TALLINN UNIVERSITY OF TECHNOLOGY
DOCTORAL THESIS
46/2019

The Effect of Maternal Microbial Communities and Preterm Birth on the Development of Infant Gut Microbiota

TIINA DRELL

TAL TECH PRESS

TALLINN UNIVERSITY OF TECHNOLOGY
School of Science
Department of Chemistry and Biotechnology
This dissertation was accepted for the defence of the degree 15/03/2019

Supervisor: Epp Sepp, MD, PhD
Institute of Biomedicine and Translational Medicine
University of Tartu
Tartu, Estonia

Co-supervisor: Professor Madis Metsis, PhD
School of Natural Sciences and Health
Tallinn University
Tallinn, Estonia

Opponents: Anne Salonen, PhD
Department of Bacteriology and Immunology
University of Helsinki
Helsinki, Finland

Elin Org, PhD
Institute of Genomics
University of Tartu
Tartu, Estonia

Defence of the thesis: 03/10/2019, Tallinn

Declaration:
Hereby I declare that this doctoral thesis, my original investigation and achievement, submitted for the doctoral degree at Tallinn University of Technology has not been submitted for doctoral or equivalent academic degree.

Tiina Drell

---
signature

European Union
European Regional
Development Fund

Investing
in your future

ISSN 2585-6898 (publication)
ISBN 978-9949-83-463-1 (publication)
ISSN 2585-6901 (PDF)
ISBN 978-9949-83-464-8 (PDF)

TALLINNA TEHNIKAÜLIKOOL
DOKTORITÖÖ
46/2019

Ema mikroobikoosluste ja enneaegse sünni mõju imiku seedetrakti mikrobioota kujunemisele

TIINA DRELL

TAL TECH
KIRJASTUS

Table of Contents

LIST OF PUBLICATIONS **7**

AUTHOR'S CONTRIBUTION TO THE PUBLICATIONS **8**

INTRODUCTION **9**

ABBREVIATIONS **10**

1 LITERATURE OVERVIEW **11**

1.1 NORMAL HUMAN MICROBIOTA IN ADULTS 11
1.2 FUNCTIONS OF THE NORMAL GUT MICROBIOTA 12
1.3 THE DEVELOPMENT OF THE GUT MICROBIOTA 13
1.3.1 Factors influencing the development of the gut microbiota *14*
1.3.1.1 Mode of delivery 14
1.3.1.2 Feeding regimen 14
1.3.1.3 Antibiotic treatment 15
1.3.2 The effect of preterm birth on the development of infant gut microbiota *17*
1.4 THE COMPOSITION OF MOTHERS' VAGINAL MICROBIOTA AND PREGNANCY OUTCOMES 19
1.5 VERTICAL TRANSMISSION OF MICROBES FROM THE MOTHER'S NORMAL MICROBIOTA OF DIFFERENT BODY SITES TO THE INFANT'S GUT MICROBIOTA 20

2 THE AIM OF THE THESIS **21**

3 MATERIALS AND METHODS **22**

3.1 STUDY COHORT AND SAMPLING 22
3.2 METHODS USED FOR BACTERIAL COMMUNITY PROFILING 23
3.3 ADDITIONAL METHODS 25
3.4 STATISTICAL ANALYSIS 25

4 RESULTS AND DISCUSSION **26**

4.1 THE VARIATION OF VAGINAL MICROBIOTA OF WOMEN OF REPRODUCTIVE AGE (PUBLICATION I AND III, UNPUBLISHED DATA) 26
4.1.1 The variation of vaginal bacterial community *26*
4.1.2 The variation of vaginal fungal community *29*
4.2 THE EFFECT OF PRETERM BIRTH IN COMBINATION WITH EXTREMELY LOW BIRTH WEIGHT ON THE DEVELOPMENT OF THE GUT MICROBIOTA (PUBLICATION II AND III) 29
4.2.1 Development of the gut microbiota in preterm infants *29*
4.2.2 Differences between the development of the gut microbiota in pre- and full term infants *32*
4.3 THE INFLUENCE OF DIFFERENT MATERNAL MICROBIAL COMMUNITIES ON THE DEVELOPMENT OF INFANT GUT AND ORAL MICROBIOTA (PUBLICATION III, UNPUBLISHED DATA) 33
4.3.1 Diversity and similarity of different maternal microbial communities with infant gut and oral microbiota *33*
4.3.2 The influence of different maternal microbial communities on the development of infant gut microbiota *34*
4.3.2.1 The similarity of bifidobacterial communities observed in different maternal microbial communities and infant gut microbiota 36
4.3.3 The influence of different maternal microbial communities on the development of infant oral microbiota *37*

5 CONCLUSIONS 39
REFERENCES 41
ACKNOWLEDGEMENTS 53
ABSTRACT 54
LÜHIKOKKUVÕTE 56
APPENDIX 59
CURRICULUM VITAE 96
ELULOOKIRJELDUS 97

List of Publications

The list of publications, by which the present thesis has been prepared:

I Drell, T., Lillsaar, T., Tummeleht, L., Simm, J., Aaspõllu, A., Väin, E., Saarma, I., Salumets, A., Donders, G. G. G., and Metsis, M. (2013). Characterization of the vaginal micro- and mycobiome in asymptomatic reproductive-age Estonian women. PLoS One 8, e54379.

II Drell, T., Lutsar, I., Štšepetova, J., Parm, Ü., Metsvaht, T., Ilmoja, M., Simm, J., and Sepp, E. (2014). The development of gut microbiota in critically ill extremely low birth weight infants assessed with 16S rRNA gene-based sequencing. Gut Microbes 5, 304 – 503.

III Drell, T., Štšepetova, J., Simm, J., Rull, K., Aleksejeva, A., Antson, A., Tillmann, V., Metsis, M., Sepp, E., Salumets, A., and Mändar, R. (2017). The influence of different maternal microbial communities on the development of infant gut and oral microbiota. Sci Rep 7, 9940.

Author's Contribution to the Publications

Contribution to the papers in this thesis are:

I In the first publication (publication I), I was one of the main authors who carried out the experiments and drafted the manuscript.

II In the second publication (publication II), I was the main author who carried out the experiments and biostatistical analysis, prepared the figures and drafted the manuscript.

III In the third publication (publication III), I was the main author who carried out the experiments and biostatistical analysis, prepared the figures and drafted the manuscript.

Introduction

Humans have on average $4x10^{13}$ bacteria harboring their body with niches varying from skin to gastrointestinal tract. The ratio of bacteria to human cells is estimated to be approximately 1:1 with the majority of the bacteria residing in the gastrointestinal tract (Sender *et al.*, 2016). The gut microbiota is in a mutualistic relationship with the host, and it is essential for digestive processes such as fermentation of carbohydrates, participating in the development of the immune system, and providing resistance to pathogens. The composition of the gut microbiota is dynamic over time, and it develops during the first years of infant's life, with colonization beginning immediately after birth and the development lasting approximately three years.

