|--------|--------------------------------------------------|
| 2'FL   | 2'-fucosyllactose                                |
| 3FL    | 3-fucosyllactose                                 |
| 3'SL   | 3'-sialyllactose                                 |
| 6'SL   | 6'-sialyllactose                                 |
| AA     | Arachidonic acid                                 |
| Abx    | Antibiotics                                      |
| ALA    | $\alpha$ -linolenic acid                         |
| ANOVA  | Analysis of variance                             |
| ASV    | Amplicon sequence variant                        |
| AUROC  | Area under the receiver operation curve          |
| BBE    | Bacteroides bile esculin agar                    |
| BEEF   | Cooked meat agar                                 |
| BF     | Breastfeeding                                    |
| BHI3   | Brain heart infusion agar                        |
| BMI    | Body Mass Index                                  |
| BSM    | Bifidobacterium Selective Media                  |
| CBA    | Columbia blood agar                              |
| CFA    | Confirmatory factor analysis                     |
| CFI    | Comparative fix index                            |
| CHILD  | Canadian Healthy Infant Longitudinal Development |
| CI     | Confidence interval                              |
| CLA    | Conjugated linoleic acid                         |
| C/S    | Caesarean section                                |
| DFLac  | Difucosyllactose                                 |
| DFLNH  | Difucosyl-lacto-N-hexaose                        |
| DFLNT  | Difucosyl-LNT                                    |
| DHA    | Docosahexaenoic acid                             |
| DHGLA  | Dihomo- $\gamma$ -linolenic acid                 |
| DPA    | Docosapentaenoic acid                            |
| DSLNH  | Disialyl-lacto-N-hexaose                         |
| DSLNT  | Disialyl-LNT                                     |
| FDSLNH | Fucosyl-disialyl-lacto-N-hexaose                 |
| FAA    | Fastidious anaerobe agar                         |
| FDR    | False discovery rate                             |
| FFQ    | Food Frequency Questionnaire                     |
| EPA    | Eicosapentanoic acid                             |
| FLNH   | Fucosyl-lacto-N-hexaose                          |
| FUT2   | Fucosyltransferase 2                             |
| GLA    | $\gamma$ -linoleic acid                          |
| HMO    | Human milk oligosaccharide                       |
| ITS    | Internal transcribed spacer                      |
| LA     | Linoleic acid                                    |
| LASSO  | Least absolute shrinkage and selection operator  |
| LDA    | Linear discriminant analysis                     |
| LefSe  | Linear discriminant analysis effect size         |
| LNFP   | Lacto-N-fucopentaose                             |
| LNH    | Lacto-N-hexaose                                  |

|           |                                              |
|-----------|----------------------------------------------|
| LNT       | Lacto-N-tetraose                             |
| LNTnT     | Lacto-N-neotetraose                          |
| LST       | Sialyl-LNT                                   |
| MAC       | MacConkey agar                               |
| MFA       | Milk fatty acids                             |
| MICE      | Multivariate imputation by chained equations |
| Min       | Minutes                                      |
| ML        | Maximum likelihood                           |
| MSA       | Mannitol slat agar                           |
| MRS       | de Man Rogosa Sharpe agar                    |
| MUFA      | Monounsaturated fatty acids                  |
| NDVI      | Normalized difference vegetation index       |
| NICU      | Neonatal Intensive Care Unit                 |
| NVD       | Normal vaginal delivery                      |
| OTU       | Operational taxonomic unit                   |
| PC        | Principal component                          |
| PCR       | Polymerase chain reaction                    |
| PERMANOVA | permutational ANOVA                          |
| PUFA      | Polyunsaturated fatty acids                  |
| RDA       | Redundancy analysis                          |
| RSMEA     | Root mean square error of approximation      |
| S         | Seconds                                      |
| SD        | Standard Deviation                           |
| SEM       | Structural equation modelling                |
| SFA       | Saturated fatty acid                         |
| SNP       | Single nucleotide polymorphism               |
| SRMR      | Standardized root mean residuals             |
| Subsp.    | Subspecies                                   |
| THY       | tryptic soy agar with yeast extract          |
| TVA       | Trans vaccenic acid                          |

# **1. INTRODUCTION**

Breastfeeding has many established benefits for maternal and child health (1), including shaping the infant gut microbiome development (2-5) and protecting against infectious diseases in infancy (6, 7) and non-communicable diseases/conditions such as asthma and obesity (8-14). These associations might vary according to the mode of breastfeeding i.e. among infants fed at the breast compared to those receiving pumped breast milk (10). This suggests that the act of suckling and/or skin-to-skin contact may contribute to the protective effect of breastfeeding. Alternatively or in addition, this finding could reflect a role for bioactive components including milk microbiota, which may be altered during the pumping and storage process.

Although previously considered sterile, breast milk is now known to contain a complex community of bacteria (15). While their exact function is not known, it is speculated that milk microbiota help establish the infant's gut microbial community and influence development of the immune system (15). If this process is disrupted, the infant may develop a dysbiotic microbial community, causing predisposition to a range of chronic diseases such as allergy, asthma and obesity (16). Previous studies on human milk microbiota have confirmed the existence of a highly diverse bacterial community using culture-dependent and -independent techniques (17) and suggested that its composition might be affected by local pathologies of the breast, mode of delivery, antibiotics, maternal health, and gestational age (18). However, these findings have not yet been reproduced in large-scale studies and the determinants of milk microbiota are still mostly unknown. Additionally, the compositions of other microbial components of the milk such as mycobiota are less frequently studied. Throughout this thesis, microbiota and mycobiota are used to refer to bacterial and fungal communities, respectively.

## 1.1. Milk Microbiota

Milk bacteriology has been mostly studied from an infectious disease and food safety perspectives and any bacteria in the milk were considered harmful contaminants introduced during handling, processing, and storage (19). In the 21<sup>st</sup> century, medical microbiology has undergone a paradigm shift from the exclusive focus on pathogenic bacteria and infectious diseases to appreciating the role of non-pathogenic bacteria and microbial communities in both infectious and noncommunicable diseases, as well as physiologic homeostasis and health (20). This shift in perspective along with the availability of next generation sequencing has resulted in an exponential increase in the study of host-associated microbiota (21). Specifically, given the central role of breastfeeding in shaping the infant gut microbiota composition (4), there is ample interest in identifying milk components which mediate this effect on the gut microbiota. While the focus has mostly been on the human milk oligosaccharides (HMOs), there is an open question about the role of milk microbiota in shaping the infant gut microbiota. Unlike other environmental sources of bacteria to the infant, breastfeeding and milk composition is evolutionarily conserved (22) suggesting that infant gut microbiota and milk components may have co-evolved to facilitate healthy microbial ecosystem development.

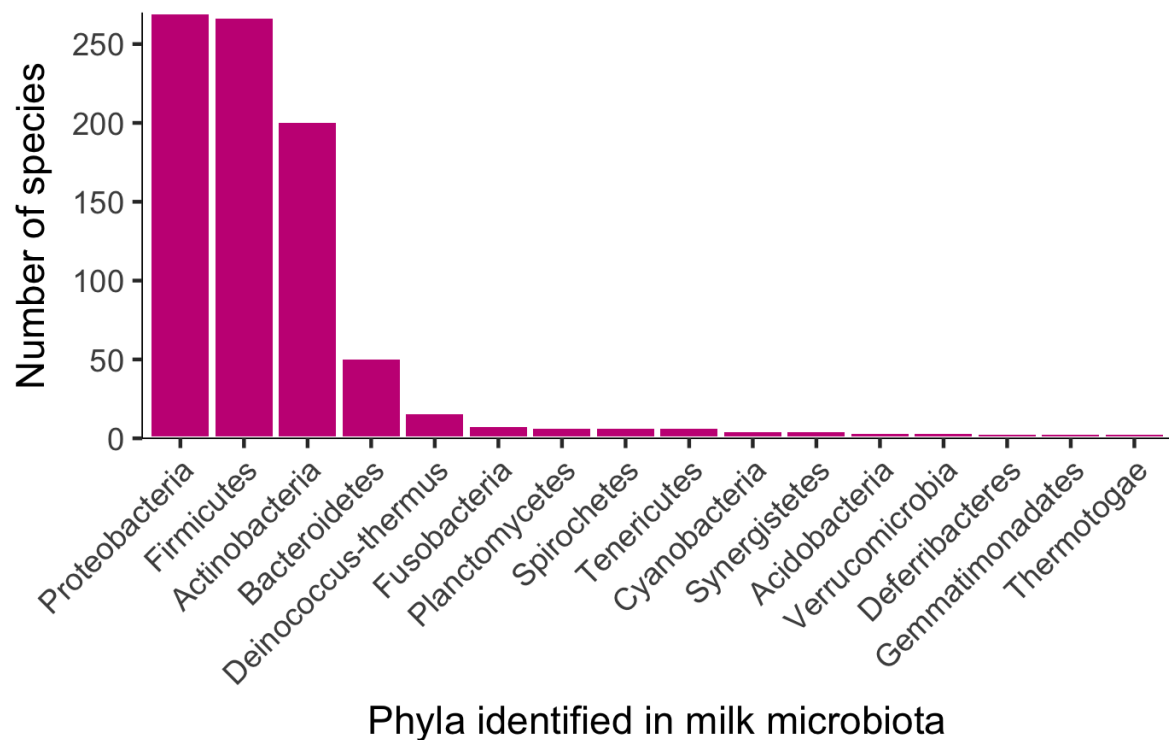
Culture-dependent and -independent studies have confirmed that human and dairy milk are not sterile and rather contain a community of live bacteria, fungi, and viruses (23). Earlier culture-based studies of the milk bacteria in 1970s suggest that the majority of milk samples have low bacterial cell concentrations (<2,500 bacteria per mL) while 20% of samples had high bacterial counts (>5,000) which were considered unacceptable for donor milk (24). Potential pathogens such as *Staphylococcus aureus* were isolated from expressed milk of healthy lactating mothers, albeit at varying concentrations (100-40,000 cells per mL) (25). More recent targeted culture-based studies in the last decade have estimated that breastfed

infants receive  $10^4$  to  $10^6$  bacteria per day (based on an average daily consumption of 800 mL of milk) with most isolated species belong to the genera *Staphylococcus*, *Streptococcus*, *Lactobacillus*, and *Bifidobacterium* (26-33). The bacterial load delivered via breastmilk is approximately similar to the number of live bacteria in fermented food including yogurt and fermented vegetables (34). It is not known whether bacteria are present in the breastmilk in all mothers and if so whether they are at sufficiently high quantities to be detected. Approximately 5% of milk samples did not yield any bacterial isolates in a culture-dependent study (33) suggesting that some milk samples may not contain detectable bacteria. The prevalence of bacterial detection will likely be influenced by sample storage and the choice of bacterial detection technique especially for bacteria requiring specialised culture conditions (e.g. anaerobic incubation). For example, while the frequency of isolation for dominant bacteria including *Staphylococcus* and *Streptococcus* was similar for culture and qPCR, *Bifidobacterium* was identified at higher frequency in molecular assessment (80% by qPCR vs. 10% by culture) (30).

The first sequencing-based human milk microbiota study was reported by Hunt et al. in 16 samples (35) followed by an increased interest in the topic by multiple groups in different countries (36-45). Proteobacteria, Firmicutes, and Actinobacteria constituted the majority of detected bacteria (Figure 1, (23)).

*Staphylococcus* and *Streptococcus* were the most abundant bacteria in the majority of these studies (23). However, there was a considerable variability in the milk composition specifically for the less abundant taxa. A number of these studies attempted to identify core bacterial species in the milk samples (35, 38, 40, 41, 46). However, differences in the definition of core microbiota along with differences in methodology, reagents, geography, as well as the small sample size in the majority of the studies resulted in discordant core microbiota representation. The majority of milk bacteria were first isolated from other

environmental or human sources including human gut, respiratory tract, and oral cavity (Figure 2, (23)).

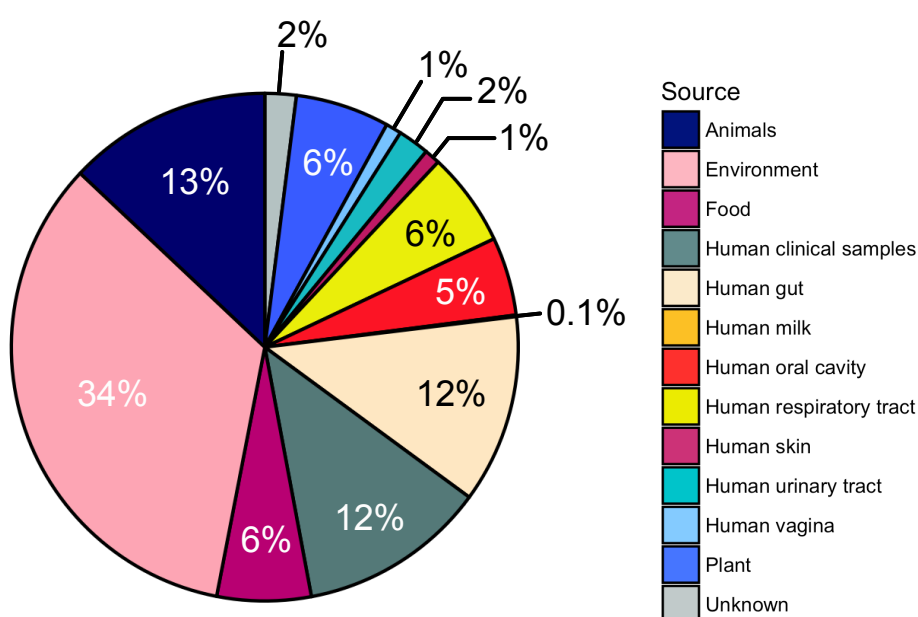


**Figure 1. Repertoire of milk bacteria at the phylum level based on systematic review of the literature.** Adapted from Togo et al. (23).

Overall, milk microbiota demonstrated a high degree of taxonomic inter-individual variability with only a minority of taxa shared across the study population (18). Taxonomic inter-individual variability (i.e. functional redundancy) is an established phenomenon in microbial ecology and human microbiome (47, 48). However, milk microbiome inter-individual variability could also be attributed to the variation in sample collection (e.g. use of pump vs. manual expression, fore- vs. hind-milk, time to the last feed, etc.) (49).

Milk is considered a low biomass sample and thus prone to be contaminated through sample collection, handling, and processing (including DNA extraction and sequencing library preparation) (50). Specifically, reagent contamination with environmental bacterial

DNA pose an important challenge in sequencing-based microbiome studies resulting in domination of the microbiome composition of low biomass samples such as milk by the contaminants (51-53). In a recent study directly addressing this concern, dairy milk samples were both cultured on non-selective media and assessed by 16S rRNA gene sequencing (51). As expected from the previous reports on reagent contaminants described above, the relative abundances of potential reagent contaminants were negatively associated with the total bacterial growth in culture. In other words, samples with higher bacterial loads had considerably lower relative abundances of the potential contaminants (51). Therefore, it is crucial to minimise the potential contaminants by avoiding batch variation for the DNA extraction and other major reagents, minimise potential environmental contamination by perform experiments in a biosafety cabinet, including appropriate negative controls, and statistically identifying potential contaminants. Previous sequencing-based milk microbiome studies have not included and/or reported negative controls and have not attempted to identify potential reagent contaminants.



**Figure 2. Proportion of bacterial species identified from breast and breast milk according their first isolation source.** Adapted from Togo et al. (23).



## **1.2. Culture-dependent analysis of milk microbiota**

Culture-independent (DNA sequencing-based) approaches have recovered a high diversity of bacteria in breast milk including with a high degree of inter-individual variability (23, 54). Although, culture-independent approaches considerably expanded our understanding of the human microbiome (55), inherent limitations of the technology include inability to distinguish live or dead bacteria, low taxonomic resolution, and lack of access to the bacterial isolates for downstream phenotypic analysis and therapeutic assessment. As a result, there is a renewed interest in employing culture-based approaches to the study of human microbiota (56-58). Although the majority of the human microbiome was considered uncultivable (59-61), recent simple modifications in the culture conditions including diversification of culture media, incubation at various levels of Oxygen availability, co-culture with an appropriate host for obligate intracellular bacteria has resulted in successful isolation of increasing numbers of bacteria from different human microbiomes (62-66).

Cultured-enriched molecular profiling refers to the method of using multiple culture conditions followed by 16S rRNA gene sequencing (63). Using this approach on stool samples, 95% of the abundant taxa ( $> 0.1\%$  relative abundance in culture-independent direct sequencing) were isolated (63). Similar approach was successfully applied to low biomass sputum samples from cystic fibrosis patients (67). Although, culture-enriched molecular profiling could isolate a large proportion of human bacteria (63), culture- and sequencing-based approaches are complementary with some non-overlapping taxa identified by both methods (68, 69). This observation highlights the need to undertake a dual strategy in the study of human microbiome. Additionally, with low biomass samples such as milk, sequencing-based approaches are more prone to amplify the signal from potential reagent contaminants (70). In this case, culture-informed results could potentially guide the data pre-processing to identify and remove potential reagent contaminants.

Breastmilk has been historically studied using targeted culture approaches to identify potential pathogens (41, 71-75). However, culture-enriched molecular profiling has not been commonly employed to investigate the diversity of the milk microbiota (54).

### **1.3. Bacterial profile of different milk fractions**

The majority of milk fat and protein content are in the form of fat globules (>90% of the total fat) and casein micelles (80% of the total protein) (76, 77). To our knowledge, all sequencing-based human milk microbiota studies to-date have been conducted on the cell/casein pellets following removal of the milk fat as a major PCR inhibitor (78). However, it has previously been shown that different milk fractions have different profiles of other milk components including miRNAs (79) and bacteria (80-83). Milk bacteria potentially exist in both planktonic and cell-associated states. In mature milk, the majority of bacterial cells were found to be associated with human cells using flow cytometry (83). In a study from 1909, cream layer of the dairy milk separated by gravity or centrifugation contained higher culturable bacterial load compared to the sediments. However, the difference in the bacterial content of cream and sediment was less when the milk was centrifuged (14-fold vs. 5-fold difference for gravity vs. centrifugation cream separation) (81). Although, it is suggested that most of the milk bacteria aggregate with the casein micelles (84), specific bacteria such as *Mycobacterium avium* subsp. *paratuberculosis* and *Staphylococcus* show affinity to the fat globule (85-90). In addition, different milk fractions demonstrate different levels of inhibition on *Staphylococcus aureus*, the major pathogen in mastitis (91).

In the only study on the microbiome profile of different milk fractions in 54 dairy milk samples, whole milk, fat, and pellet were separately assessed. Overall, 31 bacterial families were shared between the fractions while there were 6, 11, and 24 unique families present in fat, pellet, and whole milk fractions, respectively (80). However, alpha and beta diversity were not different between different fractions (80). Additionally, the variability of the

microbiome of different milk fractions is not completely understood.

#### **1.4. Milk Mycobiota**

Fungi constitute an important yet frequently neglected component of the human microbiota (92, 93) with an underappreciated role in health and disease (94, 95). Fungal and bacterial colonisation occur in parallel during early life (96), yet most infant microbiome studies have overlooked fungi. Milk is a source of diverse and viable bacteria (97), but only a few studies have assessed fungi in human milk (38, 98-101). *Candida* which is a common human commensal fungi was isolated from 11/529 (2%) of milk samples from mothers experiencing painful breastfeeding (31) and 0.9% from milk of mothers with premature infants (33). While the majority of animal milk samples contained detectable levels of fungi (102), the prevalence of fungi in human milk has varied from 40% (by culture), to 35-80% (qPCR), and 70-100% (sequencing) (98, 99, 101). Thus, the prevalence and profile of human milk mycobiota is not completely elucidated.

#### **1.5. Origins of milk microbes**

It was a long-standing belief that milk bacteria were exogenously-derived contaminants. Culture-based studies measuring the bacterial load throughout the course of a feeding showed a gradual decline over the first 15ml of expressed milk (19), providing evidence that the milk closer to the outside environment contains more bacteria. More recently, studies have shown that bacteria consumed orally by lactating women or animals can later be isolated from their milk (72, 103, 104), indicating that bacteria may also reach the mammary gland through an ‘internal’ pathway. These two hypotheses regarding the origins of milk microbiota, initially suggested by Fernandez et al. (105), are commonly referred to as the “retrograde transfer” of external bacteria, and the “entero-mammary pathway” for translocation of internal bacteria (105).

The proposed entero-mammary pathway involves immune cell-mediated bacterial translocation from the mother's gastrointestinal tract into the mammary gland, where some of these bacteria are able to colonize the available niche (105). Bacterial translocation from the gastrointestinal tract frequently occurs under normal conditions (106, 107). Pathological (diabetes mellitus, obesity, inflammatory bowel disease; (108, 109)) or physiological (pregnancy, (110-112)) conditions resulting in compromised intestinal barrier integrity lead to enhanced bacterial translocation (113). Impaired intestinal permeability is then suggested to result in an increased rate of bacterial translocation especially in the late stages of pregnancy (113, 114) providing a mechanism for the widely held hypothesis that maternal gut microbiota is vertically transferred to the infant via breast milk. Supportive evidence for the entero-mammary pathway includes the presence of a microbial community in the breast tissue of non-lactating women (115), pre-colostrum before the infant birth (116), and colostrum collected even before the first infant feeding (29), vertical transmission of pathogenic bacteria and viruses through breastfeeding (117), and the presence of orally-administered probiotic strains in the milk of lactating mothers in both human and animal studies (72, 103, 104, 118). Presence of both aerobic and anaerobic bacteria to the mammary gland has been confirmed in mice. Although a baseline bacterial translocation was observed in non-pregnant mice, the prevalence of viable bacteria in the mammary gland was ~16-fold higher immediately after delivery (75). In accordance, oral administration of probiotic *Lactobacillus* to lactating women with *Staphylococcus* mastitis were detected in the breastmilk of the recipients resulting in alleviation of the symptoms (118, 119). Similar trends were observed when the probiotics were administered in the late stages of pregnancy to prevent infective mastitis (120). Additionally, orally administered live bacteria are commonly investigated for targeted anti-tumour drug delivery (121) confirming that bacterial translocation occurs from the gastrointestinal tract.

The proposed “retrograde transfer” of external bacteria into the mammary gland is another relatively new area of research. In this case, bacteria originating from the areola skin and/or the infant oral cavity are transferred to the intra-mammary environment via the backwash of milk during suckling (122). Additionally, bacteria could also originate from breast pumps. This mechanism is less appreciated because historically the majority of milk microbiota studies were concerned with intramammary cow’s milk with the primary goal of understanding predisposition to mastitis, which is the most costly disease of dairy animals (123). However, especially in human studies, the focus is now shifting to the offspring and the milk that s/he receives, whether that milk is pumped or provided directly from the breast. Reverse flow of milk from the infant mouth back into the breast occurs during breastfeeding (122), and perhaps pumping as well, allowing for infant mouth- or pump-associated bacteria to enter the milk duct. Given the substantial and increasing use of breast pumps worldwide (124), commercial availability of a range of infant feeding equipment (125), and evidence that the bacterial content of pump-expressed milk varies considerably depending on the setting, handling, and cleaning of the pump equipment (126), it is crucial to understand the impact of nursing directly at the breast vs. pump expression and bottle feeding on the milk microbiota composition.

Overall, it is currently not known which of the proposed pathways described above contribute more strongly to the milk microbiota composition. Although understanding the origins of milk microbiota requires sophisticated experimentation; statistical approaches within well-phenotyped population cohorts will facilitate hypothesis generation.

## **1.6. Determinants of milk microbiota**

There are limited studies on determinants of the bacterial load in the breastmilk. One study comparing bacterial load between samples collected in hospital or at home did not find any difference in the bacterial load based on the place of sample collection (24). The setting

of milk expression influenced the bacterial content of the milk depending on where the milk was expressed. For example, milk expressed at home had higher levels of potential pathogens including *Staphylococcus aureus* and environmental bacteria compared to samples collected in the hospital (24) suggesting that milk handling process and hygiene are plausibly important in milk bacterial composition and concentration. Depending on the source of bacteria, different factors could contribute to shaping the milk microbiota (Figure 3). The subset of milk bacteria originating from maternal gut microbiota are plausibly influenced by the factors shaping the mother's gut microbiota such as obesity (127) or diet (128, 129). On the other hand, factors influencing the infant's microbiota such as mode of delivery (130), older siblings, complementary feeding, and mode of breastfeeding (directly at the breast vs. expressed and bottled breast milk) could potentially alter the bacteria derived from the infant's oral cavity. Use of breastfeeding aids such as pump could also independently affect milk bacteria originating from the pump apparatus (131).

Other milk components such as HMOs, milk fatty acids, hormones, immune cells, and antibodies could also modulate the milk "microenvironment" and create a niche constraint affecting the composition of microbial community. Additionally, factors that could impact the overall milk composition such as mode of delivery (132), circadian rhythm (133), lactation stage, breastfeeding exclusivity, time since the last feed, and volume of the last feed could potentially indirectly affect the milk microbial community (30, 134). Previous studies have suggested that the composition might be affected by local pathologies of the breast (e.g. mastitis & breast cancer), mode of delivery, antibiotics, maternal weight, gestational age, and mode of delivery (16, 18, 31, 135-138). Additionally, geographic location was also associated with milk microbiota composition in multi-country studies (39, 46).

| Maternal microbiota                                                                                                                                                                                                                                               |                                                                                                                                                                                               | Exogenous microbiota                                                                                                                                                                                                                     |                                                                                                                                                                                                |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Mammary gland colonisation                                                                                                                                                                                                                                        |                                                                                                                                                                                               | Intra-mammary milk inoculation                                                                                                                                                                                                           | Expressed milk inoculation                                                                                                                                                                     |
| <b>Entero-mammary<br/>Oro-mammary</b><br><br><br><br><b>Maternal factors</b><br>Diet & Lifestyle<br>Medical history<br>Gut permeability<br>Body mass index<br>Periodontal health | <b>Breast microbiota</b><br><br><br><br><b>Local factors</b><br>Mastitis<br>Prior pregnancy<br>Breast cancer | <b>Retrograde</b><br><br><br><br><b>Infant mouth</b><br>Infant Sex, Birth mode<br><br><b>Maternal skin</b><br>Eczema, topical creams<br><br><b>Pump</b> | <b>Contamination</b><br><br><br><br><b>Milk handling</b><br>Pumping<br>Bottle-feeding<br>Cleaning practices |

**Figure 3. Origins of milk microbiota.** Two pathways for the origins of milk microbiota have been hypothesized: the entero-mammary pathway and retrograde translocation (105). We propose that translocation also occurs from the maternal oral cavity (oro-mammary pathway). Further, our study suggests that direct breastfeeding at the breast vs. pump expression could influence the milk microbiota composition (139) potentially indicating the pump apparatus as an additional source of retrograde inoculation and/or contamination of the expressed milk. Resident breast microbiota could be another source of mammary gland colonisation. Depending on the origin of the bacteria, different factors could be associated with the composition of the milk microbiota. Reproduced from Moossavi and Azad (131).

Storage temperature and duration are important factors to consider, as they can variably influence the integrity and composition of nutrients, bioactives and bacteria in pumped milk (140). Research in the dairy industry and human milk has shown that suboptimal storage could lead to shifts in the milk microbiota composition or viability, and dominance of spoilage bacteria (140). Additionally, the cleaning standards of the pump and bottle could act as an additional source of bacteria to the expressed milk. It has been shown that pumping milk could also affect the bacterial composition of milk resulting in higher frequency and relative abundance of *Enterobacteriaceae*, other Gram-negative bacteria, and *Candida* while other bacterial groups including *Staphylococcus* and *Streptococcus* were unaffected by the method of milk expression (31, 141) possibly due to different origins of milk bacteria. The effect of pumping could partly be attributed to the setting, handling and storage of the milk,

and cleaning of the pump equipment (126) or the type or other qualitative aspects of pumps. In a recent systematic review, there was no difference in milk contamination with known pathogens between manual pump compared to large electric pump (142). However, multiple different products exist within each of these broad categories, and new technologies are emerging as the popularity of pumping increases (e.g. hands-free suction-based manual pumps, and small wearable electric pumps). In addition, the qualitative aspects of pumps and pump accessories (such as breast shield size and materials used in tubing and bottles) could be important for the potential bacterial transfer and thus requires further investigation.

Other microbes present in milk could also influence the composition of milk bacteria. For example, milk samples containing HIV RNA in HIV positive mothers had higher bacterial diversity and prevalence of *Lactobacillus* spp. despite the bacterial load not being affected by the presence of HIV (30). In a recent study, correlations were assessed between fungal and bacterial relative abundances in human milk (99). However, the ecological influence of milk mycobiome on milk bacterial community has not been investigated.

While the majority of milk microbiome studies have been on bacterial communities and their determinants, fewer studies have examined milk mycobiome and the determinants of milk fungal prevalence and abundance. In earlier culture-dependent studies, pump expression was associated with higher concentration of yeast and *Candida* in the milk (31). In the same study, no association with infant oral Candidiasis was observed (31). Although geographical differences were observed in a recent study of 80 women from 4 countries (99), the maternal, infant, and environmental determinants of milk fungi (mycobiota) are still mostly unknown. One study of 76 mother-infant dyads found that maternal age, blood type, antibiotics, vaginal delivery, and infant sex were associated with *Candida* colonisation of the infant, and in a subset, maternal vaginal and rectal samples were identified as potential origins of this taxon (143). However, mothers' milk was not assessed and the home environment (another



potential source of milk fungi) was not examined. In addition, bacteria-fungi interactions are ubiquitous in microbial communities (144), but have rarely been investigated in milk.

### **1.7. Milk environment as a microbial niche**

Microbial community composition is regulated by the spatio-temporal dynamics of the microenvironment (niche). Relevant niche characteristics include nutrient resources or lack thereof, pH, salinity, predation, antibiotics, hydrodynamics, and physicochemical conditions (145). In classic microbial ecology, the impact of these factors are assessed on the community dynamic and composition; however, these variables are more challenging to measure and less commonly studied for human microbiota. Rather, the focus of human microbiota studies is often to identify host and/or environmental factors that influence the microbiota composition.

Milk serves as a niche for these microbes, and has likely evolved to select and enrich beneficial bacteria and likely fungi for delivery to the infant. Compared to the growing body of evidence identifying modifiable extrinsic factors shaping the milk microbiota (18, 139), the influence of the intrinsic milk environment and its major components is relatively unexplored.

HMOs constitute 20% of milk carbohydrates (146). Most studies to-date have focused on HMOs and the infant gut microbiota, especially *Bifidobacterium* and *Bacteroides* (147-150). However, it is less clear whether milk microbiota in general or milk *Bifidobacterium* specifically could also utilize HMOs. A few observational studies have detected positive correlations between total HMO concentration and milk bacteria including *Bifidobacterium* spp. and *Staphylococcus* (151-153). On the other hand, some individual HMOs including lacto-N-fucopentaose (LNFP) I, 2'-fucosyllactose (2'FL), 3'-sialyllactose (3'SL), 6'-sialyllactose (6'SL), and lacto-N-tetraose (LNT) were negatively correlated with Actinobacteria (including *Bifidobacteria*), *Staphylococcus* and other taxa in milk. Some

HMOs have also been positively correlated with potential pathogens such as *Enterobacter/Klebsiella* (153-155). In a small metagenomic study of human milk, HMO metabolism genes were not detected, which was attributed to the absence of *Bifidobacterium* spp. (44). However, other reports observed differences in the prevalence of *Bifidobacterium* according to the maternal FUT2 secretor status (presence or absence of fucosyltransferase 2) (156), which is a major determinant of HMO synthesis. Notably, none of these studies accounted for the mode of breastfeeding (directly at the breast or pumped and bottled). Overall, these divergent results from relatively small studies highlight the need to examine HMOs and milk microbiota in larger studies while accounting for secretor status, mode of breastfeeding and other potential confounding factors.

Milk lipids are a source of energy for the infant and primarily consist of triglycerides (146), which are fatty acids bound to glycerol. While the relative amounts of some milk fatty acids are fairly consistent across studies, geographies and individual women (157), others such as docosahexaenoic acid (DHA) are strongly influenced by maternal diet (158). *In vitro* studies have shown that polyunsaturated fatty acids (PUFAs) can have inhibitory effects on bacteria, including *Lactobacillus*, in a dose-dependent manner (159) while 12:0, cis-9/trans-11 and trans-10/cis-12 conjugated linoleic acid (CLA) can inhibit growth of *Staphylococcus aureus in vitro* (160). PUFAs and saturated fatty acids (SFAs) have also been shown to influence the microbiota composition and diversity in the gut (161-163), but their impact on milk microbiota has not been widely studied. Some correlations have been observed between milk fatty acids and milk microbiota (154). Specifically, negative correlations were reported for milk monounsaturated fatty acids (MUFAs) and PUFAs with Proteobacteria, *Streptococcus*, *Acinetobacter*, *Lactobacillus*, and *Bifidobacterium* while SFAs were positively correlated with *Pseudomonas* (39). However, these correlations have not been examined while controlling for maternal fish oil supplement use, which modifies milk n3-

PUFA concentrations (164) and might also influence microbiota (165). The potential associations of these major milk components are milk microbiota while controlling for relevant confounding factors are not clear.

### **1.8. Influence of the choice of 16S rRNA gene sequencing processing algorithm on milk microbiota analysis**

Amplicon-sequencing studies targeting bacterial 16S rRNA gene are currently the most cost-effective sequencing method for microbiome studies and are widely applied to different ecosystems. While the decreasing cost of sequencing has resulted in expansion of sequencing-based microbiome studies, computational methods for fast and accurate processing of the sequencing data and taxonomy assignment are becoming increasingly more crucial to ensure an accurate reconstruction of the microbial community (166). In recent years there have been novel developments in several aspects of sequencing processing including a shift from clustering-based methods of operational taxonomic unit (OTU) designation based on sequence similarity (commonly  $> 97\%$ ) (167, 168) to denoising algorithms which identify exact amplicon sequence variants (ASV) (169-175); thereby increasing the ecological precision. Performance of the denoising methods has been assessed mostly on mock communities and some real samples in terms of taxonomy composition (173-175). Performance of the denoising methods has been assessed mostly on mock communities and some real samples in terms of taxonomy composition (173-178). However, the impact of these different methods on downstream analysis of biological associations is less clear.

There is an emphasis on reproducibility in microbiome research (179, 180). Several aspects of microbiome study including sample collection and storage, DNA extraction method, choice of primers and sequencing platform could influence the reproducibility of the results (181-183). However, challenges of reproducibility post-sequencing are the results of the choice of bioinformatics pipeline and statistical methods. Prior studies have examined

the effect of bioinformatics pipelines on the microbiome composition of mock communities (176-178). The true biological and ecological variations appear to be robust to the choice of sequencing processing method in a few studies on soil, mouse feces, and intestinal biopsy (178, 184, 185). As the majority of studies increasingly use denoising methods, a major question regarding the validity of previous studies using OTU-based methods remains unanswered.

These challenges are especially of concern for low biomass samples where the signal-to-noise ratio is very low (53, 186). Therefore, reproducible downstream analyses will depend on the precision of sequencing processing coupled with potential contaminant identification and removal. In a recent study, performances of different sequencing processing pipelines were compared on various dilutions of a mock community (176). Expectedly, the false positive detection rate increased with decreasing concentration which were mainly due to increased detection of contaminants by both clustering-based and denoising methods (176). In another study of a mock community consisting of nine species frequently isolated from dairy milk, both clustering-based and denoising methods overestimated the number of OTUs/ASVs (187). The performance of different methods on real samples (high biomass: soil and mouse feces) and (low biomass: intestinal biopsy from Crohn's Disease) was assessed. While beta diversity assessment based on abundances (Bray-Curtis or weighted UniFrac) were mostly influenced by the biological sample type, measures based on presence/absence (unweighted UniFrac) were sensitive to the choice of the pipeline (178). In a separate study on soil microbiota along elevation gradient, the choice of the bioinformatics algorithm (OTU vs. ASV) did not alter ecological conclusions (185). As low biomass samples are more sensitive to contaminations, the influence of using different sequencing processing on the biological observations is unclear. While the performance of

sequence processing methods on low biomass mock community has been assessed, little is known about their performance on real world low biomass sample.

Milk microbiome studies are gaining considerable interest as it is suggested to be a continual source of bacteria to the infant gut (188, 189). As milk is a low biomass sample, it is plausible that the milk microbiota profile generated by sequencing contains reagent contaminants and thus identification and removal of potential reagent contaminants is recommended. Recently, probabilistic approaches have been developed to statistically identify potential contaminants based on the prevalence in the negative controls compared to samples and/or correlation of the abundance of them with DNA concentrations (190). Currently it is not known if or how the choice of algorithm for sequencing processing and/or contaminant removal would impact the downstream analysis of biological observations in low biomass samples (e.g. milk).

## **1.9. Gaps in knowledge**

Major gaps in our knowledge regarding the milk microbiota composition and determinants exist as the majority of the previous studies have been i) culture-independent targeting exclusively bacteria, ii) limited by small sample size, low sequencing depth, and lack of control for confounding factors, and iii) unable to account for other major milk components. Current gaps in knowledge include:

- 1) What are the profiles of milk microbiota and mycobiota and their inter-kingdom relation in a large population?
- 2) What is the culture-dependent profile of milk microbiota using culture-enriched molecular profiling?
- 3) What are the determinants of milk microbiota and mycobiota composition while taking into account their potential origins and controlling for confounding factors?

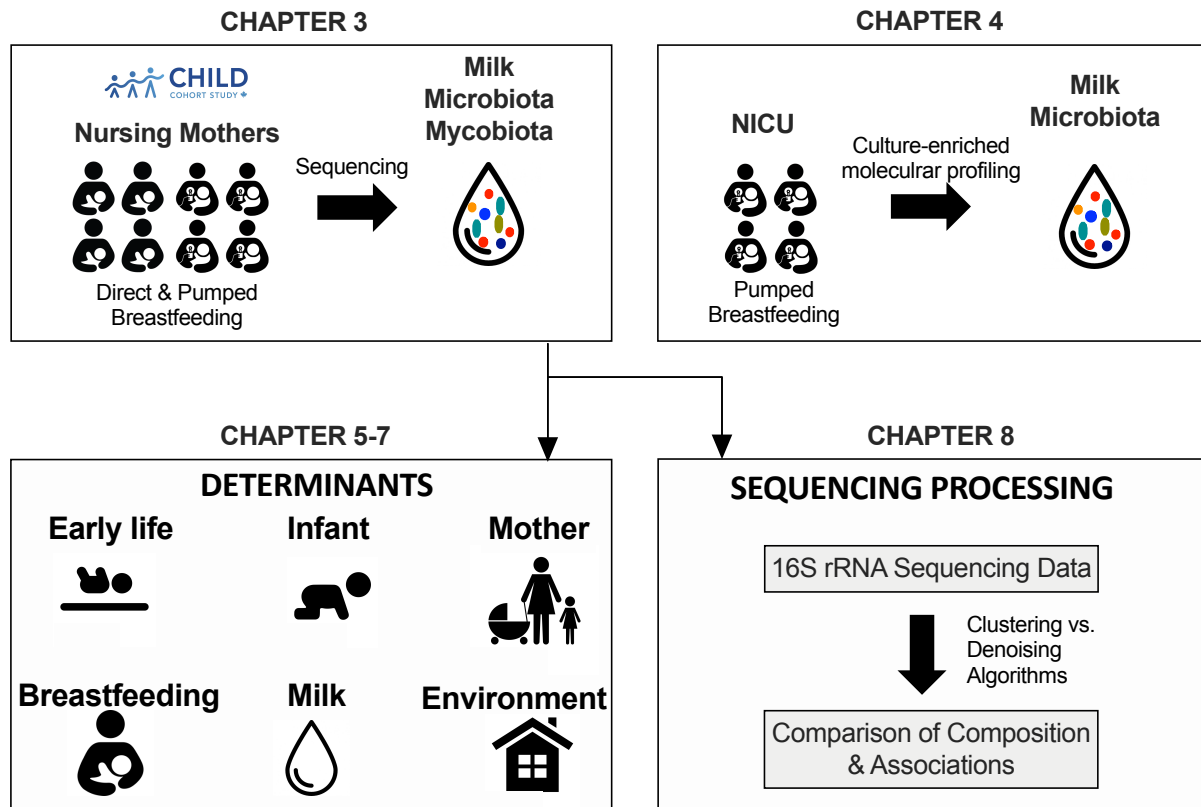
- 4) Do other major milk components influence the milk microbiota composition via shaping the milk microenvironment?
- 5) What is the impact of upstream sequencing processing on downstream statistical analyses and conclusions in low biomass samples?

#### **1.10. Hypotheses and objectives**

The main hypotheses of this thesis are:

- 1) Milk microbiota and mycobiota compositions are influenced by maternal, infant, early life, and environmental factors based on the potential routes of milk colonisation/inoculation.
  - 1a. Milk bacteria are most strongly influenced by factors affecting the infant microbiome.
  - 1b. Milk fungi are most strongly influenced by environmental factors.
- 2) Culture-enriched molecular profiling can capture the diversity of milk microbiota.
- 3) Biologically plausible statistical conclusions are robust to the choice of sequencing processing method.

The objectives of this study are to profile milk microbiota using culture-independent (bacteria and fungi) and culture-enriched molecular profiling (bacteria) (Chapters 3 and 4); assess the inter-kingdom association of milk bacteria and fungi (Chapter 3); identify factors associated with the composition of milk bacteria (Chapters 5 and 6) and fungi (Chapter 7); and examine the impact of different sequencing processing approaches on downstream statistical analysis (Chapter 8). An overview of this thesis is illustrated in Figure 4.



**Figure 4. Schematic overview of this thesis.** Milk microbiota will be studied in the CHILD Cohort Study and in samples collected from NICU at McMaster University. Milk microbiota (bacteria and fungi) will be assessed in the CHILD Cohort Study using amplicon sequencing (Chapter 3). Milk bacteria will be explored in a pilot study using culture-enriched molecular profiling (Chapter 4). The determinants of the milk microbiota will be assessed in the CHILD Cohort Study (Chapters 5-7). The impact of the choice of data processing on downstream analysis will be investigated in the CHILD Cohort Study (Chapter 8).

## 2. MATERIALS AND METHODS

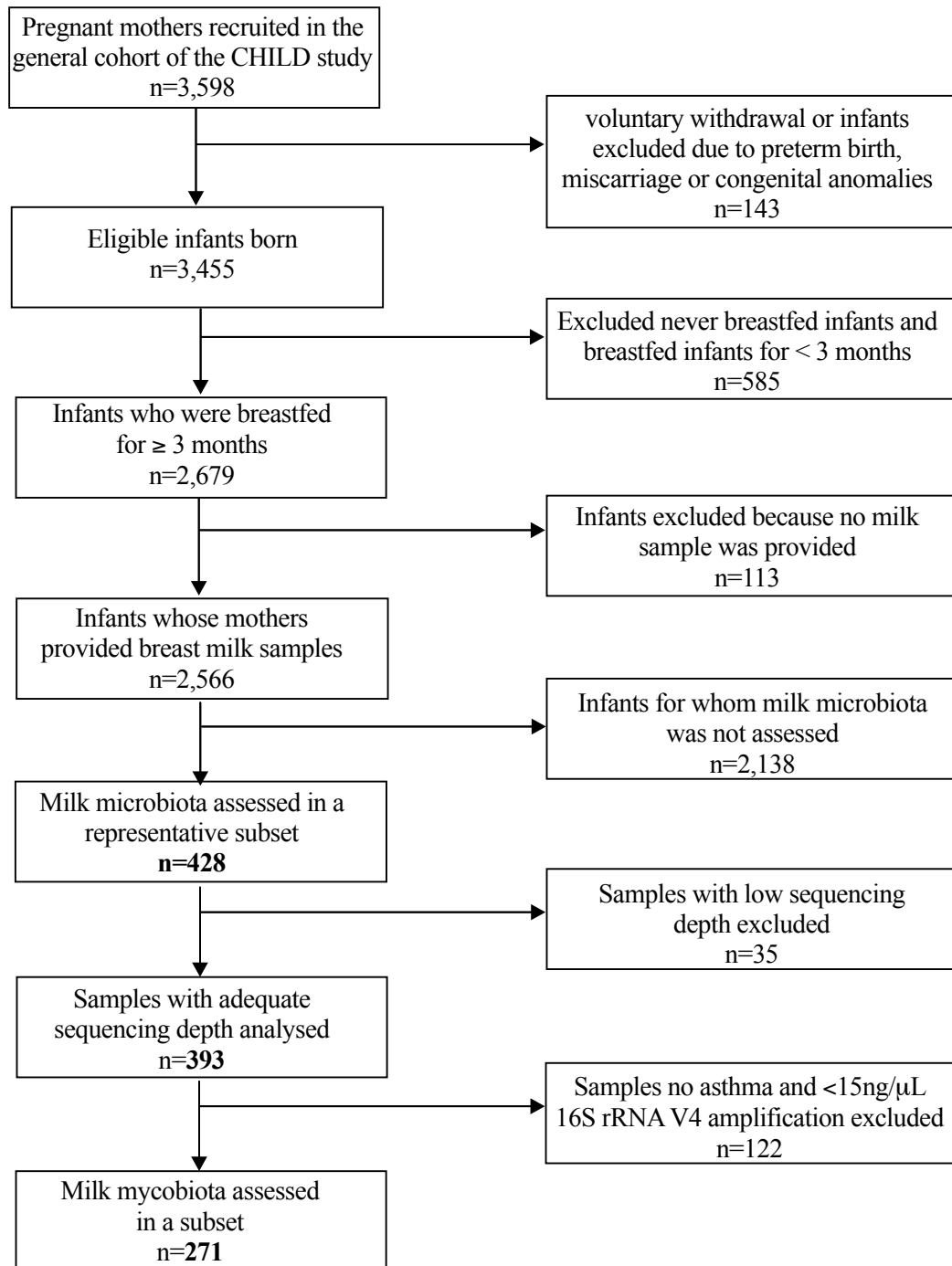
### 2.1. Culture-independent study

Milk microbiota and mycobiota was assessed in milk samples from the CHILD Cohort Study as described below.

#### 2.1.1. Study Design

We studied a representative subset of 428 mothers in the Canadian Healthy Infant Longitudinal Development (CHILD) Cohort Study ([www.childstudy.ca](http://www.childstudy.ca)) designed to study gene-environment interactions in the developmental origins of pediatric asthma and allergy and other non-communicable diseases (191). Pregnant women in the second or third trimester were recruited from the general population from four sites across Canada (Vancouver, Edmonton, Manitoba (Winnipeg, Morden, Winkler) and Toronto) between 2008 and 2012 and remained eligible if they delivered a healthy singleton infant >35 weeks gestation (n=3455). We selected a representative subset of 428 mother-infant dyads with available milk samples, ensuring equal representation across the 4 study sites, excluding dyads missing key maternal (*e.g.* diet, BMI) or infant (1y clinical) data, and then randomly selecting among the rest. Next a further subset of 271 mother-infant dyads was selected from the 428 mother-infant dyads for fungal analysis. This subset was enriched for breastmilk samples with available 16S rRNA bacterial sequencing data (139) and children diagnosed with possible or probable asthma. Samples with high concentration of DNA based on bacterial 16S rRNA V4 PCR amplification ( $\geq 15$  ng/ $\mu$ L) were prioritised (Figure 5). This study was approved by the Human Research Ethics Boards at McMaster University, the Hospital for Sick Children, and the Universities of Manitoba, Alberta, and British Columbia. Characteristics of the participants are summarised in Table 1.





**Figure 5. Flowchart of the participant selection for microbiota and mycobiota analysis in CHILD Cohort Study.**

### ***2.1.2. Data accessed from the CHILd database***

Comprehensive environmental, nutrition and health questionnaires were administered during pregnancy and at 3, 6, 12, 18, 24, 30 months, 3, 4 and 5 years. A Food Frequency Questionnaire (FFQ) was administered during pregnancy(191). We obtained permission to access the CHILd data as required for this study. Additional data including infant secretor status and environmental features were generated by collaborating members within the CHILd Cohort Study. Other milk components including fatty acids, human milk oligosaccharides, and milk hormones were previously measured in the same subset of milk samples (192-194) (section 2.2). List and definitions of variables used in this study are summarised in Table 2.

#### ***2.1.2.1. Maternal factors***

Maternal age, tobacco smoking, method of birth, parity, and intrapartum antibiotic use were documented from hospital records. Mode of delivery was categorised as normal vaginal delivery (NVD), emergency Caesarean section (C/S), or elective C/S. Maternal ethnicity, pre-pregnancy weight and height, maternal history of asthma diagnosis, tobacco smoking during pregnancy or postpartum, and intravenous and/or oral antibiotic use prior to milk sample collection were reported by standardized questionnaire. Maternal dietary patterns were determined by principal components analysis (PCA) from a modified food frequency questionnaire (195). Maternal secretor status was defined by the presence of 2'-fucosyllactose (2'-FL) or lacto-N-fucopentaose I (LNFP I) in the milk HMO profile (196).

#### ***2.1.2.2. Infant factors***

Infant sex, birth weight, and gestational age were documented from hospital records. Infant oral thrush in the first 3 months of life and intravenous and/or oral antibiotic use were reported by standardized questionnaire.

**Table 1. Characteristics of mother-infant dyads from the CHILD Cohort Study included in the microbiota analysis (n=393) in comparison with all eligible dyads (n=2536). Reproduced from Moossavi et al. (139).**

| Factor        | Characteristics                                      | Mean $\pm$ SD or n (%) <sup>a</sup> |                                           |
|---------------|------------------------------------------------------|-------------------------------------|-------------------------------------------|
|               |                                                      | Subset for this study<br>N=393      | All eligible dyads<br>N=2536 <sup>b</sup> |
| Maternal      | Age (years)                                          | 33.0 $\pm$ 4.2                      | 32.7 $\pm$ 4.2                            |
|               | Pre-pregnancy BMI (Kg/m <sup>2</sup> )               | 24.3 $\pm$ 5.2                      | 24.5 $\pm$ 6.3                            |
|               | History of atopy                                     | 251 (63.9)                          | 1624 (65.4)                               |
|               | Secretor status                                      | 279 (71.7)                          | 901 (74.7) <sup>c</sup>                   |
|               | Ethnicity                                            |                                     |                                           |
|               | Caucasian                                            | 287 (73.0)                          | 1870 (74.3)                               |
|               | Asian                                                | 73 (18.6)                           | 403 (16.0)                                |
|               | First Nations                                        | 15 (3.8)                            | 87 (3.5)                                  |
|               | Other                                                | 18 (4.6)                            | 158 (6.3)                                 |
| Infant        | Birth weight (g)                                     | 3469 $\pm$ 469                      | 3450 $\pm$ 479                            |
|               | Female sex                                           | 192 (48.9)                          | 1198 (47.2)                               |
|               | Gestational age (weeks)                              | 39.2 $\pm$ 1.3                      | 39.2 $\pm$ 1.2                            |
| Early life    | Mode of delivery                                     |                                     |                                           |
|               | Elective C/S                                         | 46 (11.9)                           | 264 (10.6)                                |
|               | Emergency C/S                                        | 47 (12.1)                           | 339 (13.6)                                |
|               | Vaginal                                              | 294 (76.0)                          | 1893 (75.8)                               |
|               | Maternal intrapartum antibiotics                     | 139 (35.8)                          | 996 (39.9)                                |
|               | Maternal postpartum antibiotics before 3-4 months    | 41 (10.6)                           | 265 (10.8)                                |
|               | Child antibiotics before 3-4 months                  | 11 (2.8)                            | 75 (3.1)                                  |
|               | Older siblings                                       |                                     |                                           |
|               | None                                                 | 211 (53.7)                          | 1345 (53.1)                               |
| Breastfeeding | One                                                  | 126 (32.1)                          | 860 (33.9)                                |
|               | Two or more                                          | 56 (14.2)                           | 330 (13.0)                                |
|               | Lactation stage at sample collection (weeks)         | 17.3 $\pm$ 5.3                      | 16.5 $\pm$ 5.0                            |
|               | Exclusive BF (breast milk only) at sample collection | 190 (48.3)                          | 1393 (55.9)                               |
|               | Direct BF (at the breast) only                       | 162 (41.9)                          | 1018 (40.1)                               |
| Milk          | Duration of BF (months)                              | 13.2 $\pm$ 5.8                      | 12.5 $\pm$ 5.6                            |
|               | Duration of exclusive BF (months)                    | 3.5 $\pm$ 2.3                       | 3.7 $\pm$ 2.3                             |
| Milk          | HMO concentration (mg/mL)                            | 10.2 $\pm$ 2.1                      | 10.3 $\pm$ 2.1 <sup>c</sup>               |
|               | HMO Simpson's diversity                              | 4.9 $\pm$ 1.4                       | 4.9 $\pm$ 1.3 <sup>c</sup>                |
|               | Insulin (pg/mL)                                      | 760 $\pm$ 664                       | Unknown                                   |
|               | Leptin (pg/mL)                                       | 556 $\pm$ 601                       | Unknown                                   |

<sup>a</sup> Percentages are calculated after excluding dyads with missing data. BF, breastfeeding; BMI, body mass index; C/S, Caesarean section; HMO, human milk oligosaccharide; NVD, normal vaginal delivery

<sup>b</sup> Those who breastfed  $\geq$ 12 weeks and provided a milk sample

<sup>c</sup> available for N=1206

Infant secretor status was determined from the infant genetics (Qingling Duan, Queen's University). Briefly, infant DNA was isolated from peripheral blood samples. Genotyping of single nucleotide polymorphisms (SNPs) was performed using the HumanCoreExome BeadChip (Illumina, San Diego, CA, USA)) comprising >240K tagSNPs and >240K exome-focused markers. A stringent process of quality control at infant and SNP levels was applied using PLINK v.1.9 (<http://pngu.mgh.harvard.edu/purcell/plink/>) (197). Samples were excluded if demonstrated low genotype call rates (missingness > 0.10), discrepancies with reported sex, were ethnic outliers based on PCA, or were related as determined by identity by descent. SNPs were excluded with missingness > 0.10, had significant departures from Hardy-Weinberg equilibrium ( $p$  value <  $10^{-6}$ ), or had a minor allele frequency < 0.01. IMPUTE2 software (198) was used to impute genotypes based on the 1000 Genomes Project reference panel (199), which resulted in >23,000,000 single nucleotide variants of high quality (> 0.7). Using this imputed dataset, infant secretor status was defined based on FUT2 rs601338 or rs1047781 SNPs (200, 201).

#### 2.1.2.3. *Breastfeeding practices*

Infant feeding was reported by standardized questionnaire at 3, 6, 12, 18 and 24 months. At the time of sample collection (3-4 months), breastfeeding status was classified as exclusive (human milk only) or partial (human milk supplemented with infant formula or solid food). Duration of any breastfeeding and exclusive breastfeeding was also evaluated (age in months at cessation of breastfeeding and introduction of formula or solid food, respectively). The mode of breast milk feeding was reported for infants at three months of age and classified as "all direct" (human milk only with no feeding of expressed milk), or "some pumped" (human milk only with at least one serving of expressed milk in the past two weeks) (10).

#### 2.1.2.4. *Home environment*

Participants were recruited from four Canadian cities: Vancouver (49.2827° N, 123.1207° W, British Columbia), Edmonton (53.5444° N, 113.4909° W, Alberta), Manitoba (49.8951° N, 97.1384° W, Winnipeg and two rural towns), and Toronto (43.6532° N, 79.3832° W, Ontario). Average climate information from 1981-2010 was obtained from Environment and Climate Change Canada (<http://climate.weather.gc.ca>; access date 13 March 2019; used in descriptive but not statistical analyses). Population density was estimated for CHILd participants' homes using 2006 census data on population counts within the smallest available census geographical boundary (250m) and was categorized as high or low (above or below the median). Census data also provided the rural vs. urban residence. "An urban area was defined as having a population of at least 1,000 and a density of 400 or more people per square kilometre. All territory outside an urban area was defined as rural area" (202)(202)(202)(202). Green space exposure was quantified using the derived normalized difference vegetation index (NDVI) in a 250 m buffer around the mother's residential addresses. NDVI, with values ranging from -1 to 1 is an indicator of overall greenness based on land surface reflectance of visible and near infrared parts of spectrum. Time weighted averages across the 12 months postpartum were assigned and categorized into grey (-1.0 to 0.2), moderate green (0.2 to 0.3), and green (0.3 to 1.0) (Unpublished, generated and provided by Hind Sbihi). Home environment characteristics including dust, moisture, and visible mould levels were determined based on questionnaires completed by mothers and a walk-through home assessment by trained research staff from the CHILd Cohort Study (203); each exposure was categorized into high or low (above or below median).

**Table 2. List and definition of variables used in this study**

| Variable                                           | Definition                                                                                                                                                                                                            |
|----------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Age (years)                                        | Maternal self-reported age from hospital record                                                                                                                                                                       |
| Pre-pregnancy BMI (Kg/m <sup>2</sup> )             | Calculated from maternal self-reported pre-pregnancy weight and height                                                                                                                                                |
| History of atopy (yes vs. no)                      | Maternal self-reported                                                                                                                                                                                                |
| Secretor status (secretor vs. non-secretor)        | Presence or near absence of 2'FL or LNFP I (196)                                                                                                                                                                      |
| Ethnicity (Caucasian, Asian, First Nations, Other) | Maternal self-reported                                                                                                                                                                                                |
| Prenatal smoking (yes vs. no)                      | Any tobacco smoking prior to pregnancy                                                                                                                                                                                |
| Birth weight (g)                                   | Infant weight at birth measured by health professionals                                                                                                                                                               |
| Sex (female vs. male)                              | Infant biological sex at birth                                                                                                                                                                                        |
| Gestational age (weeks)                            | Obtained from hospital records                                                                                                                                                                                        |
| Infant oral thrush (yes vs. no)                    | History of oral thrush postpartum until sample collection reported by parents                                                                                                                                         |
| Infant secretor status (secretor vs. non-secretor) | Determined based on FUT2 rs601338 or rs1047781 SNPs genotype status                                                                                                                                                   |
| Mode of delivery                                   | Vaginal vs. C/S (elective or emergency)                                                                                                                                                                               |
| Maternal intrapartum Abx (yes vs. no)              | Use of intrapartum antibiotic prophylaxis obtained from hospital record                                                                                                                                               |
| Maternal postpartum Abx (yes vs. no)               | Self-reported use of oral or intravenous antibiotics after birth until sample collection                                                                                                                              |
| Child Abx (yes vs. no)                             | Use of oral or intravenous antibiotics after birth until sample collection reported by the mother                                                                                                                     |
| Parity (yes vs. no; 0, 1, ≥2)                      | Number of previous pregnancies (older siblings)                                                                                                                                                                       |
| Lactation stage (weeks)                            | Age in weeks at the time of sample collection                                                                                                                                                                         |
| BF exclusivity (yes vs. no)                        | Human milk only or partial (human milk supplemented with infant formula or solid food)                                                                                                                                |
| Mode of BF (all direct vs. some pumped)            | Self-reported method of breastfeeding either by nursing directly at the breast or pumping and bottle-feeding in the last 2 weeks prior to milk sample collection (10)                                                 |
| Duration of BF (months)                            | Age in months at cessation of breastfeeding                                                                                                                                                                           |
| Duration of exclusive BF (months)                  | Age in months at introduction of formula or solid food                                                                                                                                                                |
| HMO concentration (mg/mL)                          | Total sum of all concentrations of all measured HMOs (192)                                                                                                                                                            |
| HMO Simpson's diversity                            | Diversity of the HMOs (192)                                                                                                                                                                                           |
| Insulin (pg/mL)                                    | Concentration of insulin in milk (193)                                                                                                                                                                                |
| Leptin (pg/mL)                                     | Concentration of leptin in milk (193)                                                                                                                                                                                 |
| Maternal dietary pattern                           | Based on principal components analysis from a modified food frequency questionnaire (195)                                                                                                                             |
| Population density*                                | Population counts within the smallest available census geographical boundary (250m) estimated for CHILDP participants' homes using 2006 census data on.                                                               |
| Area type*                                         | Urban vs. rural defined based on 2006 census data                                                                                                                                                                     |
| Greenness*                                         | Average NDVI in a 250 m buffer around the mother's residential addresses across the 12 months postpartum; categorized into grey (-1.0 to 0.2), moderate green (0.2 to 0.3), and green (0.3 to 1.0)                    |
| Home dust, moisture, and visible mould levels      | Based on questionnaires completed by mothers and a walk-through home assessment by trained research staff from the CHILDP Cohort Study (203). Each exposure was categorized into high or low (above or below median). |

\* Generated and provided by Hind Sbihi.