Several factors influence the development of the gut microbiota. From these factors, mode of delivery, feeding regimen, antibiotic treatment and preterm birth have been studied most extensively. Nevertheless, many studies have found contradictory results, and thus the knowledge about the effect these factors have on the development of the gut microbiota is still inconclusive. Also, relatively little is known about the mother-to-infant transmission of microbes, which has been hypothesized to be a significant source of bacteria for infant's microbial communities, especially in the case of vaginal birth. Studies that are focusing on vertical transmission of bacteria, especially using large-scale molecular methods, have only lately been carried out.

A better understanding of the development of the gut microbiota is essential to fully comprehend the functions this community has in the host and the disorders that have been associated with disruptions of the development of the gut microbiota. Several of these disorders develop in infancy, such as necrotizing enterocolitis and neonatal sepsis, but have also been suggested to develop later in life, such as diabetes, obesity, and allergies.

In current thesis, the establishment and development of infant gut microbiota were studied during the first months of life by analyzing the similarities between the microbial communities of full term infants and their mothers; and characterizing the development pattern of the gut microbiota in extremely preterm infants. The aim was to expand the knowledge about the effect that different maternal microbial communities and extremely preterm birth have on the development of the gut microbiota.

Abbreviations

BW	birth weight
DGGE	denaturing gradient gel electrophoresis
ELBW	extremely low birthweight
EOS	early onset sepsis
GA	gestational age
HITChip	human intestinal tract chip
HMO	human milk oligosaccharides
IAP	intrapartum antimicrobial prophylactic treatment
ITS-1	internal transcribed spacer-1
IQR	interquartile range
LOS	late onset sepsis
NEC	necrotizing enterocolitis
NICU	neonatal intensive care unit
OTU	operational taxonomic unit
PCA	principal component analysis
RAPD	random amplification of polymorphic DNA
SCFA analysis	Short chain fatty acid analysis
SD	standard deviation
sIgA	secretory immunoglobulin A
TTGE	temperature gradient gel electrophoresis

1 Literature overview

1.1 Normal human microbiota in adults

In humans, each body surface colonized with microbes harbors a distinct microbial community that differs in composition from other body habitats. For instance, adult oral and gut microbiota are very diverse communities whereas vaginal microbiota has low and skin microbiota intermediate microbial diversity (The Human Microbiome Project Consortium, 2012). In addition, specific dominant microbes are characteristic to each of these habitats – lactobacilli usually dominate the vaginal microbiota; members from genera *Staphylococcus*, *Propionibacterium* and *Corynebacterium* the skin and nasal microbiota; streptococci the oral microbiota; and variety of members from genus *Bacteroides* and phylum *Firmicutes* dominate the gut microbiota (The Human Microbiome Project Consortium, 2012).

Generally, the microbial communities colonizing different body sites in adults are relatively stable over time but differ between individuals (The Human Microbiome Project Consortium, 2012, Lloyd-Price *et al.*, 2017). Nevertheless, members from different taxa may follow different patterns of variation. For instance, the abundance of members from phylum *Bacteroidetes*, more specifically genus *Bacteroides,* tends to vary in the gut microbiota between individuals, while the abundance of members from phylum *Firmicutes* is more temporally dynamic within individuals (Lloyd-Price *et al.*, 2017). Variation pattern may also differ between body habitats. Microbial community colonizing the saliva, while having high microbial diversity, is similar between different individuals but skin microbiota with intermediate microbial diversity differs significantly between individuals (The Human Microbiome Project Consortium, 2012, Lloyd-Price *et al.,* 2017).

Several factors may influence the variation of microbial communities colonizing different body habitats, but they are not well understood yet (The Human Microbiome Project Consortium, 2012) and may have a cumulative indistinguishable effect. One of the environmental factors that have been shown to have a strong correlation with the microbial community colonizing the vagina is vaginal pH. Higher vaginal pH is often associated with increased diversity and reduced levels of genus *Lactobacillus* in the vaginal environment (The Human Microbiome Project Consortium, 2012). Additionally, differing community composition of vaginal, nasal, skin, gastrointestinal and oral microbiota has been associated with person's ethnicity, and shifts in skin microbiota with person's age (The Human Microbiome Project Consortium, 2012).

Besides different environmental factors, host's genetics also plays a role in this variation as the inter-individual differences of communities (skin, nasal, oral and vaginal cavity, and the gut) have been shown to correlate with genetic variation of the host (Blekhman *et al*., 2015). Host's genetics has a taxon-specific effect on the gut microbiota with some members of the families *Christensenellaceae*, *Bifidobacteriaceae, Ruminococcaceae* and *Lachnospiraceae* having a strong heritability score (indicative of the proportion of variance in abundance attributable to genetic effects as opposed to common environment), while members from phylum *Bacteroidetes* are mostly environmentally determined (Goodrich *et al.*, 2014, Goodrich *et al.*, 2016). Members of highly heritable taxa have higher levels of stability than those belonging to environmentally determined taxa (Goodrich *et al.*, 2016). However, the heritability of the microbes tends to be slightly lower than the heritability of other complex traits measured in the same population (e.g. systolic blood pressure, diseases of the digestive system etc.) (Goodrich *et al.*, 2016).

1.2 Functions of the normal gut microbiota

The gut microbiota has a mutualistic relationship with the host forming a symbiosis that is sometimes collectively called as a superorganism (Sleator, 2010). The stability of this symbiotic relationship is dependent on the competitive interactions between members of the microbial community (Coyte *et al.*, 2015), and selective mechanisms of the host, such as the growth-promoting secretions produced by epithelial cells (Schluter & Foster, 2012) and compartmentalizing species within gut communities (Coyte *et al.*, 2015).

Microbes colonizing the gastrointestinal tract are involved in the production of metabolites that have different biological functions in the host, such as modulating intestinal permeability, immune system and nutrient adsorption, and providing an energy source to epithelial cells (Li *et al.*, 2008, Nicholson *et al.*, 2012). Members of the gut microbiota also produce vitamins biotin (B8), cobalamin (B12), folate (B9), niacin (B3), pantothenate (B5), pyridoxine (B7), riboflavin (B2), and thiamin (B1) (Magnusdottir *et al*., 2015). A critical group of intestinal microbial products are short chain fatty acids, such as butyrate and propionate, that are the end-product of saccharolytic fermentation of non-digestible carbohydrates (Nicholson *et al.*, 2012). Butyrate is suggested to be a considerable source of energy for colonic epithelial cells (Roediger, 1982), and harbor extensive immunomodulatory properties by alternating immune responses of intestinal macrophages and blocking the generation of dendritic cells via the inhibition of histone deacetylases (Singh *et al.*, 2010, Chang *et al.*, 2014). Butyrate has also been shown to promote the differentiation of regulatory T cells in the colon (Furusawa *et al.*, 2013). All these mechanisms most likely contribute to the maintenance of tolerance to commensals in the gut.

Both butyrate and propionate have been shown to affect glucose and energy homeostasis of the host by activating intestinal gluconeogenesis (the process of glucose synthesis from non-carbohydrates) (De Vadder *et al.*, 2014). Butyrate activates gluconeogenesis in a direct manner through a cAMP-dependent mechanism and propionate through initiating a gut-brain neural circuit that induces the gluconeogenesis (De Vadder *et al.*, 2014). Butyrate-producing bacteria are widely distributed within *Clostridium* cluster XIVa (such as *Eubacterium rectale* and *Roseburia cecicola*) (Barcenilla *et al*., 2000), but they also include *Faecalibacterium prausnitzii*, among others from *Clostridium* cluster IV (Louis *et al.*, 2010, Louis *&* Flint, 2017). Propionate-producing bacteria belong to phylum *Bacteroidetes* and phylum *Firmicutes*, more specifically to subclass *Negativicutes* and *Clostridium* cluster XIVa (members of the family *Lachnospiraceae*, such as *Coprococcus catus, Ruminococcus obeum* and *Roseburia inulinivorans*) (Reichardt *et al*., 2014).

In addition to immunomodulation, the gut microbiota modulates the intestinal permeability by affecting the mucus layer, which acts as a barrier between bacteria and intestinal epithelium. For instance, the permeability of the inner mucus layer is transmissible in germ-free mice when exposing them to fecal contents originating from conventional mice (Jakobsson *et al*., 2015). Some *Bifidobacterium* strains improve intestinal barrier functions by protecting the epithelial barrier against TNFα-induced injury and promoting the restoration of epithelial tight junction barrier (Hsieh *et al*., 2015). Also, a recent study by Chelakkot *et al.*, (2018) has suggested that the gut permeability may be regulated by *Akkermansia municiphila* through the secretion of extracellular vesicles that enhance the tight junction function (Chelakkot *et al.*, 2018).

The gut microbiota also affects the glycosylation and transcription of mucins, which are heavily glycosylated proteins and primary component of the mucus layer (Comelli *et al.*,

2007, Arike *et al*., 2017). The concentration of glycosyltransferases is decreased in the intestinal epithelial cells and glycans tend to be shorter in the mucus layer of germ-free mice than in conventional mice (Arike *et al.*, 2017). Also, mucin expression is up-regulated in germ-free mice indicating to a possibility that increased mucin synthesis acts as a defense mechanism to compensate for the lack of endogenous microbiota acting as a first-line defense against luminal aggression (Comelli *et al*., 2007).

1.3 The development of the gut microbiota

It has been common knowledge over the years that the development of the normal gut microbiota begins straight after birth with maternal vaginal and fecal microbiota being the critical source of colonizing microbes. After the initial colonization process, the infant gut microbiota changes throughout the first three years of life (Yatsunenko *et al.*, 2012). These changes are accompanied with the decrease in inter-individual variation of the community composition (Jakobsson *et al.*, 2014) as the gut microbiota is more diverse between infants than between adults (Palmer *et al.*, 2007, Yatsunenko *et al.*, 2012). The infant gut microbiota prevailingly harbors representatives from four phyla, *Actinobacteria* (mainly genus *Bifidobacterium*), *Firmicutes* (streptococci, staphylococci, clostridia etc.), *Proteobacteria* (family *Enterobacteriaceae*) and *Bacteroidetes* (genus *Bacteroides*) (Koenig *et al.*, 2011, Turroni *et al.*, 2012, Azad *et al.*, 2013, Abrahamsson *et al.*, 2014). Bifidobacteria dominate the infant gut microbiota throughout the first year of life (Yatsunenko *et al.*, 2012, Turroni *et al.*, 2012), but abundant members also include staphylococci and gammaproteobacteria (e.g., *Enterobacteriaceae*) (Palmer *et al.*, 2007, Eggesbø *et al.*, 2011). The latter are facultative anaerobes present in the gut microbiota at the very early age and they are hypothesized to play a role in the consumption of oxygen in the infant's gut, thereby creating an anaerobic environment so that the members from obligately anaerobic taxa, such as bifidobacteria, can become the dominant representatives of the niche (Adlerberth & Wold, 2009). In line with this hypothesis, the levels of both staphylococci and gammaproteobacteria decrease substantially during the first months of life (Eggesbø *et al*., 2011, Abrahamsson *et al.*, 2014).

After the first year of life, the composition of infant gut microbiota gradually becomes more adult-like with the dominance of members from phylum *Firmicutes*, namely clostridia (Palmer *et al.*, 2007) and the levels of bifidobacteria declining proportionally with the increasing age (Yatsunenko *et al.*, 2012, Bergström *et al.*, 2014, Yassour *et al*., 2016). Nevertheless, members of different taxa follow distinct patterns of maturation. For instance, the levels of members from *Clostridium* cluster XIVa and *Proteobacteria* increase to adult-like levels already in 1–2 years old children, but members of some of the taxa, such as phylum *Actinobacteria* and class *Bacilli* (e.g. streptococci and lactobacilli), are stably more abundant in children as old as 5 years than in adults. Thus, the maturation of the gut microbiota may not be entirely over even at five years of age (Cheng *et al.*, 2016).

Another genus regularly detected in infant gut microbiota is *Bacteroides*. Members of this genus are only observed dominating in a subset of infants and can be identified at a very early age of a few days to months (Palmer *et al.*, 2007, Yassour *et al.*, 2016). *Bacteroides* is one of the most abundant genus colonizing adult gut microbiota, but the abundance also varies significantly between individuals (Arumugam *et al*., 2011). The reason for members of this genus to thrive in the early gut microbiota of some of the infants is unknown. The ecological interactions between members of different bacterial taxa (e.g. competition and commensalism among microbes), which influences the

abundance of members from phyla *Proteobacteria* (e.g. *Enterobacteriaceae*) and *Firmicutes* (e.g. clostridia), are most probably not having an effect on the levels of members from genus *Bacteroides* from phylum *Bacteroidetes* as shown by Trosvik *et al.*, (2010).