Abx, antibiotics; BF, breastfeeding; HMO, human milk oligosaccharide; NDVI, Normalized difference vegetation index

### **2.1.3. Milk sample collection**

Each mother provided one sample of milk at 4 months postpartum [mean (SD) 17 (5) weeks postpartum] in a sterile milk container provided by CHILD Cohort Study. To control for differences in the milk composition of fore- and hindmilk (204) as well as the diurnal variation (133); a mix of foremilk and hindmilk from multiple feeds during a 24-hour period was collected. Hand expression was recommended, but pumping was also acceptable. The sample was not collected aseptically. Samples were refrigerated at home for up to 24 hours before being collected and processed by study staff (205). Milk samples were collected from 2008 to 2012 and were stored at -80 °C until analysis in 2016. The time of first and last milk collection at home and the time between sample collection from home and processing by the CHILD laboratory personnel were recorded. We did not observe any associations of sample processing time with milk microbiota richness, diversity, taxa relative abundance, overall composition, and cluster membership.

### **2.1.4. DNA extraction and microbiota analysis**

Genomic DNA was extracted from 1 ml breastmilk using Quick-DNA Fungal/Bacterial extraction kit following the manufacturer's instructions (Zymo Research, USA). The samples were centrifuged (13,000 g at 4°C for 20 min), the fat rim was carefully removed using a sterile swab, and the supernatant stored for future analysis. Total DNA was extracted from the pellet.

V4 hypervariable region of the 16S rRNA gene (bacteria) and internal transcribed spacer (ITS) 2 region of the eukaryotic ribosomal genes (fungi) were amplified with modified F515/R806 primers (206) and modified ITS3/ITS4 primers (102), respectively. For each sample, the PCR reaction was performed in duplicate, consisting of an initial denaturing step at 95 °C for 5 min followed by 33 amplification cycles at 95 °C for 40 s, 51.5 °C for 40 s, and 72 °C for 50 s, with a final extension step at 72 °C for 5 min (bacteria) and an initial

denaturing step at 95 °C for 5 min followed by 32 amplification cycles at 95 °C for 40 s, 52 °C for 45 s, and 72 °C for 40 s, with a final extension step at 72 °C for 5 min (fungi) in an Eppendorf Mastercycler pro (Eppendorf, Hamburg, Germany). Sterile DNA-free water was used as negative controls in sequencing library preparation. Successful PCR amplification was assessed by agarose gel electrophoresis of a subset of samples and negative controls (bacteria) and for all samples (fungi).

Duplicate PCR products were pooled and the concentration of double-stranded DNA was quantified by Quanti-iT PicoGreen dsDNA Assay kit (Invitrogen, CA, USA). The machine was first calibrated on the serial dilution of the standard DNA provided by the kit. The concentrations of milk double-stranded DNA were used to normalise the samples 250 ng/μL of DNA and samples were pooled accordingly. For samples with very low DNA concentration, maximum 15 μL (bacteria) and 10 μL (fungi) of sample were pooled. Based on previous dairy mycobiome work in the lab, we anticipated that a subset of samples to be negative for fungi (102). For fungi, we pooled samples into three pools based on DNA concentration and PCR amplification (high, intermediate, and low DNA concentration). Pooled DNA was then cleaned using DNA Clean & Concentrator (Zymo Research, USA) and the concentration was measured using Qubit dsDNA HS Assay Kit (Invitrogen, CA, USA). To minimise the potential impact of well-to-well contamination for low DNA samples (207), the final pooled DNA consisted of high, intermediate, and low DNA with proportions of 80:10:10.

Following quantification, library was diluted in the DNA Clean & Concentrator elution buffer to the concentration of 2 nM. Subsequently 10 μL of the sample was denatured by adding 10 μL NaOH (0.2M) and diluted to the concentration of 20 pM. The library was incubated at room temperature for 5 minutes. Next, the sample was further diluted in Illumina HBY buffer to the concentration of 14 pM. Final sequencing library was prepared by mixing



510 µL of sample and 90 µL of PhiX (20 pM). Primers (forward, reverse, and index) and the library were loaded on the sequencing cartridge and were sequenced on a MiSeq platform using v3 reagent kit (Illumina, San Diego, CA, USA). Sterile DNA-free water was used as negative controls in sequencing library preparation. A mock community consisting of DNA extracted from 8 bacteria and 2 fungi species with known theoretical relative abundances was used as positive control (Zymo Research, USA).

#### ***2.1.5. 16S rRNA and ITS gene sequencing Analysis***

Overlapping paired-end reads were processed with DADA2 pipeline (173) using the open-source software QIIME 2 v.2018.6 (<https://qiime2.org>) (208). Unique amplicon sequence variants (ASVs) were assigned a taxonomy and aligned to the 2013 release of the Greengenes reference database at 99% sequence similarity (209). Fungal taxonomy assignment were achieved by aligning the denoised sequences to the 2017 release of the UNITE v.7 reference database at 99% sequence similarity (210). Demultiplexed sequencing data was deposited into the Sequence Read Archive (SRA) of NCBI and can be accessed via accession number PRJNA481046 (bacteria) and PRJNA536254 (fungi).

#### ***2.1.6. Contamination removal from the sequencing output***

Data analysis was conducted in R (211). Initial preprocessing of the ASV table was conducted using the Phyloseq package (212). Potential reagent contaminants (52) were identified using decontam package based on either the frequency of the ASV in the negative control or the negative correlation with DNA concentration (190). Overall, 9,711 unique bacterial ASVs were detected for bacterial microbiome and 173 were identified as contaminants and excluded. We did not identify any potential fungal contaminant ASVs in the fungal analysis.

#### ***2.1.7. Sequencing output processing following contaminant removal***

Samples were rarefied to 25000 and 1000 for bacterial and fungal datasets, respectively.

Subsequently, samples with less than the rarefaction threshold were excluded. Additionally, ASVs with zero abundance across the samples; non-bacterial (belonging to phylum Cyanobacteria, family of mitochondria, and class of chloroplast) or non-fungal ASVs (belonging to plants) were removed. Additional threshold filtering steps were applied as appropriate (Chapter 3).

## **2.2. Analysis of other milk components**

HMOs, milk fatty acids, and milk metabolic hormones (insulin and leptin) were measured in the same milk samples.

### **2.2.1. HMO Analysis**

HMOs were previously measured in the same milk samples by high performance liquid chromatography (192). The following 19 HMOs were detected on the basis of retention time comparison with commercial standard oligosaccharides and mass spectrometry analysis: 2'FL, 3-fucosyllactose (3FL), 3'SL, 6'SL, LNT, lacto-N-neotetraose (LNnT), LNFPI, LNFPII, LNFPIII, sialyl-LNT (LST) b, LSTc, difucosyl-LNT (DFLNT), disialyl-LNT (DSLNT), fucosyl-lacto-N-hexaose (FLNH), difucosyl-lacto-N-hexaose (DFLNH), fucosyl-disialyl-lacto-N-hexaose (FDSLNH), and disialyl-lacto-N-hexaose (DSLNH), difucosyllactose (DFLac), and lacto-N-hexaose (LNH) and were quantified based on response factors. These 19 HMOs typically account for >90% of total HMO content; their concentrations were summed to estimate total HMO concentration. The relative abundance of each HMO was calculated. The amount of HMO-bound fucose and HMO-bound sialic acid were calculated. HMOs were also grouped according to common structural elements as previously described (213) (Table 3). Maternal secretor status was defined by the presence or near absence of 2'FL or LNFP I. HMOs were studied as both absolute concentrations ( $\mu\text{g/mL}$ ) and relative proportions.

**Table 3. Human milk oligosaccharides (HMOs) and fatty acids structural categories.**  
Reproduced from Moossavi et al. (214).

| Category                                                    | Components                                                                                                                                                                   |
|-------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>HMOs</b>                                                 |                                                                                                                                                                              |
| Small HMOs                                                  | 2'FL, 3FL, 3'SL, 6'SL                                                                                                                                                        |
| Type 1 (Terminal galactose attached in $\beta$ 1-3 linkage) | LNT, LNFP I, LNFP II, LSTb, DSLNT                                                                                                                                            |
| Type 2 (Terminal galactose attached in $\beta$ 1-4 linkage) | LNnT, LNFP III, LSTc                                                                                                                                                         |
| $\alpha$ -1,2                                               | LNFP I, 2'FL                                                                                                                                                                 |
| $\alpha$ -1,3                                               | LNFP III, 3FL                                                                                                                                                                |
| $\alpha$ -2,6                                               | LSTb, LSTc, 6'SL                                                                                                                                                             |
| Fucosylated/Sialyated                                       | FDSLNH                                                                                                                                                                       |
| Fucosylated/Non-Sialyated                                   | 2'FL, 3FL, DFLac, LNFP I, LNFP II, LNFP III, DFLNT, FLNH, DFLNH                                                                                                              |
| Non-Fucosylated/Sialyated                                   | 3'SL, 6'SL, LSTb, LSTc, DSLNT, DSLNH                                                                                                                                         |
| Non-Fucosylated/Non-Sialyated                               | LNnT, LNT, LNH                                                                                                                                                               |
| <b>Fatty acids</b>                                          |                                                                                                                                                                              |
| Saturated Fatty Acids (SFA)                                 | 10:0, 12:0, 14:0, 15:0, 16:0, 17:0, 18:0, 20:0, 24:0                                                                                                                         |
| Monounsaturated Fatty Acids (MUFA)                          | 14:1n9, 16:1n9, 18:1n9, 24:1n9                                                                                                                                               |
| Polyunsaturated Fatty Acids (PUFA)<br>n-3<br>n-6            | 18:3n3, 20:4n3, 20:5n3, 22:5n3, 22:6n3,<br>18:2n6, 18:3n6, 20:2n6, 20:3n6, 20:4n6, 22:5n6,<br>22:4n6                                                                         |
| De novo synthesis                                           | 10:0, 12:0, 14:0                                                                                                                                                             |
| Diet, liver synthesis, or body storage                      | 15:0, 16:0, 17:0, 18:0, 20:0, 24:0<br>14:1n9, 16:1n9, 18:1n9, 24:1n9<br>18:3n3, 20:4n3, 20:5n3, 22:5n3, 22:6n3,<br>18:2n6, 18:3n6, 20:2n6, 20:3n6, 20:4n6, 22:5n6,<br>22:4n6 |

### 2.2.2. Fatty acids analysis

Fatty acids were also previously analysed by gas liquid chromatography at University of Alberta as previously described (215). The following 29 fatty acids were detected: capric acid (10:0), lauric acid (12:0), myristic acid (14:0), physeteric acid (14:1n9), pentadecylic acid (15:0), palmitic acid (16:0), 7-hexadecenoic acid (16:1n9), margaric acid (17:0), stearic acid (18:0), trans vaccenic acid (TVA, 18:1t11), oleic acid (18:1n9), cis-vaccenic acid (18:1c11), linoleic acid (LA, 18:2n6), arachidic acid (20:0),  $\gamma$ -linoleic acid (GLA, 18:3n6),  $\alpha$ -linolenic acid (ALA, 18:3n3), eicosadienoic acid (20:2n6), dihomo- $\gamma$ -linolenic acid (DHGLA, 20:3n6), arachidonic acid (AA, 20:4n6), conjugated linoleic acid (CLA, family of isomers of linoleic

acid), eicosapentanoic acid (EPA, 20:5n3), nervonic acid (24:1n9), docosaheptaenoic acid (DHA, 22:6n3), lignoceric acid (24:0), osbond Acid (22:5n6), docosapentaenoic acid (DPA, 22:5n3), eicosatetraenoic acid (20:4n3), adrenic acid (C22:4n6). Fatty acids were grouped into SFAs, MUFAs, and PUFAs. Additionally, they were classified as synthesized de novo in the mammary gland, or acquired from the diet, liver synthesis, or body fat storage. PUFAs were also divided into n-3 and n-6 series (Table 3).

### **2.2.3. Metabolic hormones**

Milk metabolic hormones (leptin and insulin) were measured using the Mesoscale Discovery System at University of Alberta as previously described (193). Briefly, measurements were done in duplicate using kits precoated with the antibody to each hormone following the manufacturer's instructions (MesoScale Discovery, Gaithersburg, MD, USA). Results were measured on the MesoScale Discovery Sector Imager 2400 plate reader and the Discovery Workbench 3.0 software was used to generate standard curves and calculate analyte concentrations in each sample.

## **2.3. Statistical analysis**

Statistical analysis was performed in R (211). List of statistical methods, tools, and R packages used in this study are summarised in Table 4.

### **2.3.1. Handling Missing Data**

Missing data in HMOs (n=4) and fatty acids (n=5) matrices was imputed using multivariate imputation by chained equations (MICE) package (216). Dimension reduction of ASVs, HMOs, and fatty acids was achieved by principal component analysis (PCA) (217). Samples with missing data for other covariates were excluded from multivariable analyses.

### **2.3.2. Exclusion of data**

Sequencing reads were excluded if they were deemed to be potential reagent contaminants. Samples were excluded if they had less than the rarefying threshold (25,000

and 1,000 for bacteria and fungi, respectively) sequencing reads. Additionally, samples were excluded from multivariable analyses if they were missing essential covariate data.

### **2.3.3. *Linear association of alpha diversity with maternal, infant, and milk factors***

Alpha diversity was assessed by the observed ASVs (richness; (218) and inverse Simpson Index (diversity (219). Association of alpha diversity with maternal, infant, early life, and milk factors was assessed by linear regression adjusting for factors with p value of <0.05 in univariate analysis.

### **2.3.4. *Linear relationships between taxa abundance and maternal, infant, and milk factors***

To control for the compositional nature of the data, ASV counts were centre log-ratio transformed following zero-replacement (220, 221). After this transformation, taxa relative abundances were compared at species, genus, and phylum levels by one-way analysis of variance (ANOVA). P values were corrected with Benjamini-Hochberg's false discovery rate (FDR) method (222).

### **2.3.5. *Redundancy analysis***

The association of maternal, infant, early life, and milk factors with milk microbiota composition was assessed by redundancy analysis (RDA) with 1000 permutations using the vegan package (223).

### **2.3.6. *Hierarchical clustering***

Hierarchical clustering was performed on Bray-Curtis dissimilarity matrix of core ASVs (n=12, defined as being present in at least 95% of samples with minimum of 1% mean relative abundance), with ward sum-of-square algorithm (224). The optimal number of clusters was determined using Gap statistics, which compares the observed change in within-cluster dispersion versus the expected change under an appropriate reference null distribution (225). Dissimilarity (beta diversity) of clusters was assessed by permutational ANOVA

(PERMANOVA) using the vegan package (223). Association of factors with cluster membership was assessed using ANOVA (Tukey post hoc) for continuous and  $\chi^2$  (post hoc) for categorical variables.

#### **2.3.7. Structural equation modeling**

Structural equation modeling (SEM) was performed using confirmatory factor analysis (CFA) to assess the direct vs. indirect association of determinant factors of the milk microbiota. SEM was conducted using the lavaan package (226) and path diagrams were visualized using semPlot package (227). Variable selection was informed by the results of statistical tests explained above. The milk environment was modelled as a latent construct with principal component (PC) 1 axes of HMOs, fatty acids, and hormones as indicators. The milk microbiota PC1 was used as the milk microbiota variable. Multivariate normality was assessed by Henze-Zirkler's tests (228). Given the non-normality of the data and inclusion of dichotomous categorical variables, the model was estimated using maximum likelihood (ML) parameter estimation with bootstrapping (n=1000) (229). The latent variable was scaled to have variance of one. Model fit was assessed by  $\chi^2$  test, the comparative fit index (CFI), root mean square error of approximation (RMSEA) and its 90% confidence interval (CI), and the standardized root mean residuals (SRMR). Non-significant  $\chi^2$  test, CFI $\geq$ 0.9, RMSEA $<$ 0.05, and SRMR $<$ 0.08 were considered as indications of good model fit (229). Equivalents of the final model were explored and defined by changing the directionality of associations or implying co-variation instead of causation.

#### **2.3.8. Linear discriminant analysis**

Taxa enrichment was assessed by linear discriminant analysis (LDA) effect size (LEfSe) with default parameters and logarithmic LDA score threshold of three (230).

#### **2.3.9. Presence/absence analysis**

Binary outcomes (*Bifidobacterium* and fungi prevalence) were tested by  $\kappa^2$  and logistic

regression. Unsupervised variable selection by least absolute shrinkage and selection operator (LASSO) was performed using *glmnet* package with default parameters to define the optimum lambda with 10-fold cross-validation. Using the optimum lambda, coefficients were estimated for model parameters (231). Performances of logistic regression and LASSO were assessed by area under the receiver operation curve (AUROC) analysis.

#### **2.3.10. Inter-kingdom bacteria fungi correlation network analysis**

Inter-kingdom network analysis Spearman correlation scores were calculated using the CoNet app (232) for Cytoscape (233). Only edges with correlation scores of  $> 0.5$  were kept ( $p < 0.05$ , Bonferroni-corrected). The analysis was conducted 100 bootstrap samples were used to infer pseudo-p values. The inferred correlations were restricted to those having correlations  $> |0.6|$  ( $p < 0.05$ , two-sided).

### **2.4. Comparing sequencing processing approaches**

The milk microbiota data described in section 2.1.5 was re-processed with an OTU-picking algorithm to examine the impact on downstream statistical analysis.

#### **2.4.1. Milk microbiota sequencing data re-processing using UCLUST**

As previously described, I processed the milk microbiota data using DADA2 (section 2.1.5). Next, to compare the denoising algorithm (DADA2) to a clustering-based OTU picking approach (UCLUST), I re-processed the data. Briefly, paired-end reads were merged using the PANDAseq assembler (234). Sequences with low quality base calling scores ( $< 20$ ) as well as those containing ambiguous bases in the overlapping region were discarded. The subsequent fastq file was processed using the open-source software QIIME v1.9.1 (208). Assembled reads were demultiplexed according to the barcode sequences and chimeric reads were filtered using UCHIME (235). Reads were clustered into operational taxonomic units (OTUs) using closed-reference OTU picking based on 97% similarity using UCLUST (236). Representative sequences from each OTU were assigned a taxonomy using RDP Classifier

(237) and aligned to the 2013 release of the Greengenes reference database at 97% sequence similarity (209) using PyNAST (238).

**Table 4. Key resource table for the culture-independent study of milk microbiota in the CHILD Cohort Study.** Reproduced from Moossavi et al. (139).

| REAGENT or RESOURCE                                                                                              | SOURCE                                  | IDENTIFIER                                                                                         |
|------------------------------------------------------------------------------------------------------------------|-----------------------------------------|----------------------------------------------------------------------------------------------------|
| <b>Biological samples</b>                                                                                        |                                         |                                                                                                    |
| Breastmilk (1ml)                                                                                                 | Moraes et al. 2015                      | CHILD Cohort Study                                                                                 |
| <b>Commercial kits</b>                                                                                           |                                         |                                                                                                    |
| Quick-DNA™ Fungal/Bacterial extraction kit                                                                       | Zymo Research                           | D6005                                                                                              |
| Illumina MiSeq reagent kit v3                                                                                    | Illumina                                | MS-102-3003                                                                                        |
| <b>Primers</b>                                                                                                   |                                         |                                                                                                    |
| 16S rRNA-Forward Primer<br>515F: GTGCCAGCMGCCGCGGTAA<br>16S rRNA-Reverse Primer<br>806R: GGACTACHVGGGTWTCTAAT    | Caporaso, et al., 2012                  |                                                                                                    |
| ITS2 rRNA-Forward Primer<br>ITS3: GCATCGATGAAGAACGCAGC<br>ITS2 rRNA-Reverse Primer<br>ITS4: TCCTCCGCTTATTGATATGC | Derakhshani, et al. 2018                |                                                                                                    |
| <b>Software and algorithm</b>                                                                                    |                                         |                                                                                                    |
| QIIME2                                                                                                           | Caporaso, et al. 2010                   | v. 2018.6                                                                                          |
| Greengenes                                                                                                       | DeSantis, et al. 2006                   | v. 13_8                                                                                            |
| UNITE                                                                                                            | Nilsson, et al. 2015                    | v. 7                                                                                               |
| Phyloseq                                                                                                         | McMurdie and Holmes 2013                | v. 1.19.1                                                                                          |
| Decontam                                                                                                         | Davis, et al. 2018                      | v. 1.1.0                                                                                           |
| MICE                                                                                                             | van Buuren and Groothuis-Oudshoorn 2011 | v. 2.30                                                                                            |
| CoDaSeq                                                                                                          | Gloor and Reid 2016                     | v. 0.99.1                                                                                          |
| Vegan                                                                                                            | Oksanen, et al. 2017                    | v. 2.4-4                                                                                           |
| LEFSe                                                                                                            | Segata, et al. 2011                     | <a href="http://www.huttenhower.sph.harvard.edu/galaxy">www.huttenhower.sph.harvard.edu/galaxy</a> |
| Lavaan                                                                                                           | Rosseel 2012                            | v. 0.5-23.1097                                                                                     |
| Glmnet                                                                                                           | Friedman, et al. 2010                   | v. 2.0.18                                                                                          |
| Psych                                                                                                            | Koo and Li 2016                         | v. 1.8.12                                                                                          |
| R                                                                                                                | R Core Team                             | v. 3.3.3 and 3.5.1                                                                                 |
| <b>Deposited Data</b>                                                                                            |                                         |                                                                                                    |
| Raw and analysed data (bacteria)                                                                                 | This thesis                             | PRJNA481046                                                                                        |
| Raw and analysed data (fungi)                                                                                    | This thesis                             | PRJNA536254                                                                                        |



#### **2.4.2. *Comparison of clustering-based and denoising algorithms***

Depth of sequencing and alpha diversity (observed OTUs/ASVs and inverse Simpson index) were compared between methods (n=4 datasets) using Student's t test. Within the 4 most abundant phyla, the prevalence (percentage of samples containing the taxa) and average relative abundance of classified genera were compared between UCLUST and DADA2 prior to contaminant removal using Pearson correlation.

#### **2.4.3. *Interclass correlation analysis***

Agreement and consistency of community alpha diversity and relative abundances of the most abundant families were assessed by interclass correlation by 2-way random and fixed single measurement models using Psych package (239, 240).

### **2.5. *Culture-enriched molecular profiling***

I performed a pilot study of culture-enriched molecular profiling of milk microbiota.

#### **2.5.1. *Sample collection***

Milk samples were collected from mothers of premature infants admitted to the neonatal intensive care unit (NICU), McMaster University. Milk samples were not handled or stored anaerobically. Long-term frozen milk samples (n=3) previously collected at approximately 60 days postpartum were obtained from Department of Pediatrics, McMaster University. These samples were stored for more than 6 months at -80 °C prior to analysis. In addition, fresh leftover milk samples (n=5) supplemented with various concentrations of different fortifiers were collected from NICU. Fresh milk samples were collected by pump expression and immediately transferred to the lab. Fresh samples were immediately processed and then transferred to -80 °C freezer for 2 weeks. The study was considered exempt from ethics approval by Ethics Review Board at McMaster University.

#### **2.5.2. *Culturing of milk samples***

100 µL of milk sample was diluted in 900 µL of Brain heart infusion (BHI) broth (BD,

Sparks, MD, USA) with 0.05 % L-cysteine hydrochloride hydrate. 100 µl of 10<sup>-1</sup> dilution was plated on pre-reduced 100-mm agar plates. The choice of culture conditions were modified from Lau et al. (63) based on anticipated bacterial profile of the human milk according to our previously published 16S rRNA sequencing-based study (139). Culture media included are summarised in Table 5.

**Table 5. Culture media used in the culture-enriched molecular profiling of human milk samples.**

| Culture media    | Company             | Supplement                                              | Target bacterial groups       |
|------------------|---------------------|---------------------------------------------------------|-------------------------------|
| BHI3             | BD <sup>a</sup>     | 0.5 g/L L-cysteine<br>10 mg/L hemin<br>1 mg/L Vitamin K | Anaerobes                     |
| BBE <sup>b</sup> | -                   | -                                                       | Bacteroides                   |
| BEEF             | Fluka <sup>c</sup>  | -                                                       | Strict anaerobes              |
| BSM              | Fluka               | -                                                       | Bifidobacterium               |
| CBA              | BD                  | 5% sheep's blood                                        | Non-selective                 |
| FAA              | Neogen <sup>d</sup> | -                                                       | Fastidious anaerobe           |
| MAC              | BD                  | -                                                       | Gram negative                 |
| MRS              | BD                  | -                                                       | Lactobacillus                 |
| MSA              | BD                  | -                                                       | Gram positive; Staphylococcus |
| THY              | BD                  | 5 g/L yeast extract                                     | Fastidious bacteria           |

<sup>a</sup> Sparks, MD, USA

<sup>b</sup> Prepared according to the protocol in (241).

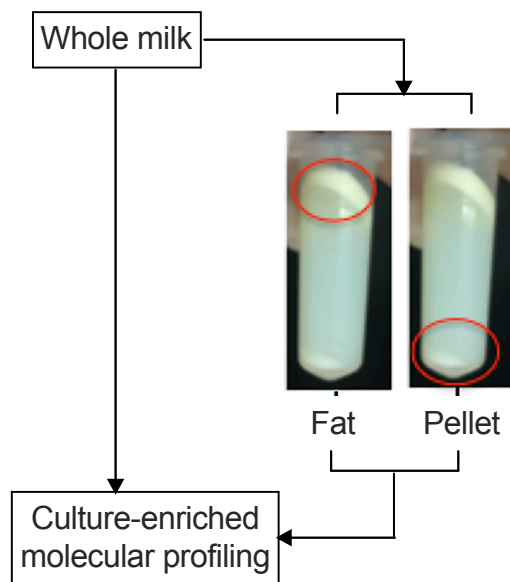
<sup>c</sup> St. Louis, MO, USA

<sup>d</sup> Lansing, MI, USA

BHI, Brain heart infusion; BEEF, Cooked meat agar; BBE, Bacteroides bile esculin agar; BSM, Bifidobacterium Selective Media; CBA, Columbia blood agar; FAA, Fastidious anaerobe agar; MAC, MacConkey agar; MRS, de Man Rogosa Sharpe agar; MSA, Mannitol slat agar; THY, Tryptic soy agar.

One set of media was incubated at 37 °C for 5 days in an anaerobic chamber and another set was incubated at 37 °C in 5 % CO<sub>2</sub> for 3 days. In a subset of fresh samples, the fat and pellet fractions were separated by centrifugation at 12,500g for 10 minutes (Figure 6). Each fraction was resuspended in 1ml of BHI and plated and incubated aerobically as described. For the comparison of fresh and short-term frozen samples, one aliquot was immediately cultured and another was frozen at -80 °C for 2 weeks after the initial plating.

100 µL of the short-term frozen samples were cultured both aerobically and anaerobically as described above. After incubation, the colonies on each plate were pooled and collected by adding 1 mL BHI broth and scraping the surface of the plate with a cell scraper. The harvested colonies were frozen in 50 % glycerol at -80 °C as stocks and 300 µl were used for DNA extraction and sequencing.



**Figure 6. Milk culture-enriched molecular profiling of different milk fractions.** Fat and pellet fractions were separated following centrifugation (see section 2.5.2). Adapted from Lima et al. (80).

### 2.5.3. DNA extraction and 16S rRNA gene sequencing

DNA extraction and purification were performed by Michelle Shah in Surette Lab as previously described (242). Briefly, 500 µl of harvested colonies was mechanically homogenized with 0.2 g of 0.1 mm glass beads (MoBio, Carlsbad, CA, USA; an additional 0.2 g of 2.8 mm glass beads was added to fecal samples) in 800 µl of 200 mM NaPO<sub>4</sub>, pH 8 and 100 µl guanidine thiocyanate-EDTAN-lauroyl sarcosine. Enzymatic lysis with 50 µl lysozyme (100 mg/ml), 50 µl mutanolysin (10 U/µl), and 10 µl RNase A (10 mg/ml) for 1 h at 37 °C was followed by the addition of 25 µl 25 % sodium dodecyl sulfate (SDS), 25 µl Proteinase K (20 mg/ml,) and 75 µl 5 M NaCl and further incubated for 1 h at 65 °C. Supernatants were collected and DNA extracted with phenolchloroform-isoamyl alcohol (25:24:1; Sigma, St. Louis, MO, USA) and further purified using DNA Clean and

Concentrator-25 columns (Zymo, Irvine, CA, USA) as per the manufacturer's instructions. Isolated DNA was stored at  $-20^{\circ}\text{C}$ .

PCR amplification of the V3 region of the 16S rRNA gene was performed by Laura Rossi in Surette Lab using F341/R518 primers as previously described (243) with the following modifications: a 60  $\mu\text{l}$  reaction containing 1.25 mM  $\text{MgCl}_2$ , 2.5 mM of each dNTP, 100 nM of each barcoded primer, and 1.25 U Taq was divided into  $3 \times 20 \mu\text{l}$  reactions for amplification. PCR conditions consisted of an initial denaturation at  $94^{\circ}\text{C}$  for 2 min, 30 cycles of  $94^{\circ}\text{C}$  for 30s,  $50^{\circ}\text{C}$  for 30s,  $72^{\circ}\text{C}$  for 30s, followed by a final elongation at  $72^{\circ}\text{C}$  for 10 min. Purified PCR products were sequenced using the Illumina MiSeq platform by the McMaster Genome Facility (Hamilton, ON, Canada). Sterile DNA-free water was used as negative controls in sequencing library preparation.

#### ***2.5.4. Microbial data pre-processing and reagent contaminant removal***

Demultiplexed forward and reverse sequencing reads were imported in R (211). Primers and their reverse complements were identified using Biostrings package (244) and subsequently removed from the sequencing data using cutadapt (245). Overlapping paired-end reads were processed using DADA2 (173) in R. Unique ASVs were assigned taxonomy and aligned to the SILVA reference database version 132 at 99% sequence similarity (246). Demultiplexed sequencing data will be deposited into the Sequence Read Archive (SRA) of NCBI.

#### ***2.5.5. Sequencing output processing following contaminant removal***

Data analysis was conducted in R (211) using the Phyloseq package (212). ASVs with zero abundance across the samples; non-bacterial (belonging to phylum Cyanobacteria, family of mitochondria, and class of chloroplast) were removed. Additional filtering steps were applied as appropriate (Chapter 4).

### 2.5.6. Statistical analysis

Statistical analysis was performed in R (211). List of statistical methods, tools, and R packages used in this study are summarised in Table 6.

**Table 6. Key resource table for the culture-enriched molecular profiling of milk microbiota.**

| REAGENT or RESOURCE                                                                                      | SOURCE                   | IDENTIFIER         |
|----------------------------------------------------------------------------------------------------------|--------------------------|--------------------|
| <b>Biological samples</b>                                                                                |                          |                    |
| Frozen breastmilk (50ml)                                                                                 | Department of Pediatrics | NA                 |
| Fresh breastmilk (10ml)                                                                                  | NICU                     | NA                 |
| <b>Culture media</b>                                                                                     |                          |                    |
| Brain heart infusion agar                                                                                | BD                       |                    |
| Columbia blood agar                                                                                      | BD                       |                    |
| Tryptic soy agar                                                                                         | BD                       |                    |
| Fastidious anaerobe agar                                                                                 | Neogen                   |                    |
| Cooked meat agar                                                                                         | Fluka                    |                    |
| Bifidobacterium Selective Media                                                                          | Fluka                    |                    |
| Mannitol slat agar                                                                                       | BD                       |                    |
| de Man Rogosa Sharpe agar                                                                                | BD                       |                    |
| Bacteroides bile esculin agar                                                                            | Atlas 2004               |                    |
| MacConkey agar                                                                                           | BD                       |                    |
| <b>Primers</b>                                                                                           |                          |                    |
| 16S rRNA-Forward Primer<br>F341: CCTACGGGAGGCAGCAG<br>16S rRNA-Reverse Primer<br>R518: ATTACCGCGGCTGCTGG | Bartram, et al. 2011     |                    |
| <b>Software and algorithm</b>                                                                            |                          |                    |
| Biostrings                                                                                               | Pagès, et al. 2019       | v. 2.48            |
| Cutadapt                                                                                                 | Martin, et al. 2011      | v. 1.18            |
| DADA2                                                                                                    | Callahan, et al. 2016    | v. 1.8             |
| SILVA                                                                                                    | Yilmaz, et al. 2014      | v. 132             |
| Phyloseq                                                                                                 | McMurdie and Holmes 2013 | v. 1.19.1          |
| CoDaSeq                                                                                                  | Gloor and Reid 2016      | v. 0.99.1          |
| Vegan                                                                                                    | Oksanen, et al. 2017     | v. 2.4-4           |
| R                                                                                                        | R Core Team              | v. 3.3.3 and 3.5.1 |
| <b>Generated Data</b>                                                                                    |                          |                    |
| Raw and analysed data (bacteria)                                                                         | This thesis              | NA                 |

2.5.6.1. *Linear association of alpha diversity with milk storage, Oxygen level for incubation, and milk fractions*

Alpha diversity was assessed by the observed ASVs (richness). Association of alpha diversity with factors of interest was assessed by ANOVA and Tukey's post hoc pairwise comparison whenever appropriate.

2.5.6.2. *Linear relationships between taxa abundance and milk storage, Oxygen level for incubation, and milk fractions*

As described above (section 2.3.4), to control for the compositional nature of the data, ASV counts were centre log-ratio transformed following zero-replacement (220, 221). After this transformation, taxa relative abundances were compared at the family level by ANOVA. P values were corrected with Benjamini-Hochberg's FDR method (222).

2.5.6.3. *Presence/absence analysis*

Frequency of isolation of bacterial families was tested by  $\kappa^2$  test.

### **3. MILK MICROBIOTA AND MYCOBIOTA PROFILE IN THE CHILD COHORT STUDY**

#### **3.1. Introduction**

The majority of culture-dependent and -independent studies of the milk microbiota have been focused on bacteria and were conducted in relatively small number of samples. In addition, bacteria-fungi interactions are ubiquitous in microbial communities (144), but have rarely been investigated in milk. To address these shortcomings, milk microbiota and mycobiota were profiled in breastmilk in the CHILD Cohort Study by 16S rRNA gene sequencing in 393 mother-infant dyads and ITS2 gene sequencing in 271 from the 393 dyads with available bacterial data (Figure 5). The section on milk microbiota profile (section 3.2) is published in Cell Host Microbe (139). Kelsey Fehr helped with fungal ITS2 gene sequencing.

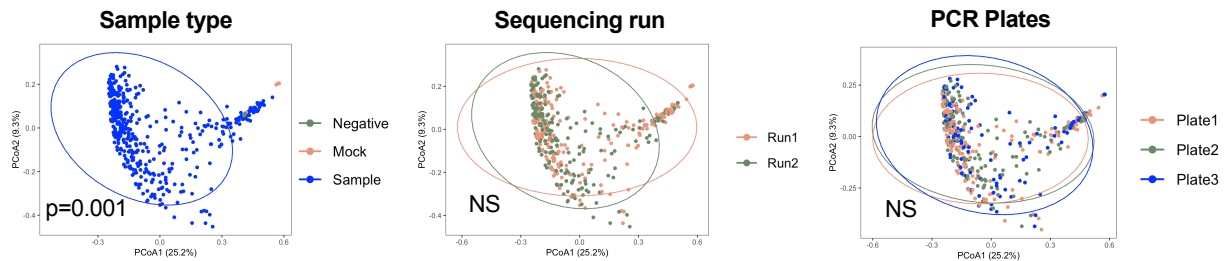
#### **3.2. Milk microbiota profile in human milk from CHILD Cohort Study**

We studied a representative subset of 393 breastfeeding dyads from the CHILD Cohort Study (Figure 5 and Table 1). The majority were Caucasian (74%), about half (54%) were primiparous and a quarter (24%) delivered by Cesarean section. The mean $\pm$ SD duration of any breastfeeding was 13 $\pm$ 6 months and the mean lactation stage at sample collection was 17 $\pm$ 5 weeks.

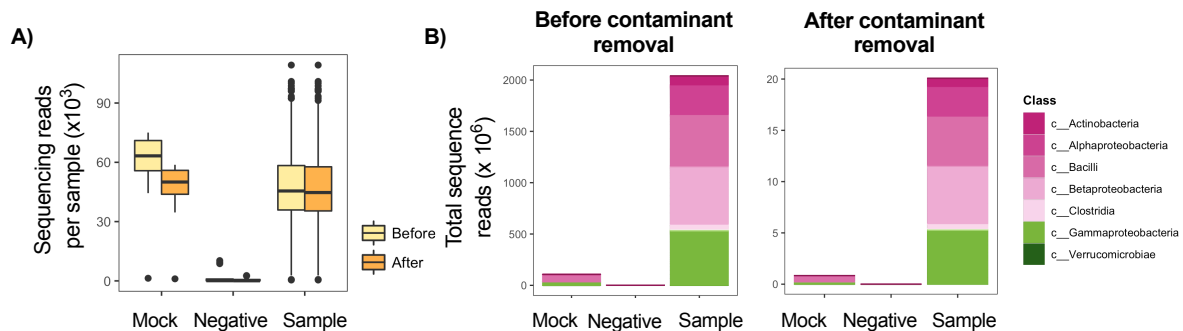
##### ***3.2.1. Batch effect and sequencing inaccuracy were ruled out by quality control assessment***

I obtained a mean of 47,710 $\pm$ 18,643 high-quality sequencing reads per sample, compared with 46,770 $\pm$ 13,479 reads from the mock community, and 627 $\pm$ 1,034 reads in sequencing negative controls. Batch effect was ruled out by comparing beta diversity

between PCR plates and sequencing run (Figure 7). Sequencing contaminants (n=173 from 9884 total ASVs) were identified and removed using the *decontam* package (190); this did not measurably affect sequencing depth (Figure 8A) or microbiota structure (Figure 8B).



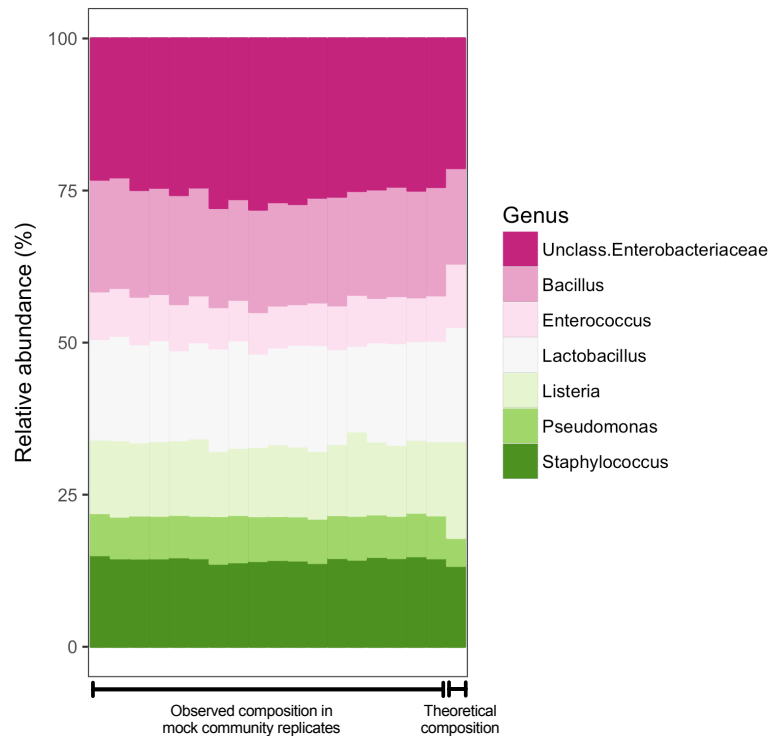
**Figure 7. Milk microbiota beta diversity by sample type, sequencing run, and PCR plates, prior to removal of potential contaminant ASVs using the *decontam* package in 393 human milk samples from the CHILD Cohort Study.** Beta diversity was assessed on Bray-Curtis dissimilarity and tested by PERMANOVA. There was a significant difference based on sample type but not sequencing runs or PCR plates. Reproduced from Moossavi et al. (139).



**Figure 8. Data quality control in human milk samples from the CHILD Cohort Study.** Comparison of A) Sequencing depth and B) Composition of the top 500 most abundant ASVs before and after contaminant removal. N samples = 428, N mock =18, N negative control =15. Reproduced from Moossavi et al. (139).

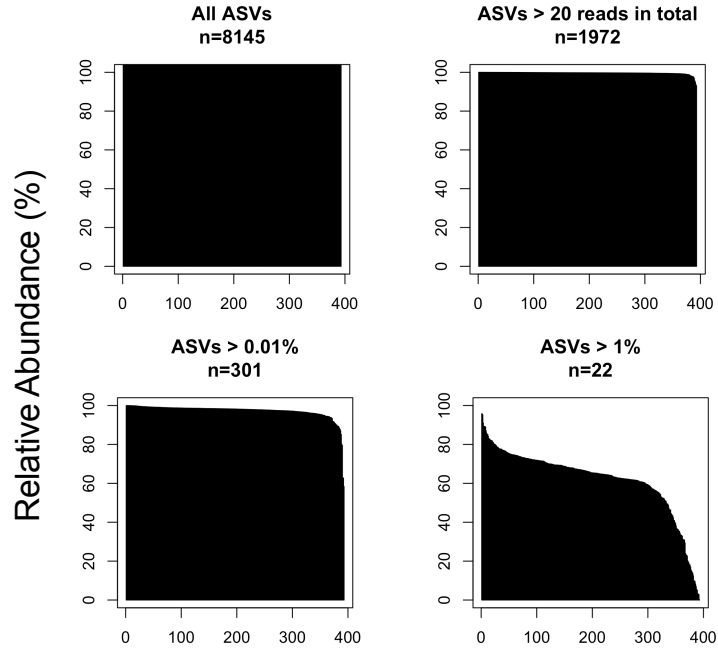
Mock community composition and agreement with theoretical composition was assessed for bacteria. We observed strong consistency between the observed and expected composition of the mock community (Figure 9). The mock community was dominated by *Cryptococcus neoformans* (Basidiomycota phylum) suggesting sequencing bias against Ascomycota in ITS2 sequencing.





**Figure 9. Relative abundance of most abundant bacterial genera (>1% mean relative abundance) in the mock community in comparison with the theoretical composition.** Each bar represents a replicate. Reproduced from Moossavi et al. (139)

Samples with less than 25,000 sequencing reads were excluded (n=35) and the remaining samples (n=393) were rarefied to the minimum 25,000 sequencing reads per sample. ASVs with zero abundance across the samples (n=894) and ASVs belonging to phylum Cyanobacteria, family of mitochondria, and class of chloroplast (n=240) were removed. ASVs with less than 20 reads across the entire dataset (n=6173) were also removed, resulting in 1972 remaining ASVs. The numbers of sequencing reads of taxa were then relativized to the total sum of 25,000. This dataset was used for analysis unless otherwise specified. The contribution of the excluded rare ASVs to the total reads per sample was deemed negligible (Figure 10).

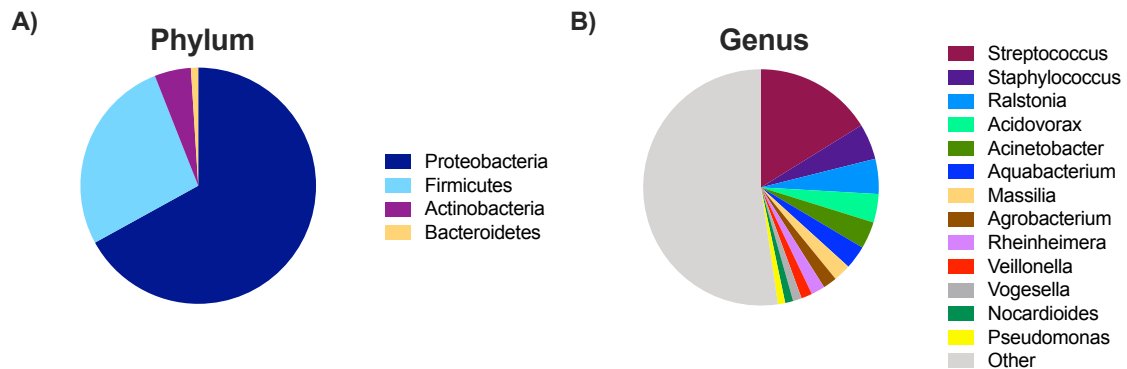


**Figure 10. Sequencing output processing to remove low abundance ASVs from milk microbiota data in the CHILD Cohort Study.** The 393 individual samples are on the X axis and relative abundances of the sums of sequence reads per sample are plotted along the Y axis. The initial ASV table following contaminant removal contained 8145 ASVs. Following exclusion of ASVs with less than 20 reads across the entire dataset, 1,972 ASVs remained. Exclusion of ASVs with less than 0.01% and 1% relative abundance resulted in 301 and 22 remaining ASVs, respectively. Analyses have been performed on the 1,972 ASVs with >20 reads in total, unless stated otherwise. Reproduced from Moossavi et al. (139)

### ***3.2.2. Milk microbiota was dominated by inversely correlated Proteobacteria and Firmicutes with high inter-individual variability***

Bacteria were present in all milk samples. Following removal of rare taxa with <20 reads in total and negligible total read counts, 18 unique phyla comprising 1972 ASVs amplicon sequence variants (ASVs) were detected. The majority of taxa detected (92%) belonged to 4 phyla: Proteobacteria (mean±SD relative abundance: 67±24%, range 3-99%), Firmicutes (26±22%, range 0.1-91%), Actinobacteria (4±4%, range 0-61%), and Bacteroidetes (1±3%, range 0-44%; Figure 11A). Proteobacteria and Firmicutes relative abundances were inversely correlated (Pearson  $r=-0.97$ ,  $p<0.001$ ). At the genus level (Figure 11B), the most abundant taxa were *Streptococcus* (16±17%), *Ralstonia* (5±3%), and *Staphylococcus* (5±11%). Four genera had ranges of higher than 70%, indicating very high

inter-individual variability: *Staphylococcus* (87%), *Streptococcus* (87%), *Acinetobacter* (87%), and *Pseudomonas* (72%; Table 7).



**Figure 11. Milk microbiota profile at A) phylum and B) genus levels in 393 human milk samples from the CHILD Cohort Study.** Genera with at least 1% mean relative abundance are presented. The less abundant genera are aggregated and visualised as Other.

**Table 7. Most abundant bacterial genera (>1% mean relative abundance) in the human milk microbiota among 393 mothers in the CHILD Cohort Study related to Figure 11.** Reproduced from Moossavi et al. (139)

| Lineage                                | Genus                 | Prevalence (%) <sup>a</sup> | Relative abundance (%) |     |       |
|----------------------------------------|-----------------------|-----------------------------|------------------------|-----|-------|
|                                        |                       |                             | Mean±SD                | CV  | Range |
| Firmicutes - Streptococcaceae          | <i>Streptococcus</i>  | 98.7                        | 16.2±16.7              | 1.0 | 87.4  |
| Firmicutes - Staphylococcaceae         | <i>Staphylococcus</i> | 99.8                        | 4.9±11.5               | 2.4 | 87.5  |
| Proteobacteria - Oxalobacteraceae      | <i>Ralstonia</i>      | 99.2                        | 4.8±2.8                | 0.6 | 9.4   |
| Proteobacteria - Comamonadaceae        | <i>Acidovorax</i>     | 99.5                        | 3.9±2.3                | 0.6 | 13.3  |
| Proteobacteria - Moraxellaceae         | <i>Acinetobacter</i>  | 48.6                        | 3.7±12.0               | 3.2 | 87.3  |
| Proteobacteria - Comamonadaceae        | <i>Aquabacterium</i>  | 99.5                        | 3.2±1.9                | 0.6 | 7.6   |
| Proteobacteria - Oxalobacteraceae      | <i>Massilia</i>       | 99.8                        | 2.4±1.4                | 0.6 | 6.5   |
| Proteobacteria - Rhizobiaceae          | <i>Agrobacterium</i>  | 99.2                        | 1.9±1.1                | 0.6 | 4.5   |
| Proteobacteria – Uncl. Alteromonadales | <i>Rheinheimera</i>   | 99.5                        | 1.9±1.2                | 0.6 | 4.7   |
| Firmicutes - Veillonellaceae           | <i>Veillonella</i>    | 78.1                        | 1.5±2.7                | 1.8 | 21.0  |
| Proteobacteria - Neisseriaceae         | <i>Vogesella</i>      | 99.2                        | 1.2±0.7                | 0.6 | 3.0   |
| Actinobacteria - Nocardioideaceae      | <i>Nocardioidea</i>   | 99.2                        | 1.1±0.6                | 0.6 | 2.6   |
| Proteobacteria - Pseudomonadaceae      | <i>Pseudomonas</i>    | 33.3                        | 1.0±7.2                | 6.9 | 72.8  |

<sup>a</sup> Percentage of samples where the taxa was present.  
CV, coefficient of variation; Uncl, unclassified.

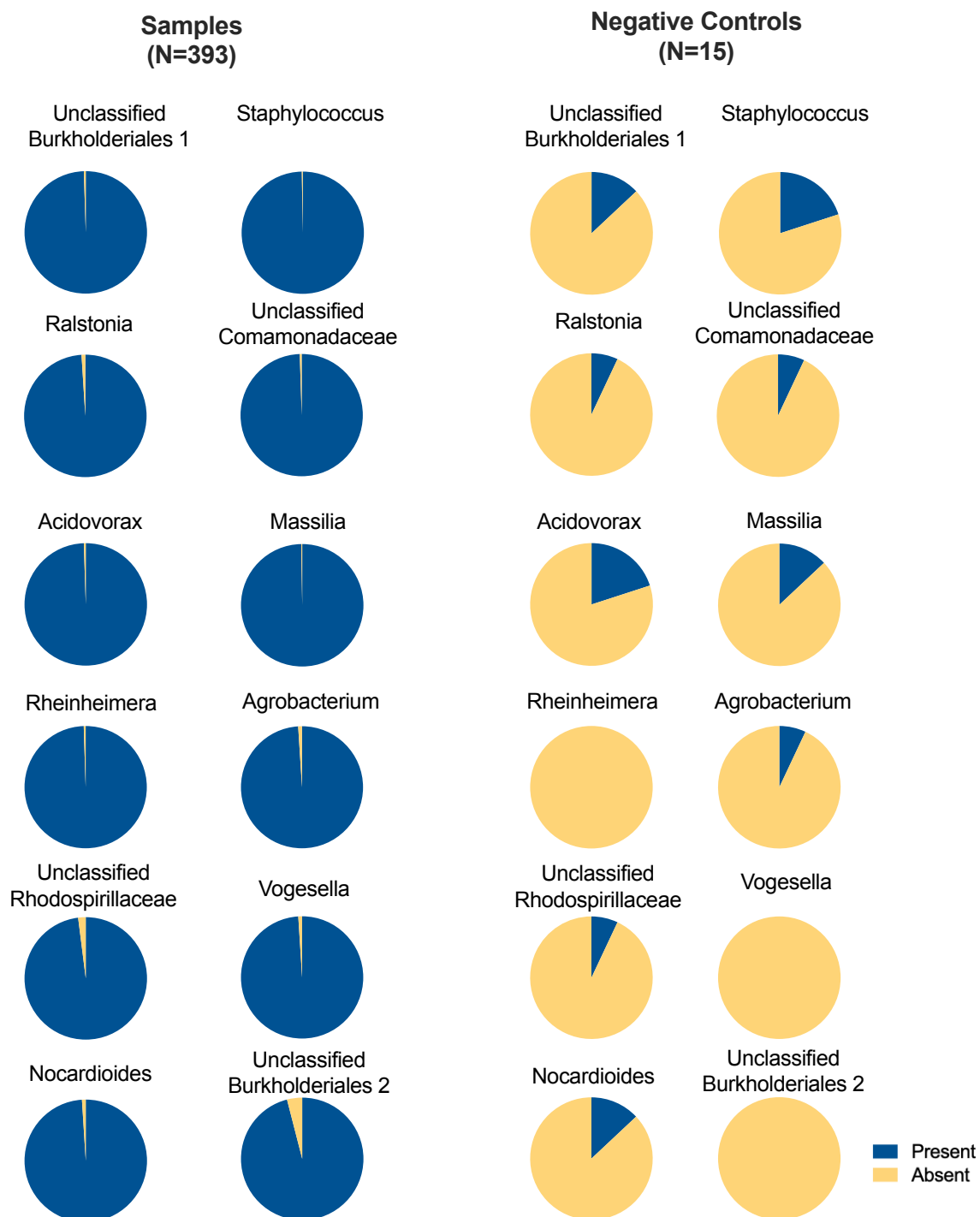
### **3.2.3. Core milk microbiota**

Core microbiota refers to taxonomy or microbial functions that are common to the majority of the hosts or environments (247). We adapted a stringent approach and defined core as ASVs with more than 1% mean relative abundance which were present in more than 95% of the samples (247, 248).

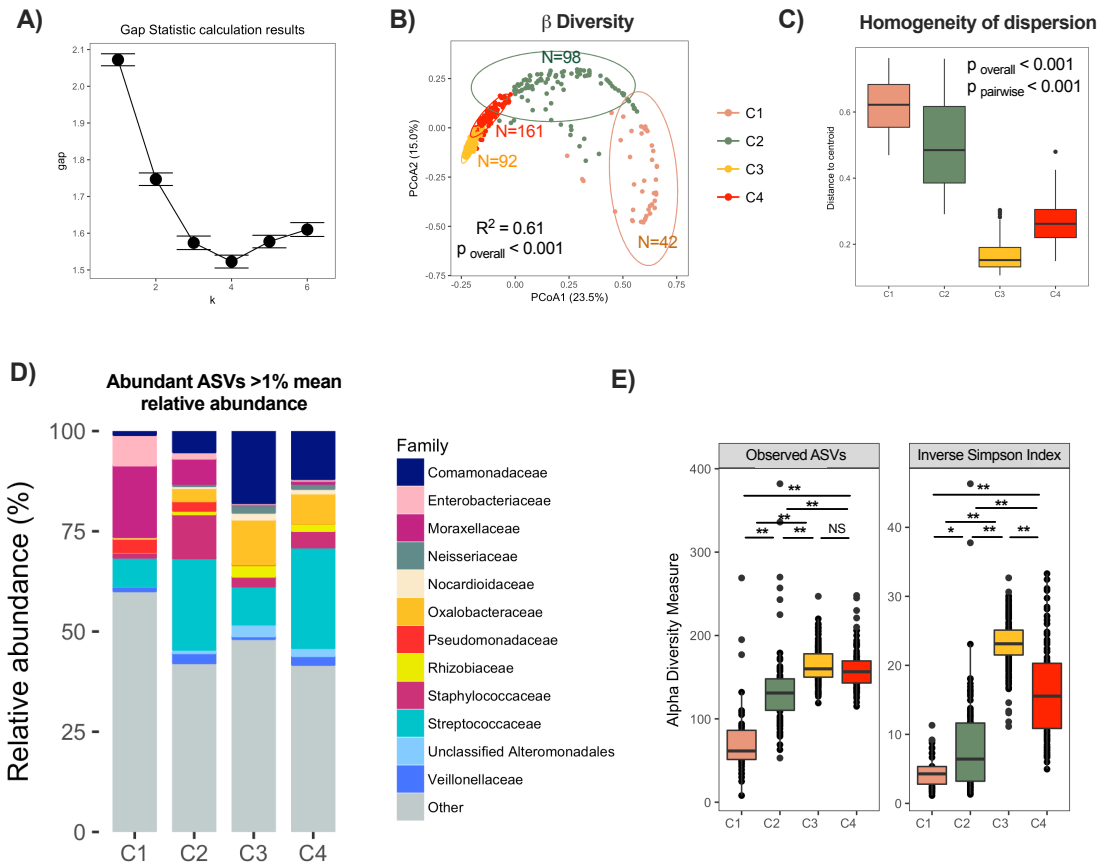
Overall, 12 core ASVs were identified which were present in the majority of samples but almost completely absent in the negative controls (Figure 12). The top five most abundant core ASVs belonged to unclassified Burkholderiales (6%±3%), *Staphylococcus* (5%±12%), *Ralstonia* (5%±3%), unclassified Comamonadaceae (4%±3%), and *Acidovorax* (4%±2%). The identified core ASVs were present in 96-99.7% of samples and 0-20% of negative controls (Figure 12). Although, some of these core taxa are environmental bacteria and potential reagent contaminants, our contamination removal confirms that these ASVs were not commonly found in the negative controls.

### **3.2.4. Milk microbiota profiles exhibited discrete compositional patterns**

Though we acknowledge that discrete clusters might not exist within the milk microbiota (249), we used hierarchical clustering to identify inherent patterns in the milk microbial community, finding four main clusters. To identify the most optimal strategy, clustering was performed on three datasets: the entire dataset, ASVs with more than 0.01% mean relative abundance, and the core ASVs. The optimum number of clusters was assessed using gap statistics (225). While no optimum clustering was detected when examining the entire dataset (ASVs with more than 0.01% mean relative abundance), the optimum number of clusters was four using the core ASVs (Figure 13A) demonstrating clear separation on PCoA plot (Figure 13B).