1.3.1 Factors influencing the development of the gut microbiota

1.3.1.1 Mode of delivery

It is hypothesized that the important source of initial colonizers of the infant gut microbiota is mothers' vaginal and perineal microbiota and these colonizers are acquired during a vaginal birth (Adlerberth & Wold, 2009). Cesarean section (C-section) rates (percentage of surgical deliveries from the total number of live births) have been growing worldwide during the past few decades reaching 23% in Europe in recent years (Betrán *et al*., 2016). Thus, the development of the gut microbiota of a substantial number of infants may be affected by this type of delivery, during what the contact with the vaginal and perineal environment is lacking. In line with this hypothesis, several studies have observed differences in the composition of the gut microbiota between vaginally delivered and C-section infants. The gut microbiota of vaginally delivered infants has been shown to have a higher resemblance to their own mothers' gut microbiota than to other mothers, which is not the case for the gut microbiota of C-section infants (Jakobsson *et al.*, 2014, Bäckhed *et al.*, 2015). The microbial diversity of the gut microbiota (Jakobsson *et al.*, 2014, Bokulich *et al.*, 2016) and variety of phylum *Bacteroidetes* (Jakobsson *et al.*, 2014) are lower in C-section infants than in vaginally born infants. Studies have also observed the lower abundance of members of genus *Bacteroides* in the gut microbiota of C-section infants (Azad *et al.*, 2013, Jakobsson *et al.*, 2014, Bäckhed *et al.*, 2015, Yassour *et al.*, 2016).

Although many studies have observed differences between the composition of the gut microbiota of C-section and vaginally delivered infants, others have failed to see these differences. For instance, a large-scale study by Chu *et al.*, (2017) which applied whole-genome shotgun sequencing, was unable to observe differences between these two groups (Chu *et al.*, 2017). Some studies have also found lower levels of bifidobacteria in the gut microbiota of C-section infants (Penders *et al.*, 2006, Biasucci *et al.*, 2010, Bäckhed *et al.*, 2015), while others have failed to confirm these results (Jakobsson *et al.*, 2014, Yassour *et al.*, 2016). In addition, Yassour *et al.*, (2016) observed low levels of members from genus *Bacteroides* coupled with lower microbial diversity in the gut microbiota of subset of vaginally born infants indicating that low microbial diversity may be associated with lower levels of members from genus *Bacteroides* in the infant gut microbiota regardless of mode of delivery (Yassour *et al.*, 2016). Thus, further studies using high-throughput methods and larger cohorts are needed to fully understand the possible effect different delivery modes have on the development of the gut microbiota.

1.3.1.2 Feeding regimen

Breast milk is the natural first food for infants that also protects them from disease while the infant's immune system matures. The protection is provided through immunological factors present in the breast milk which include, among others, macrophages and lymphocytes, cytokines (mainly TGF-β family), chemokines, and antibodies (primarily secretory IgA (SIgA)) (Ballard & Morrow, 2013). In addition to immunological protection of infants, the components of the breast milk, such as SIgA, also affect the development of the infant gut microbiota. Animal studies have shown that the number of taxa which

members change in abundance from weaning to adulthood is considerably higher in offspring from mice, who do not receive passive SIgA with breast milk, than in offspring of mice who do, and this difference increases with age, which indicates that SIgA may affect the development of more homogeneous gut microbiota (Rogier *et al*., 2014). Also, breast milk is thought to be a source for prebiotics called human milk oligosaccharides (HMOs; a family of structurally diverse unconjugated glycans) that encourage the growth of certain strains of beneficial bacteria in infant gut microbiota. *Bifidobacterium infantis* and also two members from genus *Bacteroides* (*Bacteroides fragilis* and *Bacteroides vulgatus*) have been shown to be able to metabolize free glycans from breast milk, unlike other species inhabiting the gut microbiota (such as clostridia, eubacteria, enterococci and *Escherichia coli*, but also other species of genus *Bifidobacterium*) (Ward *et al.*, 2007, Marcobal *et al.*, 2010). Being able to metabolize free glycans from breast milk may give these species a competitive advantage in the gut microbiota of breastfed infants.

Although breast milk has beneficial properties for infant's health (including to the development of the gut microbiota as described above) and more than 80% of the infants receive breast milk at least to some extent in their early lives, less than half of these infants are exclusively breastfed (Victora *et al.*, 2016). Thus, many newborns are only shortly exposed to the positive effects of breastfeeding. The duration of breastfeeding has been correlated with higher abundance of members from genus *Bifidobacterium* in the gut microbiota of breastfed infants (Yassour *et al*., 2016) with formula-fed infants having lower levels of bifidobacteria in their gut microbiota (Harmsen *et al.*, 2000, Fallani *et al.*, 2010, Bezirtzoglou *et al.*, 2011). It has to be noted that this has only been observed when infants have been fed a with a formula that does not contain pre- or probiotic supplement (Harmsen *et al.*, 2000, Fallani *et al.*, 2010, Bezirtzoglou *et al.*, 2011). When infants receive formula that does contains oligosaccharides and/or probiotics, the gut microbiota is not markedly different from breastfed infants (Rinne *et al.*, 2005, Kaassens *et al*., 2009) and usually high levels of members from genus *Bifidobacterium* can be observed (Rinne *et al.*, 2005, Penders *et al.*, 2006). Nevertheless, the gut microbiota of breastfed infants is more stable over time than in formula-fed infants receiving a formula containing oligosaccharides and/or probiotics (Kaassens *et al*., 2009).

Following the effect that breastfeeding has on the development of the gut microbiota, introduction to solid foods also impacts the development of the gut microbiota. Based on the World Health Organization recommendations, infants should be introduced to weaning when the need for energy and nutrients exceeds what can be provided through breastfeeding and/or formula, usually at around six months of age (Pan American Health Organization/ World Health Organization (PAHO/WHO), 2001). After the initiation of weaning, the composition of the gut microbiota starts to change rapidly and transitions to an "adult-like" community profile (Palmer *et al*., 2007, Koenig *et al*., 2011, Bäckhed *et al.*, 2015), which is characterized by a significant decrease in the proportion of bifidobacteria and enterobacteria; and an increase in members of clostridia (Fallani *et al.*, 2011).

1.3.1.3 Antibiotic treatment

Antibiotics are administered to infants to cure bacterial infections that occur either during a hospital stay or in an outpatient setting. In hospitals, antibiotics are often administered also as a prophylactic measure to avoid bacterial infections, such as early onset sepsis (EOS). Prophylactic treatment is initiated when there are risk factors present

for disease, such as mother's colonization with *Streptococcus agalactiae* (B-group), membranes ruptured more than 18 h before delivery, mother's fever and abnormal amniotic fluid (World Health Organization, 2013). Sepsis is the primary clinical indication for antibiotic treatment in a pediatric setting as for around 66% of antibiotic prescriptions made for neonates in hospitals are for treatment of sepsis or suspected sepsis (indicates to prophylactic treatment) (Porta *et al*., 2012). The incidence of sepsis is substantially higher among preterm than in full term infants (Schrag *et al.*, 2016). Thus, the exposure to antibiotics in a hospital setting can be considered higher for these infants. Usually, ampicillin or penicillin and gentamicin are used both for prophylaxis and treatment, as these antibiotics are the recommended first-line medicines for these purposes in infants (World Health Organization, 2013). In the outpatient setting, around 40% of infants receive antibiotic treatment at least once during their first year of life for upper respiratory tract infections, *otitis media acuta*, gastroenteritis and vomiting (Stam *et al.*, 2012).

In full term infants, antibiotic treatment during the first years of life has been associated with decreased microbial diversity (Fouhy *et al.*, 2012, Bokulich *et al*., 2016, Yassour *et al.*, 2016), delayed microbiota maturation (Bokulich *et al.*, 2016) and lower stability of the gut microbiota (Yassour *et al*., 2016). Antibiotic treatment results in deficits in members from order *Clostridiales*, such as genus *Ruminococcus* (Bokulich *et al.*, 2016), and genus *Bifidobacterium* (Tanaka *et al.*, 2009, Fouhy *et al.*, 2012), whereas there is an overgrowth of members from family *Enterobacteriaceae* (Tanaka *et al.*, 2009, Fouhy *et al.*, 2012) and *Enterococcus* (Tanaka *et al.*, 2009). Among preterm infants, antibiotic usage has been linked to increased abundance of members from gammaproteobacteria (e.g. family *Enterobacteriaceae*) (LaRosa *et al.*, 2014) and genus *Enterococcus* (Zwittink *et al.*, 2018), and decreased abundance of members from class *Clostridia* (LaRosa *et al*., 2014) as well as genus *Bifidobacterium* (Zwittink *et al.*, 2018). It has to be noted that preterm infants participating in these studies have had a relatively mature gestational age (>26 weeks) and thus, all these observations resemble the effect of antibiotic treatment observed in full term infants.

Many infants, both full term and preterm, have also been exposed to intrapartum antimicrobial prophylactic treatment (IAP), which has been shown to affect the development of the gut microbiota. IAP (ampicillin or penicillin) is administered during labor to pregnant women, who are colonized with group B *Streptococcus* and have an elevated risk for the transmission of this microbe because this colonization is considered as a predictive factor for EOS in infants (World Health Organization, 2015). Approximately 10–35% of pregnant women are colonized with this microbe and about 40–60% of group B *Streptococcus* carriage by mothers leads to the colonization of an infant. In 1 to 2% of the cases, the colonization leads to the disease (Di Renzo *et al.*, 2014).

The gastrointestinal tract of infants exposed to IAP has been shown to harbor lower levels of members from genus *Bifidobacterium* during the first week of life when compared to infants whose mothers have not received this treatment (Corvaglia *et al.*, 2016, Mazzola *et al.*, 2016). IAP also seems to result in qualitative shift within the genus *Bifidobacterium* as the gut microbiota of infants exposed to IAP harbor *B. pseudocatenulatum*, *B. pseudolongum*, and *B. longum* with higher and *B. breve*, *B. bifidum*, and *B. dentium* with lower frequency than infants whose mothers have not received IAP (Aloisio *et al.*, 2014). In preterm infants, however, the effect of IAP on the development of the gut microbiota can be observed later in life (at 30 days of age) than in full term infants (Arboleya *et al.*, 2015). IAP may have a prolonged influence of the gut

microbiota in preterm infants as the initial effect can be overshadowed by several coinciding factors present especially during the first weeks of life in preterm infants (such as less mature gestational age, lower birth weight, neonatal intensive care unit stay and parenteral feeding).

1.3.2 The effect of preterm birth on the development of infant gut microbiota

Preterm birth is defined as a delivery before 37 completed weeks of gestation (March of Dimes *et al.*, 2012). In developed countries, approximately 9% of infants born alive are born preterm (Blencowe *et al.*, 2012). Extremely preterm birth is defined as the delivery before 28 (approx. 5% from all preterm births); very preterm birth as the delivery before 28 – <32 (approx. 10% from all preterm births); and late preterm birth as the delivery before 32 – <37 completed weeks of gestation (approx. 85% from all preterm births) (March of Dimes *et al.*, 2012). These infants (especially extremely low birth weight infants) are often staying in the neonatal intensive care unit (isolated from their parents and the environment) accompanied with parenteral feeding and antibiotic treatment to improve preterm birth outcomes as suggested by World Health Organization (March of Dimes *et al.*, 2012). All these factors may affect the development of the gut microbiota.

Early composition of the gut microbiota in preterm infants has been studied relatively extensively over the years (Table 1). One of the reasons for this is that the complications of preterm birth are one of the leading causes of neonatal death in Europe (Liu *et al*., 2012) and some of the complications, such as necrotizing enterocolitis (NEC) and late onset sepsis (LOS), have been related to the development of the gut microbiota (Stewart *et al*., 2012). Although the exact mechanisms behind the correlation between NEC and LOS and the development of the gut microbiota in preterm infants is not yet clear, previous studies have observed different community structure of the gut microbiota in infants suffering NEC and LOS when compared to healthy infants (Stewart *et al*., 2012). Lower diversity and increased levels of members from genus *Escherichia* (family *Enterobacteriaceae*) have been observed to precede NEC (Wang *et al.*, 2009, Mai *et al.*, 2011, Stewart *et al.*, 2013). Differences in the composition of the gut microbiota can be observed as much as three weeks before the diagnosis (Morrow *et al.*, 2013, Claud *et al.*, 2013). Additionally, a recent study analyzing the gut microbiota of pre- and full term infants using a large-scale deep shotgun metagenomic sequencing identified uropathogenic *Escherichia coli* as a possible NEC pathogen and a risk factor for death (Ward *et al.*, 2016). In LOS patients, several studies have observed bacteria cultured from blood being also abundant in the gut microbiota (Soeorg *et al.*, 2013, Stewart *et al.*, 2017b).

Table 1. A selection of studies published from 2011 to 2017 analyzing the composition of gut microbiota of preterm infants using molecular methods.