**Figure 12. Prevalence of core ASVs in 393 human milk samples from the CHILD Cohort Study compared to negative controls.**



**Figure 13. Hierarchical clustering of core milk microbiota among 393 mothers in the CHILD Cohort Study.** The optimum numbers of clusters in our data were determined by Gap statistics. According to the gap statistics, there are four main clusters. A) Gap statistics plot, B) PCoA plot on Bray-Curtis dissimilarity, C) Heterogeneity of clusters in terms of the milk microbiota compositional dispersion. P values for overall and pairwise comparisons are provided, D) Relative abundance of dominant bacterial families in different clusters, E) Comparison of alpha diversity between clusters. Results originally reported in Moossavi et al. (139).

Cluster (C) 1 and C2 had significantly higher heterogeneity compared to C3 and C4 (Figure 13C). The four clusters (C1-4) were predominantly separated based on the relative abundances of *Moraxellaceae*, *Enterobacteriaceae*, and *Pseudomonadaceae*, (enriched in C1, n=42), *Streptococcaceae*, *Staphylococcaceae*, and *Oxalobacteraceae* (C2, n=98), *Oxalobacteriaceae* and *Comamonadaceae* (C3, n=161), and *Streptococcaceae* and *Comamonadaceae* (C4, n=92) (Figure 13D). Clusters 1 and 3 had the highest proportion of

Proteobacteria and the clusters did not separate along the gradient of Proteobacteria-to-Firmicutes ratio. The clusters were also different in terms of  $\alpha$  diversity (Figure 13E). While C1 and C2 had the lowest richness and diversity, C3 was the most diverse. C4 had high richness but the diversity was quite low, reflecting the uneven distribution of ASVs due to dominance of *Streptococcaceae* and *Comamonadaceae* in this cluster.

### **3.3. Milk mycobiota profile in human milk from CHILD Cohort Study**

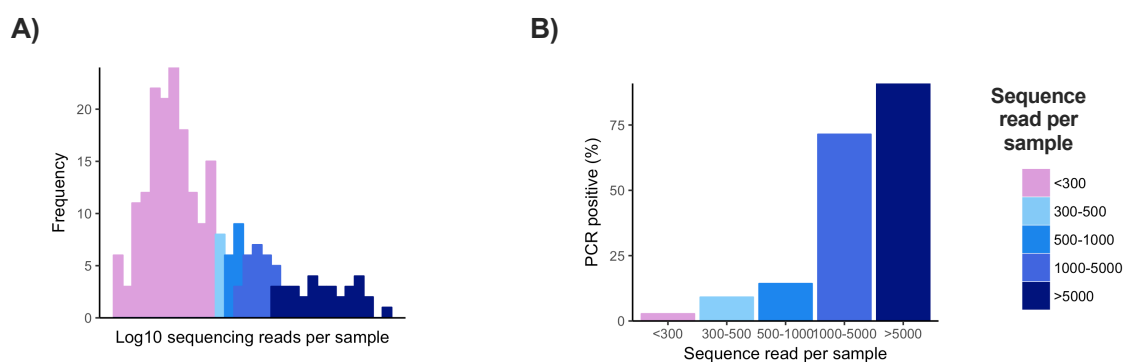
Previous dairy studies have suggested that milk also contains a fungal community (102). In a few human study, fungi were identified by both culture-dependent and -independent methods (98, 99). Several unanswered questions remain about the milk fungi: 1) how frequently are fungi present in the milk, 2) what are the dominant fungi present in milk, and 3) what factors influence the milk fungi (Chapter 7). To address these questions we profiled milk mycobiota in 271 mother-infant dyads in CHILD Cohort Study. The samples for mycobiota profiling were selected from the 393 samples with available microbiota profile while enriching for 1) samples with potentially high DNA content indirectly measured by DNA concentration following 16S rRNA gene PCR amplification and 2) asthma in the infant (Figure 5).

#### **3.3.1. Quality control**

We obtained a mean (SD) of 16,520 (78,489) and median (IQR) of 52 (13-709) high quality fungal sequencing reads per milk sample, compared with 129,100 (69,044) reads from the positive controls (stool sample and mock community), and 9 (17) reads in negative controls. The mock community was dominated by *Cryptococcus neoformans* (Basidiomycota phylum) suggesting sequencing bias against Ascomycota. We did not identify any fungal contaminant ASVs using decontam (190). Overall, 1,421 unique ASVs were detected. ASVs belonging to kingdoms other than Fungi (e.g. plants, n=82 ASVs) were removed.

### 3.3.2. Detectable fungi were absent in the majority of milk samples

We obtained an average (SD) of 16520 (78489) and median (IQR) of 52 (696) high-quality sequencing reads per sample, compared with 129,100 (69,044) reads from the positive controls (stool sample and mock community), and 9 (17) reads in negative controls. The majority of samples had very low sequencing reads (Figure 14A). The threshold applied to define presence of fungi was informed by the presence of visible PCR bands on gel electrophoresis. There was a clear trend between the presence of a PCR band and the number of sequencing reads, with 9% of samples containing 300-500 reads having a band vs. 71% of samples with 1000-5000 and 91% of samples with more than 5000 reads (Figure 14B). To optimize the retention of samples that were PCR-positive by visual inspection, as well as the elimination of samples with very low sequencing depth, samples with at least 1000 sequencing reads were considered positive for fungi. Using this threshold, 61/271 (22.5%) of samples were considered to be positive for fungi. To minimise the contribution of cross-bleeding of samples with no or very small quantities of DNA during the sequencing (250), we selected 1000 as the threshold of identifying detectable fungi resulting in 22% of samples considered as positive for fungi presence.

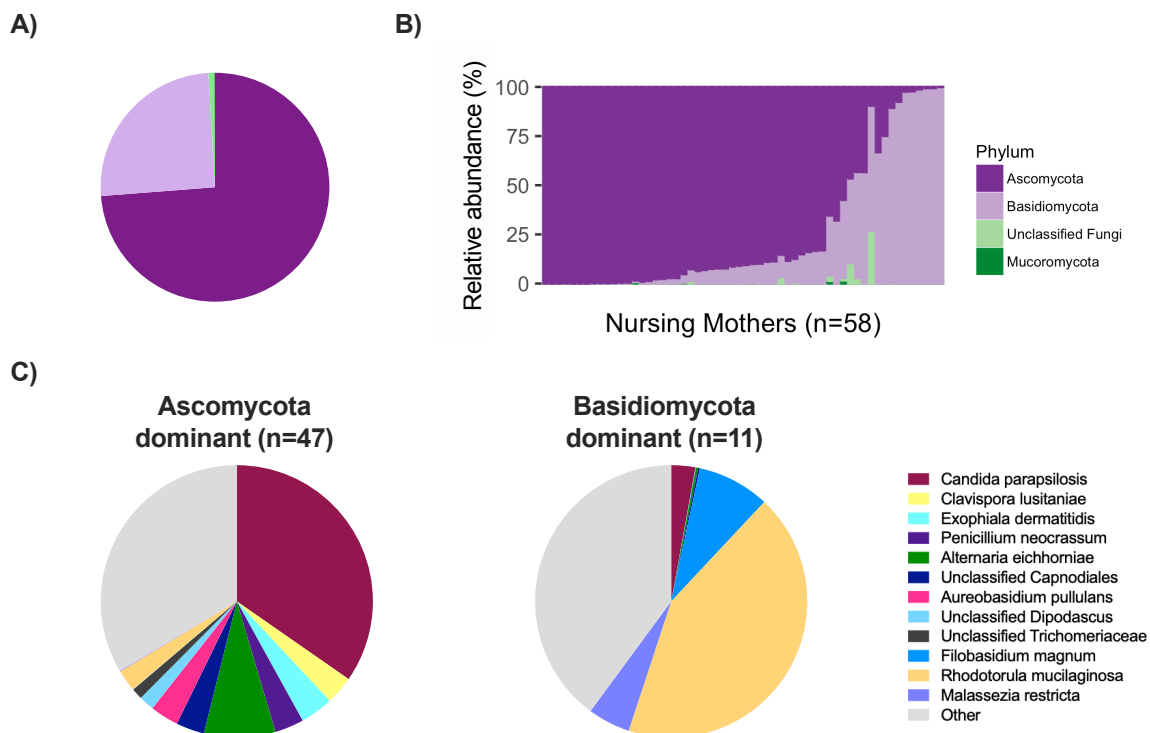


**Figure 14. Defining fungi presence in milk according to the depth of sequencing informed by the positive PCR bands in 271 human milk samples from the CHILD Cohort Study.**



### 3.3.3. Taxonomic composition of milk mycobiota

Next, among the 58 samples with detectable fungi, we assessed fungal composition and diversity. Ascomycota and Basidiomycota were the dominant fungal phyla in the milk (Figure 15A) in agreement with previous reports (98, 99). Samples were either dominated by Ascomycota (n=47) or Basidiomycota (n=11) (Figure 15B). Overall, 12 species were identified to have a minimum average relative abundance of 1%. *Candida parapsilosis* was the most abundant genera in Ascomycota-dominated samples whereas *Rhodotorula mucilaginosa* was dominant in the Basidiomycota-dominated samples (Figure 15C).



**Figure 15. Milk fungi taxonomy composition at A-B) phylum and C) species levels in 271 human milk samples from the CHILD Cohort Study.**

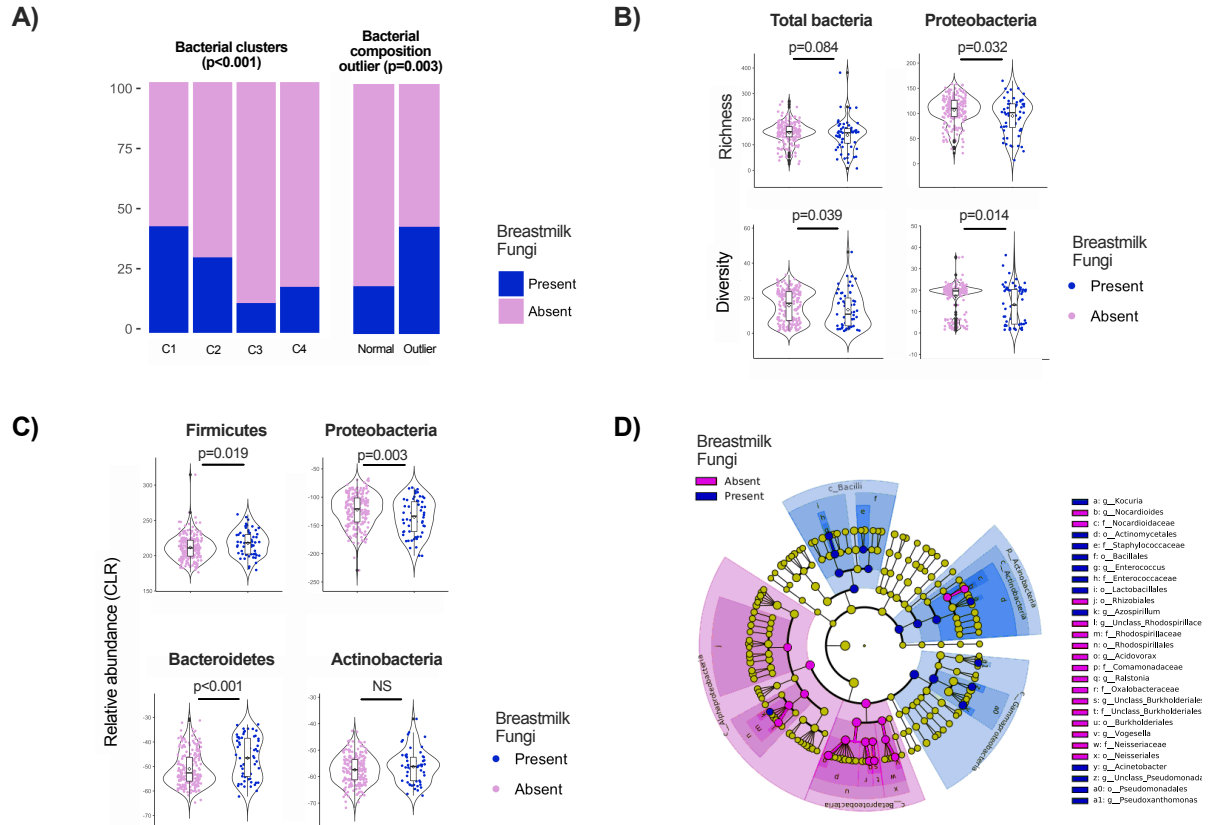
The most prevalent fungi were *Candida parapsilosis* (present in 60% of samples containing fungi, mean  $32.3\% \pm 40.9\%$  relative abundance), *Alternaria eichhorniae* (50%, mean  $7.0\% \pm 20.8\%$ ), and *Rhodotorula mucilaginosa* (43%, mean  $10.9\% \pm 26.2\%$ ). Some

samples were dominated by one species - most frequently belonging to *Candida parapsilosis* (n=17, 29% of samples), *Alternaria eichhorniae* (n=4, 7%), or *Rhodotorula mucilaginosa* (n=6, 10%). Less prevalent yet dominant species within some samples included *Clavispora lusitaniae* (n=2, 3%), *Exophiala dermatitidis* (n=2, 3%), and *Penicillium neocrassum* (n=2, 3%). Other samples contained multiple less prevalent species. Dominance of milk mycobiota by *Candida* (a common oral fungi (251)) in a subset of samples in our study and others (98, 100, 101) suggest the infant oral cavity as a potential source of milk fungi especially when directly nursed at the breast, while dominance of environmental fungi such as *Penicillium*, *Exophiala*, and *Rhodotorula* in other samples suggest contribution from maternal skin or other environmental surfaces.

### **3.4. Inter-kingdom association of milk bacteria and fungi**

Inter-kingdom interaction of fungi in bacteria in microbial communities is well known (144). The importance of a holistic approach to human microbiota taking into account all the different components including fungi is also receiving increasing attention. Currently, nothing is known about the association of bacteria and fungi in breastmilk and whether they have ecological significance in shaping the milk microbiota (bacteria, fungi, and viruses) or indirectly affect the dispersion from infant oral cavity by modulating the oral microbiota and mycobiota. We assessed the association of milk bacteria with the prevalence and composition of fungi.

First I examined the association of bacterial parameters in association with fungi prevalence. Bacterial taxonomic clusters, defined previously on the basis of hierarchical clustering of core ASVs (Figure 13)(139), were significantly associated with presence of fungi ( $p < 0.001$ ; Figure 16A).

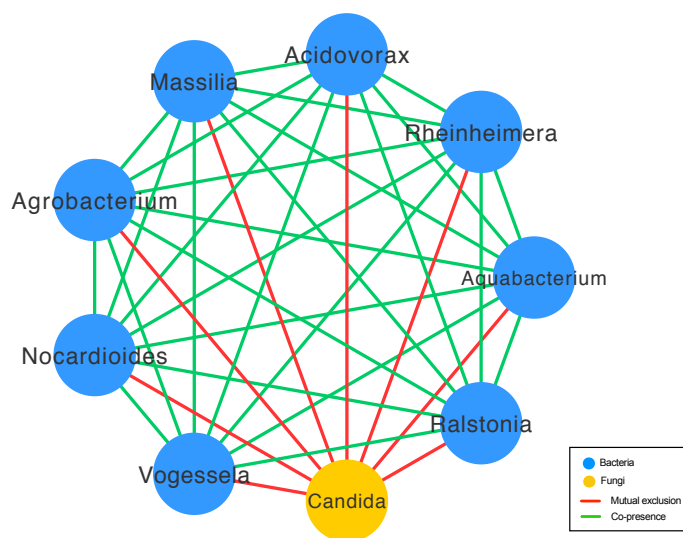


**Figure 16. Association of human milk bacteria and fungi in 271 human milk samples from the CHILD Cohort Study.** A) Prevalence of fungi according to bacterial taxonomic clusters and bacterial outliers. B) Association of total bacteria and Proteobacteria richness and diversity with presence or absence of milk fungi tested by Wilcoxon rank sum test. C) Differential abundance of milk major bacterial phyla with presence or absence of milk fungi. D) Linear discriminant analysis of milk bacteria according to presence or absence of fungi.

In the analysis of milk bacteria we identified outlier samples based on variance of the first principal component scores (179). These outliers were not associated with any technical or biological variables assessed (see Chapter 4), but we found that they had significantly higher prevalence of fungi compared to the rest of the samples (43% vs. 19%, OR=3.17 95%CI 1.49-6.65,  $p=0.002$ ) (Figure 16A). Additionally, Proteobacteria richness and diversity were lower in samples containing fungi (Figure 16B), and a similar trend was observed for total milk bacteria. Overall, the relative abundance of Proteobacteria was lower while Firmicutes and Bacteroidetes were higher in the presence of fungi (Figure 16C). Using linear discriminant analysis (230), members of Actinobacteria, Bacilli, and  $\gamma$ -Proteobacteria were

enriched in the presence of fungi while  $\alpha$ -Proteobacteria and  $\beta$ -Proteobacteria were depleted (Figure 16D). Overall, these results could reflect one or more of the following biological relationships: a) some milk bacterial communities are more permissive to fungal presence and proliferation, b) milk bacterial dynamics are influenced by fungi when they are present, and/or c) milk bacterial composition is influenced by the milk microenvironment (milk components such as HMOs and fatty acids), which also independently facilitates fungal inoculation or colonisation (252).

Interestingly, bacterial richness was positively correlated with fungal richness ( $r=0.48$ ,  $p<0.001$ ) though bacterial and fungal diversities were not significantly correlated ( $r=0.21$ ,  $p=0.11$ ), suggesting that certain factors may impact microbial dispersal in general, and thus enhance both bacterial and fungal richness, whereas the diversity of each kingdom is separately determined by the niche relevant biotic and abiotic factors. While several bacterial genera were positively associated with each other, indicative of co-presence, we did not detect any associations between different fungal genera. *Candida* was the only fungal genus associated with bacterial genera; these inter-kingdom associations were negative, suggesting a mutual exclusion relationship (Figure 17).



**Figure 17. Co-occurrence analysis of most abundant bacteria and fungi (>1% mean relative abundance) in 271 human milk samples from the CHILD Cohort Study based on Spearman rank correlation. Only significant edges with  $r>|0.5|$  are visualised. Nodes with no connections were removed, including 5 bacterial and 11 fungal taxa.**

To our knowledge, only one previous study has examined the potential interaction of fungi and bacteria in milk, finding a positive correlation between the skin inhabitant genus *Malassezia* and bacterial load (98). Our results provide preliminary evidence for antagonistic bacteria-fungi interactions, consistent with evidence that some milk bacteria demonstrate antifungal properties (253, 254), as observed in other environments (255). More research is needed to determine whether this observation is due to active fungi-bacteria ecological interactions within the milk environment, or perhaps the infant oral cavity.

### **3.5. Conclusion**

We profiled milk microbiota and mycobiota in a large population of healthy Canadian mother-infant dyads. Dominant milk bacteria in our cohort were in accordance with previously published results (23, 97). Also in agreement with previous reports (18), milk microbiota demonstrated a high degree of inter-individual variability with only a minority of taxa shared across the study population. Though we acknowledge that discrete clusters might not exist within the milk microbiota (249), we used hierarchical clustering to identify inherent patterns in the milk microbial community, finding four main clusters. In a previous report, three clusters were identified in the milk microbiota of Chinese women with dominance of *Streptococcaceae*, *Staphylococcaceae*, and *Pseudomonadaceae* defining the three clusters (40). Differences in the cluster-dominating species compared to our study could be the consequence of different clustering methods, inclusion criteria (e.g. restriction to ‘core’ taxa in our study), prevalence of key exposure variables (e.g. indirect breastfeeding), milk collection methods (e.g. aseptic protocol with electric pump in the Chinese study), or geographic variations in the milk microbiota. Further research is warranted to explore the existence and biological relevance of compositional patterns in milk microbiota.

In the general population CHILD Cohort Study, we found that approximately 20% of milk samples contained detectable fungi. Milk fungal taxonomy was consistent with previous

reports (98, 99). Ascomycota and Basidiomycota were the dominant fungal phyla in the milk and three species of *Candida*, *Alternaria*, and *Rhodotorula* were the most prevalent taxa. We observed associations between bacterial composition and presence and relative abundance of fungi, suggesting that some milk bacterial communities were more permissive to fungal presence, or vice-versa.

To our knowledge, this is one of the largest studies of human milk microbiota and mycobiota performed to date. Our results expand considerably upon existing knowledge about milk bacteria and fungi, providing evidence for bacterial-fungal interactions within the human milk.

### **Box 1. Summary of Chapter 3**

- This is one of the largest studies of human milk microbiota and mycobiota performed to date.
- All samples contained detectable bacteria whereas fungi were detected in just 20%.
- Milk microbiota was dominated by inversely correlated Proteobacteria and Firmicutes with high inter-individual variability.
- Ascomycota and Basidiomycota were the dominant fungal phyla in the milk.
- Milk microbiota profiles exhibited discrete compositional patterns. Clusters were different in the level of compositional heterogeneity, alpha diversity, and dominant taxa relative abundances.
- Bacterial taxonomic clusters with the highest compositional heterogeneity and the lowest diversity had higher prevalence of fungi.
- *Candida* (the most prevalent and abundant milk fungi) was negatively correlated with abundant milk bacteria.

## **4. CAPTURING THE DIVERSITY OF THE HUMAN MILK MICROBIOTA THROUGH CULTURE-ENRICHED MOLECULAR PROFILING**

### **4.1. Introduction**

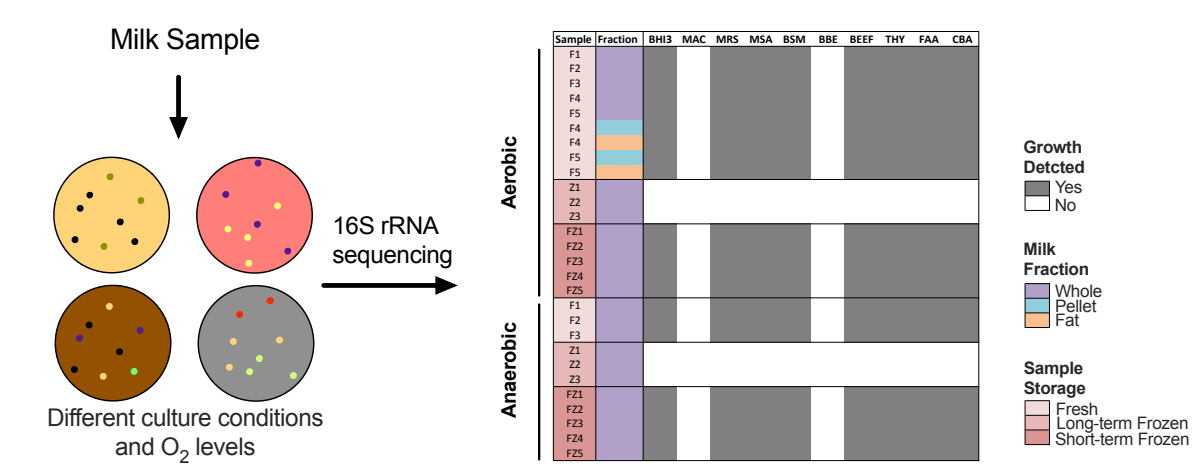
Culture-independent (DNA sequencing-based) approaches have recovered a high diversity of bacteria in breast milk including with a high degree of inter-individual variability (23, 54). Breastmilk has been historically studied using targeted culture approaches to identify potential pathogens (41, 71-75). However, culture-enriched molecular profiling has not been employed to investigate the diversity of the milk microbiota (54). Additionally, the variability of the microbiota of different milk fractions is not completely understood. We assessed the human milk microbiota through culture-enriched molecular profiling in a small pilot study on milk samples of mothers of premature infants admitted to the NICU at McMaster University (section 2.5, Figure 18). This work was conducted in Prof. Michael Surette's laboratory. DNA extraction and library preparation were performed by Michelle Shah and Laura Rossi, respectively. 16S rRNA gene sequencing was performed by personnel of Farncombe Metagenomic Facility

### **4.2. Bacteria were not isolated from long-term frozen milk samples**

We could not grow any bacterial isolates from long term frozen samples plated with ( $10^{-1}$ ) or without dilution on any of the culture conditions. In contrast, fresh and short-term frozen samples yielded detectable colonies on all culture media incubated aerobically and anaerobically except BBE and MAC (Figure 18).

Overall from 183 plate pools (generated from 8 milk samples, each cultured under 20 conditions), we obtained a mean (SD) of 10,730 (12,969) and median (IQR) of 5172 (2-56,467) high quality sequencing reads per plate pool, compared with 263 reads in the

negative control. Overall, 251 unique ASVs were detected. There was no difference in the sequencing depth of fresh vs. short-term frozen samples ( $10,694 \pm 12,831$  vs.  $11,908 \pm 13,402$ ,  $p > 0.05$ ); while plate pools of long-term frozen samples without visible colonies had depth of sequencing of  $385 \pm 763$ . Therefore, long-term frozen samples were considered negative and subsequently excluded from downstream analysis.

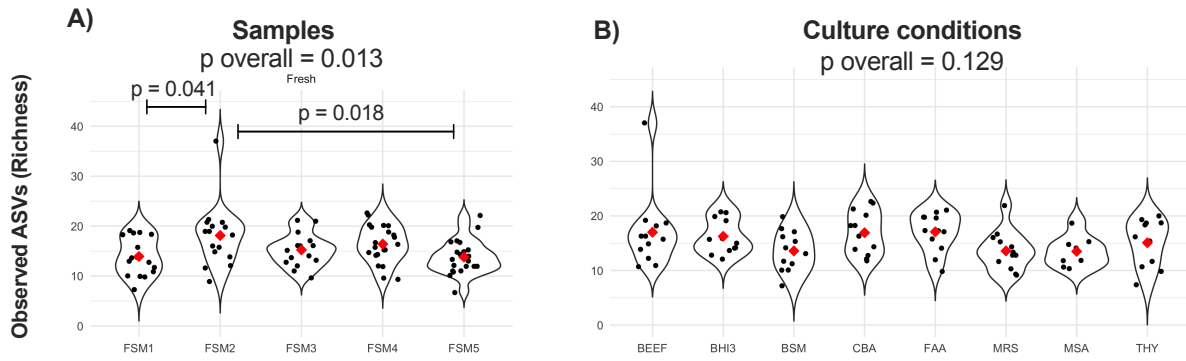


**Figure 18. Schematic overview of milk microbiota culture-enriched molecular profiling in 8 human milk samples from McMaster University.** Milk samples were cultured on 10 different media under both aerobic and anearobic conditions. Figure is adapted and modified from Lau et al. (63).

#### 4.3. Milk microbiota richness is impacted by short-term storage

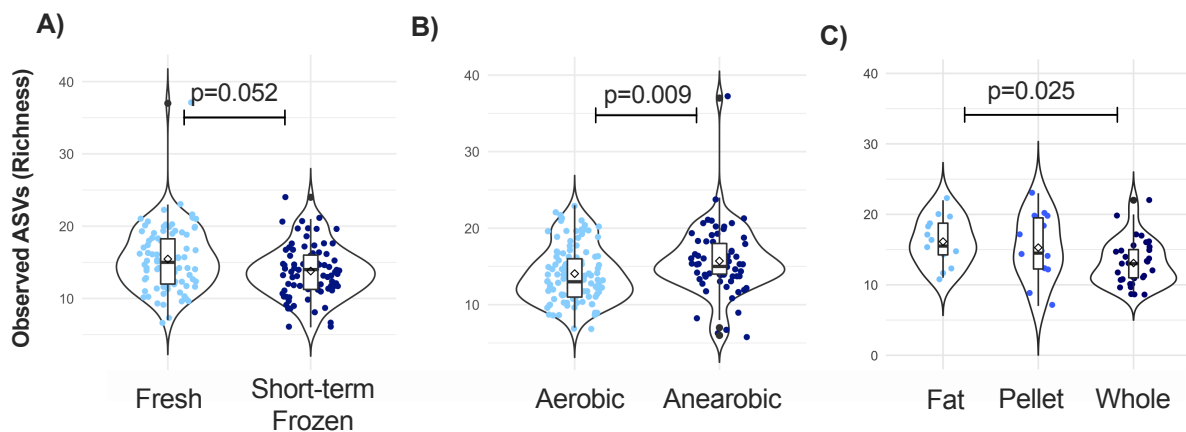
Overall,  $14.6 \pm 4.1$  unique ASVs were identified for each sample (mean $\pm$ SD). The bacterial richness was significantly different between fresh samples ( $p = 0.013$ ) driven mostly by the lower bacterial richness of milk sample #5 (Figure 19A). There was no difference in the number of detected ASVs between culture conditions for the fresh samples (Figure 19B).





**Figure 19. Comparison of the observed ASVs between A) Samples and B) Culture media in 5 human milk samples from McMaster University assessed by culture-enriched molecular profiling.**

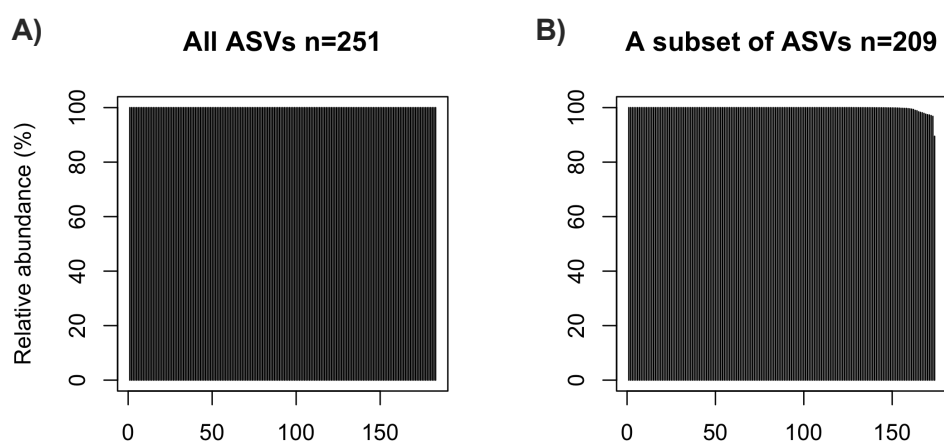
The number of detected ASVs was lower after short-term freezing ( $15.5 \pm 4.4$  vs.  $13.8 \pm 3.6$ ,  $p=0.007$ , Figure 20A) and was higher in anearobic vs. aerobic incubation ( $15.7 \pm 4.6$  vs.  $14.0 \pm 3.8$ ,  $p=0.009$ , Figure 20B). While there was no difference in the number of observed ASVs between the fat and pellet fractions, fat fraction had significantly higher bacterial richness compared to the whole milk (Figure 20C).



**Figure 20. Comparison of the number of detected amplicon sequence variants (ASVs) in 5 human milk samples from McMaster University using culture-enriched molecular profiling. A) Fresh vs. short-term frozen sample (N=5), paired t test analysis p value is reported, B) Aerobic vs. Anearobic incubation (N=5), and C) Different milk fractions (N=2).**

#### 4.4. Milk microbiota prevalence and relative abundance are influenced by short-term storage and milk fractions

Overall, 18 families constituted 219 out of 251 detected ASVs. These families constituted the majority of the sequencing reads (Figure 21). The most prevalent and abundant bacteria isolated from milk samples were *Staphylococcaceae* and *Streptococcaceae* (Table 8) in accordance with previous studies (36-45).



**Figure 21. Sequencing output processing to remove low abundance ASVs in 5 human milk samples from McMaster University.** The individual plate pools are on the X axis and relative abundances of the sums of sequence reads per sample are plotted along the Y axis. The initial ASV table following pre-processing contained 251 ASVs. Following exclusion of low prevalence ASVs, 209 ASVs remained. Analyses have been performed on the 209 ASVs, unless stated otherwise. N samples = 5, N plate pools =174.

The composition of milk bacteria was significantly influenced by short-term storage, aerobic vs. anearobic incubation, and milk fraction analysed compositions (Figure 22). *Pseudomonadaceae* and *Xanthomonadaceae* relative abundances were significantly decreased while *Staphylococcaceae* relative abundance was increased after short-term freezing at -80 °C (Figure 22). Anearobic bacteria belonging to *Bifidobacteriaceae*, Family XI, *Propionibacteriaceae*, and *Veillonellaceae* were more abundant in anearobic culture while aerobic bacteria belonging to *Corynebacteriaceae*, *Moraxellaceae*, *Rhizobiaceae*, and

*Xanthomonadaceae* were more abundant in aerobic culture (Figure 22); confirming that the expressed milk contains both aerobic and anaerobic bacteria.

**Table 8. Most frequently isolated bacterial families in the human milk microbiota in 5 human milk samples from McMaster University assessed by culture-enriched molecular profiling related to Figure 22.**

| Phylum         | Family               | Prevalence (%) <sup>a</sup> | Relative abundance (%) |      |         |
|----------------|----------------------|-----------------------------|------------------------|------|---------|
|                |                      |                             | Mean±SD                | CV   | Maximum |
| Firmicutes     | Staphylococcaceae    | 100                         | 29.8±37.9              | 1.3  | 99      |
| Firmicutes     | Family XI            | 100                         | 9.9±22.2               | 2.3  | 88      |
| Firmicutes     | Streptococcaceae     | 100                         | 15.1±31.0              | 2.1  | 98      |
| Proteobacteria | Xanthomonadaceae     | 100                         | 8.1±19.9               | 2.5  | 80      |
| Proteobacteria | Moraxellaceae        | 100                         | 9.0±22.2               | 2.5  | 99      |
| Actinobacteria | Corynebacteriaceae   | 100                         | 5.0±12.6               | 2.5  | 76      |
| Actinobacteria | Bifidobacteriaceae   | 100                         | 2.5±13.7               | 5.6  | 98      |
| Firmicutes     | Veillonellaceae      | 100                         | 2.3±7.5                | 3.2  | 45      |
| Actinobacteria | Propionibacteriaceae | 100                         | 4.9±15.3               | 3.1  | 96      |
| Firmicutes     | Enterococcaceae      | 80                          | 9.9±26.5               | 2.7  | 99      |
| Actinobacteria | Actinomycetaceae     | 60                          | 0.9±5.2                | 6.1  | 64      |
| Proteobacteria | Rhizobiaceae         | 80                          | 0.3±3.9                | 12.1 | 53      |
| Proteobacteria | Enterobacteriaceae   | 100                         | 0.5±2.4                | 4.7  | 20      |
| Firmicutes     | Lactobacillaceae     | 40                          | 0.6±4.3                | 6.9  | 44      |
| Bacteroidetes  | Bacteroidaceae       | 40                          | 0.5±7.4                | 13.4 | 100     |
| Proteobacteria | Pseudomonadaceae     | 20                          | 0.005±0.05             | 10.9 | 0.69    |
| Firmicutes     | Bacillaceae          | 0                           | 0.5±6.3                | 13.5 | 85      |
| Proteobacteria | Pasteurellaceae      | 20                          | 0.05±0.70              | 13.5 | 9.5     |

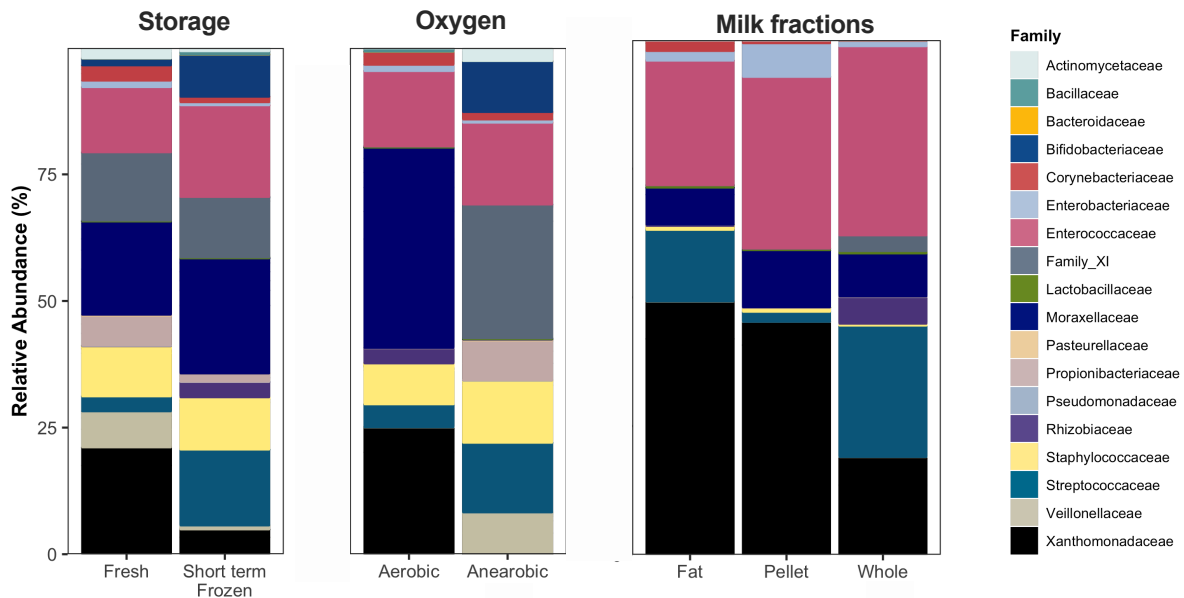
<sup>a</sup> Percentage of samples where the taxa was present.

<sup>b</sup> Only fresh samples are included.

CV, coefficient of variation

Almost all sequencing-based milk microbiome studies are conducted on the pellet following the removal of the milk fat through centrifugation (36-45). Acknowledging the small sample size (N=2), I observed that milk fractions are different in bacterial composition using culture-enriched molecular profiling (Figure 22). The overall compositions of the fat and pellet fractions were more similar compared to the whole milk. Relative abundances of the most prevalent and abundant isolated bacterial family i.e. *Streptococcaceae* (Table 8) was

higher in the whole milk compared to the pellet while *Xanthomonadaceae* was less abundant in the whole milk and *Staphylococcaceae* was not influenced by milk fractions (Figure 22). In contrast, *Corynebacteriaceae* and *Enterobacteriaceae* were more abundant in the pellet (Figure 22).

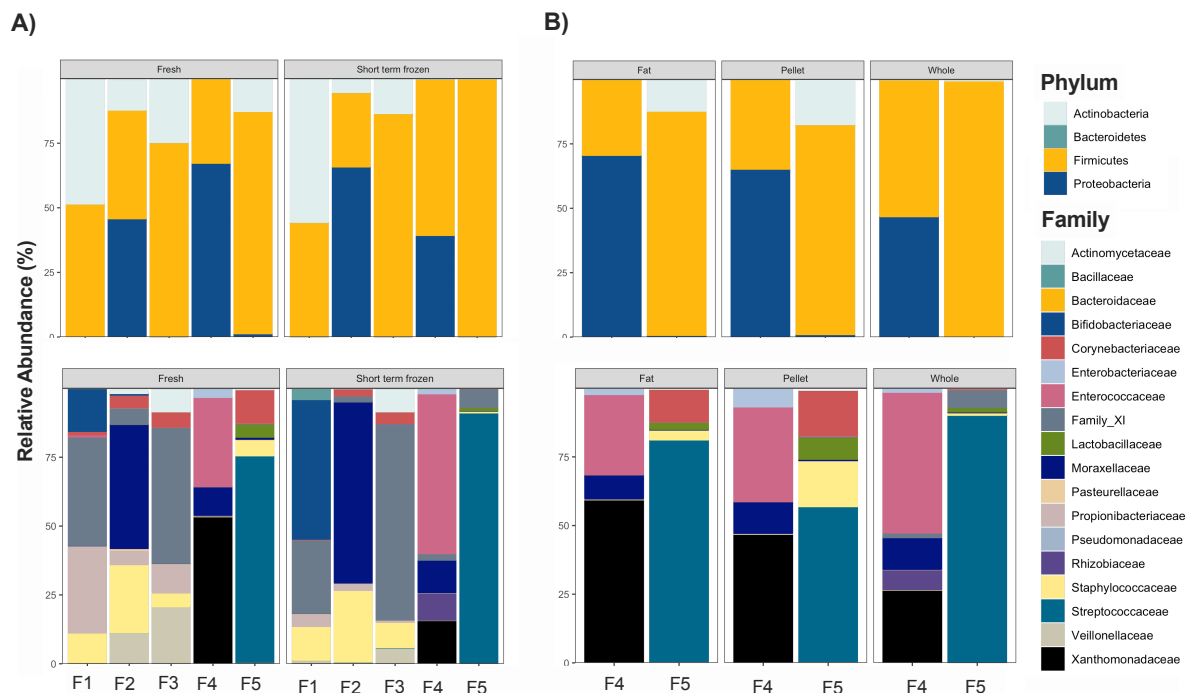


**Figure 22. Comparison of milk microbiota composition in 5 human milk samples from McMaster University using culture-enriched molecular profiling at the family level based on storage (Fresh vs. Short-term Frozen, N samples=5), Oxygen availability during incubation (Aerobic vs. Anaerobic, N samples=5), and in different milk fractions (N samples=2).**

#### 4.5. Milk microbiota composition exhibit inter-individual variability

Using culture-independent 16S rRNA gene sequencing, I observed inter-individual variability in the composition of the milk microbiota (section 3.1.2). Similar observation was made using culture-enriched molecular profiling. The inter-individual variability of bacterial composition of samples remained consistent across in fresh vs. short-term frozen (Figure 23A) and in comparison of different milk fractions (Figure 23B). However, notable differences were also observed for example expansion of *Bifidobacteriaceae* in sample 1

after short-term freezing or expansion of *Staphylococcaceae* in the pellet compared to the fat fraction or the whole milk.



**Figure 23. Milk microbiota composition at phylum and family levels in human milk samples from McMaster using culture-enriched molecular profiling in individual milk samples comparing A) Fresh vs. Short-term Frozen samples (N=5), and B) Between different milk fractions (N=2).**

#### 4.6. Conclusion

Although, targeted culture-based studies of milk have been previously reported (41, 71-75), to our knowledge, this is among the first studies on culture-enriched molecular profiling of the milk microbiota. Similar to the culture-independent results (Chapter 3), Firmicutes and Proteobacteria were the dominant bacterial phyla in human milk and several known milk bacteria including *Staphylococcus* and *Streptococcus* were commonly isolated. Milk microbiota richness was considerably lower in culture-enriched vs. culture-independent molecular profiling potentially highlighting the high prevalence of reagent contaminants or

non-viability of the milk bacteria detected in culture-independent studies. On the other hand, it is plausible that the choice of culture conditions was not sufficient to capture the true diversity of milk microbiota. Although milk microbiota was successfully profiled in long-term frozen samples from the CHILD study using the culture-independent approach (Chapter 3), culture-enriched molecular profiling of a limited number of long-term frozen milk samples did not yield any bacterial isolates. This observation might be due to very low bacterial load in those milk samples. Alternatively, it could be due to major differences between CHILD and McMaster milk samples including: 1) infant health status, 2) time of sample collection, and 3) history of direct breastfeeding at the breast vs. pump expression (Table 9).

**Table 9. Comparison of milk samples assessed in the CHILD Cohort Study to the milk samples collected at McMaster and assessed by culture-enriched molecular profiling**

| <b>Features</b>                 | <b>CHILD</b>            | <b>McMaster</b>           |
|---------------------------------|-------------------------|---------------------------|
| N                               | 393                     | 5                         |
| Infants                         | Term                    | Pre-term/admitted to NICU |
| Time postpartum                 | 3-4 months              | <1 month                  |
| Expression method               | Manual and pump         | Pump                      |
| Sample collection               | Multiple times over 24h | Once                      |
| Supplementation with fortifiers | No                      | Yes                       |

Importantly, short-term storage shifted the composition of milk microbiota from the original composition in fresh samples. Nevertheless, the inter-individual variability of milk microbiota composition was observed even after short-term storage. In addition, different bacteria were isolated from different milk fractions. We are aware of another ongoing study of milk microbiota culture-enriched molecular profiling (256) and we eagerly await the results.

Fortifiers are commonly added to the milk consumed by infants in NICU to meet various nutritional needs (257). Although the impact of fortifiers on the milk microbiota composition has not been studied, higher nutrient availability in the milk could conceivably affect bacterial metabolic activity. However, as the fortifier is added to the milk immediately before feeding to the infant, any potential effect on the milk microbiota will likely be minimised. Further studies are required to investigate the impact of fortifiers on the milk microbiota composition.

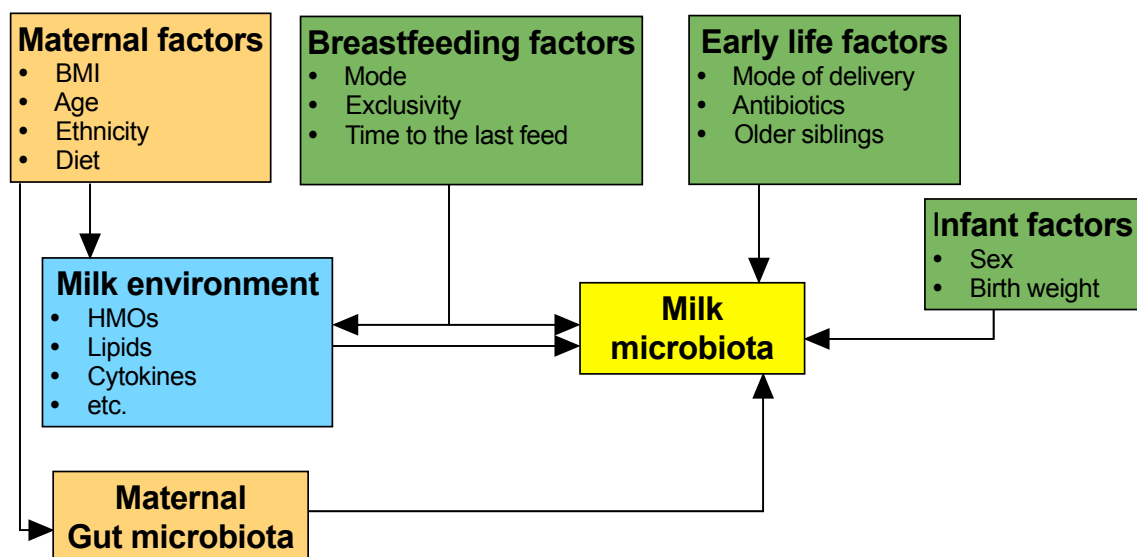
#### **Box 2. Summary of Chapter 4**

- We did not observe bacterial growth from long-term frozen milk samples stored at -80 °C for more than 6 months.
- Milk bacteria were isolated using culture-enriched molecular profiling.
- Both aerobic and anearobic bacteria were isolated from milk samples.
- Milk microbiota richness decreased following short-term storage at -80 °C.
- Bacteria relative abundance were impacted by short-term storage at -80 °C.
- Different milk fractions exhibited different bacterial composition.

## 5. DETERMINANT OF HUMAN MILK MICROBIOTA

### 5.1. Introduction

The focus of human microbiota studies is to identify host and/or environmental factors that influence the microbiota composition. Previous studies addressing the potential determinants of milk microbiota have been limited by small sample size, low sequencing depth, and lack of control for confounding factors. Milk microbiota composition was profiled in 393 mother-infant dyads (Chapter 3). Association of maternal, infant, and early life factors, breastfeeding practices, and other milk components with milk microbiota diversity, taxonomic structure, and overall composition is assessed in this chapter. The choice of variables examined was informed by our theoretical framework of the determinants of milk microbiota (Figure 24). This chapter is published in *Cell Host Microbe* (139).



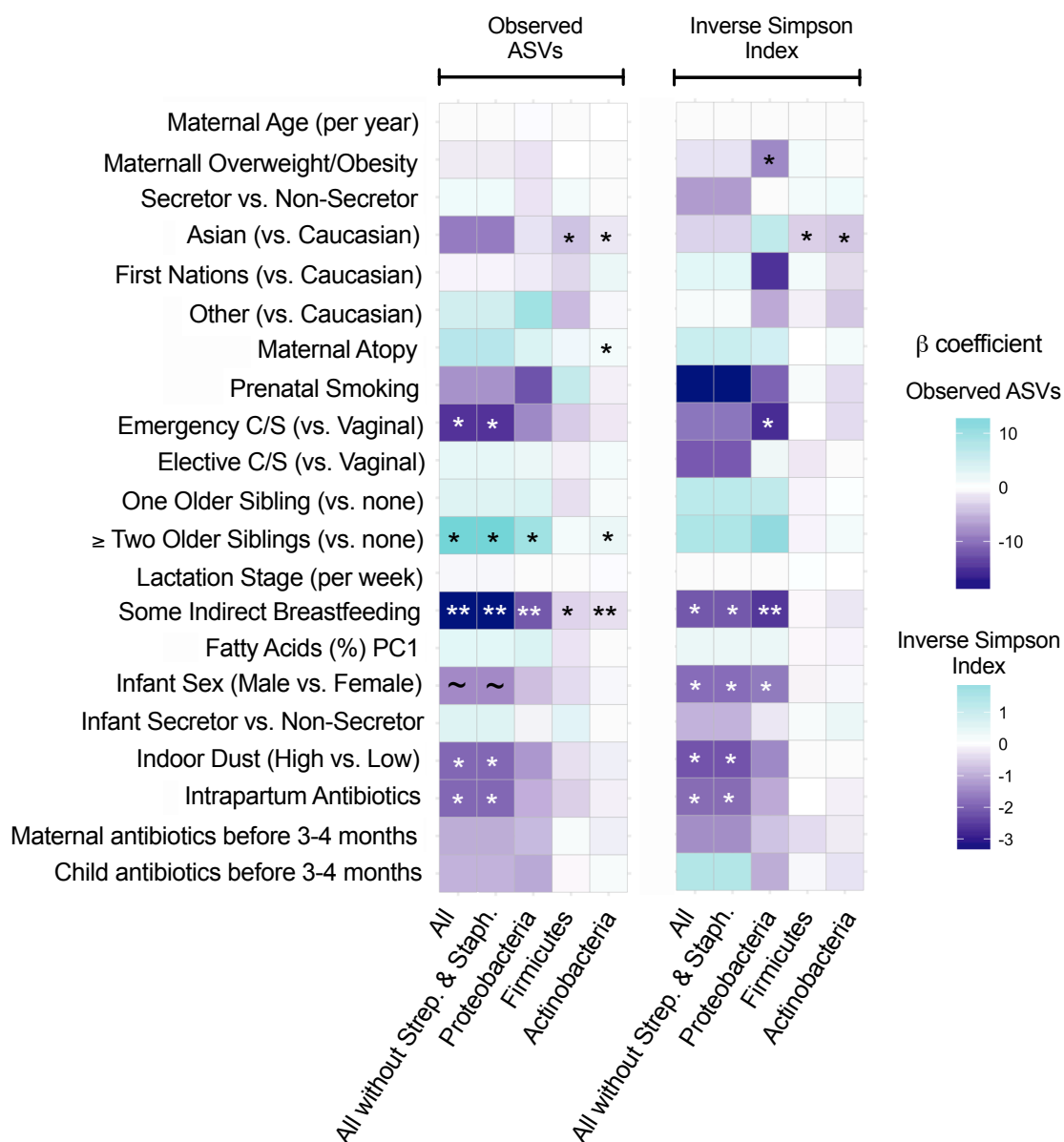
**Figure 24. Theoretical framework of how different factors are associated with milk microbiota composition.** Maternal factors are more likely to indirectly influence the milk microbiota via modulating other milk components or the maternal gut microbiota, while early life and infant factors could more directly shape the milk microbiota by influencing the infant oral microbiota. Reproduced from Moossavi et al. (139)



## **5.2. Milk microbiota diversity is associated with mode of breastfeeding, method of milk expression and other maternal factors**

Overall, milk samples had a mean $\pm$ SD richness (observed ASVs) of 147 $\pm$ 44 and diversity (inverse Simpson index) of 15.8 $\pm$ 8.7. We explored the association of alpha diversity with multiple maternal and infant factors (Figure 25 and Table 10).

Using univariate linear regression, indirect breastfeeding and male sex were associated with lower richness and diversity. In multivariable linear regression simultaneously adjusting for these factors plus mode of delivery, and parity (number of older siblings), indirect breastfeeding remained independently associated with lower milk bacterial richness (adjusted  $\beta$ =-18.9, 95%CI, -27.9, -9.9,  $p$ <0.001) and diversity (adjusted  $\beta$ =-2.08, 95%CI, -3.91, -0.25,  $p$ <0.05, Table 10). Milk expression method (manual vs. pump) was available for 111 individuals. In this subset, pump expression was associated with significantly lower bacterial richness ( $\beta$  =-39.6, 95%CI -60.5, -18.7,  $p$ <0.001) but was not associated with diversity ( $p$ >0.05). Within the range of lactation that we studied (17 $\pm$ 5 weeks postpartum, interquartile range 14-19 weeks), we did not detect significant associations of lactation stage with alpha diversity. In addition, we found no association between overall alpha diversity and mode of delivery, maternal ethnicity, history of atopy, smoking, BMI, secretor status, HMO diversity, or total HMO concentration.



**Figure 25. Factors associated with milk microbiota alpha diversity in 393 human milk samples from the CHILD Cohort Study.** Beta coefficients of univariate associations with linear regression are visualized for observed ASVs and Inverse Simpson Index for all microbiota, all excluding *Streptococcus* (Strep.) and *Staphylococcus* (Staph.), and major milk microbiota phyla. P value ~ < 0.1, \* < 0.05, \*\* < 0.001. Adapted from results original reported in Moossavi et al. (139).

**Table 10. Factors associated with milk microbiota alpha diversity in 393 human milk samples from the CHILD Cohort Study.**  
Reproduced from Moossavi et al. (139)

| Factor                       | n (%)      | Richness: Observed OTUs |                                           |                                          | Diversity: Inverse Simpson |                                           |                                          |                                          |
|------------------------------|------------|-------------------------|-------------------------------------------|------------------------------------------|----------------------------|-------------------------------------------|------------------------------------------|------------------------------------------|
|                              |            | Univariate <sup>a</sup> |                                           | Model 1 <sup>b</sup>                     | Univariate <sup>a</sup>    |                                           | Model 1 <sup>b</sup>                     | Model 2 <sup>c</sup>                     |
|                              |            | Richness<br>(Mean±SD)   | Crude $\beta$<br>(95% CI)                 | Adjusted $\beta$<br>(95% CI)             | Diversity<br>(Mean ± SD)   | Crude $\beta$<br>(95% CI)                 | Adjusted $\beta$<br>(95% CI)             | Adjusted $\beta$<br>(95% CI)             |
| Breastfeeding mode           |            |                         |                                           |                                          |                            |                                           |                                          |                                          |
| All Direct                   | 162 (41.9) | 157.8±37.0              | Ref.                                      | Ref.                                     | 17.1±8.2                   | Ref.                                      | Ref.                                     | Ref.                                     |
| Some Indirect                | 225 (58.1) | 138.7±46.9              | <b>- 19.1 (-27.8, -10.4) <sup>f</sup></b> | <b>-18.7 (-27.7, -9.64) <sup>f</sup></b> | 14.9±9.2                   | <b>- 2.2 (-3.9, -0.36) <sup>e</sup></b>   | <b>-2.04 (-3.88, -0.20) <sup>e</sup></b> | <b>-2.20 (-4.01, -0.40) <sup>e</sup></b> |
| Milk expression <sup>d</sup> |            |                         |                                           |                                          |                            |                                           |                                          |                                          |
| Manual                       | 20 (18.8)  | 170.1±40.0              | Ref.                                      | -                                        | 15.8±7.9                   | Ref.                                      | -                                        | -                                        |
| Pump                         | 91(81.2)   | 130.5±43.2              | <b>-39.6 (-60.5, -18.7) <sup>f</sup></b>  |                                          | 13.8±9.0                   | -2.08 (-6.40, 2.25)                       |                                          |                                          |
| Mode of delivery             |            |                         |                                           |                                          |                            |                                           |                                          |                                          |
| Vaginal                      | 294 (76.0) | 148.5±43.9              | Ref.                                      | Ref.                                     | 16.3±8.8                   | Ref.                                      | Ref.                                     | Ref.                                     |
| C/S emergency                | 47 (12.1)  | 133.5±47.3              | <b>-14.9 (-28.5, -1.5) <sup>e</sup></b>   | -11.3 (-25.0, 2.5)                       | 14.5±9.2                   | -1.74 (-4.46, 0.99)                       | -1.02 (-3.81, 1.78)                      | -1.29 (-4.01, 1.44)                      |
| C/S elective                 | 46 (11.9)  | 151.1±39.0              | 2.6 (-11.0, 16.3)                         | 2.5 (-11.2, 16.1)                        | 14.2±8.6                   | -2.09 (-4.84, 0.67)                       | -2.03 (-4.81, 0.76)                      | -1.89 (-4.65, 0.88)                      |
| Infant sex                   |            |                         |                                           |                                          |                            |                                           |                                          |                                          |
| Female                       | 192 (48.9) | 151.0±43.3              | Ref.                                      | Ref.                                     | 16.7±8.8                   | Ref.                                      | Ref.                                     | Ref.                                     |
| Male                         | 201 (51.1) | 142.8±43.9              | -8.2 (-16.9, 0.4)                         | -7.3 (-16.0, 1.4)                        | 14.9±8.9                   | <b>- 1.81 (-3.56, -0.06) <sup>e</sup></b> | -1.65 (-3.43, 0.12)                      | -1.62 (-3.39, 0.15)                      |
| Older siblings               |            |                         |                                           |                                          |                            |                                           |                                          |                                          |
| None                         | 211 (53.7) | 143.7±47.5              | Ref.                                      | Ref.                                     | 15.2±9.1                   | Ref.                                      | Ref.                                     | -                                        |
| One                          | 126 (32.1) | 147.1±39.7              | 3.4 (-6.3, 13.0)                          | -2.2 (-12.3, 7.8)                        | 16.4±8.6                   | 1.22 (-0.74, 3.18)                        | 0.68 (-1.37, 2.73)                       |                                          |
| Two or more                  | 56 (14.2)  | 157.7±35.8              | <b>13.9 (1.1, 26.8) <sup>e</sup></b>      | 7.5 (-6.0, 20.9)                         | 16.7±8.5                   | 1.47 (-1.15, 4.08)                        | 1.11 (-1.64, 3.85)                       |                                          |

N=393 dyads from the CHILD Cohort Study. P values <0.05 are in bold. Data is presented as mean±SD. BMI, body mass index; C/S, Caesarean section

<sup>a</sup> All factors showing an association (p<0.05) on univariate analysis are shown. Maternal BMI, lactation stage (interquartile range 14-19 weeks), maternal age, maternal atopy, maternal secretor status (defined based on the presence or absence of 2'-FL or LNFP HMOs), maternal prenatal smoking, maternal ethnicity, milk fatty acids PC1, HMO diversity, and total HMO concentration were not associated with milk microbiota alpha diversity.