Study	**Method**	**No. of infants, BW (g) and GA[1] (w)**	**Sample collection**	**Key findings**
Jacquot *et al.*, 2011	TTGE[2]	29 760–1060 g 27–29 w	Between 3–56 days or NICU[7] discharge (after every three days)	Diversity increased with time, and it was lower in infants with GA[1] < 28. High prevalence of staphylococci during the first 3-4 weeks. Low incidence of bifidobacteria.
Moles *et al.*, 2013	Culture RAPD[3] DGGE[4] HITChip[5]	14 600–2190 g 24–32 w	Weekly, starting from birth until NICU[7] discharge	Initial dominance of staphylococci replaced by enterococci and *Enterobacteriaceae*. *Bifidobacterium* and *Bacteroides* were detected at low frequency.
LaTuga *et al.*, 2011	16S seq. (V6–8) Shotgun seq.	11 490–980 g 24–27 w	One or several samples collected during the first month	*Enterobacteriaceae, Pseudomonas, Staphylococcus*, and *Enterococcus* were the most abundant taxa. Lack of lactobacilli and bifidobacteria, despite breastfeeding.
Madan *et al.*, 2012	16S seq. (V6)	6 870–1080 g 24–27 w	Weekly, starting from birth until discharge	Infants who developed sepsis had higher abundance of *Staphylococcus* and lowered microbial diversity.
Barrett *et al.*, 2013	16S seq. (V4)	10 25–33 w	On 2nd and 4th week	*Enterobacteriaceae* dominated in most of the infants. Low abundance of *Bifidobacterium* and *Lactobacillus.*
LaRosa *et al.*, 2014	16S seq. (V3–5)	58 Median (IQR): 960 g (800, 1220) 27.1 w (25.6, 25.3)	All stool samples from birth until NICU[7] discharge	The pace of the colonization process was associated with GA[1], being the slowest for infants born most prematurely. Bacilli and gammaproteobacteria (e.g. *Enterobacteriaceae*) dominated with bacilli (e.g. staphylococci) being observed in the earliest samples.
Gibson *et al.*, 2016	Shotgun seq. 16S seq. (V4)	84 Median (IQR): 865 g (718, 1141) 27 w (25, 29)	Longitudinally, during the hospitalization	*Klebsiella, Escherichia*, and *Enterobacter* gradually replace *Enterococcus* and *Staphylococcus*.
Arboleya *et al.*, 2017	qPCR SCFA analysis[6]	63 1085–1580 g 28–33 w	On 2nd, 10th and 30th day after birth	The levels of *Staphylococcus* and *Enterococcus* were negatively associated with weight gain.

[1] GA – gestational age
[2] TTGE – temperature gradient gel electrophoresis
[3] RAPD – random amplification of polymorphic DNA
[4] DGGE – denaturing gradient gel electrophoresis
[5] HITChip – human intestinal tract chip
[6] SCFA analysis – short fatty acid analysis
[7] NICU – neonatal intensive care unit

1.4 The composition of mothers' vaginal microbiota and pregnancy outcomes

The normal vaginal microbiota of majority of women is dominated by one or more species of *Lactobacillus*, although substantial number of women (approximately 27%) harbor heterogeneous vaginal microbiota that is characterized by higher diversity and high levels of anaerobic bacteria, such as members from genera *Prevotella, Dialister, Atopobium*, and *Gardnerella* (Ravel *et al.*, 2011). Most women harbor a single prevailing vaginal community type, and the fluctuations between these community types tend to be transient (Gajer *et al.*, 2012). The stability of vaginal microbiota is dependent on the composition, as the community types dominated by lactobacilli (in particular *L. crispatus* and *L. gasseri*) tend to be more stable over time when compared to the heterogeneous community (Gajer *et al.*, 2012). The exception is a community type dominated by *L. iners*, which is also more stochastic than other community types dominated by lactobacilli (Gajer *et al.*, 2012). Detection of *Lactobacillus*-poor heterogeneous vaginal microbiota (characterized by the presence of microbes such as *Gardnerella vaginalis* and *Atopobium vaginae*) has previously been related to vaginal microbiota that is considered to be shifted towards a diseased state, more specifically towards bacterial vaginosis (Donders, 2007, Zozaya-Hinchliffe *et al.*, 2010). However, studies have also shown that, for instance, *Gardnerella vaginalis* could be detected in approximately 50% of the vaginal microbial communities belonging to healthy women (Mikamo *et al.*, 2000, Fredricks *et al.*, 2005) and that *Atopobium vaginae* is also prevalent in healthy vaginal environment (Ravel *et al.*, 2011).

During pregnancy, heterogeneous vaginal microbial community type is rarely observed among pregnant women when compared to non-pregnant counterparts (Romero *et al.*, 2014a, MacIntyre *et al.*, 2015). The results of previous studies using microscopic assessment of the vaginal microbiota (e.g. Gram-stain evaluation) have shown that during pregnancy, heterogeneous vaginal microbiota, either accompanied with symptoms of bacterial vaginosis (such as excess amount of white, grey, or green malodorous discharge) or without, is associated with higher risk of preterm delivery (Leitich & Kiss, 2007, Donders *et al.*, 2009). However, the results of recent studies analyzing the relationship between specific vaginal microbial composition and preterm delivery by using molecular methods for community profiling have so far been inconclusive. Some studies have shown that either persistent presence of *Lactobacillus*-poor heterogeneous vaginal microbiota, higher prevalence of *Lactobacillus iners*, or higher microbial richness and diversity can be associated with preterm birth (Petricevic *et al.*, 2014, DiGiulio *et al.*, 2015, Freitas *et al.*, 2018). Others have not observed differences with normal pregnancies (Romero *et al.*, 2014b, MacIntyre *et al.*, 2015, Avershina *et al.*, 2017). Moreover, heterogeneous vaginal microbial community type is overrepresented among women onset of labor (Avershina *et al.*, 2017) and also postpartum (MacIntyre *et al.*, 2015, DiGiulio *et al.*, 2015) regardless of pregnancy outcome. Thus, although bacterial vaginosis and high bacterial diversity (especially during the first trimester) are considered as a risk factor for preterm birth (Leitich *et al.*, 2003, Haque *et al.*, 2017), further studies are needed to confirm the association between specific vaginal microbial community composition and higher risk for preterm delivery.

1.5 Vertical transmission of microbes from the mother's normal microbiota of different body sites to the infant's gut microbiota

Microbial communities of the mother are considered to be an important source of bacteria during the development of infant gut microbiota. Individual mother-infant pairs have been shown to share bacterial strains with a 22-fold higher rate than sharing strains with other, unrelated mothers. This applies to maternal communities colonizing the intestine, oral cavity, skin, vagina, and breast milk (Ferretti *et al.*, 2018). All these maternal communities contribute to the development of infant gut microbiota, but when comparing the number of shared operational taxonomic units (OTU) or strains, mothers' gut microbiota is by far the largest donor (Bokulich *et al.*, 2016, Ferretti *et al.*, 2018) and the least important route of transmission appears to be mothers' oral cavity (Ferretti *et al.*, 2018). Strains transmitted from mother's gut to infant gut microbiota belong to *Escherichia coli*, and members from classes *Actinobacteria* (e.g. bifidobacteria), and *Bacteroidia* (e.g. *Bacteroides*) (Bäckhed *et al.*, 2015, Ferretti *et al.*, 2018, Korpela *et al.*, 2018b). Transmitted strains tend to have a better fitness in infant gut microbiota when compared to strains not transmitted from the mother (Bäckhed *et al.*, 2015, Ferretti *et al.*, 2018, Korpela *et al.*, 2018b).