<sup>b</sup> Linear regression adjusted for breastfeeding mode, infant sex, mode of delivery, older siblings, and sample processing time; N=381

<sup>c</sup> Linear regression adjusted for breastfeeding mode, infant sex, mode of delivery, and sample processing time; N=381

<sup>d</sup> n= 111. Method of milk expression was not systematically captured, but was analyzed for samples where it was noted.

<sup>e</sup> p<0.05

<sup>f</sup> p<0.001

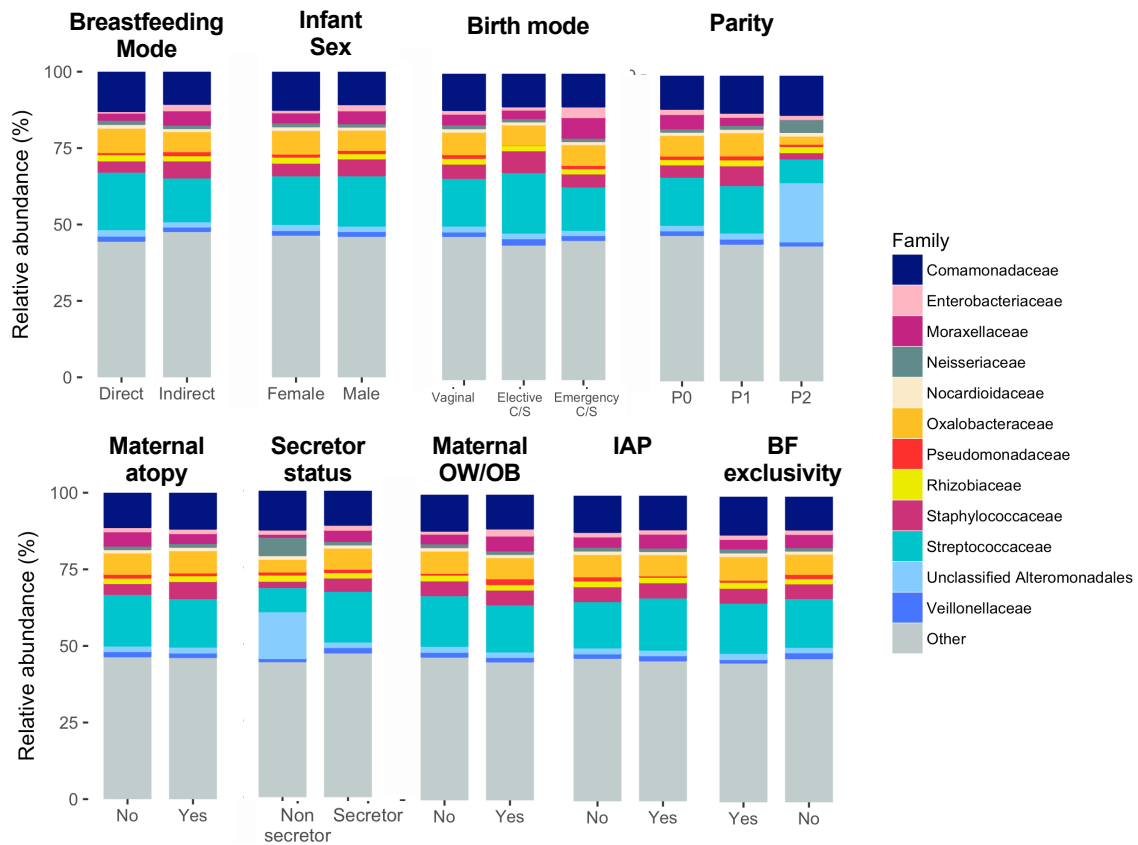
### **5.3. Associations of maternal and infant factors with milk microbiota alpha diversity are phylum-specific**

While some factors were associated with the overall milk microbiota diversity, we hypothesized the influence of mother and infant factors on the diversity would be different for different bacterial groups i.e. within-phylum alpha diversity could be differentially associated with different factors. Additionally, we considered that potential skin contamination could mask associations between these factors and milk microbiota alpha diversity. Subsequently, richness and diversity indices were measured for each major phylum in milk (Proteobacteria, Firmicutes, and Actinobacteria) and after the exclusion of *Streptococcus* and *Staphylococcus* as potentially skin-derived bacteria. There were some notable differences in the above associations when examining within-phylum alpha diversity (Figure 25). For example, emergency Caesarean section, infant sex, home dust levels, and intrapartum antibiotics were associated with lower richness of the entire community without any significant association with specific phyla. On the other hand, Asian ethnicity was associated with lower richness and diversity of Actinobacteria and Firmicutes but not the overall community. Additionally, different patterns emerged in the association of factors with richness vs. diversity. For example, while indirect breastfeeding was associated with lower richness in all three dominant phyla (Proteobacteria, Firmicutes and Actinobacteria), it was only associated with lower diversity within Proteobacteria. Additionally, while maternal overweight/obesity was not associated with overall alpha diversity, it was inversely associated with Proteobacteria diversity. Finally, maternal atopy was associated with higher Actinobacteria richness. Exclusion of *Streptococcus* and *Staphylococcus* did not impact the association of factors assessed with milk microbiota richness and diversity suggesting that despite being the dominant milk bacteria, their diversity is not the main driver of the milk microbiota diversity.

#### 5.4. Milk microbiota taxonomic composition is associated with mode of breastfeeding

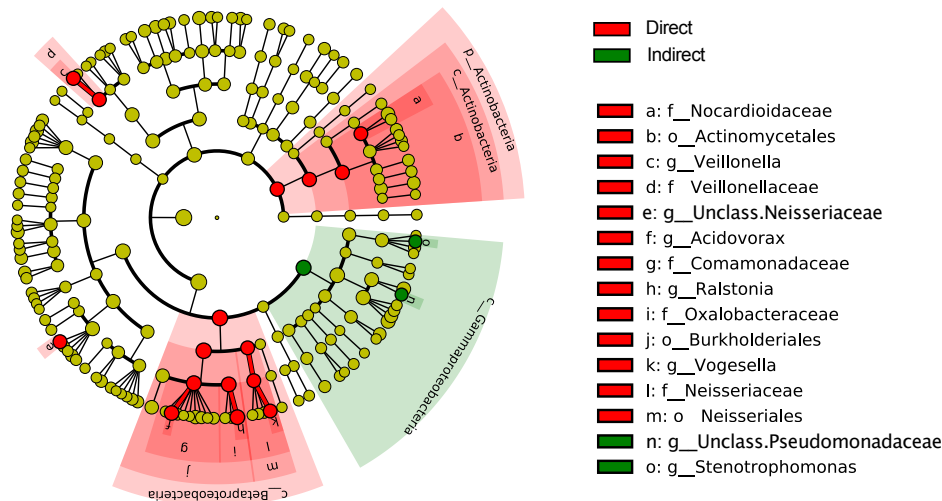
To identify maternal, infant, and early life factors associated with milk microbiota composition, we first examined the relative abundance of taxa at species and genus levels (ASVs > 0.01% mean relative abundance were included in this analysis). With the exception of mode of breastfeeding, we did not identify differential abundant taxa based on the other factors examined. Overall, only a few individual taxa were significantly differentially abundant based on breastfeeding mode using simple linear regression. *Enterobacteriaceae* and *Enterococcaceae* were enriched with indirect breastfeeding while *Gemellaceae*, *Vogesella*, and *Nocardioides* were enriched with direct breastfeeding. Although *Enterobacteriaceae* was present in 70% of samples regardless of the mode of breastfeeding, it had a 5-fold lower mean relative abundance in direct vs. indirect breastfeeding ( $1.1 \pm 3.5$  vs.  $5.0 \pm 13.5$ , respectively). The difference was even more pronounced within cluster C1 (where indirect breastfeeding was most prevalent) with a mean of  $1.2 \pm 2.1$  and  $20.2 \pm 22.9$  for direct and indirect breastfeeding, respectively. No associations were observed between individual taxa and other factors examined, including maternal secretor status, ethnicity, history of atopy, mode of delivery, number of older siblings, intrapartum antibiotics, and maternal or child antibiotics before the time of sample collection.

We next examined the most abundant taxa at the family level (ASVs with >1% mean relative abundance). With the exception of higher levels of unclassified Alteromonadales and lower levels of *Streptococcaceae* in milk of secretor or multiparous (having two or more older children) mothers ( $p > 0.05$ ), the milk microbiota taxonomic structure was not influenced by any of the factors assessed (Figure 26).



**Figure 26. Milk microbiota structure is not strongly associated with maternal, infant, and early life factors in 393 human milk samples from the CHILD Cohort Study.** Data is presented at the family level with mean relative abundance of at least 1%.

Next, we performed linear discriminant analysis (LDA) effect size (LEfSe) with default parameters (230) as an alternative method to identify taxa that were significantly enriched based on mode of breastfeeding (Figure 27). LEfSe identified several discriminant taxa that were not differentially abundant in conventional regression analysis. For example, the Actinobacteria phylum and *Veillonellaceae* (member of oral microbiota) were enriched with direct breastfeeding while *Stenotrophomonas* and *Pseudomonadaceae* (potential opportunistic pathogens) were enriched with indirect breastfeeding using discriminant analysis (Figure 27).



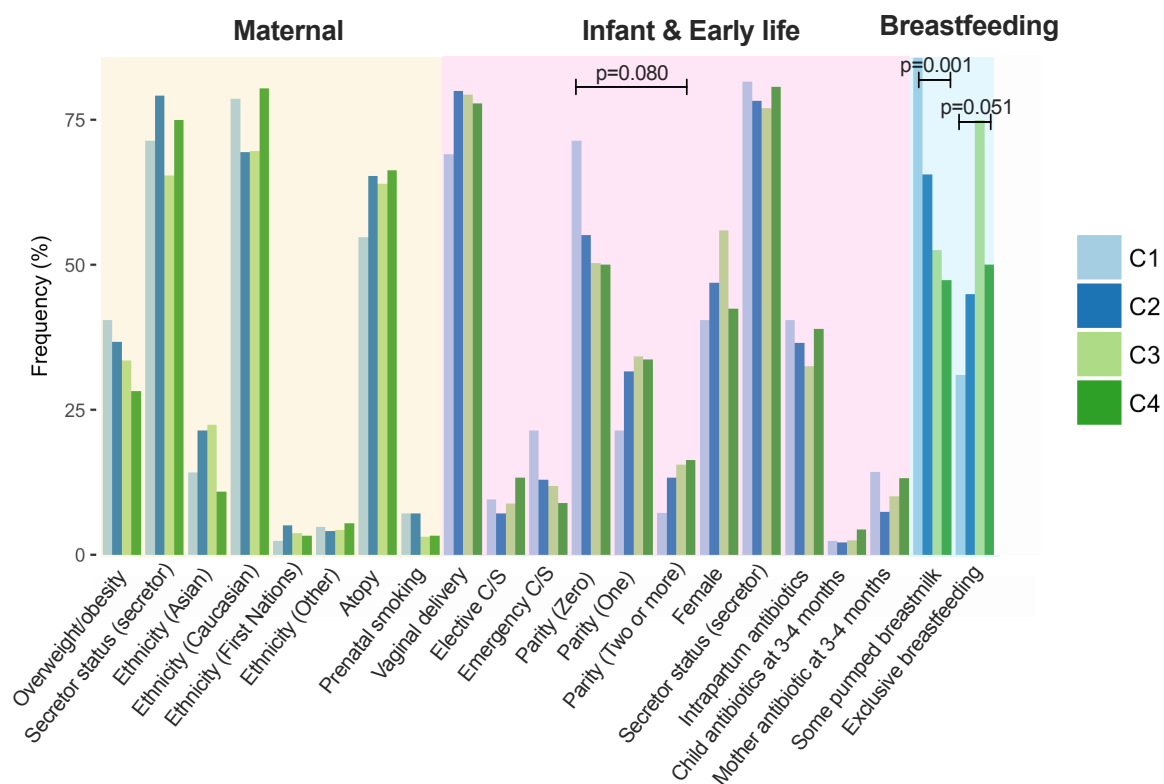
**Figure 27. Milk microbiota is associated with mode of breastfeeding by linear discriminant analysis of taxa enrichment in 393 human milk samples from the CHILD Cohort Study. Reproduced from Moossavi et al. (139).**

The current scientific consensus is that milk transfers beneficial maternal *Bifidobacterium* to the infant as *Bifidobacterium* spp. constitute the majority of shared taxa between mother's milk and infant stool (258). Although we did not detect a significant association of mode of breastfeeding with *Bifidobacterium* in the analysis of milk microbiota; we focused on this genus to examine this association further. We identified *Bifidobacterium bifidum*, and *B. animalis*, as well as two unclassified *Bifidobacterium* spp. in the milk microbiota. An unclassified *Bifidobacterium* was the most prevalent *Bifidobacterium* spp. in breastmilk, being present in 36% of samples and the prevalence was different based on the mode of breastfeeding (48% in direct vs. 30% in indirect breastfeeding,  $p < 0.001$ ) with an average relative abundance of  $0.48 \pm 0.50$  in direct and  $0.29 \pm 0.46$  in indirect.

### 5.5. Milk microbiota clusters are associated with mode of breastfeeding, method of milk expression and other maternal factors

Next, we assessed the association of maternal and infant factors with taxonomic clusters that we have identified in milk (Figure 13). Gradually decreasing proportion of mothers

within clusters C1-C4 reported feeding their infants pumped milk (indirect breastfeeding) (C1: 86%, C2: 66%, C3: 53%, and C4: 48%; Figure 28). In addition, within the subset of samples with available data on the method of milk sample collection, C1 also had the highest rate of pump vs. manual expression (94%) compared to C2 (81%), C3 (78%), and C4 (83%). Indoor dust levels were higher in C2 and C4 while there was a trend in the association of breastfeeding exclusivity (formula supplementation) and parity (number of older siblings) with milk taxonomic cluster membership. We did not detect an association of cluster membership with other maternal and infant factors (Figure 28).



**Figure 28. Univariate associations of maternal and infant factors with milk microbiota composition and cluster membership in 393 human milk samples from the CHILD Cohort Study.** Adapted from results originally reported in Moossavi et al. (139).

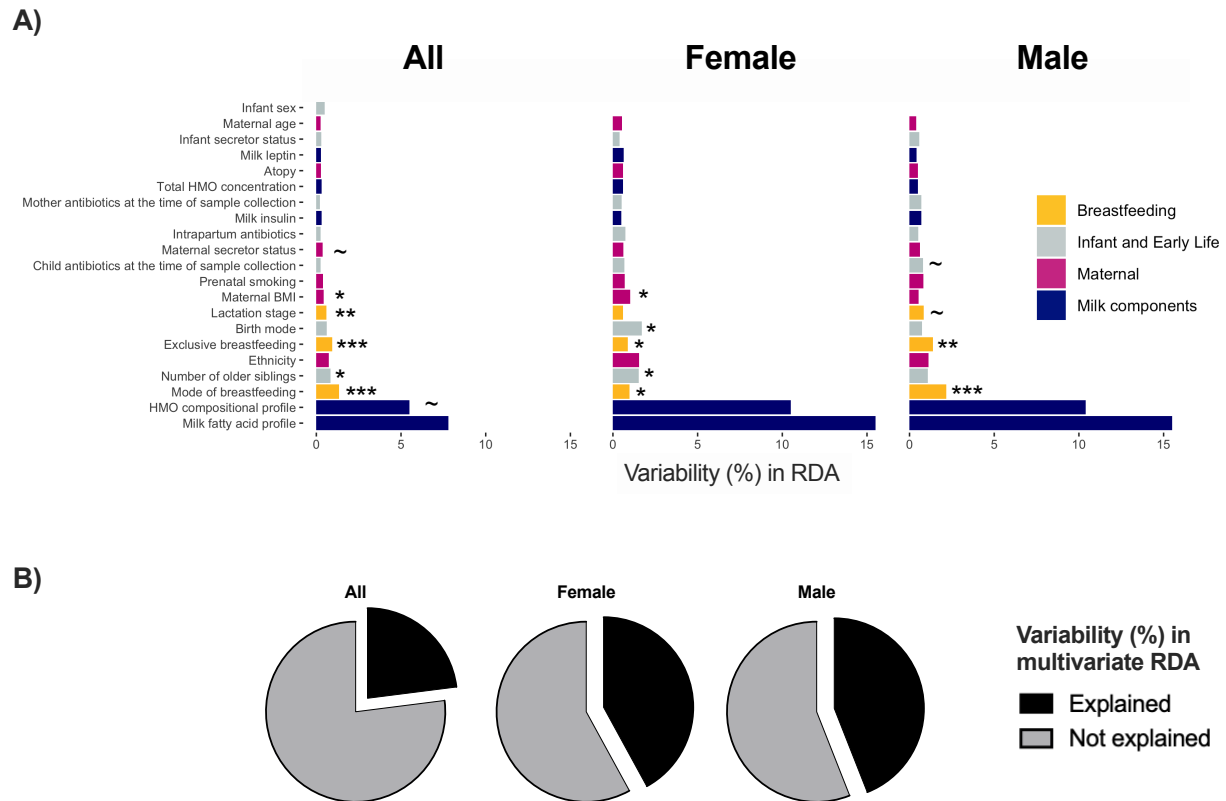


## **5.6. Overall milk microbiota composition is associated with breastfeeding practices, multiparity, and other maternal factors in a sex-specific manner**

Ecologically, it is more conceivable that factors affect the entire microbial community although the main impact might be targeted to a subset of taxa. In this case, changes in absolute abundances could be masked by an overall shift in the community composition. The overall community composition could be assessed at the community composition level by RDA. We explored the association of each factor with the entire microbiota community composition using RDA (Figure 29). Consistent with the cluster analyses, breastfeeding mode ( $p=0.001$ ) was significantly associated with overall milk microbiota composition. Several additional factors including exclusive breastfeeding, lactation stage, parity, maternal BMI, ethnicity, and infant sex were significantly associated with the milk microbiota community albeit with very low redundancy values (each accounting for  $<1\%$  of the variation in the milk microbiota). There was a trend toward association of HMO composition with milk microbiota with much larger effect size compared to the other factors assessed ( $R^2=5.5$ ,  $p=0.074$ ) (Figure 29A).

Sex-specificity was observed for univariate RDA associations for mode of breastfeeding and lactation stage (higher  $R^2$  in males), as well as parity, mode of delivery, and maternal BMI (significant in females only) in stratified analyses (Figure 29A).

The cumulative association of all the factors assessed in our study in a multivariable analysis explained approximately 25% of the total variation in the milk microbiota composition. In addition, the combined contribution of all the above-mentioned factors accounted for larger proportion of variation in sex-stratified compared to non-stratified analysis, primarily driven by a higher contribution from HMOs and milk fatty acids after stratification (Figure 29B).



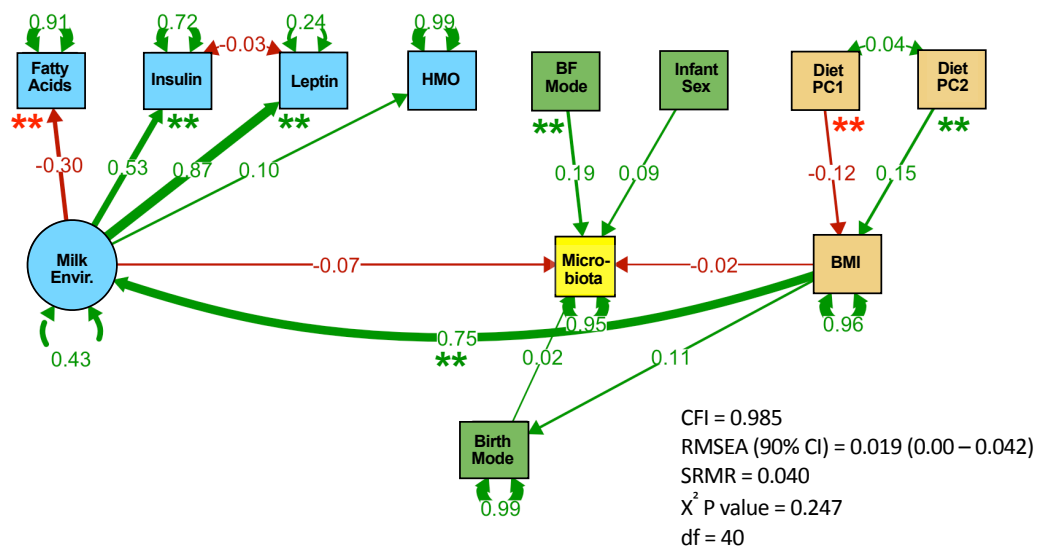
**Figure 29. Redundancy analysis of associations of maternal, infant, early life, and milk factors with the overall milk microbiota composition in 393 human milk samples from the CHILD Cohort Study.** A) Univariate analysis for all infants and stratified by infant sex. B) Cumulative amount explained by all the factors in multivariate analysis. Redundancy values ( $R^2$ ) indicate the % variation explained by each individual factor (in univariate analyses). The unexplained variability was calculated by subtracting redundancy value from 100%. BMI, body mass index; HMO, human milk oligosaccharide. P value ~ < 0.10, \* < 0.05, \*\* < 0.01, \*\*\* < 0.001. Adapted from results originally reported in Moossavi et al. (139). Reproduced from Moossavi and Azad (131).

### 5.7. Structural equation modeling identifies mode of breastfeeding as a key determinant of milk microbiota, and defines other causal pathways influencing milk composition

None of the results described so far take into account the theoretical framework of causal pathways by which different factors could potentially affect the milk microbiota. For

example, maternal factors are more likely to indirectly influence the milk microbiota via modulating other milk components or the maternal gut microbiota, while early life and infant factors could more directly shape the milk microbiota by influencing the infant oral microbiota (Figure 24). In order to examine the determinants of the milk microbiota within our theoretical framework, we performed confirmatory factor analysis (CFA) (229) which is a variant of structural equation modelling (SEM) to evaluate and compare different plausible frameworks. We hypothesized that milk components such as HMOs, hormones, and fatty acids would be potential indicators of the overall milk microenvironment, which could determine the composition and structure of the milk microbial community (152, 160). We modelled the milk environment as a latent (unobserved) construct with HMOs, fatty acids, insulin, and leptin as indicators. Our variable selection in the CFA was guided mainly by the current literature with the focus on proposed mechanisms of mammary gland colonization. Variables potentially important for the retrograde pathway are mode of delivery, infant sex, and older siblings (factors shaping the infant microbiota) as well as mode of breastfeeding (because indirect feeding eliminates contact with the infant oral cavity). Variables potentially important for the entero-mammary route include maternal BMI and dietary pattern, which could be correlated with the maternal gut microbiota.

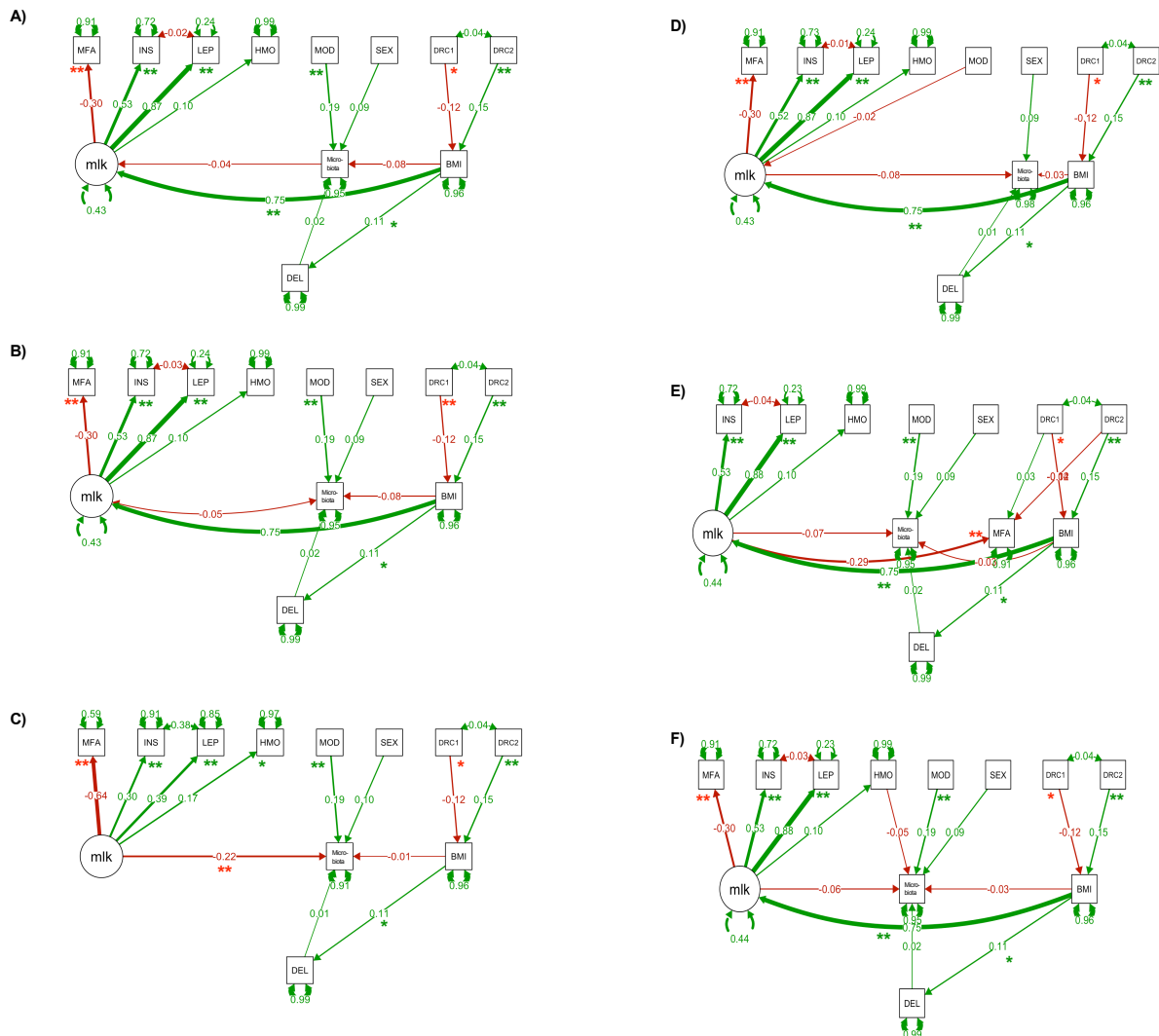
In our final model (Figure 30), based on our theoretical framework (Figure 24), breastfeeding mode was directly associated with milk microbiota with a standardized  $\beta$  coefficient of 0.19 ( $p < 0.001$ ). We also observed a trend in the direct association of infant sex with milk microbiota ( $\beta = 0.091$ ,  $p = 0.072$ ), whereas no association was detected for birth mode or maternal BMI ( $\beta < 0.10$ ,  $p$  values  $> 0.15$ ). Maternal diet influenced maternal BMI, and BMI was significantly associated with the milk environment latent construct. Overall, we did not detect a significant association between the milk environment and the milk microbiota ( $\beta = -0.049$ ,  $p = 0.45$ ) (Figure 30 and Figure 31A and B).



**Figure 30. Causal modelling of maternal, infant, and milk factors influencing the milk microbiota composition in 393 human milk samples from the CHILD Cohort Study.** Mode of breastfeeding is significantly associated with the milk microbiota in structural equation modeling. Standardized  $\beta$  coefficients are reported. BMI, maternal body mass index; CFI, comparative fix index; CI, confidence interval; RSMEA, root mean square error of approximation; SRMR, standardized root mean residuals; \*  $p < 0.05$ , \*\*  $p < 0.01$ . Green (positive), red (negative). Reproduced from Moossavi et al. (139).

We assessed other plausible causal pathways in equivalent models (Figure 31 and Table 11). For example, we removed the indirect effect of BMI on milk microbiota mediated by the milk environment (Figure 31C); however, the model was not a good fit, suggesting that BMI both directly and indirectly affects the milk microbiota. Furthermore, we observed that the effect of mode of breastfeeding on the milk microbiota was not mediated via the milk environment (Figure 31D). We also incorporated some known associations such as the direct effect of diet on milk fatty acids (Figure 31E), and potential direct effect of HMOs on milk microbiota (Figure 31F). Controlling for these additional associations, mode of breastfeeding was the only consistent factor directly associated with the milk microbiota composition. Overall, the CFA analysis suggests that 1) mode of breastfeeding directly influences the milk

microbiota, and 2) maternal diet influences BMI, which affects non-bacterial milk components that do not directly influence milk microbiota.



**Figure 31. Equivalent structural equation models assessing other potentially plausible models of associations of milk microbiota determinants in 393 milk samples from the CHILD Cohort Study. Reproduced from Moossavi et al. (139).**

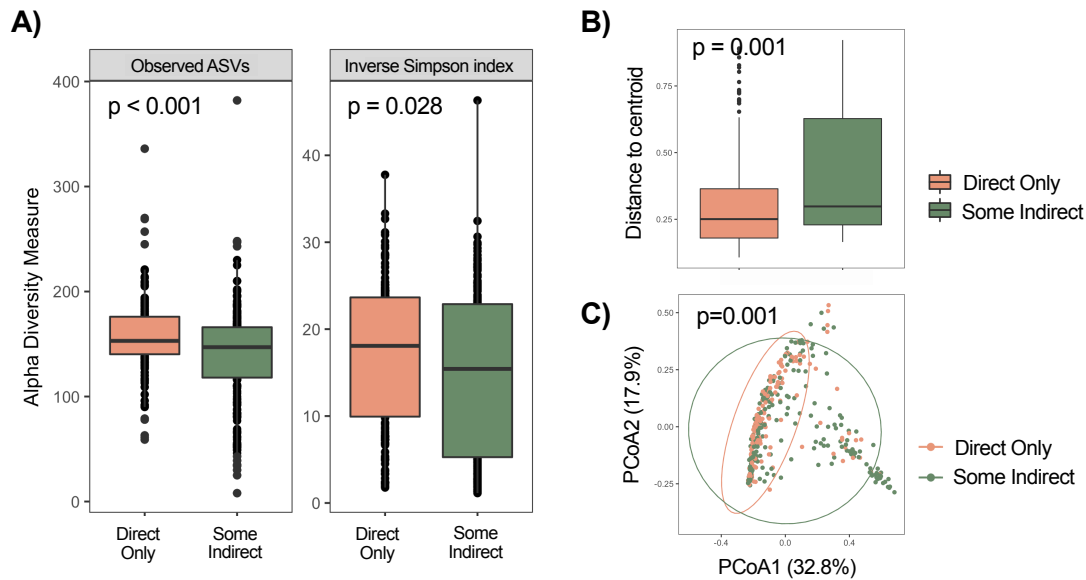
**Table 11. Comparison of the equivalent causal models of association of maternal and infant factors influencing the milk microbiota to the final model (see Figure 31). Reproduced from Moossavi et al. (139)**

| Model       | Model specification                                     | Df | $\chi^2$ p value  | CFI          | RMSEA (90% CI)               | SRMR         | AIC   |
|-------------|---------------------------------------------------------|----|-------------------|--------------|------------------------------|--------------|-------|
| Final model | Causative effect of milk on microbiota                  | 40 | 0.247             | 0.985        | 0.019 (0.00 – 0.042)         | 0.040        | 11401 |
| A           | Causative effect of microbiota on milk                  | 40 | 0.251             | 0.985        | 0.019 (0.00 – 0.042)         | 0.040        | 11401 |
| B           | Milk environment and microbiota correlation             | 40 | 0.247             | 0.985        | 0.019 (0.00 – 0.042)         | 0.040        | 11401 |
| C           | Effect of BMI on milk removed                           | 41 | <b>&lt; 0.001</b> | <b>0.404</b> | <b>0.119 (0.106 – 0.133)</b> | <b>0.108</b> | 11614 |
| D           | Effect of mode of breastfeeding on the milk environment | 40 | <b>0.028</b>      | 0.949        | 0.035 (0.012 – 0.054)        | 0.046        | 11414 |
| E           | Effect of maternal diet on fatty acids                  | 38 | 0.210             | 0.982        | 0.022 (0.00 – 0.044)         | 0.040        | 11404 |
| F           | Direct effect of HMO on microbiota                      | 39 | 0.241             | 0.984        | 0.020 (0.00 – 0.043)         | 0.040        | 11402 |

AIC, Akaike information criterion; BMI, body mass index; CFI, comparative fix index; CI, confidence interval; HMO, human milk oligosaccharide; RSMEA, root mean square error of approximation; SRMR, standardized root mean residuals. Indications of poor model fit are in **bold**:  $\chi^2$  p-value <0.05, CFI <0.9, RMSEA > 0.05, and SRMR > 0.08. Reproduced with permission from (139).

## 5.8. Indirect breastfeeding is associated with enrichment of potential pathogens and depletion of *Bifidobacteria* in milk microbiota

Given the consistent association of breastfeeding mode with different milk microbiota measures using different analytical approaches, we conducted further analyses to explore this relationship. Indirect breastfeeding was associated with lower  $\alpha$  (within sample) diversity and higher  $\beta$  (between sample) diversity (Figure 32A-C). We next examined association of mode of breastfeeding with *Bifidobacterium* spp. as they constitute the majority of shared taxa between mother's milk and infant stool (258). We identified *Bifidobacterium bifidum*, and *B. animalis*, as well as two unclassified *Bifidobacterium* spp. in the milk microbiota. An unclassified *Bifidobacterium* was the most prevalent *Bifidobacterium* spp. in breastmilk, being present in 36% of samples (48% in direct vs. 30% in indirect breastfeeding,  $p < 0.001$ ) with an average relative abundance of  $0.48 \pm 0.50\%$  in direct and  $0.29 \pm 0.46\%$  in indirect.



**Figure 32. Milk microbiota is associated with mode of breastfeeding.** A) Alpha diversity, B) Homogeneity of dispersion, and C) Beta diversity. Reproduced from Moossavi et al. (139).

## 5.9. Conclusion

Using multiple analytic approaches to profile the human milk microbiota in a general population cohort we provide evidence that its composition and diversity are influenced by maternal factors, early life events, breastfeeding practices, and other milk components. To our knowledge, this is the largest study of human milk microbiota performed to date, and the only study to apply a causal modeling approach in this context. Dominant milk bacteria in our cohort were in accordance with previously published results (97). Milk microbiota demonstrate a high degree of inter-individual variability with only a minority of taxa shared across the study population. Mode of breastfeeding was identified to be significantly associated with breastmilk microbiota composition while controlling for relationship of other confounding factors. These results considerably expand the existing knowledge on the factors influencing milk microbiota composition, providing new evidence for the importance of breastfeeding practices. Association of some factors such as maternal BMI and parity were

sex-specific. Although, some of the milk components such as total HMO were associated with the composition of milk microbiota, we did not find strong evidence of the association of the milk environment with the milk microbiota composition in a causal modeling. It is of note that while many factors were not associated with the overall composition, they were associated with within-phylum diversity (e.g. prenatal smoking and maternal atopy) suggestive of differential impact of lifestyle factors on different components of the milk microbiota.

### **Box 3. Summary of Chapter 5**

- Milk bacteria diversity is associated with mode of breastfeeding, method of milk expression and other maternal factors
- Some factors have phylum-specific effects
- Milk microbiota composition is more strongly influenced by factors affecting the exogenously-derived bacteria
- Some variations in milk microbiota are sex-specific
- Feeding method (at the breast vs. pumped) was strongly associated with milk microbiota composition



## **6. INTEGRATED ANALYSIS OF MILK MICROBIOTA WITH HUMAN MILK OLIGOSACCHARIDES AND FATTY ACIDS**

### **6.1. Introduction**

Multiple milk components including nutrients, bioactive molecules, and cellular components shape the milk microenvironment i.e. bacterial niche, which could have selective pressure on bacterial persistence and proliferation within the milk and hence influence the milk microbiota composition. In Chapter 3, we examined the association of HMOs and fatty acids with milk microbiota using summary statistics (PC1). Here, we performed a detailed analysis of the association of individual HMOs and fatty acids with the milk microbiota composition. This chapter is published in *Frontiers in Nutrition* (214).

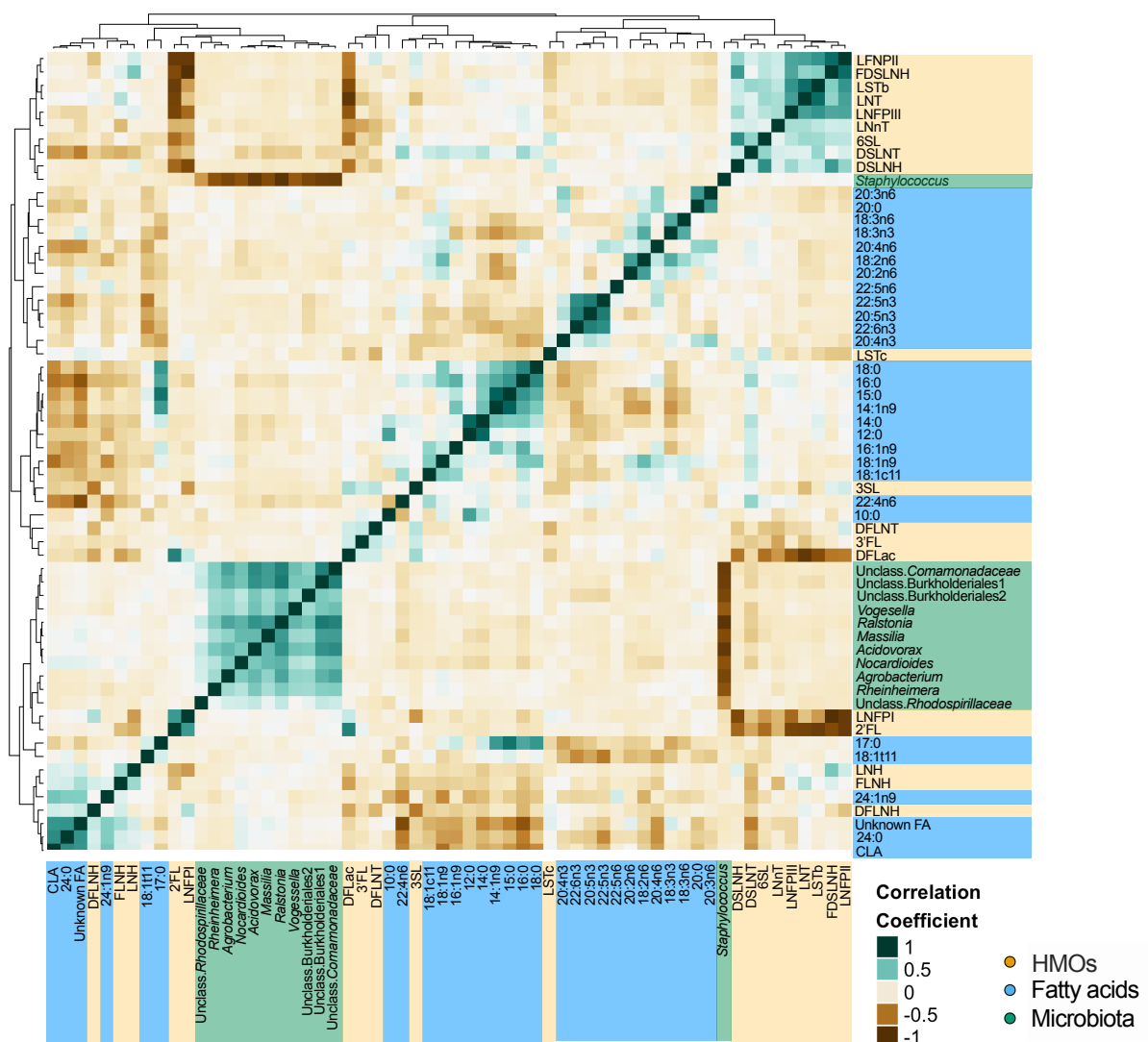
### **6.2. Overview of milk microbiota, fatty acids, and HMOs profiles in human milk in the CHILD Cohort Study**

Milk microbiota (Chapter 3), fatty acids, and HMO profiles in the CHILD Cohort Study have been previously described (139, 192, 194). Overall, *Streptococcus*, *Ralstonia*, and *Acidovorax* were the most abundant genera (Figure 33A) and Proteobacteria (mean 67%) and Firmicutes (mean 26%) were the most abundant phyla (Figure 33B). *Bifidobacteria* was present in 39% of samples at a mean relative abundance of  $0.25 \pm 0.98\%$ . 2'FL (among secretor mothers), LNFP II, and LNT were the most abundant HMOs (Figure 33C) and fucosylated/non-sialylated was the most abundant HMO group (Figure 33D). Fatty acids oleic acid (18:1n9), palmitic acid (16:0), and linoleic acid (18:2n6) were the most abundant while the majority had very low relative abundances (<1%; Figure 33E). SFAs (41%) and MUFAs (39%) were the most abundant fatty acids followed by PUFAs (17%; Figure 33F).



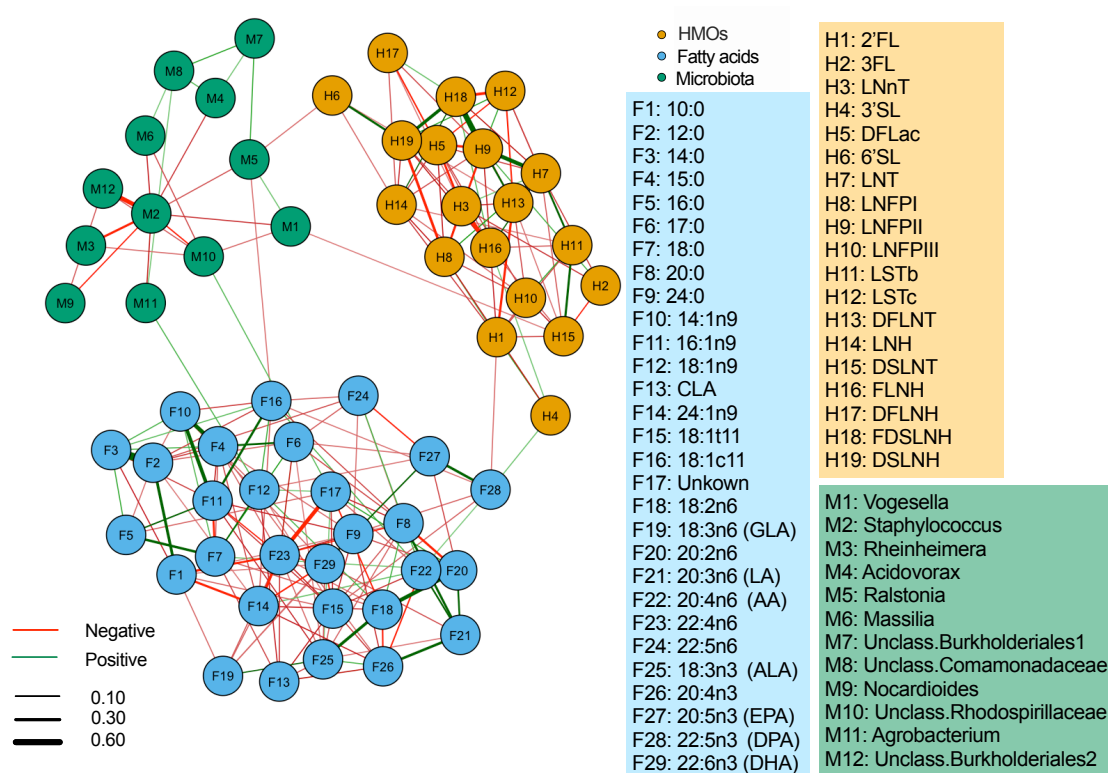
### 6.3. Correlations are stronger and more frequent within than between different types of milk components

Integrated correlations of different components were visualised in a heatmap. The strongest correlations were observed within each group (Figure 34). For example, *Staphylococcus* was strongly negatively correlated with nearly every other core taxa ( $r$  values  $< -0.60$ ,  $p < 0.05$ ), while strong positive correlations were found between LNFP II/FDSLNIH, C12:0/C14:0, 22:4n6/24:1n9 (all  $r > 0.4$ ,  $p < 0.05$ ).



**Figure 34.** Inter and intra-correlations of milk microbiota, human milk oligosaccharides (HMOs), and fatty acids by clustering heatmap based on Spearman rank correlation in 393 human milk samples from the CHILD Cohort Study. Reproduced from Moossavi et al. (214).

Next, to isolate and visualize the most significant associations, we assessed pairwise partial correlations between members of milk microbiota, HMOs, and fatty acids using Spearman correlation. Correlations with p value of less than 0.05 after Benjamini-Hochberg's correction were visualised as a network (Figure 35). Consistent with the conventional heatmap, the network shows that the number and strength of inter-group correlations were higher than intra-group correlations, which could be partly attributed to the compositional nature of all three components. Partial correlations between core microbiota and fatty acids included *Agrobacterium*/18:1n9 (positive) and *Ralstonia*/18:1c11 (negative). Correlations between microbiota and HMOs included *Vogesella*/DSLNT (negative) and *Ralstonia*/6'SL (negative) (Figure 35).

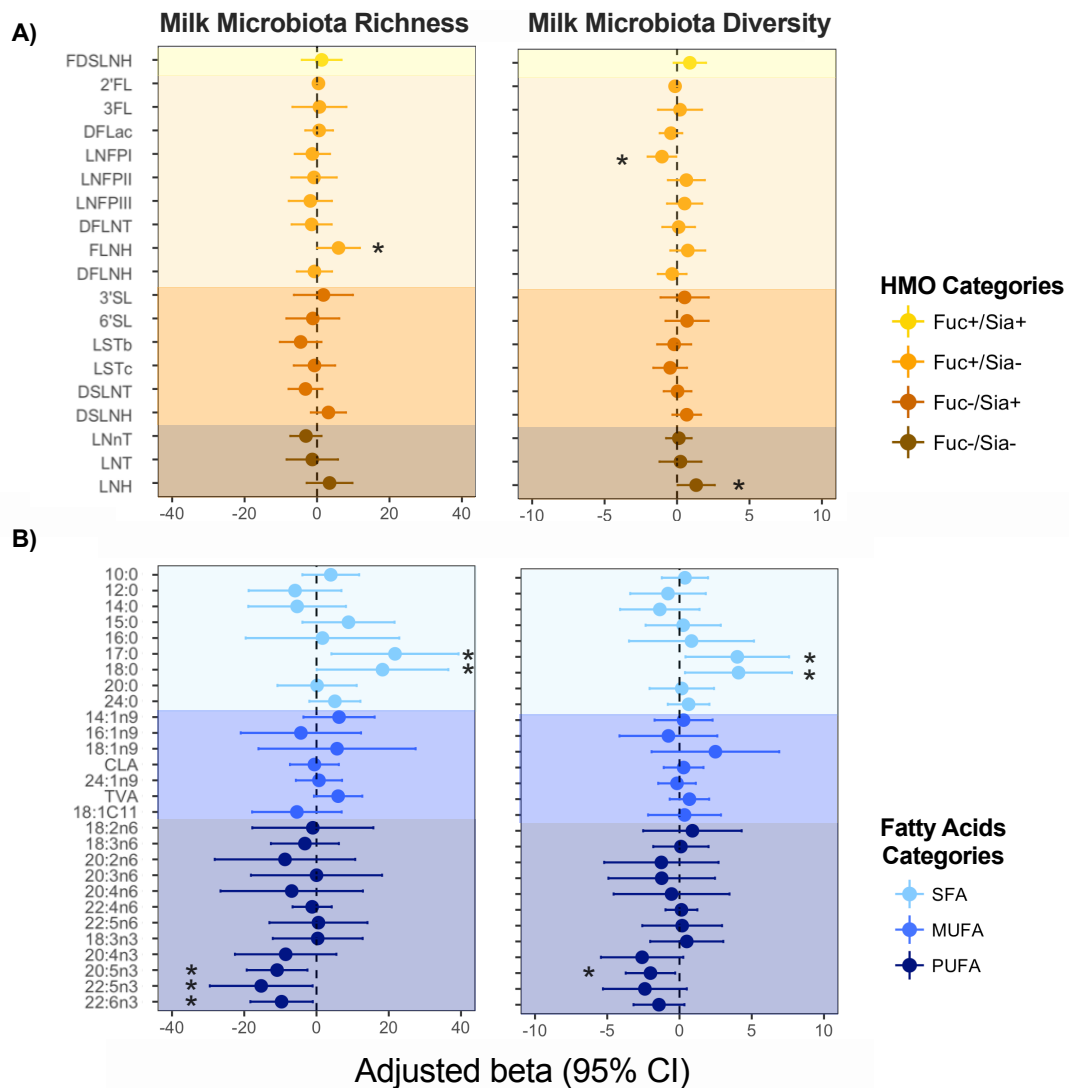


**Figure 35. Inter and intra-correlations of milk microbiota, human milk oligosaccharides (HMOs), and fatty acids based on network visualisation based on partial Spearman rank correlations in 393 human milk samples from the CHILD Cohort Study.** Line colour represents the direction of correlation: green (positive), red (negative). Line width is proportional to the correlation coefficient. Reproduced from Moossavi et al. (214).

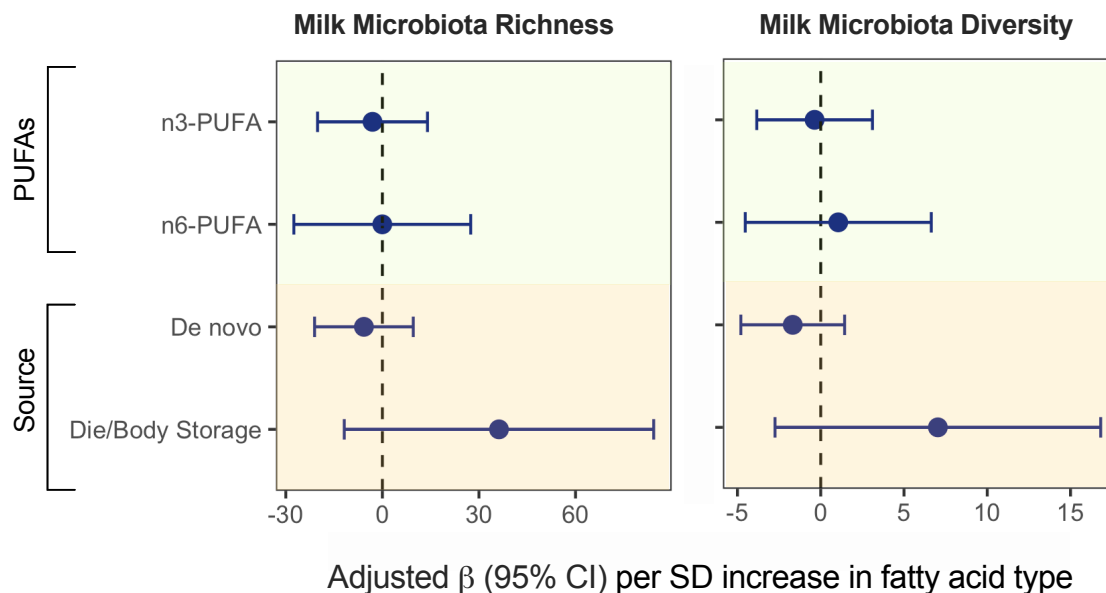
#### **6.4. Some HMOs and fatty acids are associated with milk microbiota alpha diversity and overall composition**

We previously identified mode of breastfeeding, infant sex, parity, and birth mode to be associated with microbiota richness and diversity, but found no association between milk microbiota diversity and HMO or fatty acids first principal component axes (PC1) (Chapter 4; (139)). However, when examining individual HMOs and fatty acids (expressed as relative proportions) in the current analysis, some associations were observed. Controlling for the above-mentioned covariates, FLNH was associated with microbiota richness ( $\beta=5.90$ , 95%CI (0.04 to 11.8),  $p=0.04$ ) while LNH ( $\beta=1.32$ , 95%CI (0.04 to 2.60),  $p=0.04$ ) and LNFP I ( $\beta=-1.04$ , 95%CI (-2.03 to -0.05),  $p=0.03$ ) were associated with microbiota diversity (Figure 36A). These opposite effects of different HMOs could explain the apparent lack of association of the overall HMO profile and microbiota diversity. Other HMOs were not associated with milk microbiota richness and diversity, and we did not detect any significant patterns of association when grouping HMOs according to fucosylation and sialylation (Figure 36A) or other common structural elements.

Fatty acids DHA, DPA, EPA (all n3-PUFAs) were negatively associated with milk microbiota richness (all  $p<0.05$ ), while SFAs margaric acid and stearic acid were positively associated with these metrics (Figure 36B). We did not detect any associations for MUFAs (Figure 36B), nor for fatty acids grouped based on their source/origin (i.e. synthesized de novo in the mammary gland, or acquired from the diet, liver synthesis, or body fat storage) (Figure 37).



**Figure 36. Association of A) human milk oligosaccharides (HMOs) and B) fatty acids with milk microbiota richness (observed ASVs) and diversity (inverse Simpson index) in 393 human milk samples from the CHILD Cohort Study.** Adjusted for mode of breastfeeding, infant sex, parity, and birth mode. Fatty acids and HMOs are coloured according to their categories. Reproduced from Moossavi et al. (214).



**Figure 37. Association of milk fatty acids with milk microbiota richness (observed ASVs) and diversity (inverse Simpson index) in 393 human milk samples from the CHILD Cohort Study.** Adjusted for mode of breastfeeding, infant sex, parity, and birth mode. Reproduced from Moossavi et al. (214).

Using redundancy analysis, overall HMO and fatty acid compositions were not associated with the overall composition of the milk microbiota ( $R^2=4.82$ ,  $p=0.30$  for HMOs and  $R^2=7.96$ ,  $p=0.13$  for fatty acids; Figure 29), as previously observed (139). Consistent with this finding, milk microbiota diversity and taxonomic clusters were not associated with the overall composition of HMOs or fatty acids. However, a few significant associations were observed when examining individual HMOs and fatty acids. Specifically, HMO 3'SL ( $R^2=0.58$ ,  $p=0.01$ ) and DSLNT ( $R^2=0.45$ ,  $p=0.04$ ) and fatty acid stearic acid ( $R^2=0.55$ ,  $p=0.01$ ) were significantly associated with the overall composition of the milk microbiota (Table 12).

**Table 12. Association of milk microbiota overall composition with human milk oligosaccharides (HMOs) and fatty acids in 393 human milk samples from the CHILd Cohort Study. Reproduced from Moossavi et al. (214).**

| Independent profile or component | Dependent Profile |    |                |    |                |     |
|----------------------------------|-------------------|----|----------------|----|----------------|-----|
|                                  | Microbiota        |    | HMOs           |    | Fatty acids    |     |
|                                  | R <sup>2</sup>    | p  | R <sup>2</sup> | p  | R <sup>2</sup> | p   |
| HMO compositional profile        | 4.82              |    | -              |    | 11.50          | *** |
| 3FL                              | 0.19              |    | -              |    | 0.40           | ~   |
| LNnT                             | 0.36              |    | -              |    | 0.46           | *   |
| 3SL                              | 0.58              | ** | -              |    | 0.78           | **  |
| DFLac                            | 0.23              |    | -              |    | 0.65           | **  |
| 6SL                              | 0.27              |    | -              |    | 1.44           | *** |
| LNT                              | 0.19              |    | -              |    | 0.46           | *   |
| LNFPIII                          | 0.17              |    | -              |    | 0.92           | *** |
| LNH                              | 0.28              |    | -              |    | 1.10           | *** |
| DSLNT                            | 0.45              | *  | -              |    | 3.13           | *** |
| FDSLNH                           | 0.17              |    | -              |    | 0.37           | ~   |
| Milk fatty acid profile          | 7.60              |    | 10.50          | ** | -              |     |
| 12:0                             | 0.24              |    | 0.21           | *  | -              |     |
| 16:0                             | 0.18              |    | 0.33           | ** | -              |     |
| 16:1n9                           | 0.30              |    | 0.15           | *  | -              |     |
| 18:0                             | 0.55              | *  | 0.03           |    | -              |     |
| 18:2n6                           | 0.41              | ~  | 0.02           |    | -              |     |
| 20:3n6                           | 0.30              |    | 0.12           | ~  | -              |     |
| 24:0                             | 0.39              | ~  | 0.07           |    | -              |     |
| 24:1n9                           | 0.38              | ~  | 0.08           |    | -              |     |
| 22:5n3                           | 0.41              | ~  | 0.04           |    | -              |     |

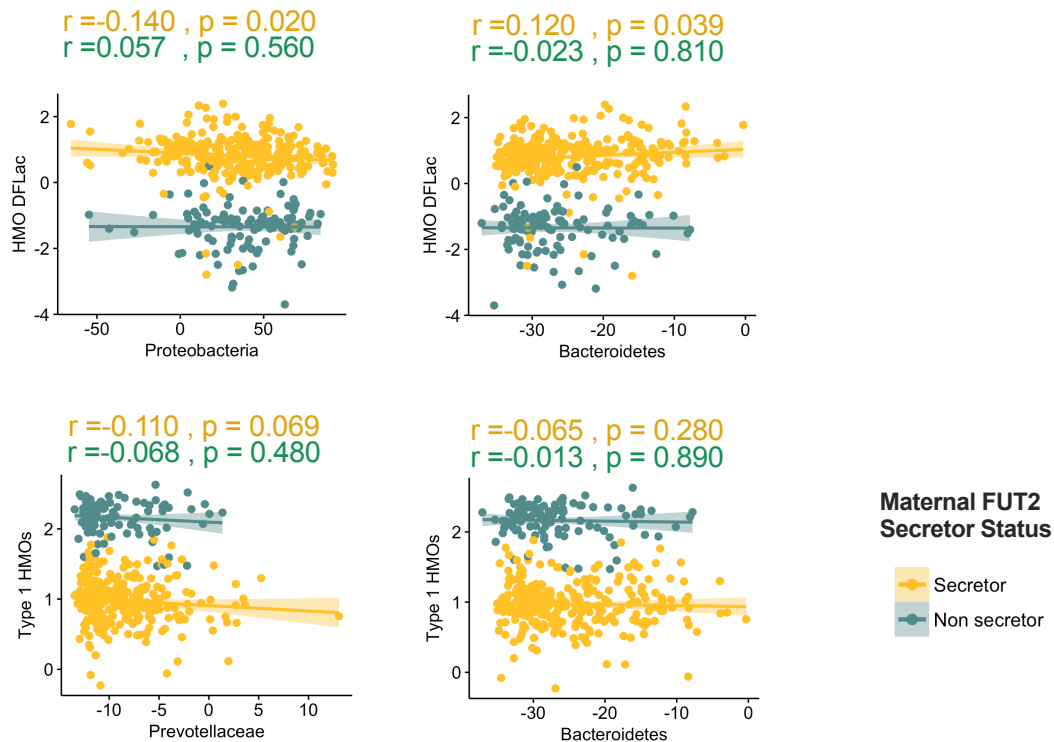
R<sup>2</sup> values reflect the percentage of variation explained using redundancy analysis. Only those with significant association with at least one component are reported. ~p<0.10, \*p<0.05, \*\*p<0.01, \*\*\*p<0.001. Reproduced from Moossavi et al. (214).

## 6.5. Some HMOs and fatty acids are weakly associated with specific milk bacteria relative abundances

Next, we assessed the correlation of individual HMOs and fatty acids with milk microbiota at phylum, family, and genus taxonomic levels (all taxa with >0.01% mean relative abundance). Only a few weak ( $r < |0.15|$ ) associations between HMOs and microbiota were identified. These associations were mostly between DFLac and type 1 HMOs (consisting of LNT, LNFP I, LNFP II, LSTb, and DSLNT) with Proteobacteria, Bacteroidetes and *Prevotellaceae* (Figure 38). When stratified by maternal secretor status, only sialylated/non-fucosylated HMOs remained significantly correlated with *Prevotella* ( $r=0.22$ ,



p=0.028). We did not detect any significant associations with HMOs expressed as absolute quantities.



**Figure 38. Association of milk microbiota with human milk oligosaccharides stratified by maternal FUT2 secretor status.**

Similarly, we identified a few weak correlations ( $r < |0.20|$ ) between fatty acids and milk bacteria including negative correlations between C18:1c11/Actinobacteria and CLA/*Porphyromonadaceae*, and positive correlations between C20:4n6/Bacteroidetes, C22:5n3/Fusobacteria, and 22:4n6/*Prevotellaceae*. No correlations with any of the fatty acid categories (SFAs, MUFAs, or PUFAs) were observed.

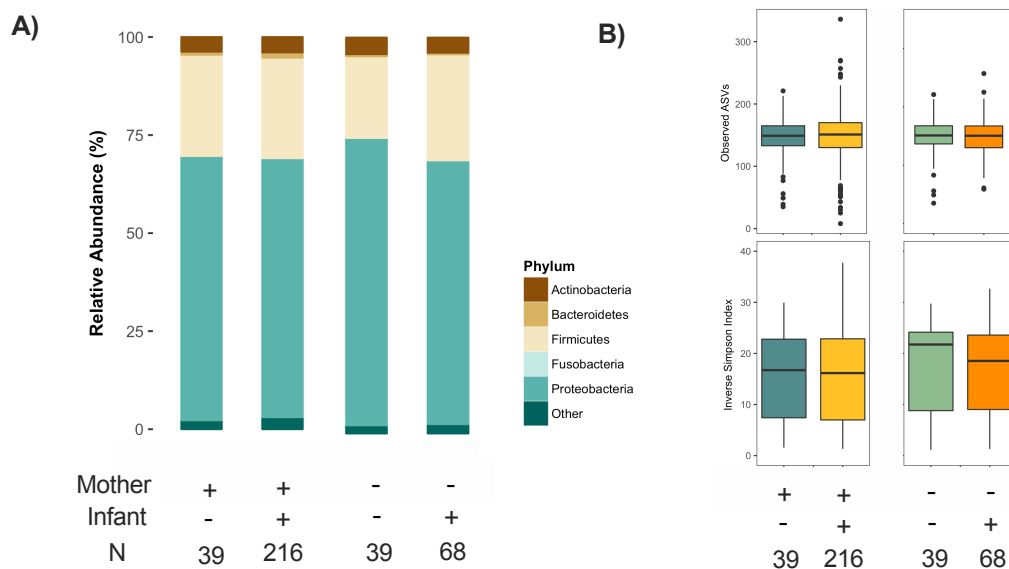
When focusing on the core milk microbiota (amplicon sequence variants present in at least 95% of samples with more than 1% mean relative abundance), some additional patterns emerged. Overall, core microbiota was weakly correlated with fatty acids and HMOs (Figure 39). *Nocardioides* was significantly associated with several fatty acids including lauric acid,

myristic acid, palmitic acid, and lignoceric acid (SFAs, correlations with  $r > |0.20|$  and  $p < 0.05$ ) (Figure 39A). The strongest individual correlations were observed for DSLNT with *Nocardioidea*, *Massilia*, and unclassified *Comamonadaceae* (negative correlations with  $r < -0.23$  and  $p < 0.05$ ) (Figure 39B). In general, the directions of correlation of core ASVs with HMOs were opposite for fucosylated vs. sialylated oligosaccharides. These correlations were directionally consistent across most taxa, with the notable exception of *Staphylococcus*: while most ASVs were positively correlated with fucosylated HMOs and negatively correlated with sialylated HMOs, these relationships were inversed for *Staphylococcus* (Figure 39B). Interestingly, when stratified by secretor status, these associations were only observed in non-secretor mothers (Figure 39C).

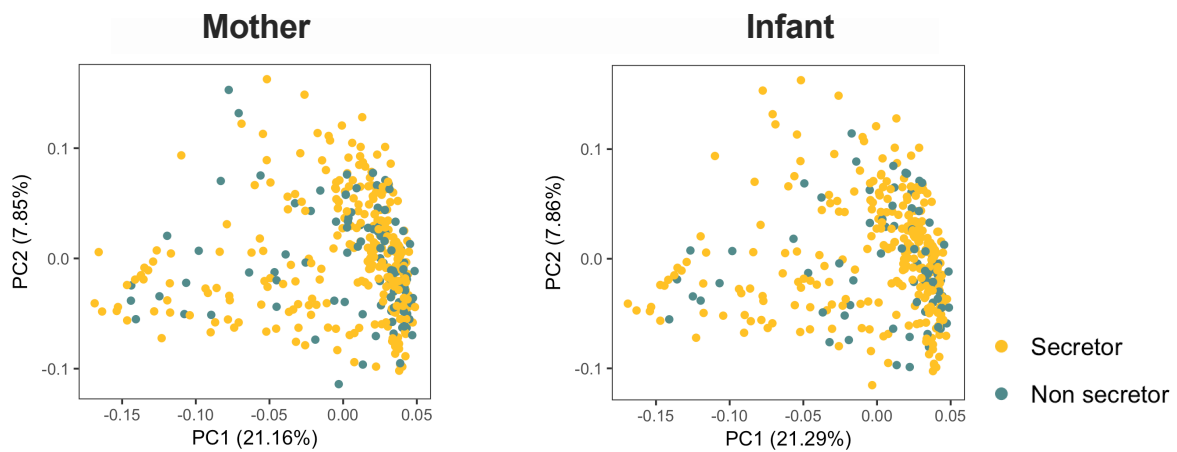
**Figure 39. Correlation of core milk microbiota (>1% relative abundance) with milk components in 393 human milk samples from the CHILd Cohort Study.** A) Fatty acids, B) human milk oligosaccharides (HMOs). Correlation of core *Staphylococcus* ASV with C) HMOs stratified by maternal secretor status and D) Fatty acids groups stratified by maternal fish oil supplement. Reproduced from Moossavi et al. (214).

## 6.6. Maternal and infant secretor status are not strongly associated with overall milk microbiota composition

As previously reported (192), maternal FUT2-secretor status was a clear determinant of HMO composition. We also assessed infant secretor status because this genotypic trait could influence the infant oral microbiota, which could potentially inoculate the mother's milk. We did not identify major differences in the milk microbiota structure or alpha diversity based on maternal or infant secretor status (Figure 40). However, Bacteroidetes relative abundance was significantly higher in milk produced by secretor mothers ( $\beta=+2.15$ , 95% CI 0.56- 3.73,  $P<0.01$ , for secretor vs. non-secretor mothers, independent of mode of breastfeeding and other relevant covariates). This difference was driven by enrichment of *Prevotella* ( $0.5\pm1.7$  vs.  $0.2\pm0.7$ ,  $p<0.05$ ). There was also a trend in the association of infant secretor status with milk Actinobacteria diversity ( $\beta=+0.37$ , 95% CI -0.031 to 0.775,  $p=0.07$  for secretor vs. non-secretor infants). Maternal and infant secretor statuses were not associated with *Bifidobacterium* relative abundance or prevalence, or milk microbiota beta diversity (Figure 41).



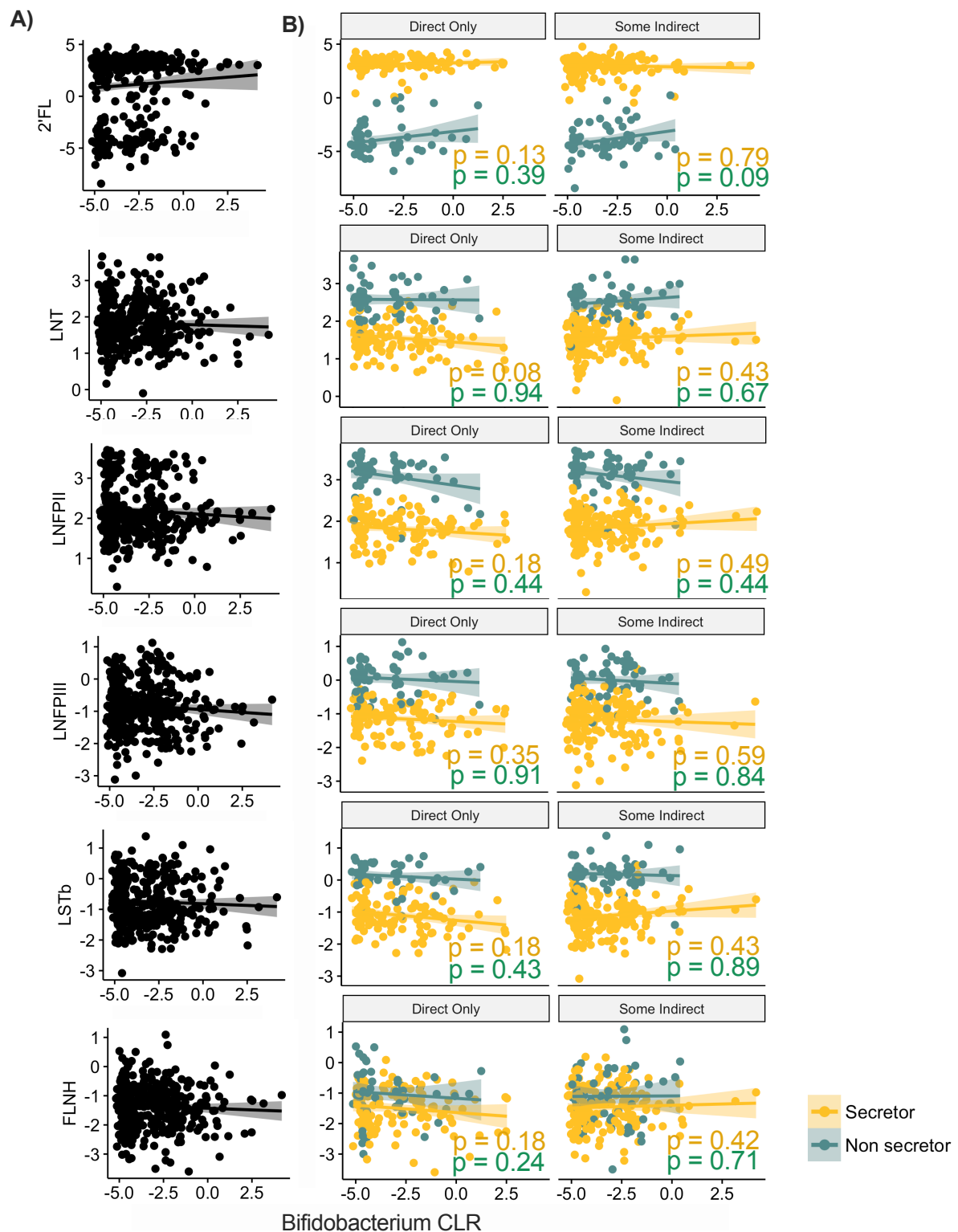
**Figure 40. Association of maternal and infant secretor status with milk microbiota alpha diversity in 393 human milk samples from the CHILD Cohort Study. Reproduced from Moossavi et al. (214).**



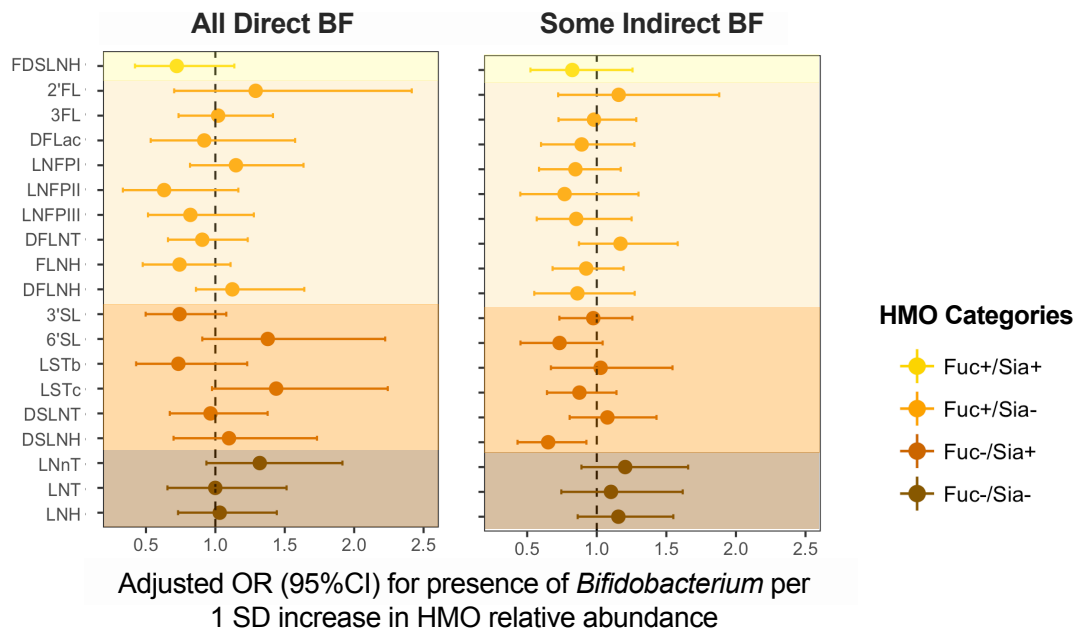
**Figure 41. Lack of association between maternal and infant FUT2 secretor status and milk microbiota overall composition in 393 human milk samples from the CHILD Cohort Study. Reproduced from Moossavi et al. (214).**

#### 6.7. HMOs are associated with milk *Bifidobacterium* prevalence and abundance

Given the established ability of *Bifidobacteria* to consume HMOs, we specifically explored the association of HMOs and *Bifidobacteria* in milk. *Bifidobacterium* was present in 154 (39%) samples with mean relative abundance of  $0.25 \pm 0.98\%$ . Overall, we did not observe significant associations for total or individual HMO concentrations or relative abundances (Figure 42); however, patterns emerged after stratification on feeding method. We undertook stratification because we have previously shown that *Bifidobacterium* is more prevalent in milk microbiota of mothers feeding exclusively at the breast (139). Among these directly-breastfeeding mothers, *Bifidobacterium* relative abundance was negatively associated with type 1 HMOs ( $r = -0.29$ ,  $p < 0.001$ ). Among mothers providing some indirect breast milk, *Bifidobacterium* was less likely to be present in milk with higher levels of DSLNH (OR=0.65, 95%CI 0.43-0.93,  $p = 0.03$  per 1 SD increase in DSLNH, Figure 43).



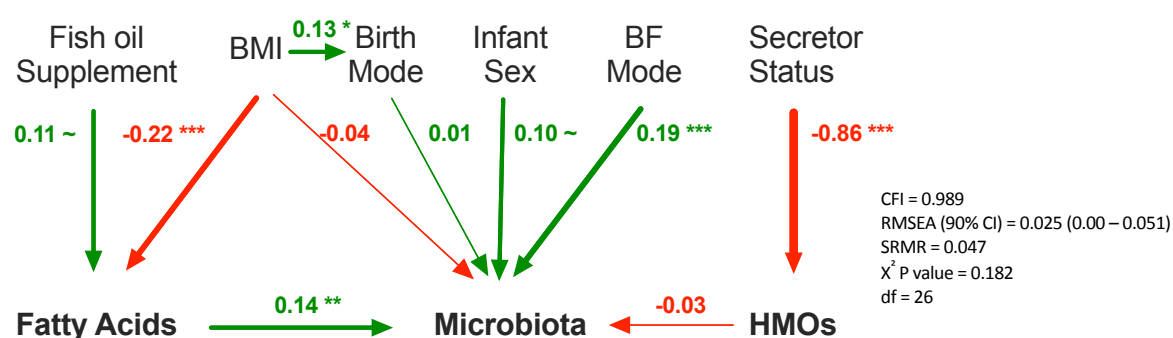
**Figure 42. Correlation of milk *Bifidobacterium* relative abundance (centre log-ratio transformed) with human milk oligosaccharides.**



**Figure 43. Association of *Bifidobacterium* prevalence with human milk oligosaccharides in 393 human milk samples from the CHILD Cohort Study.** Reproduced from Moossavi et al. (214).

#### 6.8. Fatty acids profile is associated with milk microbiota composition when controlling for relevant confounding factors

Finally, we used SEM to assess the associations between milk components using summary statistics (first principal component axes; PC1) while controlling for key covariates (Figure 44). Fatty acids PC1 was significantly associated with milk microbiota composition ( $\beta=0.14$ ,  $p<0.01$ ) when controlling for fish oil supplement use (a key dietary source of fatty acids) and maternal BMI (a key factor influencing fatty acids originating from liver synthesis or body storage), maternal secretor status, mode of breastfeeding, infant sex, and birth mode. HMO PC1 was not significantly associated with milk microbiota composition in this model. Thus, consistent with results from redundancy analysis, the SEM results suggest that fatty acids, more so than HMOs, are associated with milk microbiota composition while controlling for relevant factors (Figure 44).



**Figure 44. Structural equation modelling of the association of human milk oligosaccharides (HMOs) and fatty acids with microbiota in 393 human milk samples from the CHILD Cohort Study.** Standardized  $\beta$  coefficients are reported. BF, breastfeeding; BMI, maternal body mass index; CFI, comparative fix index; CI, confidence interval; RSMEA, root mean square error of approximation; SRMR, standardized root mean residuals; ~  $p < 0.1$ , \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ , Green (positive), red (negative). Reproduced from Moossavi et al. (214).

## 6.9. Conclusion

To our knowledge, this is the first integrative analysis of human milk microbiota with HMOs and milk fatty acids. We applied multiple statistical approaches to combine and simultaneously analyse the association of these milk components with milk microbiota while controlling for important covariates reflecting maternal genetics, diet, and breastfeeding practices. Overall, the correlations between milk components within each group were stronger and more frequent than correlations between components from different groups. Still, a few weak associations were observed between individual HMOs or fatty acids and microbial taxa, including *Bifidobacterium*. Interestingly, some of these associations were modified by maternal secretor status, fish oil supplementation, or method of breastfeeding. In structural equation models controlling for these and other key factors influencing milk composition, the overall fatty acid profile (but not the overall HMO profile) remained directly associated with milk microbiota composition. Overall, using multiple approaches for data integration to analyse three major components of human milk (fatty acids, microbiota

and HMOs), we identified strong correlations within each group, but relatively few correlations between groups. Fatty acid composition, more so than HMO profile, was associated with milk microbiota composition when controlling for relevant confounders. While the cumulative effect of HMOs on the entire microbial community appears to be small, our findings suggest that individual HMOs might promote or inhibit growth of specific milk bacteria, potentially providing a selection mechanism for vertical mother-offspring transmission of microbiota. Further research is warranted to explore these possibilities, and to determine how these and other milk components collectively modulate the infant gut microbiota and influence the risk of developing microbiota-mediated childhood diseases.

#### **Box 4. Summary of Chapter 6**

- Correlations are stronger and more frequent within than between different types of milk components
- Some HMOs and fatty acids are associated with milk microbiota  $\alpha$  diversity and overall composition
- Some HMOs and fatty acids are weakly associated with specific milk bacteria relative abundances
- Maternal and infant secretor status are not strongly associated with overall milk microbiota composition
- HMOs are associated with milk *Bifidobacterium* prevalence and abundance
- Fatty acids profile is associated with milk microbiota composition when controlling for relevant confounding factors



## **7. DETERMINANTS OF HUMAN MILK MYCOBIOTA**

### **7.1. Introduction**

Fungi constitute an important yet frequently neglected component of the human microbiota (92, 93) with an underappreciated role in health and disease (94, 95). Milk is a source of diverse and viable bacteria and fungi (38, 97-101). Although some geographical differences in milk mycobiota composition were observed in a recent study of 80 women from 4 countries (99), the maternal, infant, and environmental determinants of milk mycobiota are still mostly unknown. To address these open questions, we investigated the determinants of human milk mycobiota in a subset of 271 healthy Canadian mother-infant dyads (see Chapter 3 and Table 13) enrolled in the CHILD Cohort Study (191). Kelsey Fehr helped with fungal ITS2 gene sequencing.

### **7.2. Participant characteristics**

In our study population, the majority of dyads were urban residents (96%) although 250m radius population densities ranged widely from 0 to 36,170 (median 3,447). 67% of mothers were atopic and 25% delivered by Caesarean section (Table 13). Mean $\pm$ SD bacterial richness and diversity of milk were 145 $\pm$ 46 and 15.2 $\pm$ 9.2, respectively (Table 14).

### **7.3. Prevalence of milk fungi was significantly and independently associated with environmental characteristics, human milk oligosaccharides, and milk bacterial composition**

The origins of milk fungi are unknown, but the external (retrograde inoculation) and internal (entero-mammary) pathways hypothesized for the origins of milk bacteria (Figure 3) (105, 131) could be extended to milk fungi.

**Table 13. Characteristics of mother-infant dyads from the CHILD Cohort Study included in this analysis in comparison with previous microbiota analysis.**

|                         | Characteristics                     | Mean $\pm$ SD or n (%) <sup>a</sup> |                         |
|-------------------------|-------------------------------------|-------------------------------------|-------------------------|
|                         |                                     | Fungal study (N=271)                | Bacterial study (N=393) |
| <b>Maternal</b>         | Secretor status                     | 195 (72.8)                          | 279 (71.7)              |
|                         | Ethnicity                           |                                     |                         |
|                         | Caucasian                           | 197 (72.7)                          | 287 (73.0)              |
|                         | Asian                               | 52 (19.2)                           | 73 (18.6)               |
|                         | First Nations                       | 10 (3.7)                            | 15 (3.8)                |
|                         | Other                               | 12 (4.2)                            | 18 (4.6)                |
|                         | Prenatal smoking                    | 13 (4.8)                            | 18 (4.6)                |
|                         | History of atopy                    | 181 (67)                            | 251 (63.9)              |
| <b>Infant</b>           | Birth weight (g)                    | 3476 $\pm$ 476                      | 3469 $\pm$ 469          |
|                         | Female sex                          | 125 (46.1)                          | 192 (48.3)              |
|                         | Gestational age (weeks)             | 39.2 $\pm$ 1.3                      | 39.2 $\pm$ 1.3          |
|                         | History of oral thrush              | 33 (12.2)                           | 46 (11.8)               |
|                         | Child antibiotics before 3-4 months | 9 (3.3)                             | 11 (2.9)                |
| <b>Early life</b>       | Mode of delivery                    |                                     |                         |
|                         | Elective C/S                        | 34 (12.6)                           | 46 (11.9)               |
|                         | Emergency C/S                       | 32 (11.9)                           | 47 (12.1)               |
|                         | Vaginal                             | 203 (75.5)                          | 294 (76.0)              |
|                         | Older siblings                      |                                     |                         |
|                         | None                                | 148 (54.6)                          | 211 (53.7)              |
|                         | One                                 | 89 (32.8)                           | 126 (32.1)              |
|                         | Two or more                         | 34 (12.5)                           | 56 (14.2)               |
|                         | Maternal intrapartum antibiotics    | 96 (35.6)                           | 131 (35.4)              |
| <b>BF</b>               | Direct BF (at the breast) only      | 108 (40.0)                          | 162 (41.9)              |
|                         | Pump expression <sup>b</sup>        | 77 (81.9)                           | 91 (81.2)               |
| <b>Home environment</b> | Sample collection season            |                                     |                         |
|                         | Spring                              | 58 (21.5)                           | 72 (18.6)               |
|                         | Summer                              | 64 (27.4)                           | 103 (26.5)              |
|                         | Autumn                              | 74 (23.7)                           | 112 (28.9)              |
|                         | Winter                              | 74 (27.4)                           | 101 (26.0)              |
|                         | High Mould                          | 90 (33.3)                           | 137 (34.9)              |
|                         | High Moisture                       | 120 (44.4)                          | 183 (46.6)              |
|                         | High Dust                           | 144 (53.3)                          | 208 (52.9)              |
|                         | Dog ownership                       | 82 (30.2)                           | 119 (30.2)              |
|                         | Furry pets                          | 137 (50.6)                          | 201 (51.0)              |
|                         | Urban residence                     | 242 (95.7)                          | 91 (92.8)               |
|                         | High population density             | 126 (49.8)                          | 195 (52.7)              |
|                         | Residential vegetation <sup>c</sup> |                                     |                         |
|                         | Grey                                | 16 (6.0)                            | 20 (5%)                 |
|                         | Moderately green                    | 92 (34.3)                           | 135 (35%)               |
|                         | Green                               | 160 (59.7)                          | 233 (60%)               |
|                         | Study sites                         |                                     |                         |
|                         | Vancouver                           | 67 (24.7)                           | 96 (24.3)               |
|                         | Edmonton                            | 61 (22.5)                           | 99 (25.1)               |
|                         | Manitoba                            | 81 (29.9)                           | 96 (24.3)               |
|                         | Toronto                             | 62 (22.9)                           | 104 (26.3)              |

<sup>a</sup> Percentages are calculated after excluding dyads with missing data. BF, breastfeeding; BMI, body mass index; C/S, caesarean section; HMO, human milk oligosaccharide. <sup>b</sup> n=94; <sup>c</sup> Defined based on normalized difference vegetation index (NDVI) in a 250 m buffer around the mother's residential addresses during the 12 months postpartum.

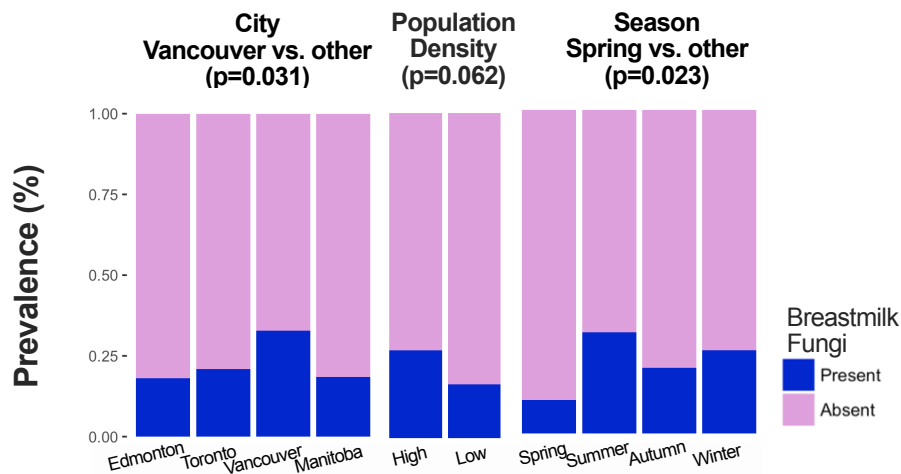
**Table 14. Milk characteristics of mother-infant dyads in milk mycobiota analysis in the CHILd Cohort Study.**

| Characteristics                             | Current Fungal study<br>(N=271) | Previous Bacterial<br>study (N=393) |
|---------------------------------------------|---------------------------------|-------------------------------------|
| Bacterial richness                          |                                 |                                     |
| Overall                                     | 145.3 ± 46.3                    | 146.8 ± 43.8                        |
| Proteobacteria                              | 104 ± 29.9                      | 105.9 ± 28.4                        |
| Firmicutes                                  | 16.4 ± 15.1                     | 15.9 ± 13.6                         |
| Actinobacteria                              | 14.0 ± 5.6                      | 14.1 ± 5.4                          |
| Bacterial diversity                         |                                 |                                     |
| Overall                                     | 15.2 ± 9.2                      | 15.8 ± 8.9                          |
| Proteobacteria                              | 16.3 ± 7.5                      | 16.8 ± 6.9                          |
| Firmicutes                                  | 3.2 ± 1.5                       | 3.2 ± 1.4                           |
| Actinobacteria                              | 3.8 ± 1.8                       | 3.7 ± 1.7                           |
| Taxonomic clusters                          |                                 |                                     |
| C1                                          | 33 (12.2)                       | 42 (10.7)                           |
| C2                                          | 73 (27.0)                       | 98 (24.9)                           |
| C3                                          | 93 (34.5)                       | 161 (41.0)                          |
| C4                                          | 71 (26.3)                       | 92 (23.4)                           |
| Bacterial composition outliers <sup>a</sup> | 35 (12.9)                       | 45 (11.5)                           |
| HMO diversity                               | 4.9 ± 1.4                       | 4.9 ± 1.4                           |
| HMO concentration (mg/mL)                   | 10.2 ± 2.1                      | 10.2 ± 2.1                          |

<sup>a</sup> Defined as those contributing greater than the median plus twice the interquartile range of the sample variance to the total (179). All dyads included in the current fungal study were also included in the previous bacterial study (139).

To address whether milk fungi were more likely acquired externally (e.g. from the environment, maternal skin and/or the infant oral cavity) or internally from the maternal gut, we assessed associations of fungi prevalence with environmental vs. maternal or other factors. Both environmental (e.g. study city, population density, and season) and maternal characteristics (e.g. antibiotics and atopy) were associated with fungi prevalence in univariate analyses (Figure 45 and Figure 46). Milk from mothers in Vancouver had the highest prevalence of fungi (32% vs. 18-21% in Edmonton, Toronto, and Manitoba; unadjusted OR=2.06, 95%CI 1.11-3.82, for Vancouver vs. other cities; p=0.021; Figure 45 and Figure 46). Vancouver is a large coastal Canadian city with more humid climate, higher annual precipitation, and milder winters compared with the other study sites, and thus likely has higher load of environmental fungi (260). In line with this, season and population density

were associated with fungal prevalence, with the lowest prevalence detected in spring and at low population density (Figure 45 and Figure 46). These findings are consistent with seasonal variations reported in goat milk fungi (261) which could be partly attributed to fluctuations in outdoor (262) and indoor (263) environmental fungi levels.

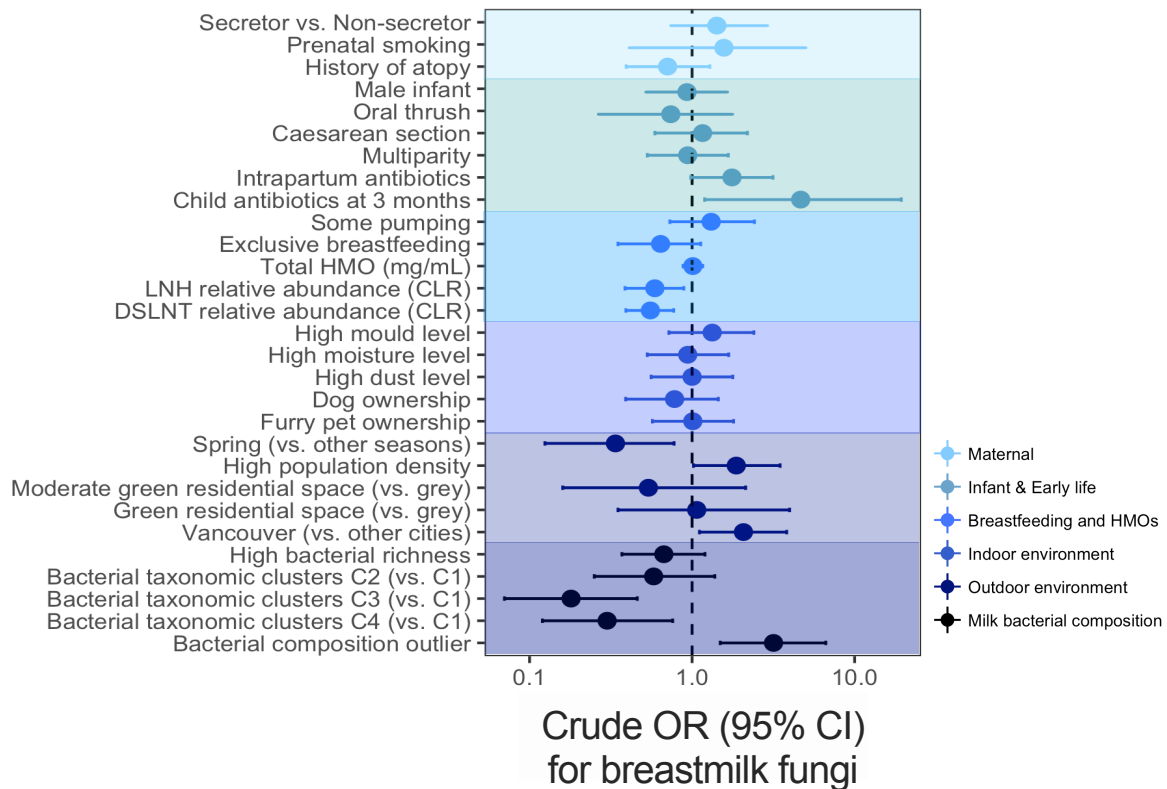


**Figure 45. Prevalence of fungi according to study site, population density, and season tested by chi-square test in 271 human milk samples from the CHILD Cohort Study.** Fungi presence was defined at 1000 sequences per sample threshold (sections 3.3.2 and Figure 14).

Among the maternal, infant, and early life factors assessed, maternal intrapartum antibiotics and infant antibiotics at the time of sample collection (3-4 months postpartum) were associated with a higher likelihood of fungi presence (Figure 46) indicating that bacteria originating from the mother or infant might be associated with the prevalence of milk fungi.

We also assessed the association of milk fungi with the 19 most abundant HMOs, which were previously measured in the same samples (192). The impact of HMOs on infant gut bacterial composition is well established, but little is known about the impact of HMOs on mycobiota. There is some evidence that certain HMOs can inhibit fungi in vitro (264), but it is also plausible that other HMOs could be metabolised by fungi and support their growth.

Here, we found that two FUT2 secretor status-independent HMOs (DSLNT and LNH) were less abundant in milk containing fungi (Figure 47), suggesting that these HMOs might suppress milk fungi via modulating the milk microbiota or fatty acids composition (214).



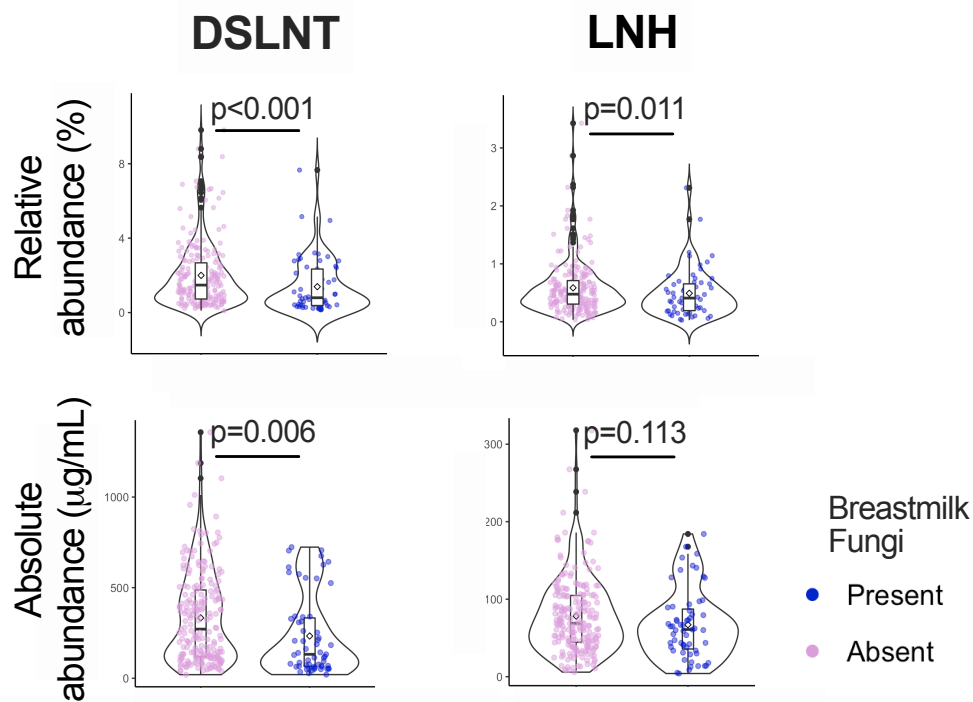
**Figure 46. Unadjusted univariate associations of fungal prevalence with maternal, infant and early life factors, breastfeeding and selected HMOs, indoor and outdoor environmental factors, and milk microbiota composition tested by logistic regression in 271 human milk samples from the CHILD Cohort Study.**

To identify the determinants of milk fungi while controlling for the above-mentioned factors, we performed unsupervised variable selection by least absolute shrinkage and selection operator (LASSO) and used the LASSO-selected variables to predict the presence of milk fungi. While LASSO does not necessarily provide evidence for biological associations, it is a good technique for feature selection as it shrinks the beta coefficients of

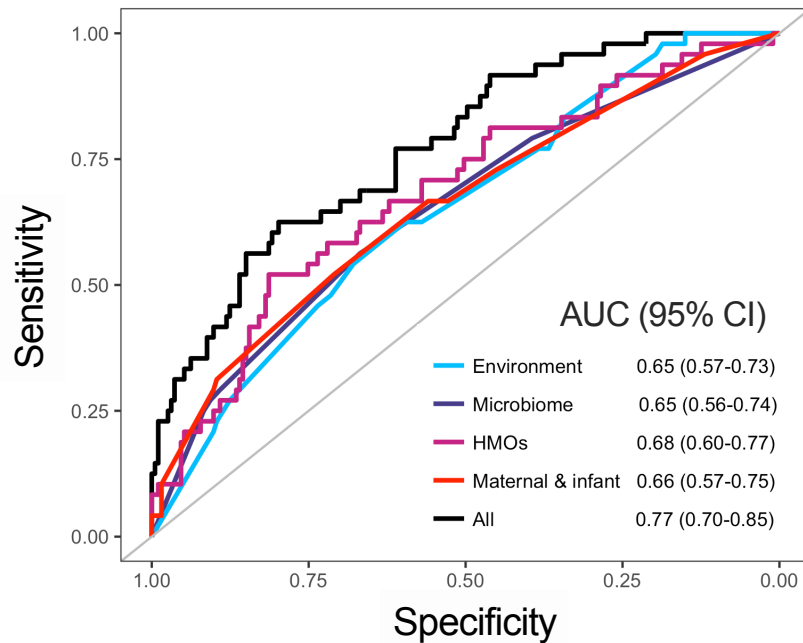
variables with minimal contribution to zero. This approach will minimally increase the bias but will enhance the interpretability of the results. LASSO is best suitable when small to moderate number of moderate-sized effect are expected (265) (as is the case in our study).

In agreement with the logistic regression results above, LASSO identified season, population density, study city, residential vegetation, bacterial clusters, bacterial outlier, DSLNT, LNH, intrapartum antibiotics, and child antibiotics at 3 months as predictors of fungi presence. In addition, LASSO identified maternal asthma and atopy, suggesting that atopic mothers may have distinct skin fungal load and/or composition (266). Combined, these selected variables had a relatively good predictive accuracy for presence of fungi (AUC=0.77, 95% CI 0.70-0.85) (

Figure 48).



**Figure 47. Relative abundances and absolute concentrations of Disialyllacto-N-tetraose (DSLNT) and Lacto-N-hexaose (LNH) based on milk fungi presence tested by ANOVA in 271 human milk samples from the CHILD Cohort Study.**

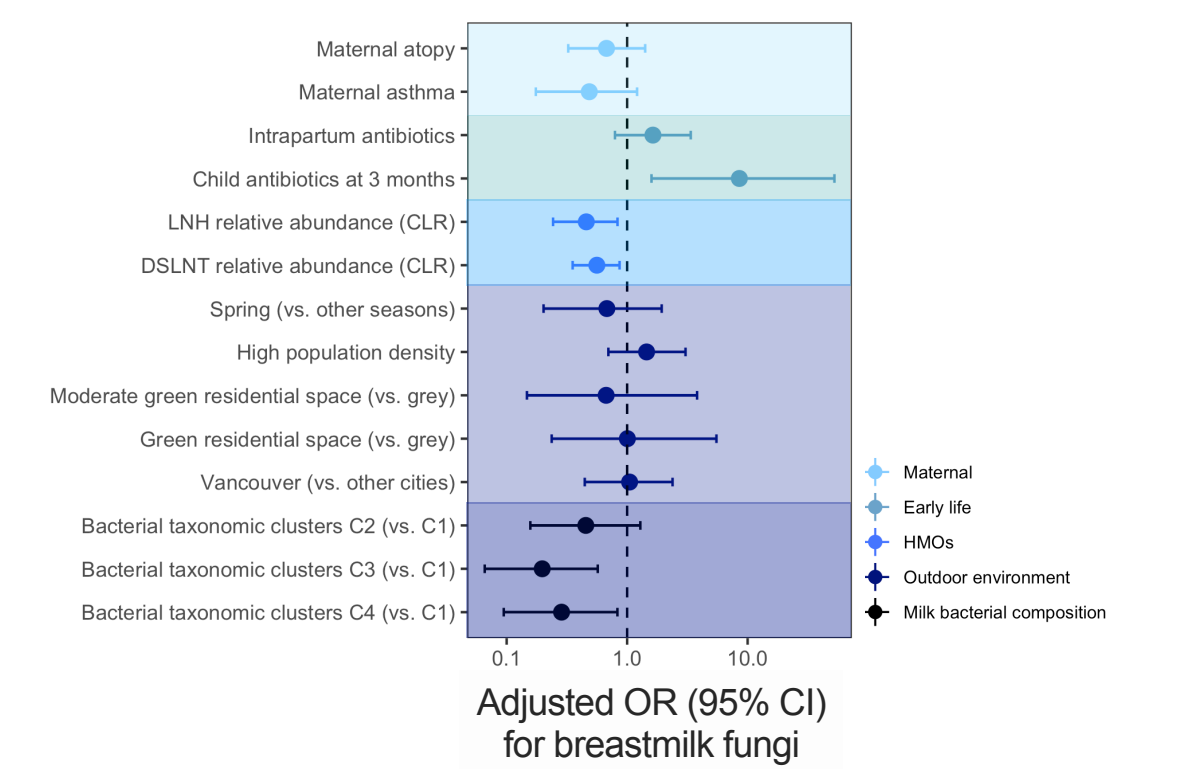


**Figure 48. Prediction accuracy of fungi presence based on unsupervised variable selection by LASSO in 271 human milk samples from the CHILD Cohort Study.** The selected variables were classified as: environmental (study site, population density, season, residential vegetation); bacterial (bacterial taxonomic cluster and bacterial composition outlier); HMOs (DSLNT and LNH); maternal and infant (intrapartum antibiotics, child antibiotics at 3 months, maternal asthma, and maternal atopy).

In a multivariable logistic regression model adjusted for all of the LASSO-selected factors, DSLNT, LNH, bacterial taxonomic clusters, and child antibiotics at 3 months were the strongest independent predictors of fungi prevalence (Figure 49).

We did not detect significant differences in fungi prevalence according to visible mould or moisture levels in the home, older siblings, pet ownership, maternal BMI or secretor status, or history of infant oral thrush (a fungal infection caused by *Candida* spp.) in the first 3 months of life (Figure 46). We lacked information on breast thrush and mastitis, which are also known to be associated with *Candida* spp. in milk (100). Additionally, although others have found that milk collection using a pump and bottle feeding were associated with increased prevalence of *Candida* spp. in milk (31), we did not observe an association

between fungi presence and mode of breastfeeding (nursing at the breast vs. pumped milk feeding at least once in the preceding two weeks) (Figure 46).



**Figure 49. Adjusted multivariable associations of host and environmental factors with the prevalence of fungi in 271 human milk samples from the CHILD Cohort Study. Variable selected by LASSO (see**

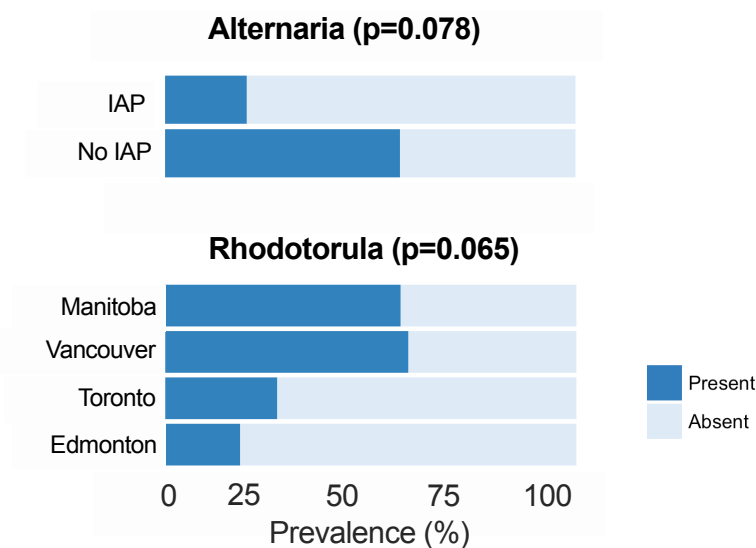
**Figure 48)** were included in the multivariable logistic regression.

**7.4. Milk mycobiota profile was not clearly associated with any available metadata characteristics**

Next, among the 58 samples with detectable fungi, we assessed fungal composition and diversity. Ascomycota and Basidiomycota were the dominant fungal phyla in the milk (Figure 15) in agreement with previous reports (98, 99). Acknowledging the low power due to small sample size, we detected borderline differences in the prevalence of *Alternaria* according to intrapartum antibiotics exposure, and *Rhodotorula* between study cities (Figure



50).



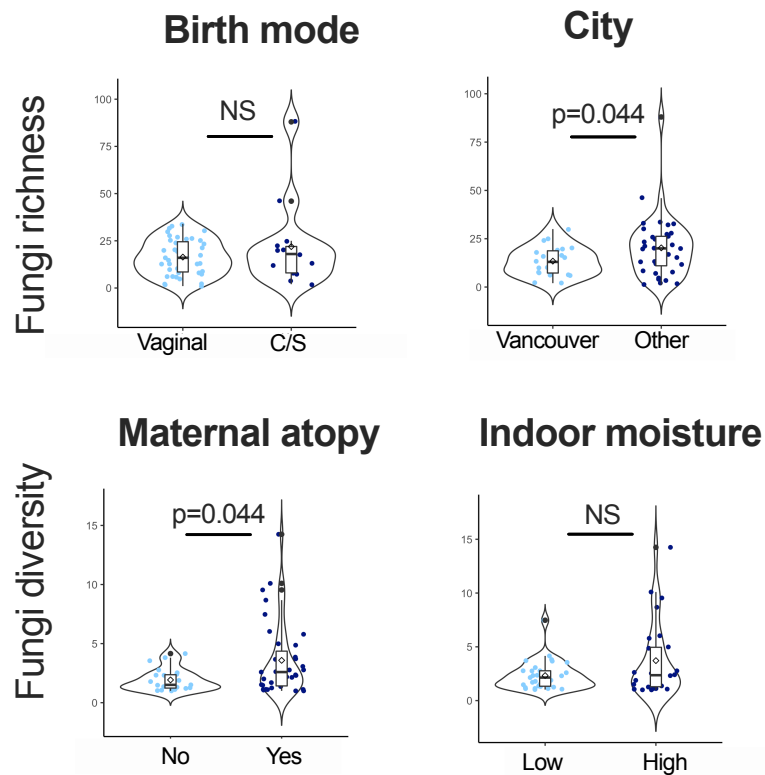
**Figure 50.** Association of intrapartum antibiotics and study city with the prevalence of fungal genera tested by chi-squared test in 58 human milk samples from the CHILD Cohort Study.

Fungal diversity was higher in atopic vs. non-atopic mothers (3.6 vs. 1.9,  $p=0.021$ ) and mothers living in homes with high vs. low indoor moisture (3.8 vs. 2.3,  $p=0.033$ ) (Figure 51) suggesting that the diversity of milk fungi is influenced by both the environment and maternal characteristics. We did not find any other significant associations between fungal taxonomic structure or diversity with other maternal, infant, early life, environmental, and milk factors.

## 7.5. Conclusion

In the general population CHILD Cohort Study, we found that approximately 20% of milk samples contained detectable fungi and provide evidence that fungal prevalence was associated with environmental characteristics and milk bacteria. Milk fungal taxonomy was consistent with previous reports (98, 99). Ascomycota and Basidiomycota were the dominant fungal phyla in the milk and three species of *Candida*, *Alternaria*, and *Rhodotorula* were the most prevalent taxa. To our knowledge, this is one of the largest studies of human milk mycobiota performed to date, and the only milk mycobiota study to examine home environmental characteristics. Our results expand considerably upon existing knowledge

about milk fungi, providing evidence for potential role of home environment, maternal characteristics, and milk components in shaping the milk mycobiota.



**Figure 51. Association of birth mode and city with fungal richness and maternal atopy and indoor moisture level with fungal diversity in 58 human milk samples from the CHILD Cohort Study.**

#### Box 5. Summary of Chapter 7

- Fungi were present in a subset of the milk samples in CHILD study
- Prevalence of milk fungi was significantly and independently associated with environmental characteristics, human milk oligosaccharides, and milk bacterial composition
- Milk fungi profile was not clearly associated with any available metadata characteristics

## **8. COMPARISON OF SEQUENCING PROCESSING APPROACHES ON DOWNSTREAM STATISTICAL RESULTS**

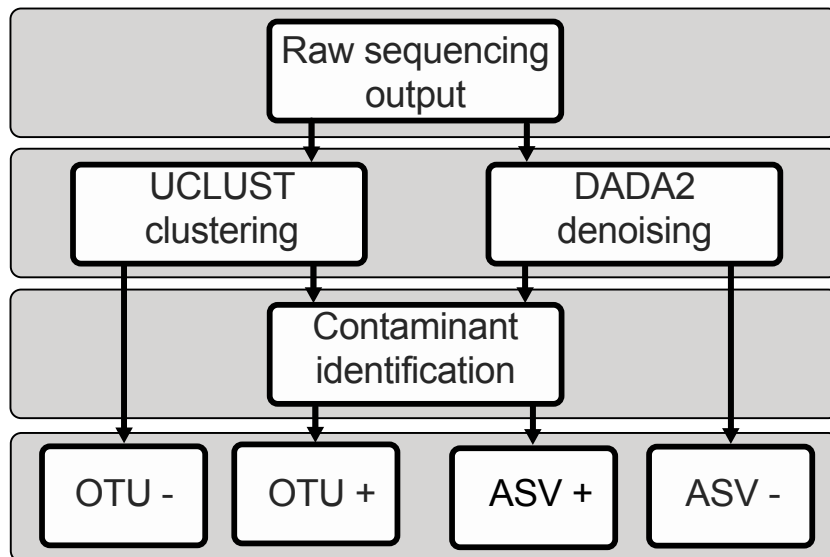
### **8.1. Introduction**

Milk microbiota in samples from the CHILD Cohort Study was processed using DADA2 denoising algorithm combined with statistical identification of potential reagent contaminants in this thesis (methods described in sections 2.1.5. and 2.1.6. and results reported in Chapters 3, 5, and 6). Although more accurate performance of the denoising methods has been confirmed on mock communities (173-178); the impact of these different methods with or without contaminant removal on characterizing real biospecimens and downstream analysis of biological associations is less clear specifically for low biomass samples such as milk. To address these knowledge gaps, we re-processed the 16S rRNA-based milk microbiota dataset of the CHILD samples (Chapters 3, 5, and 6)(139) using UCLUST and DADA2 with or without contaminant removal using *decontam* (190) using the taxonomy assignment threshold commonly used by these methods (97% for UCLUST and 99% for DADA2) . The datasets resulting from these four approaches (Figure 52) were used to test the hypothesis that denoising methods improve taxonomic accuracy compared to clustering, but biological associations are robust to the choice of upstream data processing. Kelsey Fehr generated Figures 56 and 57. Faisal Atakora performed the interclass correlation analysis (Figure 60).

### **8.2. Mock community composition is more accurately resolved using DADA2**

We analysed 18 replicates (3 per PCR plate in 2 sequencing runs) of a mock community consisting of 8 different bacterial species with a known composition (Zymo Research, USA). While the UCLUST method detected an average (SD) of 223 (50) OTUs in the mock community samples, DADA2 performance was closer to expected composition, detecting 12

(3) ASVs (Table 15). Although contaminant removal did not considerably reduce the number of OTUs, it decreased the average (SD) ASVs to 9 (3) effectively eliminating the potential contaminants (Table 15).



**Figure 52. Schematic of the study design to compare microbiota composition in a mock community and human milk samples using a clustering-based OTU method (UCLUST) and a denoising algorithm (DADA2) with and without contaminant removal in 393 human milk samples from the CHILD Cohort Study.**

Overall, there was a good agreement between the observed and expected taxonomic composition with both methods (Figure 53). However, two notable differences were observed. Most prominently, DADA2 performed better at identifying *Lactobacillus*: the actual contribution of this genus to the mock community was 19%; the estimated relative abundance was <1% using UCLUST compared to 16% using DADA2. Moreover, the proportion of identified taxa not belonging to the mock community (likely contaminants) was higher with UCLUST (~12% vs. 0.1% in DADA2). Neither method could accurately identify *Escherichia coli* or *Salmonella enterica* present in the mock community; however, for both methods, the relative abundance of taxa identified as *Enterobacteriaceae* was within the

range of the combined relative abundances of these two enteric bacteria. Thus, overall, DADA2 provided a more accurate representation of the mock community (Table 15 and Figure 53) in agreement with previous studies (173-178).

**Table 15. Comparison of the number of OTUs/ASVs in the mock community, negative controls, and milk microbiota datasets processed by clustering-based OTU method (UCLUST) and a denoising algorithm (DADA2) with or without contaminant removal in 393 human milk samples from the CHILD Cohort Study**

|                                                                                 | UCLUST     |           | DADA2      |           |
|---------------------------------------------------------------------------------|------------|-----------|------------|-----------|
| Sequence filtering <sup>a</sup>                                                 | Yes        |           | No         |           |
| Reference database                                                              | Greengenes |           | Greengenes |           |
| Similarity threshold for taxonomy assignment                                    | 97%        |           | 99%        |           |
| Contaminant identification and removal                                          | -          | +         | -          | +         |
| Sequencing depth (x1000)                                                        | 51.1±21.2  | 51.0±21.1 | 48.6±19.2  | 47.7±18.6 |
| Mock community of 8 species (N=18 <sup>b</sup> )                                |            |           |            |           |
| Average Number of OTUs/ASVs per sample                                          | 223±50     | 215±49    | 12±3       | 9±3       |
| Negative control (N=15)                                                         |            |           |            |           |
| Average Number of OTUs/ASVs per sample                                          | 58±94      | 14±20     | 42±65      | 25±35     |
| Milk samples (N=428)                                                            |            |           |            |           |
| Total number of OTUs/ASVs                                                       | 13,305     | 12,695    | 9,884      | 9,711     |
| Average Number of OTUs/ASVs per sample                                          | 364±145    | 358±132   | 170±73     | 168±70    |
| Milk samples after filtering (N=393) <sup>c</sup>                               |            |           |            |           |
| Remaining OTUs/ASVs after filtering                                             | 2,265      | 2,134     | 1,956      | 1,972     |
| Average Number of OTUs/ASVs per sample                                          | 394±91     | 445±142   | 149±44     | 147±44    |
| OTUs/ASVs > 0.01% average relative abundance <sup>d</sup>                       | 298        | 298       | 299        | 301       |
| Average number of OTUs/ASVs per sample after additional filtering <sup>de</sup> | 18±7       | 18±7      | 18±6       | 18±6      |

<sup>a</sup> Low quality base calling scores (<20) and containing ambiguous bases in the overlapping region

<sup>b</sup> 3 replicates per PCR reaction (n=3) for each sequencing run (n=2)

<sup>c</sup> Exclusion of OTUs/ASVs not present in the samples or belonging to non-bacterial taxa, rarefaction to 25000 reads per sample, filtering OTUs/ASVs with less than 20 reads across each dataset

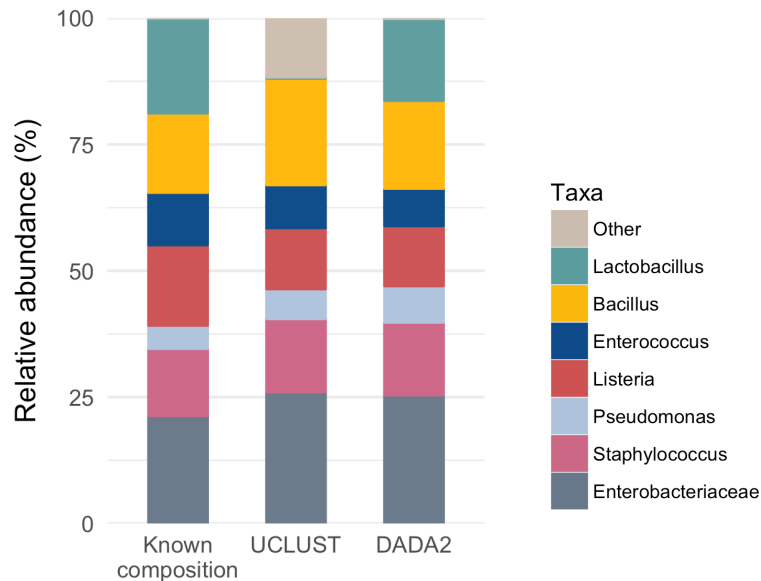
<sup>d</sup> Despite the similarity in the number of OTUs/ASVs, the composition is not identical.