Mother-to-infant vertical transmission of microbes is dependent on the mode of delivery. Immediately after birth, vaginally born infants acquire bacterial communities in their gut microbiota that resemble their mothers' vaginal microbiota whereas, in C-section infants, the gut microbiota is enriched in skin microbiota (Dominguez-Bello *et al.*, 2010). Moreover, if C-section infants are exposed to maternal vaginal fluids at birth, their gut microbiota will become enriched in vaginal bacteria during the first 30 days of life (Dominguez-Bello *et al.*, 2016). Also, 72% of the early colonizers of the vaginally born infants' gut microbiota have been shown to match species found in the stool of his or her mother, whereas in C-section infants this proportion is only 41% (Bäckhed *et al.*, 2015).

Breast milk has also been shown to be a possible source of commensal bacteria for the infant gut (Fernandez *et al.*, 2013). Breast milk harbors bacteria that belong to genera *Staphylococcus, Streptococcus, Propionibacterium* (Hunt *et al.*, 2011, Jost *et al.*, 2014, Urbaniak *et al.*, 2016), and also *Bifidobacterium* (Gueimonde *et al.*, 2007, Jost *et al.*, 2014). Although the origin of bacteria, especially bifidobacteria in the breast milk is still unknown, breastfeeding may play a role in mother-to-infant transfer of members of this genus. A study by Makino *et al.*, (2011) has shown that although the main reservoir for *Bifidobacterium longum subsp. longum* is mothers' gut microbiota, two strains of this bacterium were shared between infants' gut and breast milk microbiota while they were not observed in mothers' gut microbiota. Ferretti *et al.*, (2018) also observed *Bifidobacterium longum* and *Bifidobacterium bifidum* strains in breast milk and corresponding infant gut microbiota.

2 The aim of the thesis

This thesis aims to study the composition of vaginal microbiota in reproductive-age women and to analyze the development of infant gut microbiota by identifying the similarities between the microbial communities of full term infants and their mothers taking into account different maternal microbial communities and characterizing the development pattern of the gut microbiota in extremely preterm infants. More precisely, the thesis has the following objectives:

- to describe the variation of vaginal micro- and mycobiome in women of reproductive age to expand the knowledge of the maternal microbial community with what a newborn infant has a direct contact during vaginal delivery;
- to determine the effect of preterm birth in combination with extremely low birth weight on the development of the gut microbiota and how this developmental process differs from full term infants;
- to determine the influence of different maternal microbial communities – intestinal, vaginal, oral, breast milk, and mammary areola – on the development of infant's gut and also oral microbiota in full term infants by analyzing the similarities between these communities.

3 Materials and methods

3.1 Study cohort and sampling

The study cohort in a publication I consisted of 494 healthy pre-menopausal women, in publication II of 50 preterm infants with extremely low birth weight (ELBW), and publication III of 7 mothers with their full term infants (including two pairs of twins) (Table 2).

Table 2. The number of participants, samples and time points analyzed in the publications of this thesis. "Pub" indicates to publication number.

Subjects (no.)	Samples	No. of samples	Collection time points	Pub
Healthy women (n=494)	Vaginal brush	494	Single time point, collected during the routine gynecologist visit	I
Preterm infants (n=50)	Stool	118	Three time points: * 4-8 days after delivery * 1 month after delivery * 2 months after delivery	II
Mothers (n=7)	Vaginal swab	21	Three time points: * Before delivery * 6-8 weeks after delivery * 6 months after delivery	III
	Rectum swab	21		
	Mouthwash	21		
	Breast milk	18	Three time points: * 48-72 h after delivery * 6-8 weeks after delivery *6 months after delivery	
	Mammary areola swab	18		
Full term infants (n=9)	Stool	27		
	Oral swab	27		

In the publication I, healthy pre-menopausal Caucasian women aged 15–44 years were recruited to the study during a routine gynecological check-up visit from June to October 2010. Exclusion criteria included pregnancy, menstruation, or any signs or symptoms of urogenital disorders. The presence of urogenital infections was excluded by inspection and *in speculo* examination before a sample was taken from the vaginal fornix and cervix. Amsel criteria or Gram staining were not explicitly applied for the exclusion of bacterial vaginosis. However, the vaginal pH was measured during sampling. Participants were asked to complete a questionnaire on their medical history (such as incidence of sexually transmitted infections, vulvovaginal candidiasis and bacterial vaginosis, medication usage (incl. antibiotics and contraceptives) and regularity of menstrual cycle) and lifestyle habits, which specifically addressed their sexual behavior and feminine hygiene. The opinion of a gynecologist regarding the physical appearance of the vaginal environment (e.g. redness, characterization of vaginal discharge etc.) was included with the questionnaire.

In the publication II, infants with extremely low birth weight <1200 g were recruited from August 2006 to November 2007. The average gestational age of the participating infants was 26.64 weeks, and 22 out of 50 participating infants were born via caesarian section. The study was conducted in the neonatal intensive care units (NICUs), where infants stayed on average for 25.7 days. The patients were all receiving antibiotic treatment with ampicillin (n=25) or penicillin (n=25) combined with gentamycin and

37 receiving broad-spectrum antibiotic treatment (beta-lactam antibiotics with inhibitors, cephalosporins and carbapenems). Twelve infants received total parenteral feeding during the first week of life. Seventeen infants were breastfed while 21 were fed with formula. Twenty-four infants suffered from sepsis (infection was caused by *Staphylococcus epidermidis* in eight, *Klebsiella pneumoniae* and *Staphylococcus haemolyticus* each in four, methicillin-resistant *Staphylococcus aureus* and *Enterococcus* each in two, and *Enterobacter, Acinetobacter, Streptococcus agalactiae* and *Serratia* each in one case) and seven from necrotizing enterocolitis (NEC). Twelve mothers of the preterm infants received antenatal antibiotics, and 23 received intrapartum antibiotic prophylaxis (IAP). Eighteen mothers were suffering from chorioamnionitis and 12 had premature rupture of membranes more than 18 h prior delivery. The diagnosis of chorioamnionitis was confirmed based on clinical criteria: fever and/or high C reactive protein levels, odorous amniotic fluid, and placental inflammation.