<sup>e</sup> Excluding OTUs/ASVs with less than <1% mean relative abundance

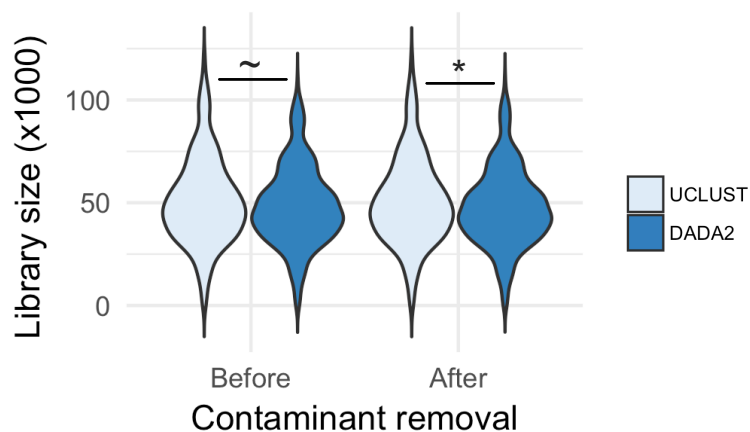
### 8.3. Observed OTUs/ASVs (richness) is influenced by the choice of sequencing processing method

Overall when comparing the four approaches, the mean depth of sequencing per sample was slightly lower in DADA2 compared to UCLUST both before and after contaminant

removal. The differences in library size within each method before and after *decontam* were negligible (Figure 54).

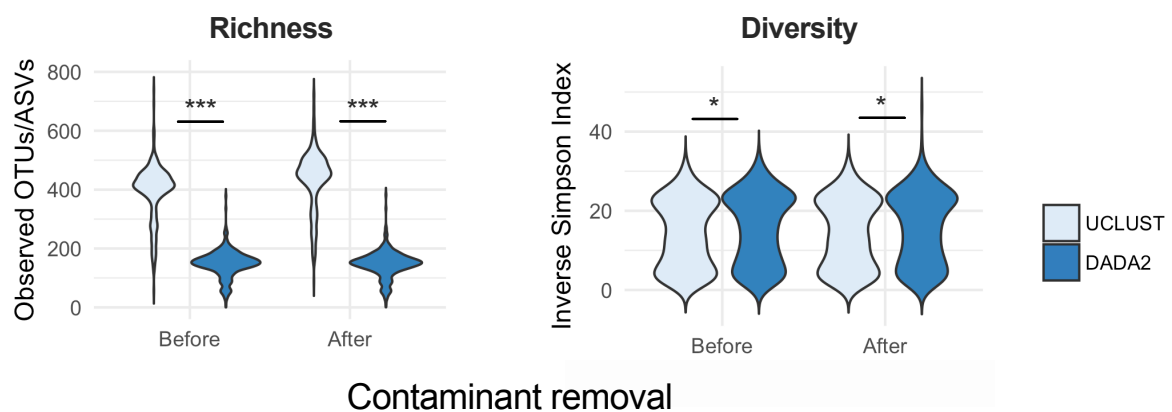


**Figure 53. Composition of the mock community by UCLUST and DADA2 prior to contaminant removal.** Each bar represents the average data from 8 replicates. Also see Figure 9. Relative abundance of taxa with >1% mean relative abundance in the mock community in comparison with the theoretical composition.



**Figure 54. Comparison of the library size on datasets processed by clustering-based OTU method (UCLUST) and a denoising algorithm (DADA2) with or without contaminant removal in 393 human milk samples from the CHILD Cohort Study.** ~  $p < 0.1$ , \*  $p < 0.05$

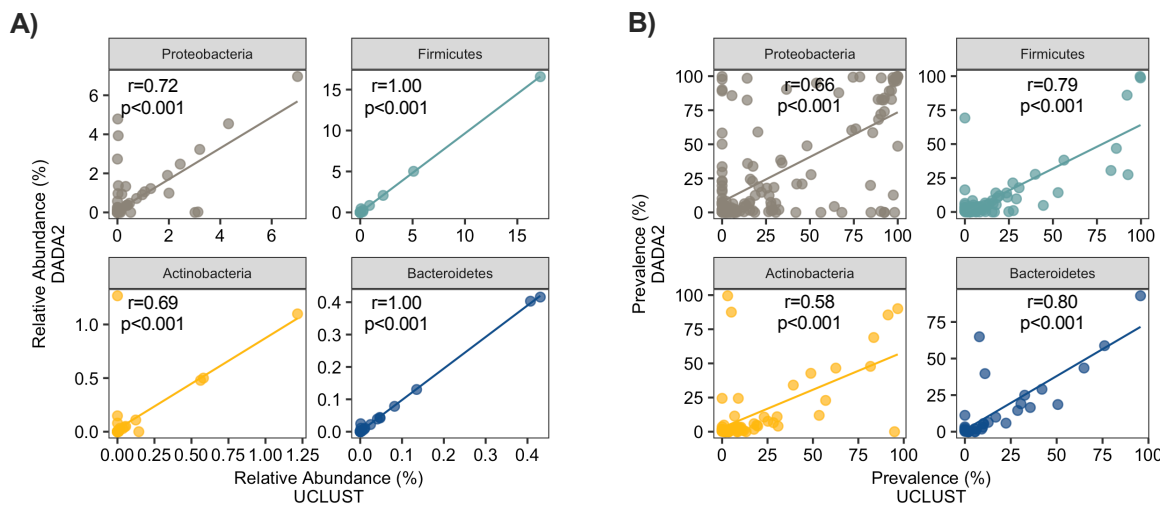
Despite differences in the initial number of OTUs/ASVs in total and on average, UCLUST and DADA2 resulted in a similar number of remaining OTUs/ASVs (298 and 299 respectively before contaminant removal) after filtering taxa with less than 0.01% mean relative abundance (Table 15). This suggests the majority of “noisy true reads” (true reads containing sequencing errors (176)) initially retained by UCLUST were eliminated by applying abundance-based filtering. In agreement with the literature (176, 178), there was a considerable difference in the number of observed OTUs/ASVs prior to filtering very low abundance taxa ( $364 \pm 145$  average OTUs vs.  $170 \pm 73$  average ASVs per sample, Table 15). The bacterial richness at OTU/ASV level remained higher in UCLUST vs. DADA2 even after data filtering, regardless of contaminant removal ( $394 \pm 91$  vs.  $148 \pm 44$ ,  $p < 0.001$ ) (Figure 55). In contrast, milk microbiota diversity was slightly but significantly higher with DADA2 regardless of contaminant removal ( $14.3 \pm 8.5$  vs.  $15.6 \pm 8.7$   $p < 0.05$ ) (Figure 55) suggesting that the community is dominated by true sequences in both methods.



**Figure 55. Comparison of milk microbiota richness (observed OTUs/ASVs) and diversity (inverse Simpson index) between UCLUST and DADA2 with and without contaminant removal in 393 human milk samples from the CHILD Cohort Study. \*  $p < 0.05$ , \*\*\*  $p < 0.001$**

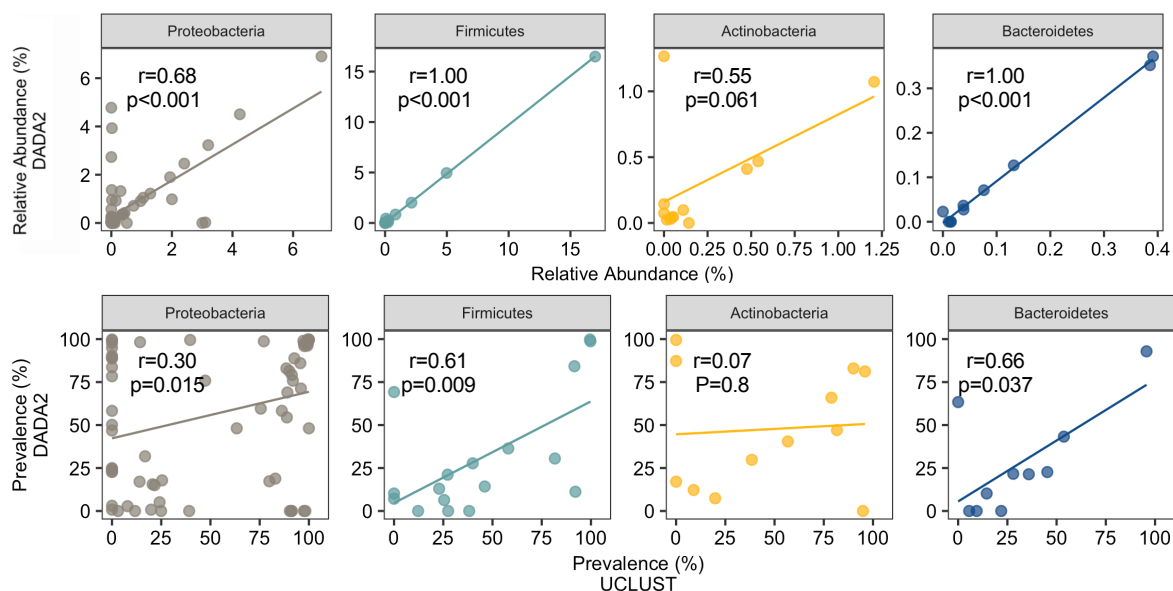
#### 8.4. Milk microbiota structure and relative abundances are not strongly impacted by the choice of sequencing processing method

Next, we compared the relative abundance (Figure 56A) and prevalence (Figure 56B) of genera belonging to the major milk phyla between methods and observed high degrees of correlation especially for Firmicutes, Actinobacteria, and Bacteroidetes. In comparing abundant genera (>0.01% mean relative abundance), relative abundances remained highly correlated (Figure 57).



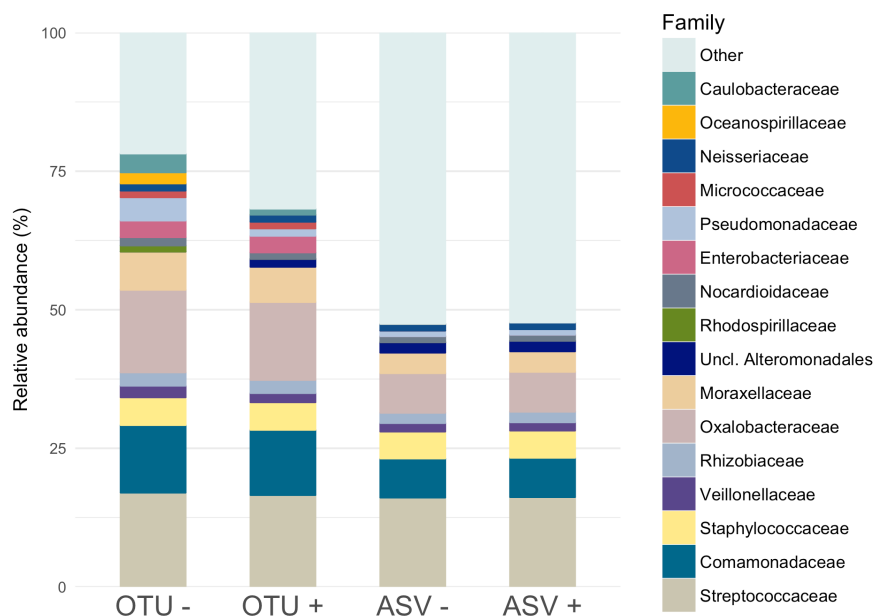
**Figure 56.** Comparison of the A) relative abundance and B) prevalence of bacterial genera on datasets processed by clustering-based OTU method (UCLUST) and a denoising algorithm (DADA2) without contaminant removal in 393 human milk samples from the CHILD Cohort Study. OTUs/ASVs with less than 20 total read count across each dataset and unclassified genera were excluded (n=234 shared genera included in the analysis). Each dot represents a classified genus.





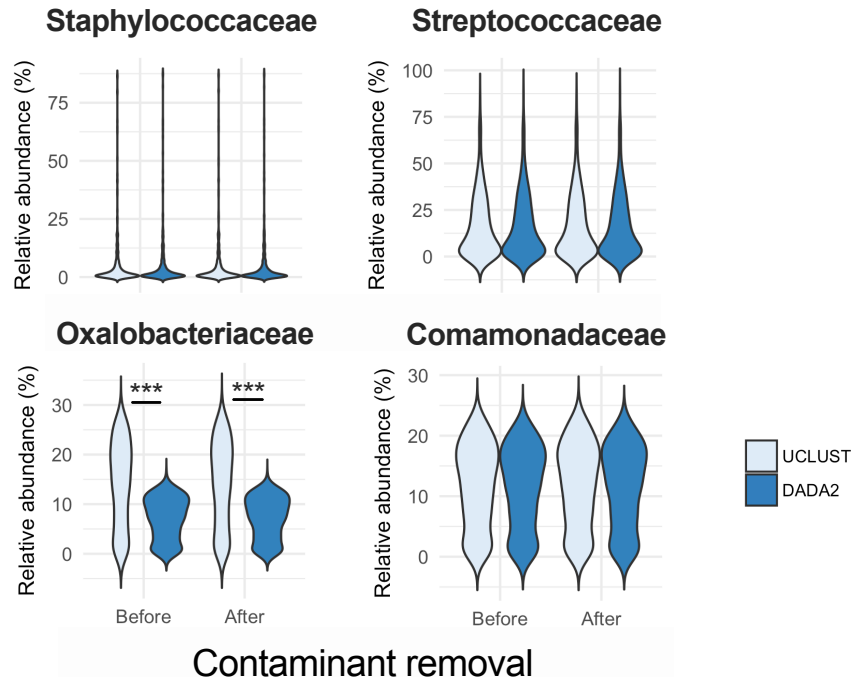
**Figure 57. Comparison of the relative abundance and prevalence of abundant bacterial genera (>0.01% mean relative abundance) on datasets processed by clustering-based OTU method (UCLUST) and a denoising algorithm (DADA2) without contaminant removal in 393 human milk samples from the CHILD Cohort Study. Unclassified genera were excluded (n=68 shared genera included in the analysis).**

Overall, the relative abundances of the most abundant families (>1% mean relative abundance) including *Streptococcaceae* and *Staphylococcaceae* were not considerably impacted by the choice of the method with and without contaminant removal (Figure 58 and Figure 59). However, there were notable differences between methods for some other taxa. For example, at the family level, *Oxalobacteriaceae* was detected at higher relative abundances by UCLUST compared to DADA2 (14.1%±8.4% vs. 7.2%±4.1%, Figure 59), while *Enterobacteriaceae* and *Caulobacteraceae* were only observed as a top abundant family using UCLUST (Figure 58). In addition, *Oxalobacteriaceae* and *Comamonadaceae* were not assigned taxonomy at genus level using UCLUST, whereas some members of these families including *Acidovorax* (family *Comamonadaceae*), *Ralstonia*, and *Massilia* (*Oxalobacteriaceae*) were resolved by DADA2.



**Figure 58. Comparison of the composition of abundant families (>1% mean relative abundance) between UCLUST and DADA2 with or without contaminant removal in 393 human milk samples from the CHILD Cohort Study.** Contaminant removal reduced the relative abundance of certain low-abundance taxa (e.g. *Caulobacteraceae* and *Rhodospirillaceae*) and proportion of Other taxa (OTUs with less than 1% mean relative abundance) estimated by UCLUST, but generally did not affect the microbiota profile estimated by DADA2.

In contrast, *Methylibium* (family *Comamonadaceae*) and *Erwinia* (family *Enterobacteriaceae*) were only identified by UCLUST as abundant taxa. Overall, the total proportion of less abundant OTUs/ASVs (<1% mean relative abundance) was higher in DADA2 (Figure 58) suggesting that the number of true abundant taxa was inflated by UCLUST possibly due to binning of multiple similar sequence variants into an OTU. In agreement with this interpretation, contaminant removal considerably increased this proportion only in UCLUST (Figure 59 and Table 15).

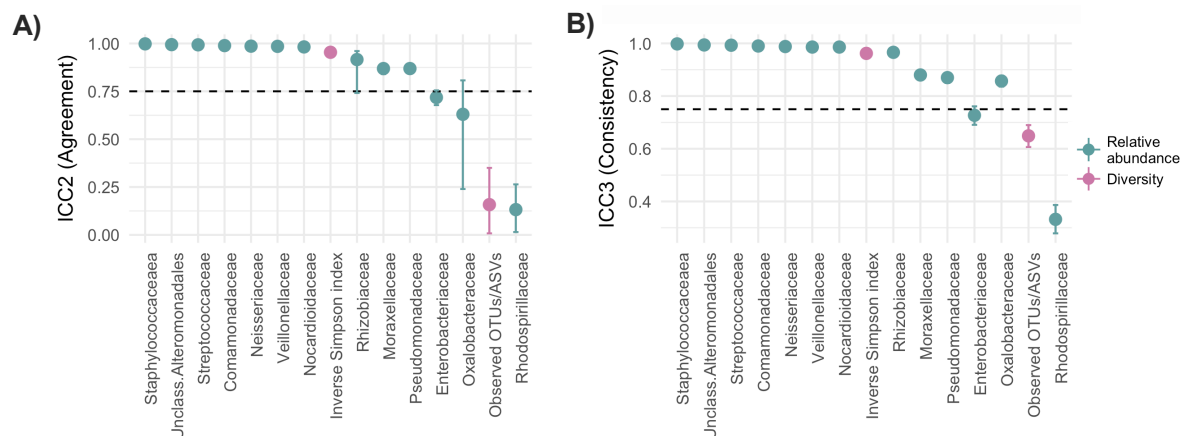


**Figure 59. Comparison of the relative abundances of four abundant families (>1% mean relative abundance) on datasets processed by clustering-based OTU method (UCLUST) and a denoising algorithm (DADA2) with or without contaminant removal in 393 human milk samples from the CHILD Cohort Study. ANOVA test. \*\*\*  $p < 0.001$**

Overall, agreement and consistency between different methods (Figure 60) were quite low for milk microbiota richness, highlighting the sensitivity of this measure to the choice of bioinformatics method. Nevertheless, a very high inter-class correlation for inverse Simpson index (0.95) and relative abundances for the majority of the abundant families (above 0.75 for 10/13 families) suggests an acceptable degree of agreement and consistency, which would be required for downstream analyses to generate comparable results.

### **8.5. Downstream statistical analyses are robust to the choice of data processing approach**

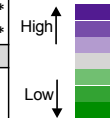
Given the differences in microbiota alpha diversity and taxonomic structure resulting from the choice of processing approaches (Figure 55 and Figure 58), we explored whether the processing approach also influenced the association of microbiota and metadata variables.



**Figure 60. A) Agreement and B) consistency between methods of sequencing processing by intraclass correlation for alpha diversity and 13 most abundant families in 393 human milk samples from the CHILD Cohort Study.**

To do this, we compared the association of maternal, infant, and early life factors, breastfeeding practices, and other milk components with milk microbiota richness, diversity, and overall composition as previously described (139). Overall, there were high degrees of concordance in the direction, strength, and significance of association between milk microbiota diversity and overall composition with the independent variables assessed using both methods, regardless of contaminant removal (Figure 61). However, some method-related differences were observed for associations with microbiota richness. While the direction and strength of associations using DADA2 without contaminant removal were comparable to the UCLUST, contaminant removal generally resulted in weaker associations and lower effect size estimates when using DADA2-processed data (e.g. for maternal atopy, infant sex, and mode of breastfeeding). Occasionally, the direction of association was also different in DADA2/decontaminated compared to the other processing methods (e.g. for prenatal smoking and fatty acids profile) (Figure 61).

| Microbiota features                                  | Richness |     |         |     |         |     |         |     | Diversity |   |         |   |         |   |         |    | Overall composition |     |                |     |                |     |                |     |
|------------------------------------------------------|----------|-----|---------|-----|---------|-----|---------|-----|-----------|---|---------|---|---------|---|---------|----|---------------------|-----|----------------|-----|----------------|-----|----------------|-----|
|                                                      | UCLUST   |     |         |     | DADA2   |     |         |     | UCLUST    |   |         |   | DADA2   |   |         |    | UCLUST              |     |                |     | DADA2          |     |                |     |
|                                                      | -        |     | +       |     | -       |     | +       |     | -         |   | +       |   | -       |   | +       |    | -                   |     | +              |     | -              |     | +              |     |
| Contaminant removal                                  | $\beta$  | p   | $\beta$ | p   | $\beta$ | p   | $\beta$ | p   | $\beta$   | p | $\beta$ | p | $\beta$ | p | $\beta$ | p  | R <sup>2</sup>      | p   | R <sup>2</sup> | p   | R <sup>2</sup> | p   | R <sup>2</sup> | p   |
| <b>Maternal factors</b>                              |          |     |         |     |         |     |         |     |           |   |         |   |         |   |         |    |                     |     |                |     |                |     |                |     |
| Age (years)                                          | -0.37    |     | -0.35   |     | -0.37   |     | -0.23   |     | 0.04      |   | 0.04    |   | 0.04    |   | -0.02   |    | 0.21                |     | 0.20           |     | 0.25           |     | 0.23           |     |
| Pre-pregnancy BMI (Kg/m <sup>2</sup> )               | -0.92    |     | -1.07   |     | -0.92   |     | -0.25   |     | 0.00      |   | -0.01   |   | 0.00    |   | -0.04   |    | 0.55                | *   | 0.60           | *   | 0.43           | *   | 0.45           | *   |
| History of atopy                                     | 22.55    |     | 20.81   |     | 22.55   |     | 7.56    | ~   | 0.99      |   | 0.94    |   | 0.99    |   | 0.97    |    | 0.27                |     | 0.25           |     | 0.26           |     | 0.25           |     |
| Secretor status                                      | 1.25     |     | 0.50    |     | 1.25    |     | 1.46    |     | -1.25     |   | -1.29   |   | -1.25   |   | -1.23   |    | 0.41                |     | 0.40           |     | 0.39           |     | 0.35           |     |
| Ethnicity                                            | -        |     | -       |     | -       |     | -       |     | -         |   | -       |   | -       |   | -       |    | 0.82                |     | 0.80           |     | 0.74           |     | 0.66           |     |
| Asian vs. Caucasian                                  | -26.13   |     | -23.93  |     | -26.13  |     | -9.31   |     | 0.15      |   | 0.18    |   | 0.15    |   | -0.53   |    | -                   |     | -              |     | -              |     | -              |     |
| First Nations vs. Caucasian                          | -28.45   |     | -27.44  |     | -28.45  |     | -0.67   |     | 0.76      |   | 0.80    |   | 0.76    |   | 0.49    |    | -                   |     | -              |     | -              |     | -              |     |
| Other vs. Caucasian                                  | -1.28    |     | -0.27   |     | -1.28   |     | 4.94    |     | -0.30     |   | -0.26   |   | -0.30   |   | 0.16    |    | -                   |     | -              |     | -              |     | -              |     |
| Prenatal smoking                                     | 16.76    |     | 17.33   |     | 16.76   |     | -7.60   |     | -3.21     |   | -3.18   |   | -3.21   |   | -3.32   |    | 0.46                | *   | 0.46           | *   | 0.39           | ~   | 0.48           | *   |
| <b>Infant factors</b>                                |          |     |         |     |         |     |         |     |           |   |         |   |         |   |         |    |                     |     |                |     |                |     |                |     |
| Birth weight (g)                                     | -0.01    |     | -0.02   |     | -0.01   |     | -0.01   |     | 0.00      |   | 0.00    |   | 0.00    |   | 0.00    |    | 0.18                |     | 0.17           |     | 0.21           |     | 0.17           |     |
| Male sex                                             | -29.59   | *   | -27.63  | ~   | -29.59  | *   | -8.24   | ~   | 0.00      | * | -1.90   | * | -1.98   | * | -1.81   | *  | 0.50                | *   | 0.43           | *   | 0.49           | *   | 0.58           | *   |
| Gestational age (weeks)                              | -2.71    |     | -3.09   |     | -2.71   |     | -0.84   |     | 0.03      |   | 0.02    |   | 0.03    |   | 0.00    |    | 0.24                |     | 0.26           |     | 0.21           |     | 0.20           |     |
| <b>Early life factors</b>                            |          |     |         |     |         |     |         |     |           |   |         |   |         |   |         |    |                     |     |                |     |                |     |                |     |
| Mode of delivery                                     | -        |     | -       |     | -       |     | -       |     | -         |   | -       |   | -       |   | -       |    | 0.64                |     | 0.62           |     | 0.59           |     | 0.63           |     |
| C/S elective vs. NVD                                 | 5.76     |     | 6.52    |     | 5.76    |     | 2.61    |     | -2.29     | ~ | -2.25   | ~ | -2.29   | ~ | -2.09   |    | -                   |     | -              |     | -              |     | -              |     |
| C/S emergency vs. NVD                                | -43.21   | ~   | -42.44  | ~   | -43.21  | ~   | -14.96  | *   | -1.49     |   | -1.44   |   | -1.49   |   | -1.74   |    | -                   |     | -              |     | -              |     | -              |     |
| Maternal intrapartum antibiotics                     | -36.81   | *   | -35.84  | *   | -36.81  | *   | -10.89  | *   | -1.67     | ~ | -1.61   | ~ | -1.67   | ~ | -1.86   | *  | 0.21                |     | 0.23           |     | 0.25           |     | 0.25           |     |
| Maternal postpartum antibiotics before 3-4 months    | -5.84    |     | -5.22   |     | -5.84   |     | -5.73   |     | -1.00     |   | -0.96   |   | -1.00   |   | -1.41   |    | 0.17                |     | 0.19           |     | 0.21           |     | 0.45           |     |
| Child antibiotics before 3-4 months                  | -6.17    |     | -5.59   |     | -6.17   |     | -5.48   |     | 1.94      |   | 1.97    |   | 1.94    |   | 1.36    |    | 0.24                |     | 0.23           |     | 0.26           |     | 0.88           | ~   |
| Number of older siblings                             | -        |     | -       |     | -       |     | -       |     | -         |   | -       |   | -       |   | -       |    | 0.88                | *   | 0.86           | *   | 0.87           | *   | 0.92           | *   |
| One vs. None                                         | 0.48     |     | -2.07   |     | 0.48    |     | 3.37    |     | 1.46      |   | 1.36    |   | 1.46    |   | 1.22    |    | -                   |     | -              |     | -              |     | -              |     |
| > Two vs. None                                       | 27.96    |     | 27.61   |     | 27.96   |     | 13.93   | *   | 1.95      |   | 1.95    |   | 1.95    |   | 1.47    |    | -                   |     | -              |     | -              |     | -              |     |
| <b>Breastfeeding</b>                                 |          |     |         |     |         |     |         |     |           |   |         |   |         |   |         |    |                     |     |                |     |                |     |                |     |
| Lactation stage at sample collection (weeks)         | -1.34    |     | -1.29   |     | -1.34   |     | -0.41   |     | -0.06     |   | -0.05   |   | -0.06   |   | -0.05   |    | 0.71                | **  | 0.56           | *   | 0.75           | **  | 0.59           | **  |
| Exclusive BF (breast milk only) at sample collection | 14.74    |     | 15.72   |     | 14.74   |     | 4.90    |     | 0.80      |   | 0.85    |   | 0.80    |   | 0.49    |    | 0.95                | **  | 0.74           | **  | 0.95           | *** | 0.82           | *** |
| Some indirect BF                                     | -52.77   | *** | -50.40  | *** | -52.77  | *** | -19.12  | *** | -1.93     | * | -1.80   | * | -1.93   | * | -2.15   | *  | 1.40                | *** | 1.27           | *** | 1.31           | *** | 1.60           | *** |
| Duration of BF (months)                              | 3.98     | **  | 3.96    | **  | 3.98    | **  | 1.64    | *** | 0.19      | * | 0.19    | * | 0.19    | * | 0.21    | ** | 1.13                | *** | 1.06           | *** | 1.14           | *** | 1.33           | *** |
| Duration of exclusive BF (months)                    | 3.87     |     | 4.25    |     | 3.87    |     | 1.59    |     | 0.37      | ~ | 0.38    | * | 0.37    | ~ | 0.33    |    | 1.16                | *** | 0.98           | *** | 1.13           | *** | 1.14           | *** |
| <b>Milk factors</b>                                  |          |     |         |     |         |     |         |     |           |   |         |   |         |   |         |    |                     |     |                |     |                |     |                |     |
| HMO concentration (mg)                               | 0.71     |     | 0.72    |     | 0.71    |     | 0.33    |     | -0.21     |   | -0.21   |   | -0.21   |   | -0.28   |    | 0.40                |     | 0.34           |     | 0.33           |     | 0.31           |     |
| HMO Simpson's diversity                              | -1.85    |     | -1.84   |     | -1.85   |     | -1.21   |     | 0.27      |   | 0.27    |   | 0.27    |   | 0.23    |    | 0.14                |     | 0.16           |     | 0.17           |     | 0.16           |     |
| HMO PC1                                              | -0.36    |     | -0.16   |     | -0.36   |     | -0.44   |     | 0.66      |   | 0.67    | ~ | 0.66    | ~ | 0.61    |    | 0.41                | ~   | 0.38           | ~   | 0.36           | ~   | 0.37           | ~   |
| MFA PC1                                              | -2.27    |     | -2.72   |     | -2.27   |     | 2.73    |     | 0.17      |   | 0.13    |   | 0.17    |   | 0.37    |    | 0.49                | *   | 0.51           | *   | 0.45           | *   | 0.43           | *   |



**Figure 61. Impact of four sequence processing approaches on observed associations of milk microbiota richness (observed OTUs/ASVs), diversity (inverse Simpson index), and overall composition with maternal, infant, early life, breastfeeding, and milk factors in 393 human milk samples from the CHILD Cohort Study.** Beta coefficients from univariate linear regression (richness and diversity) or R<sup>2</sup> from redundancy analysis (overall composition) are presented and colour coded within each microbiota feature. Results of DADA2 with contaminant removal are originally reported in Moossavi et al. (139). BF, breastfeeding; BMI, body mass index; C/S, Cesarean section; HMO, human milk oligosaccharide; MFA, milk fatty acids; NVD, normal vaginal delivery; PC1, Principal Component 1 \* p<0.05, \*\* p<0.01, \*\*\* p<0.001.

## 8.6. Conclusion

In summary, we have confirmed that a denoising algorithm (DADA2) more accurately resolves the mock community composition compared to a clustering-based method. However, there were high degrees of agreement and consistency in milk microbial features regardless of the methods. Importantly, the epidemiological associations of richness, diversity, and overall composition were not strongly impacted by the choice of the sequencing processing approach. Addition of a contaminant removal step resulted in minor shifts in the composition of the most abundant taxa as well as association of richness with a few of the variables assessed when using DADA2-processed data. The overall conclusions of the study and the main determinants of milk microbiota composition (e.g. consistent association of feeding mode with milk microbiota composition (139)) remained robust to the choice of OTU vs. ASV methods with or without contaminant removal. Our results suggest that previous studies of milk microbiota and potentially other low biomass samples using OTU picking approaches are likely valid specifically for the composition of the most abundant taxa, and the metrics which put less emphasis on richness.

### Box 6. Summary of Chapter 8

- Mock community composition is more accurately resolved using DADA2
- Observed OTUs/ASVs (richness) is influenced by the choice of sequencing processing method
- There was a high degree of agreement and consistency for inverse Simpson index and relative abundances for the majority of the abundant bacterial families
- Downstream statistical analysis was robust to the choice of data processing approach

## 9. DISCUSSION

### 9.1. Summary

I studied the milk microbiota and mycobiota composition in the CHILD Cohort Study using amplicon sequencing (Chapter 3). By accessing the rich metadata from the CHILD Cohort Study and other milk components measured in the same set of samples, I comprehensively investigated the association of maternal, infant, early life, home environment, and milk factors with milk microbiota and mycobiota composition (Chapters 5-7). Additionally, I examined the milk microbiota in a pilot study using culture-enriched molecular profiling (Chapter 4). Finally, I compared the impact of the bioinformatics processing on the results of downstream statistical analysis (Chapter 8).

Overall, there was a good agreement in the milk microbiota composition between culture-independent and culture-enriched methods (Chapter 3 and 4). We observed inter-individual variability in milk microbiota composition in both culture-independent and culture-enriched methods. Notably some abundant taxa detected by culture-independent sequencing (e.g. *Oxalobacteraceae* and *Comamonadaceae*) were not among the most frequently isolated bacteria using culture. These bacteria are generally considered to be reagent contaminants (51, 52) and although some were identified as such using statistical tools (190), the majority were not excluded. Short-term freezing considerably impacted the composition of milk bacteria while we were not able to grow any bacterial isolates from the long-term frozen samples (Chapter 4).

Using multiple analytic techniques, we provide epidemiological evidence that mode of breastfeeding (directly at the breast vs. pumped and bottle-fed) has a significant association with the composition of milk bacteria potentially highlighting the importance of exogenously-derived bacteria either from the infant mouth or the pumping and storage procedure as sources of milk bacteria (131, 139). We further show that the composition of

other major milk components such as fatty acids and HMOs could potentially be associated with the overall composition of the milk microbiota while controlling for relevant confounding factors. However, based on our results, it seems that the individual members of these bioactive molecules are less strongly associated with milk microbiota composition (214).

Additionally, I studied human milk mycobiota in a general population cohort. Defining fungal presence at minimum threshold of 1000 reads per sample, I observed that the majority of milk samples did not contain detectable fungi and I provided evidence that presence of fungi was associated with residential characteristics and milk microbiota composition. Fungi diversity was associated with mode of delivery, maternal atopy, and home environment.

Finally, I compared the milk microbiota composition and associations with maternal, infant, early life and milk factors on data processed by a clustering-based method into OTUs vs. a denoising method resulting in ASVs. Additionally, I compared the impact of contaminant removal on the statistical conclusions. Overall, richness was strongly impacted by the choice of the bioinformatics approach while statistical contaminant removal had minimal additional impact. There was an acceptable agreement and consistency in the relative abundances of the dominant milk bacteria and milk diversity. Additionally, the main conclusions remained robust to the choice of data processing.

To our knowledge, this is among the largest studies of human milk microbiota and mycobiota performed to date, and it is the only milk microbiota study to apply causal modeling and the only milk mycobiota study to examine home environmental characteristics. Moreover, this is among the first studies of milk microbiota culture-enriched molecular profiling. Together, these results considerably expand upon existing knowledge about milk microbiota and mycobiota, providing evidence for novel factors associated with the



composition. However, as the field is new, we have generated more questions that we answered (267).

## **9.2. Origins of milk bacteria**

There are many unanswered questions regarding the biological plausibility of entero-mammary or oro-mammary pathways. For example, does bacterial translocation occur through colonisation of immune cells (268) or other mechanisms (269)? Do bacteria transferred to the mammary gland remain viable? Will those bacteria be released or remain within immune cells upon reaching the mammary gland? Do specific bacteria get selected from the gut or oral microbiota for translocation, and how does this occur? Is the translocation directed specifically to the mammary gland? Is this process specifically stimulated or upregulated during pregnancy and postpartum? If so, how is this process regulated? Do these bacteria permanently colonise the mammary gland and remain viable regardless of lactation? Despite these uncertainties, we reasoned that factors affecting the maternal microbiota should be associated with milk microbiota composition if there is a sub-community originating from maternal gut (or mouth). Indeed, we found that maternal characteristics such as BMI, atopy, and smoking were associated with milk microbiota composition. However, it is equally plausible that these factors influence the mother's skin microbiota or the infant's oral microbiota, which could serve as "external" sources of milk bacteria – a theory supported by other studies demonstrating high similarity between mothers' milk, areola and infant oral microbial compositions (270, 271).

The proposed "retrograde transfer" of external bacteria into the mammary gland is another relatively new area of research. In addition to the areola skin and the infant oral cavity, exogenously-derived bacteria could also originate from breast pumps. This mechanism is less appreciated because historically the majority of milk microbiota studies were concerned with intramammary cow's milk with the primary goal of understanding

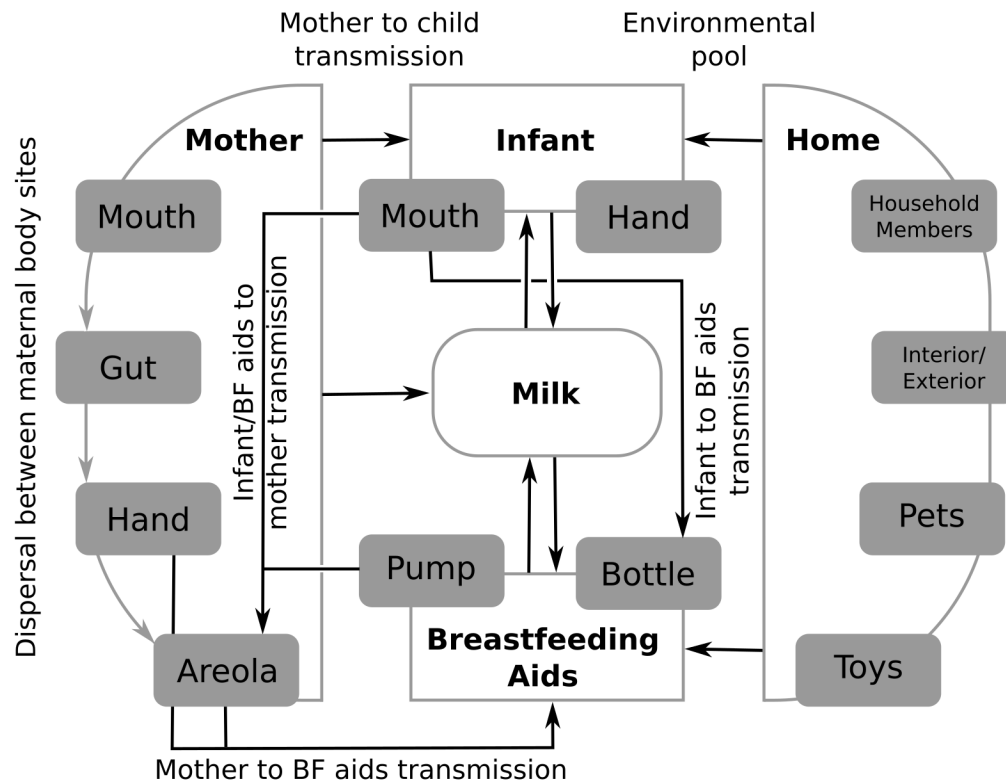
predisposition to mastitis, which is the most costly disease of dairy animals (123). However, especially in human studies, the focus is now shifting to the offspring and the milk that s/he receives, whether that milk is pumped or provided directly from the breast. Reverse flow of milk from the infant mouth back into the breast occurs during breastfeeding (122), and probably pumping as well, allowing for infant mouth- or pump-associated bacteria to enter the milk duct. Our finding that feeding mode (nursing directly from the breast vs. using a pump) was the most consistent factor associated with milk microbiota composition suggests the exogenously-derived bacteria have a stronger role in milk inoculation than the entero-mammary pathway. Further research is needed to confirm whether bacteria actually enter the mammary gland during nursing and/or pumping, or simply inoculate milk as it is expressed.

Based on our research and the existing literature, we propose a slightly modified nomenclature for the possible origins of milk bacteria: the “oro/entero-mammary pathway” (including translocation of both maternal oral and maternal gut bacteria) and the acquisition of “exogenously-derived” bacteria (including maternal skin, infant oral, and pump-associated bacteria) (Figure 3). The extent to which each mechanism influences the milk microbiota load and composition requires further investigation. Additionally, the complex inter-related dispersal routes between different potential sources of milk microbiota should be acknowledged and explored (Figure 62).

### **9.3. Are milk bacteria viable and/or metabolically active?**

It is a topic of interest whether milk bacteria are viable and/or metabolically active. Given that others and we have successfully cultured milk bacteria (either through targeted or high-throughput approaches) (26-33, 272) confirms that at least a subset of milk bacteria are viable. However, it is less clear whether milk bacteria are actively proliferating in the milk within the mammary gland. One metagenomic study of human milk samples did not identify HMO-utilizing pathways in the milk metagenome (44). It is in accordance with our finding

that seems that the individual members of the HMOs as well as fatty acids are not strongly associated with milk microbiota composition (214).



**Figure 62. Different dispersal routes of bacteria between maternal body sites, infant, and breastfeeding aids.** Bacterial dispersal can occur between different maternal body sites and between infant and breastfeeding aids within the context of the environmental pool of species. This complex dispersal network results in bacterial transmission from 1) mother to infant, 2) infant/breastfeeding aids to mother, and 3) mother to breastfeeding aids. Arrows indicate routes of bacterial transmission. BF, Breastfeeding. Reproduced from Moossavi and Azad (131).

This observation could be explained in four different scenarios: 1) milk bacteria are not metabolically active and thus are not utilising or otherwise impacted by these two milk components, 2) milk bacteria are metabolically active but the majority do not have the biochemical capacity to utilise HMO, 3) milk bacteria are metabolically active but fatty acids are bound within glycerol molecules and thus are not freely available to the bacteria, 4) free

fatty acids are available after degradation of glycerol in the infant mouth and retrograde transportation to the milk, however, either the quantities are very low to induce any effect or the effects induced are in bacterial physiology and behaviour and not growth and survival.

Milk is evolutionarily developed to provide nutrition and bioactives to the infant (273-277). Specifically, HMOs enrich for bifidobacteria in the infant gut microbiota (278, 279). It could be considered counterintuitive if the milk bacteria consume the HMOs or other milk components within the milk environment while they have been evolutionarily developed for shaping the microbiota in the infant gut. On the other hand, we speculate that milk can enrich the infant mouth bacteria and provide a protective mechanism for them to bypass the stomach acidity. In this case it is plausible that (a subset of) bacteria will be actively proliferating in the milk. Further research is required to assess the metabolic activity of milk bacteria.

#### **9.4. Determinants of milk microbiota composition**

We extensively studied the influence of maternal, infant, early life, breastfeeding, and other milk factors on milk microbiota composition using redundancy analysis (139) and identified several factors including mode of breastfeeding, lactation stage, and maternal BMI to be associated with the overall composition of the milk microbiota community, albeit with very low redundancy values (each accounting for < 2% of the variation in the milk microbiota) (Figure 29). The cumulative association of all the factors assessed in our study accounted for approximately 25% of the total variation in the milk microbiota composition.

There is great interest in identifying modifiable factors influencing the milk microbiota, with studies to date reporting inconsistent associations for mode of delivery, maternal diet and maternal BMI (40, 43, 97, 135, 280). Maternal diet and BMI are interrelated, and both can modify gut microbiota composition (127, 128), intestinal permeability (281, 282), as well as the macro- and micro-nutrient profile of human milk (213, 283), altering the niche for milk microbiota. Maternal BMI is also positively associated with Caesarean section delivery (284).

In our SEM analysis, maternal diet, maternal BMI and delivery mode did not directly affect milk microbiota. However, controlling for the maternal dietary pattern as a factor either affecting the maternal BMI or directly modulating milk fatty acid composition, we observed that BMI could indirectly (via the effect on the overall milk environment) influence the milk microbiota. Using RDA, we found that mode of delivery and maternal BMI were associated with milk microbiota in female infants only, although the effect sizes were small (<2% of variation explained). We also observed a trend towards lower bacterial richness following emergency Caesarean section after controlling for relevant confounding factors. Overall these associations were relatively subtle and sex-specific, which may explain why previous smaller studies have not found a consistent overall effect of delivery mode on milk microbiota composition.

It is notable that host factors explain only a fraction of the microbiota composition in our study and others (285) suggesting that the factors shaping these microbial communities are not well captured by the usual characteristics of interest, such as mode of delivery and antibiotic use. Although we have examined a few components of the milk niche (HMOs, fatty acids, and hormones), more ecologically-informed parameters (such as Carbon and Nitrogen availability and pH) might be of more direct relevance for future studies. It will also be important to explore the impact of host genetics and immunity. Finally, understanding biotic interactions among bacteria and between bacteria and other microorganisms including fungi and viruses will also be fundamental in understanding the determinants of the milk microbiota composition.

### **9.5. Sex-specific effect of factors on milk microbiota**

Interestingly in our study, milk microbiota diversity was influenced by infant sex and we observed sex-specific differences in the strength and significance of associations between microbiota and other factors. For example, maternal BMI was only significantly associated

with milk microbiota in milk produced for daughters while the variability explained by mode of breastfeeding was higher for sons. In addition, the combined contribution of all the factors accounted for larger proportion of variation in sex-stratified compared to non-stratified analysis, primarily driven by a higher contribution from HMOs and milk fatty acids after stratification. Although not widely studied in modern humans, the Trivers-Willard hypothesis suggests that maternal investment in milk composition differs by offspring sex (286). For example, sex-dependent variations have been observed for milk calcium, cortisol, and protein in *Rhesus macaques* and *Macropus eugenii* (287-290). Fetal sex is also suggested to affect milk production volume in dairy cattle (291), and we have found sex differences in the duration of breastfeeding (possibly reflecting milk supply) in two independent birth cohorts in Canada and the United Kingdom (292). Whether sex differences in the milk microbiota could also be explained in light of the evolutionary biology or cultural practices needs further investigation.

We observed intriguing differences in milk microbiota  $\alpha$  diversity and overall community structure according to infant sex. There was a trend in direct association of infant sex with milk microbiota in our SEM analysis, and we also found sex-dependent associations of breastfeeding mode, exclusivity, and lactation stage (higher  $R^2$  in males), as well as parity, mode of delivery, and maternal BMI (significant in females only) with milk microbiota in RDA. Lack of sex-stratified analysis could potentially explain why previous studies have reported inconclusive results on the association of the above-mentioned factors with milk microbiota.

The mechanism for the observed sex-differences in milk microbiota remains to be determined. Sex differences in gut microbiota have been reported, and are usually attributed to hormonal differences between male and female hosts (293). While the host is always female in the case of milk microbiota, it is suspected (as described above) that milk

microbiota is partially derived from the infant oral cavity, which may differ in male vs. female infants (294). The sex differences we have observed in milk microbiota could therefore be interpreted as additional evidence supporting the retrograde inoculation hypothesis. Other mechanisms for these sex differences are also possible, as sex-dependent variations have been observed in other milk components including calcium, cortisol, and fat (287-289, 295).

However, sex-dependent variations have been observed in the milk calcium and cortisol in *Rhesus macaques* specifically in nulliparous animals (287-289) and milk proteins in *Macropus eugenii* (290). Human milk was found to be higher in fat concentration for daughters in economically-deprived families (295) which was suggestive of Trivers-Willard hypothesis of differences in the mother's investment based on the offspring sex (286). However, similar effect was not detected in Filipino mothers (296). Although, infant sex was not significantly associated with milk microbiota in SEM, we observed sex-dependent association of breastfeeding parameters including mode, exclusivity, duration, and HMO composition with milk microbiota variation in RDA. Whether sex differences in the milk microbiota could also be explained in light of the evolutionary biology needs further investigation.

Several aspects of pregnancy outcome is dependent on fetal sex (297) with a direct interaction with maternal genome in altering cardiometabolic parameters in the susceptible women (298). The milk composition in dairy cattle, fetus sex is suggested to affect milk production (291). We did not observe a sex-dependent difference in HMOs and metabolic hormones in the same set of samples (unpublished). However, sex-dependent variations are observed in calcium and cortisol (288, 289) in rhesus macaques specifically in nulliparous animals (287) and in x and y in *Macropus eugenii* (290). Human milk was found to be richer for daughters in poor families in Kenya (295) which was suggestive of Trivers-Willard

hypothesis (286). However, similar effect was not detected in Filipino mothers (296). Whether sex differences in the milk microbiota could also be explained in light of the evolutionary biology needs further investigation.

#### **9.6. Pumping, milk microbiota and associated controversies**

Although our study was not specifically designed to assess the effect of pumping, we compared milk from mothers who sometimes used a breast pump to express milk for their infant (at least once in the two weeks prior to sample collection) to those who did not feed pumped milk to their infant. Although we did not document nursing frequency, it is reasonable to assume it was lower among mothers who pumped; thus, one could argue that we actually (or also) assessed the impact of less vs. more direct exposure to the infant oral cavity. In other words, our results could reflect an increased exposure to pumps and/or a decreased exposure to the infant mouth. Either way, we observed that the bacterial diversity and composition were slightly different between these two scenarios, providing some evidence for the role of retrograde (exogenous) milk inoculation.

Although some bacterial families containing environmental opportunistic pathogens such as *Stenotrophomonas*, *Pseudomonadaceae*, and *Enterobacteriaceae* were enriched in the milk of mothers who pumped, our observational study could not assess the causal role of pumping, nor the clinical significance of infant exposure to these milk bacteria. It is relevant to note that while they are potential pathogens, *Stenotrophomonas* and *Pseudomonadaceae* are common environmental bacteria, frequently detected in drinking water where they do not cause harm in healthy individuals (299). Similarly, species belonging to the *Enterobacteriaceae* family are normal members of the gastrointestinal microbiota and most strains are harmless to the host (300). Overall, in the absence of strain-level information, absolute quantification, and evidence for colonisation of these bacteria in the infant gut or upper respiratory tract, it is not possible to confirm the clinical significance of these milk



bacteria. Studies designed specifically to address these questions are warranted, given the substantial and increasing use of breast pumps worldwide (124), and evidence that the bacterial content of pump-expressed milk varies considerably depending on the setting, handling and storage of the milk, and cleaning of the pump equipment (126). We are aware of one such study (clinicaltrials.gov, NCT03123874) (141) and eagerly await the results.

Our paper drew considerable media attention to the fact that pumping may impact milk microbiota composition. Some headlines inaccurately stated that pumped milk “causes diseases” or provides “nasty bacteria” to infants, raising understandable concerns from mothers who pump. These unsubstantiated claims oversimplify a complex and emotionally-charged issue. It must be acknowledged that, for a variety of reasons, pumping may be the only way for some mothers to provide their own breast milk to their infant, and this is still beneficial compared to formula feeding. The well-established benefits of breast(milk)feeding for both maternal and infant health (301) occur through multiple different and still incompletely understood mechanisms, many of which involve other non-microbial components of milk that remain intact upon pumping and storage. Still, our research (10, 12) and other studies (302) do suggest that pumped milk is not equivalent to nursing at the breast, so it is important to continue studying the impact of pumping on milk composition in order to understand and optimize this process. Finally, it is also important to undertake knowledge translation efforts with the recognition that supporting breastfeeding is a societal responsibility and not solely the duty of individual mothers.

### **9.7. Culture-dependent milk microbiota analysis**

Microbiome studies are shifting to culture-based isolation of human-associated bacteria allowing phenotypic characterisation and translational analysis (303, 304). The majority of culturomics or culture-enriched molecular profiling studies have been conducted on human stool samples (56-58, 303). In a recently published report, several members of the milk

microbiota including *Streptococcus* spp. and Enterobacteriaceae were frequently isolated using anaerobic culture (272). In our pilot, we successfully isolated major milk bacteria including *Streptococcus*, *Staphylococcus*, and *Bifidobacterium* from fresh milk samples. The number of isolates and the bacterial composition was different for different milk samples in accordance with the inter-individual variability observed in the culture-independent results from this thesis (139).

In our small culture-dependent pilot of pumped milk, we could not isolate bacteria from long-term frozen samples. In a previous targeted culture-dependent study from 2011, approximately 5% of milk samples did not yield any bacterial isolates (33) suggesting that some milk samples may not contain detectable bacteria or have low bacterial cell concentrations in accordance with previous studies (24).

Statistical methods to identify potential reagent contaminants such as decontam (190) are not optimised to detect highly prevalent contaminants, a likely phenomenon in sequencing-based study of microbiota in low biomass samples such as milk (51, 52). In this case, results of culture-enriched molecular profiling are suggested to help identify the true vs. noisy (contaminant) signals in the data (70). In comparing the culture-independent milk microbiota composition in the CHILd Cohort Study to that of the culture-enriched molecular profiling notable observations were made: 1) All the commonly isolated milk bacteria using culture-enriched molecular profiling such as *Staphylococcus* and *Streptococcus* were also highly prevalent in the culture-independent approach in agreement with a previous comparison of culture and qPCR (30); 2) *Bifidobacteriaceae* was detected at the same prevalence in culture and sequencing. This observation was in contrast to a previous report on higher frequency of detection of *Bifidobacterium* by molecular assessment (30); 3) Some of the most prevalent bacterial families in the CHILd Cohort Study such as *Pseudomonadaceae* were isolated at very low prevalence in culture; and 4) Some of the most abundant bacterial

families in the CHILD Cohort Study such as *Comamonadaceae* and *Oxalobacteraceae* were not present in culture and thus are likely reagent contaminants. This observation highlights the limitations of statistical approaches to identify potential reagent contaminants. In agreement with de Goffau et al. (70), I propose that culture-informed milk microbiota studies could enhance the success of identifying and excluding potential reagent contaminants and thereby provide a more accurate representation of the microbiota profile.

Culture-based microbiome analysis main advantage is the isolation and banking of bacterial species from samples (niches) of interest. These bacterial isolates can then be used for various downstream investigations including in-depth genetic and phenotypic characterisations of the bacteria (303, 305), screening of bacteria-bacteria interactions (306, 307) and drug/metabolite-bacteria interactions (308-311), causal studies of the role of the bacteria in human disease in model systems (304). There is a keen interest in exploring probiotic properties of human milk bacteria (119, 312-314). Isolation of bacteria is required for the in vitro and in vivo investigation of probiotic properties of milk microbiota (315-317) and has been previously applied to human milk (318).

#### **9.8. Milk microbiota analyses are robust to the choice of data processing and contaminant removal**

Sequencing-based microbiome studies are highly influenced by the various bioinformatics and data processing choices (319, 320). Specifically, with the recent shift from clustering-based OTU picking methods to denoising algorithms identifying ASVs, it is not clear how the results previously published using OTU-picking methods will compared to the more recent results including ours using ASV methods. Moreover, milk is a low biomass sample and thus highly susceptible to reagent contaminants in culture-independent sequencing-based microbiota profiling (51). To our knowledge, the majority of previously published milk microbiota studies have not assessed the potential reagent contaminants (113).

Therefore, the comparability and generalizability of different studies in terms of the milk microbiota composition and association with maternal and infant characteristics are not known.

In a head-to-head comparison of an OTU-picking method with a denoising algorithm, we observed high degrees of agreement and consistency in milk microbial features regardless of the methods. Importantly, the epidemiological associations of richness, diversity, and overall composition were not strongly impacted by the choice of the sequencing processing approach. Addition of a contaminant removal step resulted in minor shifts in the composition of the most abundant taxa as well as association of richness with a few of the variables assessed when using DADA2-processed data. The overall conclusions of the study and the main determinants of milk microbiota composition (e.g. consistent association of feeding mode with milk microbiota composition (139)) remained robust to the choice of OTU vs. ASV methods with or without contaminant removal. Our results suggest that previous studies of milk microbiota and potentially other low biomass samples using OTU picking approaches are likely valid specifically for the composition of the most abundant taxa, and the metrics which put less emphasis on richness.

## **9.9. Are fungi universally present in milk?**

It is intriguing that fungi were not present in the majority of samples in our study. Previous culture-dependent and -independent studies in dairy animals (102, 321) and humans (31, 38, 98-100) have confirmed the presence of fungi in milk. However, while the majority of animal milk samples contained detectable levels of fungi (102), the prevalence of fungi in human milk has varied from 40% (by culture), to 35-80% (qPCR), and 70-100% (sequencing) (98, 99, 101). The prevalence of fungi in our study was lower than previously reported estimates; potentially due to the higher and more conservative threshold we applied to define fungal presence, other methodological differences (e.g. sequencing depth or target region:

ITS2 in our study and 28S rRNA and ITS1 in previous reports), or true geographical variations between countries (Canada in our study vs. Spain, South Africa, Finland, and China in previous reports). However, our results are consistent with low prevalence of *Candida* in human milk (31, 33) and reports of non-universal fungal presence in the infant gut over the first year of life (95). Given that fungi constitute the rare biosphere of the human microbiota (92), it is conceivable that a greater sampling effort may improve their identification in low biomass samples such as milk.

#### **9.10. Origins and determinants of milk mycobiota**

The origins of milk fungi are unclear and it is not known whether milk is evolved to transfer maternal fungi to the infant. The external (retrograde inoculation) and internal (entero-mammary) pathways hypothesized for the origins of milk bacteria (105, 131) could potentially be extended to milk fungi. However, the supporting evidence in general is lacking. Additionally, it is not clear whether fungi exist in the intramammary milk or only following expression potentially originating from maternal skin. Dominance of milk mycobiota by *Candida* (a common oral fungi (251)) in a subset of samples in our study and others (98, 100, 101) suggest the infant oral cavity as a potential source of milk fungi even in the absence of symptomatic oral thrush (322), while dominance of environmental fungi such as *Penicillium*, *Exophiala*, and *Rhodotorula* in other samples suggest contribution from maternal skin or other environmental surfaces (Figure 63).

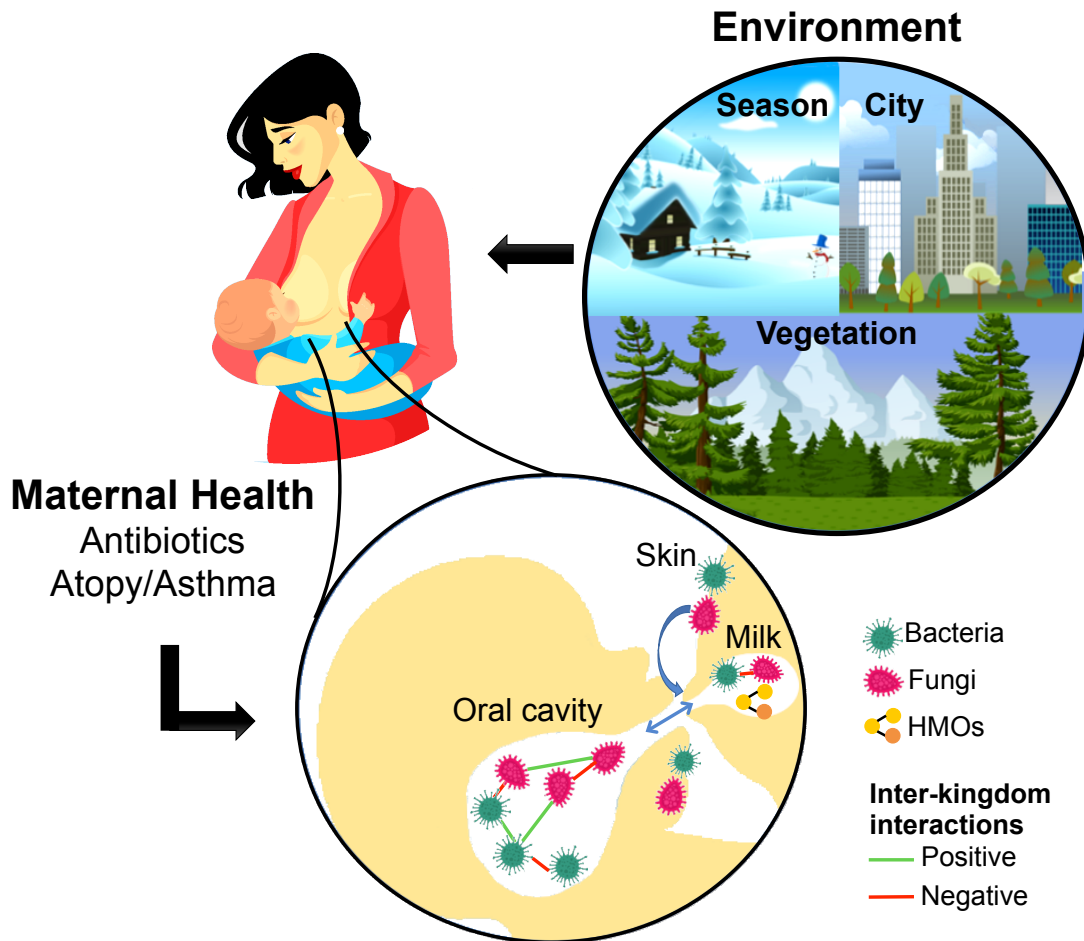
Although previous studies have found that milk collection using a pump, bottle feeding, and frequent pacifier use were associated with increased prevalence of *Candida* spp. in human milk and the infant oral cavity (31, 322), we did not observe any differences in mycobiota presence, diversity or composition according to mode of breastfeeding (nursing at the breast vs. pumped and bottle feeding at least once in the preceding two weeks). It remains

possible that pumping has a transient effect on the milk mycobiota that was not detectable in our study.

We observed that some maternal characteristics such as maternal atopy were associated with mycobiota diversity suggesting that the diversity of milk fungi is influenced by maternal characteristics. The direction of association with maternal atopy warrants further investigation, as it is possible that host immunity influences both atopic sensitization and fungal colonization of the skin and/or milk, or alternatively, that environmental fungi influence both maternal atopy and milk mycobiota (266). We have also observed associations between two specific HMOs and milk fungi, which warrants further exploration. There is preliminary evidence that milk protein content might be associated with milk fungi (98), but the extent to which other milk components including HMOs and immunomodulatory factors influence milk fungi remains to be elucidated.

### **9.11. Regional and seasonal differences in the prevalence of fungi**

The environmental pool of available species is an important factor influencing the formation and composition of microbial communities (323). Similar to Boix-Amorós et al., we observed geographical variations in prevalence of milk fungi with the highest prevalence of fungi observed in milk from mothers residing in Vancouver (a large coastal Canadian city with humid climate, high annual precipitation, and mild winters) compared with smaller Canadian cities in the Prairies with lower average temperature precipitation. This difference was partly explained by differences in population density and season, but was not related to other characteristics of the outdoor environment (e.g. greenness) or home environment (e.g. mould, dust). Prevalence of fungi was lower in milk samples collected in spring consistent with seasonal variations reported in goat milk fungi (261) which could be partly attributed to fluctuations in outdoor (262) and indoor (263) environmental fungi levels.



**Figure 63. Potential sources of milk mycobiota and factors influencing their prevalence and composition.** Environmental factors including season, city, and vegetation can influence the abundance and composition of the available pool of fungal species that colonise the maternal skin and infant oral cavity. Fungi could be transferred to the milk from the skin and/or infant oral cavity via retrograde mechanism. Milk bacterial composition could also influence milk fungi either within the milk environment or via interactions on the skin or in the infant oral cavity. Maternal characteristics and components of the milk microenvironment could also influence milk mycobiota (see Chapter 7).

### 9.12. Bacteria-fungi interaction in the milk: direct or indirect interaction?

We observed associations between bacterial composition and presence and relative abundance of fungi, suggesting that some milk bacterial communities were more permissive to fungal presence, or vice-versa. Correlation of bacterial and fungal richness but not diversity suggests that certain factors may impact microbial dispersal in general, and thus enhance both bacterial and fungal richness, whereas the diversity of each kingdom is

separately determined by the niche relevant biotic and abiotic factors. To our knowledge, only one previous study has examined the potential interaction of fungi and bacteria in milk, finding a positive correlation between the skin inhabitant genus *Malassezia* and bacterial load (98). Our results provide preliminary evidence for antagonistic bacteria-fungi interactions, consistent with evidence that some milk bacteria demonstrate antifungal properties (253, 254), as observed in other environments (255). More research is needed to determine whether this observation is due to active fungi-bacteria ecological interactions within the milk environment, or perhaps the infant oral cavity.

### **9.13. Milk microbiota and other milk components**

Our study uniquely investigated the relationship of milk microbiota and other milk components such as fatty acids, maternal hormones, and HMOs, which may shape the overall milk environment and create the niche for the milk microbiota (153). Using RDA we observed a trend in the association of milk fatty acid and HMO profiles with milk microbiota composition. However, using SEM we did not detect a directional or correlational association of the overall milk environment with the milk microbiota composition, perhaps because our study lacked information on other key components of the milk environment (e.g. immune cells, cytokines and micro- and macro-nutrients). Further research is warranted to explore the interaction between milk microbiota and other milk components, and to study their combined impact on infant development.

To our knowledge, this is the first integrative analysis of human milk microbiota with HMOs and milk fatty acids. We applied multiple statistical approaches to combine and simultaneously analyse the association of these milk components with milk microbiota while controlling for important covariates reflecting maternal genetics, diet, and breastfeeding practices. Overall, the correlations between milk components within each group were stronger and more frequent than correlations between components from different groups.



Still, a few weak associations were observed between individual HMOs or fatty acids and microbial taxa, including *Bifidobacterium*. Interestingly, some of these associations were modified by maternal secretor status, fish oil supplementation, or method of breastfeeding. In structural equation models controlling for these and other key factors influencing milk composition, the overall fatty acid profile (but not the overall HMO profile) remained directly associated with milk microbiota composition.

Our study contributes new evidence on the potential interactions between HMOs and milk bacteria. Some bacterial genera are known to either utilise HMOs (e.g. *Bifidobacterium*, *Bacteroides*, and *Lactobacillus*) or benefit from their growth-promoting effect without direct utilisation (e.g. *Staphylococcus*) (152, 324, 325). We identified correlations between certain HMOs and Bacteroidetes, *Bifidobacterium*, and *Staphylococcus*. The correlations were relatively weak and occurred in both positive and negative directions, in agreement with the different HMO consumption characteristics of these bacteria (152, 324). However, these correlations were different from previous small-scale q-PCR-based (151, 156) and sequencing-based studies (153, 154, 156). Notably, Bacteroidetes relative abundance was higher in milk from secretor mothers suggesting preferential utilisation of  $\alpha$ -1,2 HMOs. Lack of major shifts in the mature milk microbiota by maternal secretor status was in accordance with previous studies (155, 156). Interestingly, while nearly all core milk microbiota demonstrated specific and inverse correlation patterns with fucosylated and sialylated HMOs (which could indicate utilization or inhibition), these relationships were inversed for *Staphylococcus*. Although *Staphylococcus* does not utilise HMOs, it could still be influenced by the presence of HMOs in the microenvironment (152). Finally, HMO 3'SL and DSLNT ( $\alpha$ 2-3-sialylated HMOs) were significantly associated with the overall composition of the milk microbiota(326), suggesting that some HMOs might indeed shape milk microbiota. Interestingly, some *Bifidobacterium* spp. found in milk have sialidase enzymes with activity

on  $\alpha$ 2-3 linkages, enabling them to release free sialic acid from these particular HMOs (326). However, whether these HMOs are metabolized by the majority of milk bacteria is not known. Given the large abundance of lactose as a more accessible carbon source in milk, it is also not known if HMOs are a major source of energy for milk bacteria.

#### **9.14. HMOs and milk *Bifidobacterium***

We specifically investigated *Bifidobacterium* because this species is predominant in the gut of breastfed infants, yet the source of this bacterium is unclear. Interestingly, only 39% of human milk samples in our study contained *Bifidobacterium*, which is curious in light of the current consensus that *Bifidobacterium* is vertically transferred via breastmilk to ‘seed’ the infant gut. Notably, our samples were collected at 3-4 months postpartum, so it remains possible that bifidobacteria are more ubiquitously present earlier in lactation. Previous studies examining strains of *Bifidobacterium* in infant gut, mother’s milk, and maternal gut have concluded that they could be transferred from mother to infant through milk (258, 327). If this is true, HMOs could provide a mechanism for selection and enrichment of *Bifidobacterium* spp. in human milk as it is being produced in the mammary gland. Indeed, we found significant associations between a few HMOs or HMO types and the presence or relative abundance of *Bifidobacterium* in milk, suggesting that certain HMO profiles could perhaps generate a favourable niche for this bacterium in milk by providing a preferential substrate for specific *Bifidobacterium* species. However, this hypothesis requires further investigation since *Bifidobacterium* and HMO metabolism pathways were not detected in a small-scale shotgun metagenomic study of the milk microbiota (44), and to our knowledge, there is currently no evidence that bacteria actively metabolize HMOs in the mammary gland.

#### **9.15. Milk fatty acids and milk microbiota**

We also provide evidence for a potential association between milk fatty acids and microbiota. The majority of studies examining fatty acids and microbial communities have

focused on microbial production of short chain fatty acids in the gastrointestinal tract (328), whereas we explored associations between medium and long chain fatty acids and microbiota in milk. We observed a significant association of the fatty acids profile with the overall microbiota composition. However, this pattern emerged only when examining the summary metric of fatty acids profile (*i.e.* PC1) and controlling for HMO profile and confounding variables, potentially indicating a subtle impact of the fatty acids on microbiota. We also identified a few specific fatty acids to be associated with microbiota  $\alpha$  diversity or overall composition. Stearic acid (18:0) was the fourth most abundant fatty acid in our study and was significantly associated with the overall composition of the milk microbiota independently of confounding factors. The causative mechanism for these associations is unclear since some fatty acids have antibacterial properties (159, 160) while others are produced or consumed by the microbial community. Pathways are described for recognition and uptake of exogenous fatty acids by bacteria. Internalised fatty acids will be metabolized and utilized as transcriptional regulators, source of energy, or membrane phospholipids (329). Moreover, some fatty acids are considered as potential prebiotics (330). Even though we did not observe many correlations of milk fatty acids with milk microbiota, it is possible that fatty acids alter bacterial physiology and virulence (331-334) rather than growth kinetics. Alternatively, milk fatty acids could modulate milk immune cell concentrations (153) and hence indirectly affect milk microbiota. Generally, it is unclear if milk fatty acids, which are largely bound to glycerol, are readily accessible to milk bacteria. It is possible that some fatty acids are released in the infant oral cavity by lingual lipases (335) and thus either locally affect the oral bacteria before retrograde inoculation of the mammary gland, or alternatively get transferred to the intramammary region with the retrograde milk flow during breastfeeding.

### **9.16. Functional importance of milk microbiota**

Breastfeeding has an important role in infant gut microbiota development and maturation (4). This effect is mainly believed to be mediated via other components of the milk, most prominently HMOs that remain intact upon pumping and storage. There is limited evidence in animals that mode of breastfeeding might impact the offspring gut microbiota composition (336). However, the extent to which milk bacteria contribute to the infant gut microbiota is currently unknown. We had previously observed higher risk of possible asthma at the age of 3 years with some pumped milk compared to direct breastfeeding (10) and thus, speculated that this observation might be due to differences in the milk microbiota composition. We are currently exploring these associations in the CHILD Cohort Study.

The functional importance of the milk microbiota is poorly understood. In the dairy industry, the focus of milk microbiota research is heavily geared towards udder health and prevention of mastitis (123). In humans, it is often argued that breast milk not only provides nutrients to the infant, but also serves as a source of probiotic bacteria contributing to the establishment of the infant gut microbiota (188). This hypothesis is largely based on comparative analyses identifying a few shared taxa between maternal milk and infant stool that are absent in the infant oral cavity (258, 327). However, computational estimations suggest that only 20% of taxa in the infant gut microbiota are shared with breast milk (189). The infant oral microbiota is another possible source of seeding bacteria for the infant gut (337, 338). It is suggested that oral microbiota could reach distal segments of the GI tract directly through swallowing saliva, or indirectly through retrograde inoculation of breast milk (258). There is also a theoretical possibility of bacterial exchange in the opposite direction, with milk bacteria seeding the oral cavity (258). Moreover, aspirations during breastfeeding could deliver bacteria to the infant respiratory system (339). Thus, through multiple possible mechanisms and origins, milk bacteria may colonize and/or functionally influence the infant

gut, oral cavity and respiratory tract. We did not assess the infant microbiota in our milk microbiota study (139), but these data are available for the CHILD Cohort Study and we are currently undertaking an integrated analysis of milk, stool and nasal microbiota.

There has been a major focus on *Bifidobacterium* as the most abundant genus in the infant gut and the best-studied bacterium in terms of HMO utilization (340). Assuming *Bifidobacterium* spp. are evolutionarily selected to be transferred to the infant (341, 342), the evidence is scarce that this transfer has evolved to be exclusively from the mother (vertical transfer) or to occur via breast milk. Acknowledging the limits of detection by current methods, *Bifidobacterium* spp. were detected at low frequency by culture and sequencing in previous milk microbiota studies (214, 343), as well as our study where only 40% of milk samples contained amplicon sequencing variants from this genus (214). While vertical transfer of obligate symbionts occurs in some holobionts (e.g. *Buchnera aphidicola* in aphids), in other cases they can be horizontally transferred from the environment (e.g. *Vibrio fischeri* in squid) or from other members of the host population (e.g. core Gram negative bacteria in honey bees) (344). Similarly in humans, horizontal transfer of bifidobacteria and other bacteria from the environment and other household members likely contributes to the development of the infant gut microbiota (345), regardless of breastfeeding. A fresh outlook based on principles of holobiont ecology (346) is required to better understand the functional importance of milk in vertical and horizontal bacterial transmission.

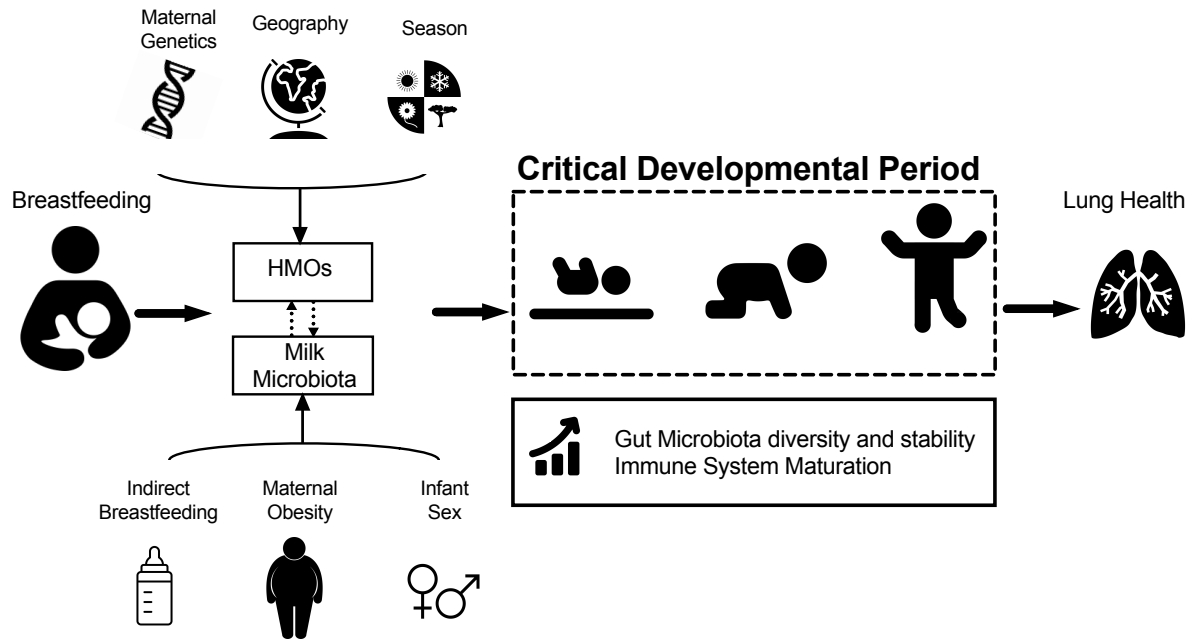
Developmental origin of health and disease hypothesis stipulates that early life events have long-lasting health impacts on the infant (347, 348). Infant gut microbiota develops through a complex process of stepwise successions beginning at birth (349). Priority effect hypothesis suggests that the final composition of the infant gut microbiota depend on the type and order in which organisms are colonised in the intestinal tract (350). Breastfeeding is by far the most influential factors shaping the infant gut microbiota (4, 258, 327). Milk

microbiota could provide the pioneering species and impact the final composition through the mechanism of priority effect (188). Strain similarities between maternal and infant gut and the milk microbiota in human studies support this hypothesis (327). It is plausible, that the distinct profile of milk microbiota could partly be responsible for modulating the infant gut microbiota taxonomic and functional diversity contributing to the overall metabolic homeostasis.

Milk bacterial isolates have different levels of resistance to conditions mimicking passage through the gastrointestinal tract (i.e. pepsin at pH 2 representing stomach and bile and pancreatin at pH 8 representing small intestine). For example, *Lactobacillus rhamnosus* and *Enterococcus* isolated from milk were sensitive to low pH while being tolerant to intestinal conditions whereas *L. gasseri* survived low pH (32). If milk bacteria are indeed seeding the infant gut, they need to be able to survive the harsh chemical conditions of different segments of the gastrointestinal tract. It is speculated that although milk bacterial isolates are vulnerable to pH and bile salt fluctuations, milk could provide a protective vehicle for safe passage of bacteria to distal segments of the infant gut.

Breastfeeding has been associated with lower risk of non-communicable diseases/conditions such as asthma and obesity (8-14). Although, the results are contradictory in different studies the overall trend seems to be protective especially in early life (8, 351). Health impacts of breastfeeding on the infant are suggested to be partly mediated via its impact on the infant gut microbiota development (352-354), which is implicated in predisposition to pediatric asthma and allergy (355). The beneficial effects of breastfeeding on infant health could potentially be mediated by HMOs and milk microbiota which modulate the infant gut microbiota composition (188); and milk immune components (e.g. cytokines, antibodies, and immune cells) which confer immunity to the infant in the absence of a mature immune function (356). Variations in HMOs have been associated with food

sensitization (357, 358) and asthma (Our group, unpublished). Currently there is no evidence on the association of milk microbiota and mycobiota with infant health outcomes including asthma. However, infant's gut and upper respiratory tract microbiota composition have been associated with asthma (352-354, 359-361). Additionally, early life exposure to fungi has been associated with pediatric asthma in farm and urban environments (94, 362, 363). If breastfeeding is indeed modulating the infant microbiota development by enriching, protecting, and delivering bacteria and potentially fungi to the infant respiratory and gastrointestinal tract (18), it is plausible that alterations in the composition of the milk microbiota and mycobiota could lead to an altered trajectories of respiratory and gut microbiota development in the infant specifically, if the milk is depleted in beneficial microbes or enriched in environmental bacteria or fungi and potential pathogens. This thesis has shown that mode of breastfeeding (pumped and bottled vs. directly at the breast) affects the composition of milk microbiota reducing the overall alpha diversity and enriching potential pathogens including *Sphingomonadaceae* and *Enterobacteriaceae* (139). Additionally, an increased risk of developing asthma due to bottle feeding has been observed in the same population (10). It is plausible that the risk-modifying effect of bottle-feeding is partially mediated via its impact on the milk microbiota. Given the discordant prevalence of atopy across study sites in CHILD cohort (364), and higher prevalence of milk fungi in Vancouver compared to other sites (267), it is also conceivable that milk fungi could influence the risk of asthma (Figure 64).



**Figure 64. Association of breastfeeding with lung health and asthma.** Breastfeeding influences infant gut microbiota development and stability during the critical developmental period in early life via two of its main components: human milk oligosaccharides (HMOs) and milk microbiota. HMOs exert a selective pressure within the infant gut microbial niche, preferentially promoting the proliferation of specific bacteria including bifidobacteria. Milk is also a source of viable bacteria originating from the maternal gut and infant oral cavity. HMO composition is influenced by maternal genetics, geography, and season while microbiota is affected by maternal weight status, mode of breastfeeding and infant sex. Variations in HMOs and milk microbiota could modulate the effect of breastfeeding on the infant gut microbiota, which in turn shapes the infant immune system and could ultimately influence lung health and development. Reproduced from Moossavi et al. (188).



## **10. OVERALL CONCLUSION AND FUTURE DIRECTION**

### **10.1. Conclusion**

In this large and comprehensive study, we profiled milk microbiota and mycobiota using culture-independent amplicon sequencing in the CHILD Cohort Study and conducted a pilot study of culture-enriched molecular profiling of milk microbiota. We have used multivariable approaches to explore many features and potential determinants of human milk microbiota and mycobiota composition. Our results suggest that there is an agreement between culture-independent and culture-enriched milk microbiota composition. Additionally, we showed that multiple maternal, infant and environmental factors interactively influence milk microbiota composition. Most strikingly, indirect breastfeeding and pump-expression were consistently associated with milk microbiota composition highlighting the importance of breastfeeding practices. Interesting sex-differences were also identified, as well as potential associations between microbiota and other milk components. Further research is warranted to replicate these findings in other populations and explore their implications for infant health and development.

The fact that human milk contains bacteria, even when collected aseptically, has been known since the 1920s (365). Yet, milk bacteria are often perceived in an unfavourable light, considered as contaminants introduced during handling, processing, and storage (19). The recent paradigm shift regarding the significance of the host-associated microbiome including milk microbiota has resulted in a new appreciation of the milk bacteria (20). The culture-independent approach to study milk microbiota is relatively new and poses unique challenges compared to other host-associated microbial niches with higher biomass, such as the gastrointestinal tract. Nevertheless, these methods offer new opportunities to study and understand the microbes present in human milk, identify their sources, mechanistically

dissect their role in maternal and infant health, and help inform new microbiome-targeted strategies for health promotion and disease prevention.

Research over the past decade has demonstrated a diverse community of bacterial species in human milk (17). The initial pioneering studies focused primarily on describing the milk microbiota rather than understanding the maternal, infant, milk, and breastfeeding factors that influence its diversity and composition. In accordance with previous studies, I showed that milk microbiota taxonomic composition was highly variable between individuals. By applying multiple statistical methods and theory-driven causal modelling, I provided evidence that mode of breast milk feeding (nursing directly at the breast vs. pumping and feeding breast milk from a bottle) was significantly associated with milk microbiota composition, more than any other factor examined. These findings have improved our understanding of milk bacteria, their origins, and their potential impact on infant health, and raised many new questions for further investigation.

Our results indicate that fixed and modifiable factors can influence the milk microbiota in a sex-specific manner. Whether milk provides bacteria to colonize the infant gut or merely provides selective nutrients to foster a permissive environment is still an open question. Further investigation is needed to determine if milk effectively transfers maternal microbiota or simply enriches and protects the infant oral microbiota for a safe passage to the infant's distal gut. Studies are also needed to identify sources of exogenously-derived bacteria, characterize other elements of the milk microbiota, including viruses, and determine the impact of pumping on the microbiota of expressed milk. The impact of milk microbiota on infant gut microbiota development and health remains to be explored, and could have important implications for microbiota-based strategies for early life prevention of chronic conditions.

As denoising methods for processing amplicon sequencing data become increasingly common, it is important to understand their potential impact on estimating microbiota composition and biological associations, and reconsider the interpretability of previous studies. Our results suggest that although microbiota richness and some members of the microbial community are strongly influenced by the choice of sequencing processing method, there is high agreement and consistency between methods for estimating microbiota diversity and quantifying the majority of the most abundant taxa. In light of our observation that the associations with metadata and the main conclusions were robust to the choice of sequence processing approaches and contamination removal, previous studies of milk microbiota and potentially other low biomass samples using OTU picking approaches are likely valid both in terms of the composition of the abundant taxa and associations, especially for metrics that put less emphasis on richness. The overall conclusions are summarised in Box 7.

#### **Box 7. Summary of the overall conclusions**

- Human milk contains a diverse community of bacteria profiled by both culture-independent (CHILD study) and culture-enriched (pilot study) molecular profiling
- Human milk also contains fungi albeit at lower prevalence in the CHILD study.
- Inter-kingdom interactions exist between milk microbiota and mycobiota.
- Milk contains viable bacteria that could be isolated by culture-enriched molecular profiling.
- Bacterial profiles of different milk fractions are different.
- There is a good agreement in the most prevalent/abundant bacteria between culture-independent (CHILD study) and culture-enriched (pilot study) microbiome.
- Several maternal, infant, milk, and environmental factors are associated with milk microbiota and mycobiota compositions in the CHILD study.
- Mode of breastfeeding has the strongest influence on milk microbiota while environmental factors are the most prominent factors shaping milk mycobiota.
- Choice of sequencing processing approach with or without contaminant removal does not impact the main statistical conclusions related to milk microbiota determinants.

## **10.2. Strengths**

The main strengths of this study are dual culture-independent profiling of microbiota and mycobiota in the CHILD Cohort Study complemented with a pilot culture-enriched molecular profiling of milk microbiota. Culture-dependent approach confirmed the composition of milk microbiota identified in the CHILD Cohort Study. The major strength of the CHILD project is our unique multi-variable and multi-method approach to assessing the effect of diverse maternal, infant, early life, environmental, and milk factors on the milk microbiota and mycobiota, using the rich data and large sample size afforded by the CHILD Cohort Study. The comprehensive statistical approach undertaken in this thesis has significantly improved our understanding of the determinants of milk microbiota and mycobiota composition. Additionally, an integrative, multi-variable approach was adapted to assess the associations between major milk components including HMOs and fatty acids and milk microbiota. In this study, we were able to control for maternal intake of fish oil supplement and secretor status as key factors influencing fatty acids and HMOs, respectively. Notably, however, the many factors we evaluated collectively explained less than a third of the total variation observed in milk microbiota composition, indicating that other unmeasured factors such as milk immune factors and host genetics are potentially contributing to the large inter-individual variation in milk microbiota profiles.

## **10.3. Limitations**

The main limitation of our study is that milk samples in the CHILD Cohort Study were pooled from multiple feeds and were not collected aseptically, in order to reflect the average composition of milk fed to each infant. While this collection protocol precludes analysis of potential diurnal variations and limits our ability to strictly study the milk microbiota (without skin or other contaminants), it provides an accurate representation of the microbiota that infants ingest, which is arguably more biologically relevant to infant health outcomes.

Although refrigeration could potentially impact the microbial profile of the milk samples (366), we did not find any association between sample processing time and milk microbiota composition. We collected a single sample from each mother, so we could not examine longitudinal changes in milk microbiota composition over time. Although we identified potential reagent contaminants at the sequencing step, sequencing results were not available for DNA extraction negative controls and thus we were not able to identify and remove potential contaminants which might have been introduced during the extraction. As 16S rRNA gene sequencing has limited capacity to resolve the taxa to species and strain levels, further metagenomic studies are required to confirm and validate the results of this study. Information on mastitis, which could be associated with shifts in milk microbiota and fatty acid compositions (42, 367), was not available in our study, and absolute concentrations of fatty acids were not available. Other milk components such as antimicrobial peptides and antibodies, which could confound the association of HMOs and fatty acids with milk microbiota, were not available for our samples. Finally, as with all culture-independent microbiota studies, we could not quantify bacterial load nor confirm the viability of bacteria identified in our samples (312).

#### **10.4. Future direction**

##### ***10.4.1. Scientific considerations***

Milk microbiota is a hot research topic in the field of developmental origins of health and disease. Building on earlier studies, our recent findings shed new light on the composition and determinants of the milk microbiota, and raise new research questions about the origins, function and analysis of milk bacteria (Box 8).

### Box 8. Unanswered questions about milk microbiota

- Do specific bacteria get selected from the maternal gut and/or oral microbiota for translocation to the mammary gland? How does this occur?
  - Is the translocation specifically directed to the mammary gland?
  - Do bacteria transferred to the mammary gland remain viable?
  - Are these bacteria released or do they remain within immune cells upon reaching the mammary gland?
  - Do these bacteria permanently colonise the mammary gland and remain viable regardless of lactation?
  - Is this process specifically stimulated or upregulated during pregnancy and postpartum? If so, how is this process regulated?
- Do external bacteria enter the mammary gland through the nipple? How does this occur?
  - Are infant oral bacteria transferred into the mammary gland during suckling?
  - Are bacteria from breast pumps transferred into the mammary gland during pumping?
  - Does the mechanical force of pumping affect the intramammary milk microbiota?
- How do breast pump cleaning practices and milk storage conditions affect the microbiota of expressed milk?
- How do the qualities and mechanics of different breast pumps and accessories affect the microbiota of expressed milk?
- What best practices should be applied to milk microbiota analysis to promote accurate and reproducible results (e.g. collection, storage, processing, DNA extraction, and sequencing quality controls)?
- How does milk microbiota composition vary throughout lactation?
- How do other microorganisms (e.g. fungi and viruses) contribute to milk microbial communities?
- How do other milk components and characteristics influence the milk microbiota?
- Do milk microbes colonize the infant gut?
- What is the relative importance of the prebiotic (e.g. milk oligosaccharides) vs. probiotic (i.e. live microorganisms) components of human milk?

So far, milk microbiota studies have mainly been observational cross-sectional studies based on 16S rRNA amplicon sequencing, a method that is susceptible to reagent contamination when analyzing low biomass samples such as milk, and cannot distinguish between live or dead bacteria. Better understanding of the dynamics and function of milk

microbiota requires a comprehensive multi-pronged strategy that 1) assesses viability, activity, and function of milk bacteria; 2) examines the temporality of the microbial load and composition throughout lactation; 3) studies other members of the microbial community including viruses, which can conceivably be vertically transferred more readily than the bacteria (341), and fungi, which could have important health implications; 4) evaluates the interaction between milk microbiota with the maternal and infant immune systems; and 5) experimentally establishes the functional significance of milk microbiota. In addition, the origins of milk bacteria should be rigorously assessed using appropriate experimental study designs and suitable animal models that distinguish between intramammary milk and the expressed milk ingested by the offspring.

#### ***10.4.2. Technical considerations***

- Best practices in microbiome studies specifically for low biomass samples should include: 1) negative controls for culture, DNA extraction, and sequencing library preparation, 2) sequencing sufficient numbers of negative controls, 3) biological replicates in different runs, and 4) limit of detection controls using serial dilution.
- I propose that culture-informed milk microbiota studies could enhance the success of identification and exclusion of potential reagent contaminants and thereby provide a more accurate representation of the microbiota profile.
- The current biobanking gold standard for microbiome research aims to preserve the original bacterial profile for downstream DNA and/or RNA-based sequencing. Frozen stool samples have been successfully used for culturomics (64). However, I could not isolate bacteria from long-term frozen milk samples which could potentially be due to variability in the bacterial load of milk samples and the overall low biomass of milk microbiota. Although, the results of this thesis should be verified in larger studies and milk samples of mothers of healthy term infant, it is recommended that milk samples be stored using cryopreservatives

such as glycerol to maintain the viability of bacteria for future culture-dependent downstream investigation.

- There is a large variation in the absolute concentration of the total bacterial load in milk (24, 25). Differences in microbiome composition based on relative vs. absolute quantification have been previously shown (368). The more accurate absolute quantification was specifically important in identifying bacteria associated with health outcomes (369) highlighting the importance of absolute quantification in the study of the milk microbiota composition specifically in relation to infant health outcomes.
- Bacteria exhibit a considerable amount of heterogeneity in metabolic activity (370, 371). The activity of bacteria could be assessed based on various cellular processes including proliferation (372). By experimentally separating bacteria into multiple activity groups using appropriate targeted labels prior to sequencing, it is possible to profile the microbiota based on the activity levels. This approach has been successfully applied to the study of gut microbiota (373).

#### ***10.4.3. Practical and translational considerations***

Study of human milk microbiota has two immediate and potentially highly important applications.

- It must be acknowledged that, for a variety of reasons, pumping may be the only way for some mothers to provide their own breast milk to their infant, and this is still beneficial compared to formula feeding. We have shown that pumping has the most consistent impact on the milk microbiota composition potentially introducing environmental opportunistic bacteria into the milk (139). While the significance of this observation for the health of the infant is not clear, it could facilitate renewed discussions about the choice of the pump (e.g. tubing material resistant to biofilm formation) and cleaning practices among the stakeholders.



- The current standard in processing donor milk in human milk banks involves culture-based bacteriological assessment for the common pathogens (including group B *Streptococcus*, *Staphylococcus aureus*, coagulase-negative *Staphylococcus*,  $\alpha$ -*Streptococcus*, *Enterococcus*, and *Bacillus* sp.) followed by pasteurization (62.5 °C for 30 min) which eliminates most milk bacteria except the spore-forming *Bacillus* species (374). Subsequently, donor milk will be excluded if any of the following criteria are met: 1) isolation of any pathogen, 2) non-pathogenic bacterial load  $> 10^4$  colony-forming units/mL before pasteurization, or 3) any positive culture after pasteurization (375). Donor milks are critically important for the premature infants admitted to NICU who are especially vulnerable to microbiome-mediated conditions, which can be fatal, including necrotizing enterocolitis (376). Breastmilk provides optimal nutrition for premature infants, and human donor milk is the best alternative when mothers cannot provide their own milk (377). While donor milk processing aims to minimise the risk of pathogen transmission to the highly vulnerable infant, it also eliminates the non-pathogenic components of the milk microbiota. While the role of milk microbiota in infant health is not elucidated, it is conceivable that milk microbiota influences the development of the infant respiratory tract and gut microbiota compositions. It is therefore an open question if reconstitution of the donor milk with the infant's own mother milk microbiota (378, 379) would impact the health outcomes in premature infants.

### 10.5. Closing remark

Based on the results of this thesis and in accordance with the previously published report, some bacterial and fungal groups are commonly detected in human milk of healthy mother-infant dyads. However, the taxonomic structure and the diversity of the milk microbial community are highly individualised. There is interest in defining a normal milk microbiota. However, as I previously argued, it might be impractical to define normal microbiota in the absence of health outcomes (380). While milk microbiology has been

studied since the 19<sup>th</sup> century (381) from an infectious disease perspective, the current appreciation of commensal microbes combined with the increasing availability of novel technologies and methodologies has set the stage for a comprehensive and theory-driven approach to the study of milk microbiota, and the translation of this knowledge to improve the health of the mothers and infants.

## 11. REFERENCES

1. Binns C, Lee M, Low WY. The Long-Term Public Health Benefits of Breastfeeding. *Asia Pac J Public Health*. 2016;28(1):7-14.
2. Azad MB, Konya T, Maughan H, Guttman DS, Field CJ, Chari RS, et al. Gut microbiota of healthy Canadian infants: profiles by mode of delivery and infant diet at 4 months. *CMAJ*. 2013;185(5):373-4.
3. Azad MB, Konya T, Persaud RR, Guttman DS, Chari RS, Field CJ, et al. Impact of maternal intrapartum antibiotics, method of birth and breastfeeding on gut microbiota during the first year of life: a prospective cohort study. *BJOG*. 2016;123:983-93.
4. Stewart CJ, Ajami NJ, O'Brien JL, Hutchinson DS, Smith DP, Wong MC, et al. Temporal development of the gut microbiome in early childhood from the TEDDY study. *Nature*. 2018;562(7728):583-8.
5. Ho NT, Li F, Lee-Sarwar KA, Tun HM, Brown BP, Pannaraj PS, et al. Meta-analysis of effects of exclusive breastfeeding on infant gut microbiota across populations. *Nat Commun*. 2018;9(1):4169.
6. listed] Na. Effect of breastfeeding on infant and child mortality due to infectious diseases in less developed countries: a pooled analysis. *The Lancet*. 2000;355(9202):451-5.
7. Duijts L, Ramadhani MK, Moll HA. Breastfeeding protects against infectious diseases during infancy in industrialized countries. A systematic review. *Maternal & Child Nutrition*. 2009;5(3):199-210.
8. Dogaru CM, Nyffenegger D, Pescatore AM, Spycher BD, Kuehni CE. Breastfeeding and childhood asthma: systematic review and meta-analysis. *Am J Epidemiol*. 2014;179(10):1153-67.
9. Azad MB, Vehling L, Lu Z, Dai D, Subbarao P, Becker AB, et al. Breastfeeding, maternal asthma and wheezing in the first year of life: a longitudinal birth cohort study. *Eur Respir J*. 2017;49:1602019.
10. Klopp A, Vehling L, Becker AB, Subbarao P, Mandhane PJ, Turvey SE, et al. Modes of Infant Feeding and the Risk of Childhood Asthma: A Prospective Birth Cohort Study. *J Pediatr*. 2017;190:192-9.
11. Azad MB. Infant Feeding and the Developmental Origins of Chronic Disease in the CHILD Cohort: Role of Human Milk Bioactives and Gut Microbiota. *Breastfeeding Medicine*. 2019;14(S1):S-22-S-4.
12. Azad MB, Vehling L, Chan D, Klopp A, Nickel NC, McGavock JM, et al. Infant Feeding and Weight Gain: Separating Breast Milk From Breastfeeding and Formula From Food. *Pediatrics*. 2018;142(4):pii: e20181092.

13. Yan J, Liu L, Zhu Y, Huang G, Wang PP. The association between breastfeeding and childhood obesity: a meta-analysis. *BMC Public Health*. 2014;14:1267.
14. Harder T, Bergmann R, Kallischnigg G, Plagemann A. Duration of breastfeeding and risk of overweight: a meta-analysis. *Am J Epidemiol*. 2005;162(5):397-403.
15. Parigi SM, Eldh M, Larssen P, Gabrielsson S, Villablanca EJ. Breast milk and solid food shaping intestinal immunity. *Front Immunol*. 2015;6:415. doi:10.3389/fimmu.2015.00415
16. Gomez-Gallego C, Garcia-Mantrana I, Salminen S, Collado MC. The human milk microbiome and factors influencing its composition and activity. *Semin Fetal Neonatal Med*. 2016;21(6):400-5.
17. Fitzstevens JL, Smith KC, Hagadorn JI, Caimano MJ, Matson AP, Brownell EA. Systematic Review of the Human Milk Microbiota. *Nutr Clin Pract*. 2016;32(3):354-64.
18. Bode L, McGuire M, Rodriguez JM, Geddes DT, Hassiotou F, Hartmann PE, et al. It's Alive: Microbes and Cells in Human Milk and Their Potential Benefits to Mother and Infant. *Adv Nutr*. 2014;5(5):571-3.
19. West PA, Hewitt JH, Murphy OM. The Influence of Methods of Collection and Storage on the Bacteriology of Human Milk. *J Appl Bacteriol*. 1979;46(2):269-77.
20. Vayssier-Taussat M, Albina E, Citti C, Cosson JF, Jacques MA, Lebrun MH, et al. Shifting the paradigm from pathogens to pathobiome: new concepts in the light of meta-omics. *Front Cell Infect Microbiol*. 2014;4:29.
21. Cani PD. Human gut microbiome: hopes, threats and promises. *Gut*. 2018;67(9):1716-25.
22. Skibił AL, Downing LM, Orr TJ, Hood WR. The evolution of the nutrient composition of mammalian milks. *J Anim Ecol*. 2013;82(6):1254-64.
23. Togo A, Dufour JC, Lagier JC, Dubourg G, Raoult D, Million M. Repertoire of human breast and milk microbiota: a systematic review. *Future Microbiol*. 2019;14:623-41.
24. Davidson DC, Poll RA, Roberts C. Bacteriological monitoring of unheated human milk. *Arch Dis Child*. 1979;54(10):760-4.
25. Foster WD, Harris RE. The incidence of *Staphylococcus pyogenes* in normal human breast milk. *J Obstet Gynaecol Br Emp*. 1960;67:463-4.
26. Heikkilä MP, Saris PEJ. Inhibition of *Staphylococcus aureus* by the commensal bacteria of human milk. *Journal of Applied Microbiology*. 2003;95(3):471-8.
27. Tušar T, Žerdoner K, Matijašić BB, Paveljšek D, Benedik E, Bratanič B, et al. Cultivable Bacteria from Milk from Slovenian Breastfeeding Mothers. *Food Technol Biotechnol*. 2014;52(2):242-7

28. Jost T, Lacroix C, Braegger C, Chassard C. Assessment of bacterial diversity in breast milk using culture-dependent and culture-independent approaches. *Br J Nutr.* 2013;110(7):1253-62.
29. Damaceno QS, Souza JP, Nicoli JR, Paula RL, Assis GB, Figueiredo HC, et al. Evaluation of Potential Probiotics Isolated from Human Milk and Colostrum. *Probiotics Antimicrob Proteins.* 2017;9(4):371-9.
30. Gonzalez R, Maldonado A, Martin V, Mandomando I, Fumado V, Metzner KJ, et al. Breast milk and gut microbiota in African mothers and infants from an area of high HIV prevalence. *PLoS One.* 2013;8(11):e80299.
31. Jimenez E, Arroyo R, Cardenas N, Marin M, Serrano P, Fernandez L, et al. Mammary candidiasis: A medical condition without scientific evidence? *PLoS One.* 2017;12(7):e0181071.
32. Kozak K, Charbonneau D, Sanozky-Dawes R, Klaenhammer T. Characterization of bacterial isolates from the microbiota of mothers' breast milk and their infants. *Gut Microbes.* 2015;6(6):341-51.
33. Schanler RJ, Fraley JK, Lau C, Hurst NM, Horvath L, Rossmann SN. Breastmilk cultures and infection in extremely premature infants. *J Perinatol.* 2011;31(5):335-8.
34. Rezac S, Kok CR, Heermann M, Hutkins R. Fermented Foods as a Dietary Source of Live Organisms. *Front Microbiol.* 2018;9:1785.
35. Hunt KM, Foster JA, Forney LJ, Schutte UM, Beck DL, Abdo Z, et al. Characterization of the diversity and temporal stability of bacterial communities in human milk. *PLoS One.* 2011;6(6):e21313.
36. Sam Ma Z, Guan Q, Ye C, Zhang C, Foster JA, Forney LJ. Network analysis reveals a potentially 'evil' alliance of opportunistic pathogens inhibited by a cooperative network in human milk bacterial communities. *Sci Rep.* 2015;5:8275. doi: 10.1038/srep08275.
37. Drago L, Toscano M, De Grandi R, Grossi E, Padovani EM, Peroni DG. Microbiota network and mathematic microbe mutualism in colostrum and mature milk collected in two different geographic areas: Italy versus Burundi. *ISME J.* 2017;11(4):875–84.
38. Jimenez E, de Andres J, Manrique M, Pareja-Tobes P, Tobes R, Martinez-Blanch JF, et al. Metagenomic Analysis of Milk of Healthy and Mastitis-Suffering Women. *J Hum Lact.* 2015;31(3):406-15.
39. Kumar H, du Toit E, Kulkarni A, Aakko J, Linderborg KM, Zhang Y, et al. Distinct Patterns in Human Milk Microbiota and Fatty Acid Profiles Across Specific Geographic Locations. *Front Microbiol.* 2016;7:1619.
40. Li SW, Watanabe K, Hsu CC, Chao SH, Yang ZH, Lin YJ, et al. Bacterial Composition and Diversity in Breast Milk Samples from Mothers Living in Taiwan and Mainland China. *Front Microbiol.* 2017;8:965.

41. Murphy K, Curley D, O'Callaghan TF, O'Shea CA, Dempsey EM, O'Toole PW, et al. The Composition of Human Milk and Infant Faecal Microbiota Over the First Three Months of Life: A Pilot Study. *Sci Rep.* 2017;7:40597.
42. Patel SH, Vaidya YH, Patel RJ, Pandit RJ, Joshi CG, Kunjadiya AP. Culture independent assessment of human milk microbial community in lactational mastitis. *Sci Rep.* 2017;7(1):7804.
43. Sakwinska O, Moine D, Delley M, Combremont S, Rezzonico E, Descombes P, et al. Microbiota in Breast Milk of Chinese Lactating Mothers. *PLoS One.* 2016;11(8):e0160856.
44. Ward TL, Hosid S, Ioshikhes I, Altosaar I. Human milk metagenome- a functional capacity analysis. *BMC Microbiol.* 2013;13:116. doi:10.1186/471-2180-13-116.
45. Williams JE, Carrothers JM, Lackey KA, Beatty NF, York MA, Brooker SL, et al. Human Milk Microbial Community Structure Is Relatively Stable and Related to Variations in Macronutrient and Micronutrient Intakes in Healthy Lactating Women. *J Nutr.* 2017;147(9):1739-48.
46. Lackey KA, Williams JE, Meehan CL, Zachek JA, Benda ED, Price WJ, et al. What's Normal? Microbiomes in Human Milk and Infant Feces Are Related to Each Other but Vary Geographically: The INSPIRE Study. *Frontiers in Nutrition.* 2019;6:45. doi: 10.3389/fnut.2019.00045. eCollection 2019.
47. Louca S, Jacques SMS, Pires APF, Leal JS, Srivastava DS, Parfrey LW, et al. High taxonomic variability despite stable functional structure across microbial communities. *Nat Ecol Evol.* 2016;1(1):15.
48. Human Microbiome Project C. Structure, function and diversity of the healthy human microbiome. *Nature.* 2012;486(7402):207-14.
49. Galante L, Vickers MH, Milan AM, Reynolds CM, Alexander T, Bloomfield FH, et al. Feasibility of Standardized Human Milk Collection in Neonatal Care Units. *Sci Rep.* 2019;9(1):14343.
50. Godden SM, Jansen JT, Leslie KE, Smart NL, Kelton DF. The effect of sampling time and sample handling on the detection of *Staphylococcus aureus* in milk from quarters with subclinical mastitis. *Can Vet J.* 2002;43(1):38-42.
51. Dahlberg J, Sun L, Persson Waller K, Ostensson K, McGuire M, Agenas S, et al. Microbiota data from low biomass milk samples is markedly affected by laboratory and reagent contamination. *PLoS One.* 2019;14(6):e0218257.
52. Salter SJ, Cox MJ, Turek EM, Calus ST, Cookson WO, Moffatt MF, et al. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biol.* 2014;12:87. doi: 10.1186/s12915-014-0087-z.
53. Eisenhofer R, Minich JJ, Marotz C, Cooper A, Knight R, Weyrich LS. Contamination in Low Microbial Biomass Microbiome Studies: Issues and Recommendations. *Trends Microbiol.* 2019;27(2):105-17.

54. Ruiz L, Garcia-Carral C, Rodriguez JM. Unfolding the Human Milk Microbiome Landscape in the Omics Era. *Front Microbiol.* 2019;10:1378.
55. Hiergeist A, Glasner J, Reischl U, Gessner A. Analyses of Intestinal Microbiota: Culture versus Sequencing. *ILAR J.* 2015;56(2):228-40.
56. Bilen M, Dufour JC, Lagier JC, Cadoret F, Daoud Z, Dubourg G, et al. The contribution of culturomics to the repertoire of isolated human bacterial and archaeal species. *Microbiome.* 2018;6(1):94.
57. Lagier JC, Hugon P, Khelaifia S, Fournier PE, La Scola B, Raoult D. The rebirth of culture in microbiology through the example of culturomics to study human gut microbiota. *Clin Microbiol Rev.* 2015;28(1):237-64.
58. Lagier JC, Dubourg G, Million M, Cadoret F, Bilen M, Fenollar F, et al. Culturing the human microbiota and culturomics. *Nat Rev Microbiol.* 2018;16:540-50.
59. Epstein SS. The phenomenon of microbial uncultivability. *Curr Opin Microbiol.* 2013;16(5):636-42.
60. Rappe MS, Giovannoni SJ. The uncultured microbial majority. *Annu Rev Microbiol.* 2003;57:369-94.
61. Hug LA. Sizing Up the Uncultured Microbial Majority. *mSystems.* 2018;3(5):pii: e00185-18.
62. Browne HP, Forster SC, Anonye BO, Kumar N, Neville BA, Stares MD, et al. Culturing of 'unculturable' human microbiota reveals novel taxa and extensive sporulation. *Nature.* 2016;533(7604):543-6.
63. Lau JT, Whelan FJ, Herath I, Lee CH, Collins SM, Bercik P, et al. Capturing the diversity of the human gut microbiota through culture-enriched molecular profiling. *Genome Med.* 2016;8(1):72.
64. Lagier JC, Khelaifia S, Alou MT, Ndongo S, Dione N, Hugon P, et al. Culture of previously uncultured members of the human gut microbiota by culturomics. *Nat Microbiol.* 2016;1:16203.
65. Forster SC, Kumar N, Anonye BO, Almeida A, Viciani E, Stares MD, et al. A human gut bacterial genome and culture collection for improved metagenomic analyses. *Nat Biotechnol.* 2019;37(2):186-92.
66. Zou Y, Xue W, Luo G, Deng Z, Qin P, Guo R, et al. 1,520 reference genomes from cultivated human gut bacteria enable functional microbiome analyses. *Nat Biotechnol.* 2019;37(2):179-85.
67. Sibley CD, Grinwis ME, Field TR, Eshaghurshan CS, Faria MM, Dowd SE, et al. Culture enriched molecular profiling of the cystic fibrosis airway microbiome. *PLoS One.* 2011;6(7):e22702.

68. Wilson KH, Blichington RB. Human colonic biota studied by ribosomal DNA sequence analysis. *Appl Environ Microbiol.* 1996;62(7):2273-8.
69. Lagier JC, Armougom F, Million M, Hugon P, Pagnier I, Robert C, et al. Microbial culturomics: paradigm shift in the human gut microbiome study. *Clin Microbiol Infect.* 2012;18(12):1185-93.
70. de Goffau MC, Lager S, Salter SJ, Wagner J, Kronbichler A, Charnock-Jones DS, et al. Recognizing the reagent microbiome. *Nat Microbiol.* 2018;3(8):851-3.
71. Martin R, Langa S, Reviriego C, Jimenez E, Marin ML, Xaus J, et al. Human milk is a source of lactic acid bacteria for the infant gut. *J Pediatr.* 2003;143(6):754-8.
72. Abrahamsson TR, Sinkiewicz G, Jakobsson T, Fredrikson M, Björkstén B. Probiotic Lactobacilli in Breast Milk and Infant Stool in Relation to Oral Intake During the First Year of Life. *J Pediatr Gastroenterol Nutr.* 2009;49(3):349-54.
73. Solis G, de Los Reyes-Gavilan CG, Fernandez N, Margolles A, Gueimonde M. Establishment and development of lactic acid bacteria and bifidobacteria microbiota in breast-milk and the infant gut. *Anaerobe.* 2010;16(3):307-10.
74. Arbolea S, Ruas-Madiedo P, Margolles A, Solis G, Salminen S, de Los Reyes-Gavilan CG, et al. Characterization and in vitro properties of potentially probiotic Bifidobacterium strains isolated from breast-milk. *Int J Food Microbiol.* 2011;149(1):28-36.
75. Perez PF, Dore J, Leclerc M, Levenez F, Benyacoub J, Serrant P, et al. Bacterial imprinting of the neonatal immune system: lessons from maternal cells? *Pediatrics.* 2007;119(3):e724-32.
76. Dalgleish DG, Corredig M. The structure of the casein micelle of milk and its changes during processing. *Annu Rev Food Sci Technol.* 2012;3:449-67.
77. Andreas NJ, Kampmann B, Mehring Le-Doare K. Human breast milk: A review on its composition and bioactivity. *Early Hum Dev.* 2015;91(11):629-35.
78. Wilson IG. Inhibition and facilitation of nucleic acid amplification. *Appl Environ Microbiol.* 1997;63(10):3741-51.
79. Li R, Dudemaine PL, Zhao X, Lei C, Ibeagha-Awemu EM. Comparative Analysis of the miRNome of Bovine Milk Fat, Whey and Cells. *PLoS One.* 2016;11(4):e0154129.
80. Lima SF, Bicalho MLS, Bicalho RC. Evaluation of milk sample fractions for characterization of milk microbiota from healthy and clinical mastitis cows. *PLoS One.* 2018;13(3):e0193671.
81. Anderson JF. The Relative Proportion of Bacteria in Top Milk (Cream Layer) and Bottom Milk (SkimMilk), and Its Bearing on Infant Feeding. *J Infect Dis.* 1909;6(3):392-400.
82. Sun L, Dicksved J, Priyashantha H, Lundh A, Johansson M. Distribution of bacteria between different milk fractions, investigated using culture-dependent methods and



molecular-based and fluorescent microscopy approaches. *J Appl Microbiol.* 2019;127(4):1028-37.

83. Boix-Amoros A, Collado MC, Mira A. Relationship between Milk Microbiota, Bacterial Load, Macronutrients, and Human Cells during Lactation. *Front Microbiol.* 2016;7:492.

84. Hickey CD, Sheehan JJ, Wilkinson MG, Auty MA. Growth and location of bacterial colonies within dairy foods using microscopy techniques: a review. *Front Microbiol.* 2015;6:99.

85. Beaver A, Cazer CL, Ruegg PL, Grohn YT, Schukken YH. Implications of PCR and ELISA results on the routes of bulk-tank contamination with *Mycobacterium avium* ssp. paratuberculosis. *J Dairy Sci.* 2016;99(2):1391-405.

86. Ali-Vehmas T, Westphalen P, Myllys V, Sandholm M. Binding of *Staphylococcus aureus* to milk fat globules increases resistance to penicillin-G. *J Dairy Res.* 1997;64(2):253-60.

87. Herthnek D, Nielsen SS, Lindberg A, Bölske G. A robust method for bacterial lysis and DNA purification to be used with real-time PCR for detection of *Mycobacterium avium* subsp. paratuberculosis in milk. *J Microbiol Methods.* 2008;75(2):335-40.

88. O'Flaherty S, Coffey A, Meaney WJ, Fitzgerald GF, Ross RP. Inhibition of bacteriophage K proliferation on *Staphylococcus aureus* in raw bovine milk. *Lett Appl Microbiol.* 2005;41(3):274-9.

89. Vidanarachchi JK, Li S, Lundh AS, Johansson M. Short communication: Lipolytic activity on milk fat by *Staphylococcus aureus* and *Streptococcus agalactiae* strains commonly isolated in Swedish dairy herds. *J Dairy Sci.* 2015;98(12):8560-4.

90. Paape MJ, Guidry AJ. Effect of fat and casein on intracellular killing of *Staphylococcus aureus* by milk leukocytes. *Proc Soc Exp Biol Med.* 1977;155(4):588-93.

91. Fijałkowski K, Peitler D, Żywicka A, Karakulska J, Czerniawska-Piątkowska E. Influence of milk, milk fractions and milk proteins on the growth and viability of mastitis-causing *Staphylococcus aureus* strain. *Ital J Anim Sci.* 2017;16(2):321-8.

92. Huseyin CE, O'Toole PW, Cotter PD, Scanlan PD. Forgotten fungi-the gut mycobiome in human health and disease. *FEMS Microbiol Rev.* 2017;41(4):479-511.

93. Laforest-Lapointe I, Arrieta MC. Microbial Eukaryotes: a Missing Link in Gut Microbiome Studies. *mSystems.* 2018;3(2):e00201-17.

94. Ege MJ, Mayer M, Normand AC, Genuneit J, Cookson WO, Braun-Fahrlander C, et al. Exposure to Environmental Microorganisms and Childhood Asthma. *N Engl J Med.* 2011;364(8):701-9.

95. Arrieta MC, Arevalo A, Stiemsma L, Dimitriu P, Chico ME, Loor S, et al. Associations between infant fungal and bacterial dysbiosis and childhood atopic wheeze in a nonindustrialized setting. *J Allergy Clin Immunol*. 2018;142(2):424-34.
96. Ward TL, Dominguez-Bello MG, Heisel T, Al-Ghalith G, Knights D, Gale CA. Development of the Human Mycobiome over the First Month of Life and across Body Sites. *mSystems*. 2018;3(3):pii: e00140-17.
97. McGuire MK, McGuire MA. Got bacteria? The astounding, yet not-so-surprising, microbiome of human milk. *Curr Opin Biotechnol*. 2017;44:63-8.
98. Boix-Amoros A, Martinez-Costa C, Querol A, Collado MC, Mira A. Multiple Approaches Detect the Presence of Fungi in Human Breastmilk Samples from Healthy Mothers. *Sci Rep*. 2017;7(1):13016.
99. Boix-Amoros A, Puente-Sanchez F, du Toit E, Linderborg KM, Zhang Y, Yang B, et al. Mycobiome profiles in breast milk from healthy women depend on mode of delivery, geographic location and interaction with bacteria. *Appl Environ Microbiol*. 2019;85(9):pii: e02994-18.
100. Amir LH, Donath SM, Garland SM, Tabrizi SN, Bennett CM, Cullinane M, et al. Does *Candida* and/or *Staphylococcus* play a role in nipple and breast pain in lactation? A cohort study in Melbourne, Australia. *BMJ Open*. 2013;3(3): e002351
101. Heisel T, Nyaribo L, Sadowsky MJ, Gale CA. Breastmilk and NICU surfaces are potential sources of fungi for infant mycobiomes. *Fungal Genet Biol*. 2019;128:29-35.
102. Derakhshani H, Plaizier JC, De Buck J, Barkema HW, Khafipour E. Association of bovine major histocompatibility complex (BoLA) gene polymorphism with colostrum and milk microbiota of dairy cows during the first week of lactation. *Microbiome*. 2018;6(1):203.
103. Treven P, Mrak V, Bogovic Matijasic B, Horvat S, Rogelj I. Administration of probiotics *Lactobacillus rhamnosus* GG and *Lactobacillus gasseri* K7 during pregnancy and lactation changes mouse mesenteric lymph nodes and mammary gland microbiota. *J Dairy Sci*. 2015;98(4):2114-28.
104. de Andres J, Jimenez E, Chico-Calero I, Fresno M, Fernandez L, Rodriguez JM. Physiological Translocation of Lactic Acid Bacteria during Pregnancy Contributes to the Composition of the Milk Microbiota in Mice. *Nutrients*. 2018;10(1):14.
105. Fernandez L, Langa S, Martin V, Maldonado A, Jimenez E, Martin R, et al. The human milk microbiota: origin and potential roles in health and disease. *Pharmacol Res*. 2013;69(1):1-10.
106. Berg RD, Garlington AW. Translocation of Certain Indigenous Bacteria from the Gastrointestinal Tract to the Mesenteric Lymph Nodes and Other Organs in a Gnotobiotic Mouse Model. *Infect Immun*. 1979;23(2):403-11.
107. Sedman PC, Macfie J, Sagar P, Mitchell CJ, May J, Mancey-Jones B, et al. The prevalence of gut translocation in humans. *Gastroenterology*. 1994;107(3):643-9.

108. Chelakkot C, Ghim J, Ryu SH. Mechanisms regulating intestinal barrier integrity and its pathological implications. *Exp Mol Med*. 2018;50(8):103.
109. Thaiss CA, Levy M, Grosheva I, Zheng D, Soffer E, Blacher E, et al. Hyperglycemia drives intestinal barrier dysfunction and risk for enteric infection. *Science*. 2018;359(6382):1376-83.
110. Kerr CA, Grice DM, Tran CD, Bauer DC, Li D, Hendry P, et al. Early life events influence whole-of-life metabolic health via gut microflora and gut permeability. *Crit Rev Microbiol*. 2015;41(3):326-40.
111. Reyes H, Zapata R, Hernandez I, Gotteland M, Sandoval L, Jiron MI, et al. Is a leaky gut involved in the pathogenesis of intrahepatic cholestasis of pregnancy? *Hepatology*. 2006;43(4):715-22.
112. Smyth SH, Doyle-McCullough M, Cox OT, Carr KE. Effect of reproductive status on uptake of latex microparticles in rat small intestine. *Life Sci*. 2005;77(26):3287-305.
113. Sakwinska O, Bosco N. Host Microbe Interactions in the Lactating Mammary Gland. *Front Microbiol*. 2019;10:1863.
114. Rodriguez JM. The origin of human milk bacteria: is there a bacterial entero-mammary pathway during late pregnancy and lactation? *Adv Nutr*. 2014;5(6):779-84.
115. Urbaniak CC, J., Brackstone MM, J. M., Gloor GB, Baban CK, Scott L, O'Hanlon DM, et al. Microbiota of human breast tissue. *Appl Environ Microbiol*. 2014;80(10):3007-14.
116. Ruiz L, Bacigalupe R, Garcia-Carral C, Boix-Amoros A, Arguello H, Silva CB, et al. Microbiota of human precolostrum and its potential role as a source of bacteria to the infant mouth. *Sci Rep*. 2019;9(1):8435.
117. Lawrence RM, Lawrence RA. Breast milk and infection. *Clin Perinatol*. 2004;31(3):501-28.
118. Arroyo R, Martin V, Maldonado A, Jimenez E, Fernandez L, Rodriguez JM. Treatment of infectious mastitis during lactation: antibiotics versus oral administration of Lactobacilli isolated from breast milk. *Clin Infect Dis*. 2010;50(12):1551-8.
119. Jimenez E, Fernandez L, Maldonado A, Martin R, Olivares M, Xaus J, et al. Oral administration of Lactobacillus strains isolated from breast milk as an alternative for the treatment of infectious mastitis during lactation. *Appl Environ Microbiol*. 2008;74(15):4650-5.
120. Fernandez L, Cardenas N, Arroyo R, Manzano S, Jimenez E, Martin V, et al. Prevention of Infectious Mastitis by Oral Administration of Lactobacillus salivarius PS2 During Late Pregnancy. *Clin Infect Dis*. 2016;62(5):568-73.

121. Cronin M, Morrissey D, Rajendran S, El Mashad SM, van Sinderen D, O'Sullivan GC, et al. Orally administered bifidobacteria as vehicles for delivery of agents to systemic tumors. *Mol Ther*. 2010;18(7):1397-407.
122. Gardner H, Kent JC, Lai CT, Mitoulas LR, Cregan MD, Hartmann PE, et al. Milk ejection patterns: an intra- individual comparison of breastfeeding and pumping. *BMC Pregnancy Childbirth*. 2015;15:156.
123. Derakhshani H, Fehr KB, Sepehri S, Francoz D, De Buck J, Barkema HW, et al. Invited review: Microbiota of the bovine udder: Contributing factors and potential implications for udder health and mastitis susceptibility. *J Dairy Sci*. 2018;101(12):10605-25.
124. Martucci J. History of Medicine: Breast Pumping. *AMA J Ethics*. 2013;15(9):791-7.
125. Johns HM, Forster DA, Amir LH, McLachlan HL. Prevalence and outcomes of breast milk expressing in women with healthy term infants: a systematic review. *BMC Pregnancy Childbirth*. 2013;13:212.
126. Haiden N, Pimpel B, Assadian O, Binder C, Kreissl A, Repa A, et al. Comparison of bacterial counts in expressed breast milk following standard or strict infection control regimens in neonatal intensive care units: compliance of mothers does matter. *J Hosp Infect*. 2016;92(3):226-8.
127. Collado MC, Isolauri E, Laitinen K, Salminen S. Distinct composition of gut microbiota during pregnancy in overweight and normal-weight women. *Am J Clin Nutr*. 2008;88(4):894-9.
128. Wu GD, Chen J, Hoffmann C, Bittinger K, Chen YY, Keilbaugh SA, et al. Linking long-term dietary patterns with gut microbial enterotypes. *Science*. 2011;334(6052):105-8.
129. Padilha M, Danneskiold-Samsoe NB, Brejnrod A, Hoffmann C, Cabral VP, Iaucci JM, et al. The Human Milk Microbiota is Modulated by Maternal Diet. *Microorganisms*. 2019;7(11): 502.
130. Chu DM, Ma J, Prince AL, Antony KM, Seferovic MD, Aagaard KM. Maturation of the infant microbiome community structure and function across multiple body sites and in relation to mode of delivery. *Nat Med*. 2017;23(3):314-26.
131. Moossavi S, Azad MB. Origins of human milk microbiota: new evidence and arising questions *Gut Microbes*. 2019;doi: 10.1080/19490976.2019.1667722.
132. Affolter M, Garcia-Rodenas CL, Vinyes-Pares G, Jenni R, Roggero I, Avanti-Nigro O, et al. Temporal Changes of Protein Composition in Breast Milk of Chinese Urban Mothers and Impact of Caesarean Section Delivery. *Nutrients*. 2016;8(8):pii: E504. doi: 10.3390/nu8080504.
133. Nozad S, Ramin AG, Asri Rezaie S. Diurnal variations in milk macro-mineral concentrations in Holstein dairy cows in Urmia, Iran. *Vet Res Forum*. 2012;3(4):281-5.

134. Jackson DA, Imong SM, Silprasert A, Ruckphaopunt S, Woolridge MW, Baum JD, et al. Circadian variation in fat concentration of breast-milk in a rural northern Thai population. *Br J Nutr.* 2007;59(03):349-63.
135. Cabrera-Rubio R, Collado MC, Laitinen K, Salminen S, Isolauri E, Mira A. The human milk microbiome changes over lactation and is shaped by maternal weight and mode of delivery. *Am J Clin Nutr.* 2012;96(3):544-51.
136. Cabrera-Rubio R, Mira-Pascual L, Mira A, Collado MC. Impact of mode of delivery on the milk microbiota composition of healthy women. *J Dev Orig Health Dis.* 2016;7(1):54-60.
137. Khodayar-Pardo P, Mira-Pascual L, Collado MC, Martinez-Costa C. Impact of lactation stage, gestational age and mode of delivery on breast milk microbiota. *J Perinatol.* 2014;34(8):599-605.
138. Hermansson H, Kumar H, Collado MC, Salminen S, Isolauri E, Rautava S. Breast Milk Microbiota Is Shaped by Mode of Delivery and Intrapartum Antibiotic Exposure. *Front Nutr.* 2019;6:4.
139. Moossavi S, Sepehri S, Robertson B, Bode L, Goruk S, Field CJ, et al. Composition and Variation of the Human Milk Microbiome is Influenced by Maternal and Early Life Factors *Cell Host Microbe.* 2019;25:324-35.
140. Peters MDJ, McArthur A, Munn Z. Safe management of expressed breast milk: A systematic review. *Women and Birth.* 2016;29(6):473-81.
141. Reyes SM, Allen DL, Williams J, McGuire MA, McGuire MK, Hay AG, et al. Sources of Bacterial Contamination in Home-Expressed Human Milk. *Breastfeeding medicine : the official journal of the Academy of Breastfeeding Medicine.* 2018;13(7):A-50.
142. Becker GE, Smith HA, Cooney F. Methods of milk expression for lactating women. *Cochrane Database Syst Rev.* 2016;9:CD006170.
143. Bliss JM, Basavegowda KP, Watson WJ, Sheikh AU, Ryan RM. Vertical and horizontal transmission of *Candida albicans* in very low birth weight infants using DNA fingerprinting techniques. *Pediatr Infect Dis J.* 2008;27(3):231-5.
144. Frey-Klett P, Burlinson P, Deveau A, Barret M, Tarkka M, Sarniguet A. Bacterial-fungal interactions: hyphens between agricultural, clinical, environmental, and food microbiologists. *Microbiol Mol Biol Rev.* 2011;75(4):583-609.
145. Saleem M. *Microbiome community ecology: Fundamentals and applications*: Springer.
146. Garwolińska D, Namieśnik J, Kot-Wasik A, Hewelt-Belka W. Chemistry of human breast milk – a comprehensive review of the composition and role of milk metabolites in child development. *J Agric Food Chem.* 2018;66(45):11881-96.

147. Bender JM, Li F, Martelly S, Byrt E, Rouzier V, Leo M, et al. Maternal HIV infection influences the microbiome of HIV-uninfected infants. *Sci Transl Med*. 2016;8(349):349ra100-349ra100.
148. Kirmiz N, Robinson RC, Shah IM, Barile D, Mills DA. Milk Glycans and Their Interaction with the Infant-Gut Microbiota. *Annu Rev Food Sci Technol*. 2018;9(1):429-50.
149. Korpela K, Salonen A, Hickman B, Kunz C, Sprenger N, Kukkonen K, et al. Fucosylated oligosaccharides in mother's milk alleviate the effects of caesarean birth on infant gut microbiota. *Sci Rep*. 2018;8(1): 13757.
150. Asakuma S, Hatakeyama E, Urashima T, Yoshida E, Katayama T, Yamamoto K, et al. Physiology of consumption of human milk oligosaccharides by infant gut-associated bifidobacteria. *J Biol Chem*. 2011;286(40):34583-92.
151. Aakko J, Kumar H, Rautava S, Wise A, Autran C, Bode L, et al. Human milk oligosaccharide categories define the microbiota composition in human colostrum. *Benef Microbes*. 2017;8(4):563-7.
152. Hunt KM, Preuss J, Nissan C, Davlin CA, Williams JE, Shafii B, et al. Human milk oligosaccharides promote the growth of staphylococci. *Appl Environ Microbiol*. 2012;78(14):4763-70.
153. Williams JE, Price WJ, Shafii B, Yahvah KM, Bode L, McGuire MA, et al. Relationships Among Microbial Communities, Maternal Cells, Oligosaccharides, and Macronutrients in Human Milk. *J Hum Lact*. 2017;33(3):540-51.
154. Gómez-Gallego C, Morales J, Monleón D, du Toit E, Kumar H, Linderborg K, et al. Human Breast Milk NMR Metabolomic Profile across Specific Geographical Locations and Its Association with the Milk Microbiota. *Nutrients*. 2018;10(10):1355.
155. Ramani S, Stewart CJ, Laucirica DR, Ajami NJ, Robertson B, Autran CA, et al. Human milk oligosaccharides, milk microbiome and infant gut microbiome modulate neonatal rotavirus infection. *Nat Commun*. 2018;9(1):5010.
156. Cabrera-Rubio R, Kunz C, Rudloff S, Garcia-Mantrana I, Crehua-Gaudiza E, Martinez-Costa C, et al. Association of Maternal Secretor Status and Human Milk Oligosaccharides With Milk Microbiota: An Observational Pilot Study. *J Pediatr Gastroenterol Nutr*. 2019; 68(2):256-263..
157. Koletzko B, Rodriguez-Palmero M, Demmelmair H, Fidler N, Jensen R, Sauerwald T. Physiological aspects of human milk lipids. *Early Hum Dev*. 2001;65(Suppl):S3-S18.
158. Urwin HJ, Miles EA, Noakes PS, Kremmyda LS, Vlachava M, Diaper ND, et al. Salmon consumption during pregnancy alters fatty acid composition and secretory IgA concentration in human breast milk. *J Nutr*. 2012;142(8):1603-10.
159. Kankaanpää PE, Salminen SJ, Isolauri E, Lee YK. The influence of polyunsaturated fatty acids on probiotic growth and adhesion. *FEMS Microbiol Lett*. 2001;194(2):149-53.

160. Kelsey JA, Bayles KW, Shafii B, McGuire MA. Fatty Acids and Monoacylglycerols Inhibit Growth of *Staphylococcus aureus*. *Lipids*. 2006;41(10):951-61.
161. Kabara JJ, Swieczkowski DM, Conley AJ, Truant JP. Fatty acids and derivatives as antimicrobial agents. *Antimicrob Agents Chemother*. 1972;2(1):23-8.
162. Churchward CP, Alany RG, Snyder LAS. Alternative antimicrobials: the properties of fatty acids and monoglycerides. *Crit Rev Microbiol*. 2018;44(5):561-70.
163. Menni C, Zierer J, Pallister T, Jackson MA, Long T, Mohny RP, et al. Omega-3 fatty acids correlate with gut microbiome diversity and production of N-carbamylglutamate in middle aged and elderly women. *Sci Rep*. 2017;7(1): 11079.
164. Dunstan JA, Roper J, Mitoulas L, Hartmann PE, Simmer K, Prescott SL. The effect of supplementation with fish oil during pregnancy on breast milk immunoglobulin A, soluble CD14, cytokine levels and fatty acid composition. *Clin Exp Allergy*. 2004;34(8):1237-42.
165. Caesar R, Tremaroli V, Kovatcheva-Datchary P, Cani PD, Backhed F. Crosstalk between Gut Microbiota and Dietary Lipids Aggravates WAT Inflammation through TLR Signaling. *Cell Metab*. 2015;22(4):658-68.
166. Muir P, Li S, Lou S, Wang D, Spakowicz DJ, Salichos L, et al. The real cost of sequencing: scaling computation to keep pace with data generation. *Genome Biol*. 2016;17:53.
167. Stackebrandt E, Goebel BM. A place for DNA-DNA reassociation and 16S rRNA sequence analysis in the present species definition in bacteriology. *Int J Syst Bacteriol*. 1994;44(4):846-9.
168. Kopylova E, Navas-Molina JA, Mercier C, Xu ZZ, Mahe F, He Y, et al. Open-Source Sequence Clustering Methods Improve the State Of the Art. *mSystems*. 2016;1(1):pii: e00003-15.
169. Callahan BJ, McMurdie PJ, Holmes SP. Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *ISME J*. 2017;11(12):2639-43.
170. Eren AM, Maignien L, Sul WJ, Murphy LG, Grim SL, Morrison HG, et al. Oligotyping: Differentiating between closely related microbial taxa using 16S rRNA gene data. *Methods Ecol Evol*. 2013;4(12):doi: 10.1111/2041-210X.12114.
171. Eren AM, Morrison HG, Lescault PJ, Reveillaud J, Vineis JH, Sogin ML. Minimum entropy decomposition: unsupervised oligotyping for sensitive partitioning of high-throughput marker gene sequences. *ISME J*. 2015;9(4):968-79.
172. Tikhonov M, Leach RW, Wingreen NS. Interpreting 16S metagenomic data without clustering to achieve sub-OTU resolution. *ISME J*. 2015;9(1):68-80.
173. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJ, Holmes SP. DADA2: High-resolution sample inference from Illumina amplicon data. *Nat Methods*. 2016;13(7):581-3.

174. Amir A, McDonald D, Navas-Molina JA, Kopylova E, Morton JT, Zech Xu Z, et al. Deblur Rapidly Resolves Single-Nucleotide Community Sequence Patterns. *mSystems*. 2017;2(2):pii: e00191-16.
175. Edgar RC. UNOISE2: improved error-correction for Illumina 16 S and ITS amplicon sequencing. *bioRxiv* 2016:081257; doi: <https://doi.org/10.1101/>.
176. Caruso V, Song X, Asquith M, Karstens L, Gibbons SM. Performance of Microbiome Sequence Inference Methods in Environments with Varying Biomass. *mSystems*. 2019;4(1):pii: e00163-18.
177. Chen W, Zhang CK, Cheng Y, Zhang S, Zhao H. A comparison of methods for clustering 16S rRNA sequences into OTUs. *PLoS One*. 2013;8(8):e70837.
178. Nearing JT, Douglas GM, Comeau AM, Langille MGI. Denoising the Denoisers: an independent evaluation of microbiome sequence error-correction approaches. *PeerJ*. 2018;6:e5364.
179. Gloor GB, Wu JR, Pawlowsky-Glahn V, Egozcue JJ. It's all relative: analyzing microbiome data as compositions. *Ann Epidemiol*. 2016;26(5):322-9.
180. Ravel J, Wommack KE. All hail reproducibility in microbiome research. *Microbiome*. 2014;2(1):8.
181. Panek M, Cipicic Paljetak H, Baresic A, Peric M, Matijasic M, Lojkic I, et al. Methodology challenges in studying human gut microbiota - effects of collection, storage, DNA extraction and next generation sequencing technologies. *Sci Rep*. 2018;8(1):5143.
182. Hahn A, Sanyal A, Perez GF, Colberg-Poley AM, Campos J, Rose MC, et al. Different next generation sequencing platforms produce different microbial profiles and diversity in cystic fibrosis sputum. *J Microbiol Methods*. 2016;130:95-9.
183. Hsieh YH, Peterson CM, Raggio A, Keenan MJ, Martin RJ, Ravussin E, et al. Impact of Different Fecal Processing Methods on Assessments of Bacterial Diversity in the Human Intestine. *Front Microbiol*. 2016;7:1643.
184. Allali I, Arnold JW, Roach J, Cadenas MB, Butz N, Hassan HM, et al. A comparison of sequencing platforms and bioinformatics pipelines for compositional analysis of the gut microbiome. *BMC Microbiol*. 2017;17(1):194.
185. Glassman SI, Martiny JBH. BROADSCALE Ecological Patterns Are Robust to Use of Exact Sequence Variants versus Operational Taxonomic Units. *mSphere*. 2018;3(4):pii: e00148-18.
186. Karstens L, Asquith M, Caruso V, Rosenbaum JT, Fair DA, Braun J, et al. Community profiling of the urinary microbiota: considerations for low-biomass samples. *Nat Rev Urol*. 2018;15(12):735-49.



187. Xue Z, Kable ME, Marco ML. Impact of DNA Sequencing and Analysis Methods on 16S rRNA Gene Bacterial Community Analysis of Dairy Products. *mSphere*. 2018;3(5):pii: e00410-18.
188. Moossavi S, Miliku K, Sepehri S, Khafipour E, Azad MB. The prebiotic and probiotic properties of human milk: implications for infant immune development and pediatric asthma. *Front Pediatr*. 2018;6:197.
189. Pannaraj PS, Li F, Cerini C, Bender JM, Yang S, Rollie A, et al. Association Between Breast Milk Bacterial Communities and Establishment and Development of the Infant Gut Microbiome. *JAMA Pediatr*. 2017;171(7):647-54.
190. Davis NM, Proctor D, Holmes SP, Relman DA, Callahan BJ. Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. *Microbiome*. 2018;6(1):226. doi: 10.1186/s40168-018-0605-2.
191. Subbarao P, Anand SS, Becker AB, Befus AD, Brauer M, Brook JR, et al. The Canadian Healthy Infant Longitudinal Development (CHILD) Study: examining developmental origins of allergy and asthma. *Thorax*. 2015;70(10):998-1000.
192. Azad MB, Robertson B, Atakora F, Becker AB, Subbarao P, Moraes TJ, et al. Human Milk Oligosaccharide Concentrations Are Associated with Multiple Fixed and Modifiable Maternal Characteristics, Environmental Factors, and Feeding Practices. *J Nutr*. 2018;148(11):1733-42.
193. Chan D, Goruk S, Becker AB, Subbarao P, Mandhane PJ, Turvey SE, et al. Adiponectin, leptin and insulin in breast milk: associations with maternal characteristics and infant body composition in the first year of life. *Int J Obes (Lond)*. 2017;doi: 10.1038/ijo.2017.189.
194. Miliku K, Duan Q, Moraes TJ, Becker AB, Mandhane PJ, Turvey SE, et al. Human milk fatty acid composition is associated with dietary, genetic, sociodemographic, and environmental factors in the Canadian Healthy Infant Longitudinal Development (CHILD) cohort study. *Am J Clin Nutr*. 2019;110(6):1370-83.
195. de Souza RJ, Zulyniak MA, Desai D, Shaikh MR, Campbell NC, Lefebvre DL, et al. Harmonization of Food-Frequency Questionnaires and Dietary Pattern Analysis in 4 Ethnically Diverse Birth Cohorts. *J Nutr*. 2016;146(11):2343-50.
196. Alderete TL, Autran C, Brekke BE, Knight R, Bode L, Goran MI, et al. Associations between human milk oligosaccharides and infant body composition in the first 6 mo of life. *Am J Clin Nutr*. 2015;102(6):1381-8.
197. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007;81(3):559-75.
198. Marchini J, Howie B, Myers S, McVean G, Donnelly P. A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat Genet*. 2007;39(7):906-13.

199. Genomes Project C, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, et al. A global reference for human genetic variation. *Nature*. 2015;526(7571):68-74.
200. Kelly RJ, Rouquier S, Giorgi D, Lennon GG, Lowe JB. Sequence and expression of a candidate for the human Secretor blood group alpha(1,2)fucosyltransferase gene (FUT2). Homozygosity for an enzyme-inactivating nonsense mutation commonly correlates with the non-secretor phenotype. *J Biol Chem*. 1995;270(9):4640-9.
201. Ferrer-Admetlla A, Sikora M, Laayouni H, Esteve A, Roubinet F, Blancher A, et al. A natural history of FUT2 polymorphism in humans. *Mol Biol Evol*. 2009;26(9):1993-2003.
202. Population Centre and Rural Area Classification 2016: Statistics Canada; [Available from: <https://www.statcan.gc.ca/eng/subjects/standard/pcrac/2016/introduction>].
203. Takaro TK, Scott JA, Allen RW, Anand SS, Becker AB, Befus AD, et al. The Canadian Healthy Infant Longitudinal Development (CHILD) birth cohort study: assessment of environmental exposures. *J Expo Sci Environ Epidemiol*. 2015;25(6):580-92.
204. Hytten FE. Clinical and Chemical Studies in Human Lactation. *Br Med J*. 1954;1(4855):175-82.
205. Moraes TJ, Lefebvre DL, Chooniedass R, Becker AB, Brook JR, Denburg J, et al. The Canadian healthy infant longitudinal development birth cohort study: biological samples and biobanking. *Paediatr Perinat Epidemiol*. 2015;29(1):84-92.
206. Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, Fierer N, et al. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J*. 2012;6(8):1621-4.
207. Walker AW. A Lot on Your Plate? Well-to-Well Contamination as an Additional Confounder in Microbiome Sequence Analyses. *mSystems*. 2019;4:e00186-19.
208. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, et al. QIIME allows analysis of high-throughput community sequencing data. *Nat Methods*. 2010;7(5):335-6.
209. DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, et al. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl Environ Microbiol*. 2006;72(7):5069-72.
210. Nilsson RH, Tedersoo L, Ryberg M, Kristiansson E, Hartmann M, Unterseher M, et al. A comprehensive, automatically updated fungal ITS sequence dataset for reference-based chimera control in environmental sequencing efforts. *Microbes Environ*. 2015;30(2):145-50.
211. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2017.
212. McMurdie PJ, Holmes S. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One*. 2013;8(4):e61217.

213. McGuire MK, Meehan CL, McGuire MA, Williams JE, Foster J, Sellen DW, et al. What's normal? Oligosaccharide concentrations and profiles in milk produced by healthy women vary geographically. *Am J Clin Nutr.* 2017;105(5):1086-100.
214. Moossavi S, Atakora F, Miliku K, Sepehri S, Robertson B, Duan QL, et al. Integrated Analysis of Human Milk Microbiota with Oligosaccharides and Fatty Acids in the CHILD Cohort. *Front Nutr.* 2019;6:58.
215. Cruz-Hernandez C, Goeuriot S, Giuffrida F, Thakkar SK, Destailats F. Direct quantification of fatty acids in human milk by gas chromatography. *J Chromatogr A.* 2013;1284:174-9.
216. van Buuren S, Groothuis-Oudshoorn K. mice: Multivariate Imputation by Chained Equations in R. *Journal of Statistical Software.* 2011;45(3):1-67.
217. Meng C, Zeleznik OA, Thallinger GG, Kuster B, Gholami AM, Culhane AC. Dimension reduction techniques for the integrative analysis of multi-omics data. *Brief Bioinform.* 2016;17(4):628-41.
218. Kroes I, Lepp PW, Relman DA. Bacterial diversity within the human subgingival crevice. *Proc Natl Acad Sci USA.* 1999;96(25):14547-52.
219. Simpson EH. Measurement of diversity. *Nature.* 1949;163(4148):688.
220. Palarea-Albaladejo J, Martin-Fernandez JA. zCompositions -- R package for multivariate imputation of left-censored data under a compositional approach. *Chemom Intell Lab Syst.* 2015;143:85-96.
221. Gloor GB, Reid G. Compositional analysis: a valid approach to analyze microbiome high-throughput sequencing data. *Can J Microbiol.* 2016;62(8):692-703.
222. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J R Statist Soc B.* 1995;57(1):289-300.
223. Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D, et al. *Vegan: Community Ecology Package.* R package version 2.4-3. 2017.
224. Ward Jr. JH. Hierarchical Grouping to Optimize an Objective Function. *J Am Stat Assoc.* 1963;58(301):236-44.
225. Tibshirani R, Walther G, Hastie T. Estimating the number of clusters in a data set via the gap statistic. *J R Statist Soc B.* 2001;63(2):411-23
226. Rosseel Y. lavaan: An R Package for Structural Equation Modeling. *J Stat Softw.* 2012;48(2):1-36.
227. Epskamp S, Stuber S. semPlot: Path Diagrams and Visual Analysis of Various SEM Packages' Output. R package version 1.1. 2017.

228. Korkmaz S, Goksuluk D, Zararsiz G. MVN: An R Package for Assessing Multivariate Normality. *The R Journal*. 2014;6(2):151-62.
229. Kline RB. Principles and practices of structural equation modeling. 4 ed. New York: The Guilford Press; 2016.
230. Segata N, Izard J, Waldron L, Gevers D, Miropolsky L, Garrett WS, et al. Metagenomic biomarker discovery and explanation. *Genome Biol*. 2011;12(6):R60.
231. Friedman J, Hastie T, Tibshirani R. Regularization Paths for Generalized Linear Models via Coordinate Descent. *J Stat Software*. 2010;33(1):1-22.
232. Faust K, Raes J. CoNet app: inference of biological association networks using Cytoscape. *F1000Res*. 2016;5:1519.
233. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*. 2003;13(11):2498-504.
234. Masella AP, Bartram AK, Truszkowski JM, Brown DG, Neufeld JD. PANDAseq: paired-end assembler for illumina sequences. *BMC Bioinformatics*. 2012;13:31.
235. Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R. UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics*. 2011;27(16):2194-200.
236. Edgar RC. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*. 2010;26(19):2460-1.
237. Wang Q, Garrity GM, Tiedje JM, Cole JR. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl Environ Microbiol*. 2007;73(16):5261-7.
238. Caporaso JG, Bittinger K, Bushman FD, DeSantis TZ, Andersen GL, Knight R. PyNAST: a flexible tool for aligning sequences to a template alignment. *Bioinformatics*. 2010;26(2):266-7.
239. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *J Chiropr Med*. 2016;15(2):155-63.
240. Revelle W. psych: Procedures for Personality and Psychological Research, Northwestern University, Evanston, Illinois, USA, <https://CRAN.R-project.org/package=psych> Version =1.8.12. 2018.
241. Atlas R. Handbook of microbiological media. Boca Raton: CRC Press; 2004.
242. Whelan FJ, Verschoor CP, Stearns JC, Rossi L, Luinstra K, Loeb M, et al. The loss of topography in the microbial communities of the upper respiratory tract in the elderly. *Ann Am Thorac Soc*. 2014;11(4):513-21.

243. Bartram AK, Lynch MD, Stearns JC, Moreno-Hagelsieb G, Neufeld JD. Generation of multimillion-sequence 16S rRNA gene libraries from complex microbial communities by assembling paired-end illumina reads. *Appl Environ Microbiol.* 2011;77(11):3846-52.
244. Pagès H, Aboyoun P, Gentleman R, DebRoy S. Biostrings: Efficient manipulation of biological strings. R package version 2.48.0. 2019.
245. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnetjournal.* 2011;17:10-2.
246. Yilmaz P, Parfrey LW, Yarza P, Gerken J, Pruesse E, Quast C, et al. The SILVA and "All-species Living Tree Project (LTP)" taxonomic frameworks. *Nucleic Acids Res.* 2014;42(Database issue):D643-8.
247. Astudillo-Garcia C, Bell JJ, Webster NS, Glasl B, Jompa J, Montoya JM, et al. Evaluating the core microbiota in complex communities: A systematic investigation. *Environ Microbiol.* 2017;19(4):1450-62.
248. Shade A, Handelsman J. Beyond the Venn diagram: the hunt for a core microbiome. *Environ Microbiol.* 2012;14(1):4-12.
249. Knights D, Ward TL, McKinlay CE, Miller H, Gonzalez A, McDonald D, et al. Rethinking "enterotypes". *Cell Host Microbe.* 2014;16(4):433-7.
250. MacConaill LE, Burns RT, Nag A, Coleman HA, Slevin MK, Giorda K, et al. Unique, dual-indexed sequencing adapters with UMIs effectively eliminate index cross-talk and significantly improve sensitivity of massively parallel sequencing. *BMC Genomics.* 2018;19(1):30.
251. Bandara HMHN, Panduwawala CP, Samaranayake LP. Biodiversity of the human oral mycobiome in health and disease. *Oral Dis.* 2019;25(2):363-71.
252. Rowan-Nash AD, Korry BJ, Mylonakis E, Belenky P. Cross-Domain and Viral Interactions in the Microbiome. *Microbiol Mol Biol Rev.* 2019;83(1):pii: e00044-18.
253. Delavenne E, Mounier J, Deniel F, Barbier G, Le Blay G. Biodiversity of antifungal lactic acid bacteria isolated from raw milk samples from cow, ewe and goat over one-year period. *Int J Food Microbiol.* 2012;155(3):185-90.
254. Allonsius CN, Vandenheuvel D, Oerlemans EFM, Petrova MI, Donders GGG, Cos P, et al. Inhibition of *Candida albicans* morphogenesis by chitinase from *Lactobacillus rhamnosus* GG. *Sci Rep.* 2019;9(1):2900.
255. Bahram M, Hildebrand F, Forslund SK, Anderson JL, Soudzilovskaia NA, Bodegom PM, et al. Structure and function of the global topsoil microbiome. *Nature.* 2018;560(7717):233-7.
256. Butcher J, Unger S, Stintzi A, O'Connor DL. Characterizing The Breast Milk Microbiome With Culture-Enriched Molecular Profiling. *Breastfeeding Med.* 2018 13(7):A-18.

257. Arslanoglu S, Boquien CY, King C, Lamireau D, Tonetto P, Barnett D, et al. Fortification of Human Milk for Preterm Infants: Update and Recommendations of the European Milk Bank Association (EMBA) Working Group on Human Milk Fortification. *Front Pediatr*. 2019;7:76.
258. Biagi E, Quercia S, Aceti A, Beghetti I, Rampelli S, Turrone S, et al. The Bacterial Ecosystem of Mother's Milk and Infant's Mouth and Gut. *Front Microbiol*. 2017;8:1214.
259. Miliku K, Becker AB, Goruk S, Subbarao P, Mandhane PJ, Turvey SE, et al. Human milk fatty acids and asthma development in preschool children. *Curr Dev Nutr*. 2018;2(11):nzy040.
260. Amend AS, Seifert KA, Samson R, Bruns TD. Indoor fungal composition is geographically patterned and more diverse in temperate zones than in the tropics. *Proc Natl Acad Sci USA*. 2010;107(31):13748-53.
261. Callon C, Duthoit F, Delbes C, Ferrand M, Le Frileux Y, De Cremoux R, et al. Stability of microbial communities in goat milk during a lactation year: molecular approaches. *Syst Appl Microbiol*. 2007;30(7):547-60.
262. Wojcik A, Blaszkowska J, Kurnatowski P, Goralska K. Sandpits as a reservoir of potentially pathogenic fungi for children. *Ann Agric Environ Med*. 2016;23(4):542-8.
263. Adams RI, Miletto M, Taylor JW, Bruns TD. Dispersal in microbes: fungi in indoor air are dominated by outdoor air and show dispersal limitation at short distances. *ISME J*. 2013;7(7):1262-73.
264. Gonias S, Tuepker M, Heisel T, Autran C, Bode L, Gale CA. Human Milk Oligosaccharides Inhibit *Candida albicans* Invasion of Human Premature Intestinal Epithelial Cells. *J Nutr*. 2015;145(9):1992-8.
265. Tibshirani R. Regression Shrinkage and Selection via the Lasso. *J R Statist Soc B*. 1996;58(1):267-88.
266. Zhang E, Tanaka T, Tajima M, Tsuboi R, Nishikawa A, Sugita T. Characterization of the skin fungal microbiota in patients with atopic dermatitis and in healthy subjects. *Microbiol Immunol*. 2011;55(9):625-32.
267. Moossavi S, Fehr K, Derakhshani H, Sbihi H, Robertson B, Bode L, et al. Human milk fungi: environmental determinants and inter-kingdom associations with milk bacteria 2019.
268. Fung TC, Bessman NJ, Hepworth MR, Kumar N, Shibata N, Kobuley D, et al. Lymphoid-Tissue-Resident Commensal Bacteria Promote Members of the IL-10 Cytokine Family to Establish Mutualism. *Immunity*. 2016;44(3):634-46.
269. Damgaard C, Magnussen K, Enevold C, Nilsson M, Tolker-Nielsen T, Holmstrup P, et al. Viable bacteria associated with red blood cells and plasma in freshly drawn blood donations. *PLoS One*. 2015;10(3):e0120826.

270. Drell T, Stsepetova J, Simm J, Rull K, Aleksejeva A, Antson A, et al. The Influence of Different Maternal Microbial Communities on the Development of Infant Gut and Oral Microbiota. *Sci Rep*. 2017;7(1):9940.
271. Williams JE, Carrothers JM, Lackey KA, Beatty NF, Brooker SL, Peterson HK, et al. Strong Multivariate Relations Exist Among Milk, Oral, and Fecal Microbiomes in Mother-Infant Dyads During the First Six Months Postpartum. *J Nutr*. 2019;149(6):902-14.
272. Schwab C, Ramirez AG, Vischer M, Lacroix C, Voney E. Characterization of the cultivable microbiota in fresh and stored mature human breast milk. *Front Microbiol*. 2019;doi: 10.3389/fmicb.2019.02666.
273. Ballard O, Morrow AL. Human milk composition: nutrients and bioactive factors. *Pediatr Clin North Am*. 2013;60(1):49-74.
274. Dugdale AE. Evolution and infant feeding. *Lancet*. 1986;1(8482):670-3.
275. Farhadian M, Rafat SA, Hasanpur K, Ebrahimi M, Ebrahimie E. Cross-Species Meta-Analysis of Transcriptomic Data in Combination With Supervised Machine Learning Models Identifies the Common Gene Signature of Lactation Process. *Front Genet*. 2018;9:235.
276. Lefevre CM, Sharp JA, Nicholas KR. Evolution of lactation: ancient origin and extreme adaptations of the lactation system. *Annu Rev Genomics Hum Genet*. 2010;11:219-38.
277. Hennet T, Borsig L. Breastfed at Tiffany's. *Trends Biochem Sci*. 2016;41(6):508-18.
278. Duranti S, Lugli GA, Milani C, James K, Mancabelli L, Turrone F, et al. *Bifidobacterium bifidum* and the infant gut microbiota: an intriguing case of microbe-host co-evolution. *Environ Microbiol*. 2019;21(10):3683-95.
279. Turrone F, Milani C, Duranti S, Mahony J, van Sinderen D, Ventura M. Glycan Utilization and Cross-Feeding Activities by *Bifidobacteria*. *Trends Microbiol*. 2018;26(4):339-50.
280. Urbaniak C, Angelini M, Gloor GB, Reid G. Human milk microbiota profiles in relation to birthing method, gestation and infant gender. *Microbiome*. 2016;4:1.
281. Gohir W, Kennedy KM, Wallace JG, Saoi M, Bellissimo CJ, Britz-McKibbin P, et al. High-fat diet intake modulates maternal intestinal adaptations to pregnancy and results in placental hypoxia, as well as altered fetal gut barrier proteins and immune markers. *J Physiol*. 2019;597(12):3029-51.
282. Morkkala K, Roytio H, Munukka E, Pietila S, Ekblad U, Ronnema T, et al. Gut Microbiota Richness and Composition and Dietary Intake of Overweight Pregnant Women Are Related to Serum Zonulin Concentration, a Marker for Intestinal Permeability. *J Nutr*. 2016;146(9):1694-700.

283. Mazurier E, Rigourd V, Perez P, Buffin R, Couedelo L, Vaysse C, et al. Effects of Maternal Supplementation With Omega-3 Precursors on Human Milk Composition. *J Hum Lact*. 2017;33(2):319-28.
284. Pettersen-Dahl A, Murzakanova G, Sandvik L, Laine K. Maternal body mass index as a predictor for delivery method. *Acta Obstet Gynecol Scand*. 2018;97(2):212-8.
285. Falony G, Joossens M, Vieira-Silva S, Wang J, Darzi Y, Faust K, et al. Population-level analysis of gut microbiome variation. *Science* 2016 Apr 29;352(6285):560-4. 2016;352(6285):560-4.
286. Veller C, Haig D, Nowak MA. The Trivers-Willard hypothesis: sex ratio or investment? *Proc Biol Sci*. 2016;283(1830):pii: 20160126.
287. Hinde K. First-time macaque mothers bias milk composition in favor of sons. *Curr Biol*. 2007;17(22):R958-9.
288. Hinde K, Foster AB, Landis LM, Rendina D, Oftedal OT, Power ML. Daughter dearest: Sex-biased calcium in mother's milk among rhesus macaques. *Am J Phys Anthropol*. 2013;151(1):144-50.
289. Sullivan EC, Hinde K, Mendoza SP, Capitanio JP. Cortisol concentrations in the milk of rhesus monkey mothers are associated with confident temperament in sons, but not daughters. *Dev Psychobiol*. 2011;53(1):96-104.
290. Robert KA, Braun S. Milk composition during lactation suggests a mechanism for male biased allocation of maternal resources in the tammar wallaby (*Macropus eugenii*). *PLoS One*. 2012;7(11):e51099.
291. Barreiro R, Regal P, Diaz-Bao M, Vazquez BI, Cepeda A. Effects of bovine pregnancy on the fatty acid composition of milk: the significance for humans needs. *Food Addit Contam Part A Chem Anal Control Expo Risk Assess*. 2017;34(4):608-16.
292. Azad MB, Hinde K, Timpson NJ, Moossavi S, Becker AB, Mandhane P, et al. More milk for daughters? Sex differentiated breastfeeding in two birth cohorts. *Breastfeeding medicine : the official journal of the Academy of Breastfeeding Medicine*. 2018;13(7):A11.
293. Markle JG, Frank DN, Mortin-Toth S, Robertson CE, Feazel LM, Rolle-Kampczyk U, et al. Sex differences in the gut microbiome drive hormone-dependent regulation of autoimmunity. *Science*. 2013;339 (6123):1084-8.
294. Takeshita T, Kageyama S, Furuta M, Tsuboi H, Takeuchi K, Shibata Y, et al. Bacterial diversity in saliva and oral health-related conditions: the Hisayama Study. *Sci Rep*. 2016;6:22164.
295. Fujita M, Roth E, Lo YJ, Hurst C, Vollner J, Kendell A. In poor families, mothers' milk is richer for daughters than sons: a test of Trivers-Willard hypothesis in agropastoral settlements in Northern Kenya. *Am J Phys Anthropol*. 2012;149(1):52-9.



296. Quinn EA. No evidence for sex biases in milk macronutrients, energy, or breastfeeding frequency in a sample of Filipino mothers. *Am J Phys Anthropol.* 2013;152(2):209-16.
297. Al-Qaraghoul M, Fang YMV. Effect of Fetal Sex on Maternal and Obstetric Outcomes. *Front Pediatr.* 2017;5:144. doi: 10.3389/fped.2017.00144. eCollection 2017.
298. Hoche B, Chen YP, Schlemm L, Burdack A, Li J, Halle H, et al. Fetal sex determines the impact of maternal PROGINS progesterone receptor polymorphism on maternal physiology during pregnancy. *Pharmacogenet Genomics.* 2009;19(9):710-8.
299. Ashbolt NJ. Microbial Contamination of Drinking Water and Human Health from Community Water Systems. *Curr Environ Health Rep.* 2015;2(1):95-106.
300. Litvak Y, Mon KKZ, Nguyen H, Chanthavixay G, Liou M, Velazquez EM, et al. Commensal Enterobacteriaceae Protect against Salmonella Colonization through Oxygen Competition. *Cell Host Microbe.* 2019;25(1):128-39 e5.
301. Victora CG, Bahl R, Barros AJD, França GVA, Horton S, Krasevec J, et al. Breastfeeding in the 21st century: epidemiology, mechanisms, and lifelong effect. *Lancet.* 2016;387(10017):475-90.
302. Soto-Ramirez N, Kar S, Zhang H, Karmaus W. Infant feeding patterns and eczema in children in the first 6 years of life. *Clin Exp Allergy.* 2017;47(10):1285-98.
303. Almeida A, Mitchell AL, Boland M, Forster SC, Gloor GB, Tarkowska A, et al. A new genomic blueprint of the human gut microbiota. *Nature.* 2019;568(7753):499-504.
304. Neville BA, Forster SC, Lawley TD. Commensal Koch's postulates: establishing causation in human microbiota research. *Curr Opin Microbiol.* 2017;42:47-52.
305. Lebeis SL. The potential for give and take in plant-microbiome relationships. *Front Plant Sci.* 2014;5.
306. Gould AL, Zhang V, Lamberti L, Jones EW, Obadia B, Korasidis N, et al. Microbiome interactions shape host fitness. *Proc Natl Acad Sci U S A.* 2018;115(51):E11951-E60.
307. Mohan R, Sanpitakseree C, Desai AV, Sevgen SE, Schroeder CM, Kenis PJA. A microfluidic approach to study the effect of bacterial interactions on antimicrobial susceptibility in polymicrobial cultures. *RSC Advances.* 2015;5(44):35211-23.
308. Sommer M, Dantas G. Antibiotics and the resistant microbiome. *Curr Opin Microbiol.* 2011;14(5):556-63.
309. Bisanz JE, Spanogiannopoulos P, Pieper LM, Bustion AE, Turnbaugh PJ. How to Determine the Role of the Microbiome in Drug Disposition. *Drug Metab Dispos.* 2018;46(11):1588-95.

310. Chankhamjon P, Javdan B, Lopez J, Hull R, Chatterjee S, Donia MS. Systematic mapping of drug metabolism by the human gut microbiome. *bioRxiv* 2019:538215; doi: <https://doi.org/10.1101/>.
311. Zimmermann M, Zimmermann-Kogadeeva M, Wegmann R, Goodman AL. Mapping human microbiome drug metabolism by gut bacteria and their genes. *Nature*. 2019;570(7762):462-7.
312. McGuire MK, McGuire MA. Human milk: mother nature's prototypical probiotic food? *Adv Nutr*. 2015;6(1):112-23.
313. Sornplang P, Piyadeatsoontorn S. Probiotic isolates from unconventional sources: a review. *J Anim Sci Technol*. 2016;58:26.
314. Lara-Villoslada F, Sierra S, Diaz-Ropero MP, Olivares M, Xaus J. Safety assessment of the human milk-isolated probiotic *Lactobacillus salivarius* CECT5713. *J Dairy Sci*. 2007;90(8):3583-9.
315. Anandharaj M, Rani RP, Swain MR. Production of High-Quality Probiotics by Fermentation. In: Gupta VK, Treichel H, Shapaval V, de Oliveira LA, Tuohy MG, editors. *Microbial Functional Foods and Nutraceuticals*. First ed: John Wiley & Sons Ltd; 2018.
316. de Melo Pereira GV, de Oliveira Coelho B, Magalhaes Junior AI, Thomaz-Soccol V, Soccol CR. How to select a probiotic? A review and update of methods and criteria. *Biotechnol Adv*. 2018;36(8):2060-76.
317. Hill C, Guarner F, Reid G, Gibson GR, Merenstein DJ, Pot B, et al. Expert consensus document. The International Scientific Association for Probiotics and Prebiotics consensus statement on the scope and appropriate use of the term probiotic. *Nat Rev Gastroenterol Hepatol*. 2014;11(8):506-14.
318. Martín R, Jiménez E, Olivares M, Marín ML, Fernández L, Xaus J, et al. *Lactobacillus salivarius* CECT 5713, a potential probiotic strain isolated from infant feces and breast milk of a mother–child pair. *International Journal of Food Microbiology*. 2006;112(1):35-43.
319. Walker AW, Martin JC, Scott P, Parkhill J, Flint HJ, Scott KP. 16S rRNA gene-based profiling of the human infant gut microbiota is strongly influenced by sample processing and PCR primer choice. *Microbiome*. 2015;3:26.
320. Siegwald L, Caboche S, Even G, Viscogliosi E, Audebert C, Chabe M. The Impact of Bioinformatics Pipelines on Microbiota Studies: Does the Analytical "Microscope" Affect the Biological Interpretation? *Microorganisms*. 2019;7(10):393.
321. Delavenne E, Mounier J, Asmani K, Jany JL, Barbier G, Le Blay G. Fungal diversity in cow, goat and ewe milk. *Int J Food Microbiol*. 2011;151(2):247-51.
322. Darwazeh AM, Al-Bashir A. Oral candidal flora in healthy infants. *Journal of oral pathology & medicine : official publication of the International Association of Oral Pathologists and the American Academy of Oral Pathology*. 1995;24(8):361-4.

323. Adair KL, Douglas AE. Making a microbiome: the many determinants of host-associated microbial community composition. *Curr Opin Microbiol.* 2016;35:23-9.
324. Craft KM, Townsend SD. The Human Milk Glycome as a Defense Against Infectious Diseases: Rationale, Challenges, and Opportunities. *ACS Infect Dis.* 2017;4(2):77-83.
325. Zúñiga M, Monedero V, Yebra MJ. Utilization of Host-Derived Glycans by Intestinal *Lactobacillus* and *Bifidobacterium* Species. *Front Microbiol.* 2018;9:1917.
326. Juge N, Tailford L, Owen CD. Sialidases from gut bacteria: a mini-review. *Biochem Soc Trans.* 2016;44(1):166-75.
327. Asnicar F, Manara S, Zolfo M, Truong DT, Scholz M, Armanini F, et al. Studying Vertical Microbiome Transmission from Mothers to Infants by Strain-Level Metagenomic Profiling. *mSystems.* 2017;2(1):e00164-16.
328. McNabney S, Henagan T. Short Chain Fatty Acids in the Colon and Peripheral Tissues: A Focus on Butyrate, Colon Cancer, Obesity and Insulin Resistance. *Nutrients.* 2017;9(12):1348.
329. Fujita Y, Matsuoka H, Hirooka K. Regulation of fatty acid metabolism in bacteria. *Mol Microbiol.* 2007;66(4):829-39.
330. Gibson GR, Hutkins R, Sanders ME, Prescott SL, Reimer RA, Salminen SJ, et al. Expert consensus document: The International Scientific Association for Probiotics and Prebiotics (ISAPP) consensus statement on the definition and scope of prebiotics. *Nat Rev Gastroenterol Hepatol.* 2017;14(8):491-502.
331. Baker LY, Hobby CR, Siv AW, Bible WC, Glennon MS, Anderson DM, et al. *Pseudomonas aeruginosa* responds to exogenous polyunsaturated fatty acids (PUFAs) by modifying phospholipid composition, membrane permeability, and phenotypes associated with virulence. *BMC Microbiol.* 2018;18(1):117.
332. Moravec AR, Siv AW, Hobby CR, Lindsay EN, Norbash LV, Shults DJ, et al. Exogenous Polyunsaturated Fatty Acids Impact Membrane Remodeling and Affect Virulence Phenotypes among Pathogenic *Vibrio* Species. *Appl Environ Microbiol.* 2017;83(22):e01415-17.
333. Eder AE, Munir SA, Hobby CR, Anderson DM, Herndon JL, Siv AW, et al. Exogenous polyunsaturated fatty acids (PUFAs) alter phospholipid composition, membrane permeability, biofilm formation and motility in *Acinetobacter baumannii*. *Microbiology.* 2017;163(11):1626-36.
334. Hobby CR, Herndon JL, Morrow CA, Peters RE, Symes SJK, Giles DK. Exogenous fatty acids alter phospholipid composition, membrane permeability, capacity for biofilm formation, and antimicrobial peptide susceptibility in *Klebsiella pneumoniae*. *Microbiologyopen.* 2018:e00635.

335. Fredrikzon B, Hernell O, Bläckberg L. Lingual lipase. Its role in lipid digestion in infants with low birthweight and/or pancreatic insufficiency. *Acta Paediatr Scand Suppl* 1982;296:75-80. 1982;296:75-80.
336. Bi Y, Cox MS, Zhang F, Suen G, Zhang N, Tu Y, et al. Feeding modes shape the acquisition and structure of the initial gut microbiota in newborn lambs. *Environ Microbiol*. 2019; 21(7):2333-2346.
337. Costello EK, Carlisle EM, Bik EM, Morowitz MJ, Relman DA. Microbiome assembly across multiple body sites in low-birthweight infants. *MBio*. 2013;4(6):e00782-13.
338. Ding T, Schloss PD. Dynamics and associations of microbial community types across the human body. *Nature*. 2014;509(7500):357-60.
339. Biesbroek G, Bosch AA, Wang X, Keijser BJ, Veenhoven RH, Sanders EA, et al. The impact of breastfeeding on nasopharyngeal microbial communities in infants. *Am J Respir Crit Care Med*. 2014;190(3):298-308.
340. Thomson P, Medina DA, Garrido D. Human milk oligosaccharides and infant gut bifidobacteria: Molecular strategies for their utilization. *Food Microbiol*. 2018;75:37-46.
341. Duranti S, Lugli GA, Mancabelli L, Armanini F, Turrone F, James K, et al. Maternal inheritance of bifidobacterial communities and bifidophages in infants through vertical transmission. *Microbiome*. 2017;5(1):66.
342. Milani C, Mancabelli L, Lugli GA, Duranti S, Turrone F, Ferrario C, et al. Exploring Vertical Transmission of Bifidobacteria from Mother to Child. *Appl Environ Microbiol*. 2015;81(20):7078-87.
343. Jost T, Lacroix C, Braegger CP, Rochat F, Chassard C. Vertical mother-neonate transfer of maternal gut bacteria via breastfeeding. *Environ Microbiol*. 2014;16(9):2891-904.
344. Haag KL. Holobionts and their hologenomes: Evolution with mixed modes of inheritance. *Genet Mol Biol*. 2018;41(1 suppl 1):189-97.
345. Lane AA, McGuire MK, McGuire MA, Williams JE, Lackey KA, Hagen EH, et al. Household composition and the infant fecal microbiome: The INSPIRE study. *Am J Phys Anthropol*. 2019;doi: 10.1002/ajpa.23843.
346. Simon JC, Marchesi JR, Mougél C, Selosse MA. Host-microbiota interactions: from holobiont theory to analysis. *Microbiome*. 2019;7(1):5.
347. Kemp MW, Kallapur SG, Jobe AH, Newnham JP. Obesity and the developmental origins of health and disease. *J Paediatr Child Health*. 2012;48(2):86-90.
348. Stiemsma LT, Michels KB. The Role of the Microbiome in the Developmental Origins of Health and Disease. *Pediatrics*. 2018;141(4):e20172437.

349. Laforest-Lapointe I, Arrieta MC. Patterns of Early-Life Gut Microbial Colonization during Human Immune Development: An Ecological Perspective. *Front Immunol.* 2017;8:788.
350. Fukami T. Historical Contingency in Community Assembly: Integrating Niches, Species Pools, and Priority Effects. *Annu Rev Ecol Evol Syst.* 2015;46(1):1-23.
351. Ek WE, Karlsson T, Hernández CA, Rask-Andersen M, Johansson Å. Breast-feeding and risk of asthma, hay fever, and eczema. *J Allergy Clin Immunol.* 2018;141(3):1157-9.e9.
352. Abrahamsson TR, Jakobsson HE, Andersson AF, Bjorksten B, Engstrand L, Jenmalm MC. Low gut microbiota diversity in early infancy precedes asthma at school age. *Clin Exp Allergy.* 2014;44(6):842-50.
353. Arrieta MC, Stiemsma LT, Dimitriu PA, Thorson L, Russell S, Yurist-Doutsch S, et al. Early infancy microbial and metabolic alterations affect risk of childhood asthma. *Sci Transl Med.* 2015;7(307):307ra152. doi: 10.1126/scitranslmed.aab2271.
354. Ruokolainen L, Paalanen L, Karkman A, Laatikainen T, von Hertzen L, Vlasoff T, et al. Significant disparities in allergy prevalence and microbiota between the young people in Finnish and Russian Karelia. *Clin Exp Allergy.* 2017;47(5):665-74.
355. Bridgman SL, Kozyrskyj AL, Scott JA, Becker AB, Azad MB. Gut microbiota and allergic disease in children. *Ann Allergy Asthma Immunol.* 2016;116(2):99-105.
356. Ardeshtir A, Narayan NR, Méndez-Lagares G, Lu D, Rauch M, Huang Y, et al. Breast-fed and bottle-fed infant rhesus macaques develop distinct gut microbiotas and immune systems. *Sci Transl Med.* 2014;6(252):252ra120.
357. Miliku K, Robertson B, Sharma AK, Subbarao P, Becker AB, Mandhane PJ, et al. Human milk oligosaccharide profiles and food sensitization among infants in the CHILD Study. *Allergy.* 2018;73(10):2070-3.
358. Seppo AE, Autran CA, Bode L, Jarvinen KM. Human milk oligosaccharides and development of cow's milk allergy in infants. *J Allergy Clin Immunol.* 2017;139(2):708-11 e5.
359. Adlerberth I, Strachan DP, Matricardi PM, Ahrne S, Orfei L, Aberg N, et al. Gut microbiota and development of atopic eczema in 3 European birth cohorts. *J Allergy Clin Immunol.* 2007;120(2):343-50.
360. Bisgaard H, Li N, Bonnelykke K, Chawes BL, Skov T, Paludan-Muller G, et al. Reduced diversity of the intestinal microbiota during infancy is associated with increased risk of allergic disease at school age. *J Allergy Clin Immunol.* 2011;128(3):646-52 e1-5.
361. Boutin S, Depner M, Stahl M, Graeber SY, Dittrich SA, Legatzki A, et al. Comparison of Oropharyngeal Microbiota from Children with Asthma and Cystic Fibrosis. *Mediators Inflamm.* 2017;2017:5047403.

362. Seo S, Choung JT, Chen BT, Lindsley WG, Kim KY. The level of submicron fungal fragments in homes with asthmatic children. *Environ Res.* 2014;131:71-6.
363. Sharpe RA, Bearman N, Thornton CR, Husk K, Osborne NJ. Indoor fungal diversity and asthma: a meta-analysis and systematic review of risk factors. *J Allergy Clin Immunol.* 2015;135(1):110-22.
364. Sbihi H, Allen RW, Becker A, Brook JR, Mandhane P, Scott JA, et al. Perinatal Exposure to Traffic-Related Air Pollution and Atopy at 1 Year of Age in a Multi-Center Canadian Birth Cohort Study. *Environ Health Perspect.* 2015;123(9):902-8.
365. Dudgeon LS, Jewesbury RC. The Bacteriology of Human Milk. *J Hyg (Lond).* 1924;23(1):64-76.
366. Sosa R, Barness L. Bacterial Growth in Refrigerated Human Milk. *Am J Dis Child.* 1987;141(1):111-2.
367. Hunt KM, Williams JE, Shafii B, Hunt MK, Behre R, Ting R, et al. Mastitis is associated with increased free fatty acids, somatic cell count, and interleukin-8 concentrations in human milk. *Breastfeeding medicine : the official journal of the Academy of Breastfeeding Medicine.* 2013;8(1):105-10.
368. Vandeputte D, Kathagen G, D'hoel K, Vieira-Silva S, Valles-Colomer M, Sabino J, et al. Quantitative microbiome profiling links gut community variation to microbial load. *Nature.* 2017;551(7681):507-11.
369. Vieira-Silva S, Sabino J, Valles-Colomer M, Falony G, Kathagen G, Caenepeel C, et al. Quantitative microbiome profiling disentangles inflammation- and bile duct obstruction-associated microbiota alterations across PSC/IBD diagnoses. *Nat Microbiol.* 2019;4(11):1826-31.
370. Ryall B, Eydallin G, Ferenci T. Culture history and population heterogeneity as determinants of bacterial adaptation: the adaptomics of a single environmental transition. *Microbiol Mol Biol Rev.* 2012;76(3):597-625.
371. Takhaveev V, Heinemann M. Metabolic heterogeneity in clonal microbial populations. *Curr Opin Microbiol.* 2018;45:30-8.
372. Maurice CF, Turnbaugh PJ. Quantifying the metabolic activities of human-associated microbial communities across multiple ecological scales. *FEMS Microbiol Rev.* 2013;37(5):830-48.
373. Carmody RN, Bisanz JE, Bowen BP, Maurice CF, Lyalina S, Louie KB, et al. Cooking shapes the structure and function of the gut microbiome. *Nat Microbiol.* 2019;4(12):2052-63.
374. Landers S, Updegrave K. Bacteriological screening of donor human milk before and after Holder pasteurization. *Breastfeeding medicine : the official journal of the Academy of Breastfeeding Medicine.* 2010;5(3):117-21.

375. Kim J, Unger S. Human milk banking. *Paediatr Child Health*. 2010;15(9):595-602.
376. Groer MW, Luciano AA, Dishaw LJ, Ashmeade TL, Miller E, Gilbert JA. Development of the preterm infant gut microbiome: a research priority. *Microbiome*. 2014;2:38.
377. American Academy of Pediatrics. Breastfeeding and the use of human milk. *Pediatrics*. 2012;129(3):e827-41.
378. Cacho NT, Harrison NA, Parker LA, Padgett KA, Lemas DJ, Marcial GE, et al. Personalization of the Microbiota of Donor Human Milk with Mother's Own Milk. *Frontiers in Microbiology*. 2017;8.
379. Fernandez L, Ruiz L, Jara J, Orgaz B, Rodriguez JM. Strategies for the Preservation, Restoration and Modulation of the Human Milk Microbiota. Implications for Human Milk Banks and Neonatal Intensive Care Units. *Front Microbiol*. 2018;9:2676.
380. Moossavi S. Normal Microbiota: Let us Change our Perspective. *Arch Iran Med*. 2019;22(4):220.
381. [No Author listed]. Bacteria in milk and its products. *Science*. 1889;14(341):116-8.