In publication III, pregnant women were enrolled during the first period of spontaneous onset of labor or up to 24 h before elective cesarean section with intact amniotic membranes between May 2012 and September 2013. All the recruited women had uncomplicated term pregnancies, they did not have any infectious diseases that required antibiotic treatment, and they had no history of diabetes or hypertensive disorders during the second half of pregnancy. The mean age of the women was 33.1± 4.5 years. Five women out of seven gave birth via cesarean section and received prophylactic cefuroxime treatment before the incision in the skin. None of the women received IAP. All infants were full term (average gestational age, 38.9 weeks) and had normal birth weights (average weight, 3379.3 g). To determine the diagnosis for bacterial vaginosis, bacterioscopic smears were made from separate vaginal swab samples. The samples were Gram stained and scored according to the Nugent method: 0–3 (normal), 4–6 (intermediate), and 7–10 (bacterial vaginosis).

3.2 Methods used for bacterial community profiling

Briefly, the experimental side of bacterial community profiling was carried out by extracting the DNA from the samples, amplifying approximately 300 bp of the 16S rRNA gene, purifying the amplicons from PCR reagents and sequencing the amplicons (Table 3).

Table 3. Methods used in the publications of this thesis. "Pub" indicates to publication number.

Pub	DNA extraction	16S rRNA region	PCR amplicon purification	Sequencing method
I	BioSprint 96 DNA Blood Kit[1]	V1-V2[2]	Pooling and extracting of amplicons from 1.5% agarose gel with QIAquick Gel Extraction kit (Qiagen)	Pyrosequencing (454 Life Sciences)
II	QiaAmp Stool DNA Mini Kit	V1-V4	Agencourt AMPure XP (Beckman Coulter)	Pyrosequencing (454 Life Sciences)
III	MoBio PowerFecal DNA Isolation Kit[4]	V1-V2	Agencourt AMPure XP	Illumina MiSeq system

[1] BioSprint 96 DNA Blood Kit protocol preceded an initial homogenization step using ceramic beads and TissueLyser II (Qiagen).
[2] Two sequential PCR reactions were carried out to extend the primers to full length. The primers, consisting of 454-specific sequencing adapters, amplicon-specific primer sequences and the eight bp unique barcode tags, were bisected to avoid the extensive length and high melting temperatures, which could have led to inhibition of PCR reaction. This study design was changed in later projects (Publications II and III) as two sequential PCR reactions as well as the 50-fold dilution of amplicons carried out between the reactions increases the risk of contamination and significant inhibition of PCR reactions was not observed.
[3] The MoBio PowerFecal DNA Isolation Kit's (Mo Bio) protocol was modified for the swab, mouthwash and breast milk samples by adding steps for extraction of bacteria.

Data obtained from the sequencing experiments (except for the experiments described in the next section) were analyzed using MOTHUR software. In all three projects, operational taxonomic units (OTUs) were generated with a 97% identity threshold defined by the average neighbor hierarchical clustering algorithm and taxonomic assignments were performed against the SILVA bacterial database using the Naive Bayesian classifier. The parameters and MOTHUR algorithms used in the data analysis are specified in Table 4.

Table 4. Data analysis carried out in the in the publications of this thesis. "Pub" indicates to publication number, "NB cutoff" indicates to Naïve Bayesian classifier confidence cutoff value.

Pub	Trimming and denoising	Removal of chimeric sequences	NB cutoff	Additional denoising
I	quality score > 25 length > 170 bp	UChime	75%	OTUs with < 5 sequences and present in only 1 sample with a rel. abundance of ≤1% were removed
II	PyroNoise length > 150 bp		90%	OTUs with < 5 sequences were removed
III	quality score > 25 length > 225 bp			OTUs with < 2 sequences and/or present in < 2 samples were removed

Additional taxonomic assignments were carried out using the BLASTN against the NCBI nt database in a publication I and NCBI 16S ribosomal RNA gene sequences database in publication III.

3.3 Additional methods

In the publication I, the fungal component of the vaginal microbiota was assessed in 251 participating women. The analysis was based on sequencing and data analysis of the internal transcribed spacer-1 (ITS-1) region of the eukaryotic ribosomal DNA. The fungal sequences were first clustered with cd-hit-est software with 97% cluster identity. cd-hit-est results were compared against the UNITE fungal database using a BLAST search with 97% identity and an e-value below 10^{-20}.

As the primer 27F, which was used to amplify the 16S rRNA gene sequences in publications II and III, contains two mismatches against the bifidobacterial 16S rRNA gene sequences (Frank et al., 2008), real-time PCR experiments were carried out in both projects to quantify members from genus *Bifidobacterium*.

In publication III, denaturing gradient gel electrophoresis (DGGE) and sequencing analysis with LI-COR DNA Sequencer 4000L (Licor) were carried out to analyze the species composition of genus *Bifidobacterium* colonizing different body sites of the mothers and their infants.

3.4 Statistical analysis

Statistical analysis was performed with the R software using the following statistical methods:

- The sequence counts were normalized in all of the projects using a package "clusterSim".
- To describe the general bacterial diversity, a Shannon diversity index was calculated using package "vegan". The calculation was based on natural logarithm.
- To analyze the similarity between the samples, Cosine similarity index and Jaccard distance values were calculated using packages "lsa" and "vegan", respectively, or a clustering analysis was carried out building a similarity tree based on a hierarchical Yue-Clayton theta value using a command "tree.shared" in the MOTHUR software.
- A linear regression analysis was used to analyze the correlation between a scalar dependent variable (such as relative abundance values of the OTUs) and the predictor variables. Mixed effect modeling was also used to analyze this type of correlation using package "lme4".
- A logistic regression analysis was used to analyze the correlation between a binary dependent variable (such as the prevalence values of OTUs) and the predictor variables.
- Categorical values were compared with the Fisher exact test.
- Continuous variables were compared with the Welch Two-sample T-test or the Wilcoxon rank-sum test.

All analyses were performed with the Holm-Bonferroni correction (using the package "multcomp"), and the overall level of a significant difference was set at 5% in all of the projects.

4 Results and discussion

4.1 The variation of vaginal microbiota of women of reproductive age (Publication I and III, unpublished data)

In our study described in detail in Publication I, we aimed to profile the composition of normal vaginal microbiota in 494 asymptomatic reproductive-age (15–44 y) women in a single time point analysis using 16S rRNA gene sequencing. Only samples with higher than 400 trimmed sequences were included to further analysis with a total of 432 samples being analyzed further.

4.1.1 The variation of vaginal bacterial community

Overall, 99% of the women were colonized with lactobacilli, and the mean relative abundance of this genus was approximately 0.7, which is consistent with the general knowledge that normal vaginal microbiota is dominated by lactobacilli. Nevertheless, members of genera suggestive of a bacterial vaginosis-like microbiota (such as members from genera *Gardnerella* and *Atopobium*) were also relatively abundant. Although there was an inter-individual variability of vaginal bacterial communities observed in our study, altogether 307 analyzed vaginal samples (71.1%) clustered into five major groups based on the similarities in relative abundance values of the OTUs (Figure 1).

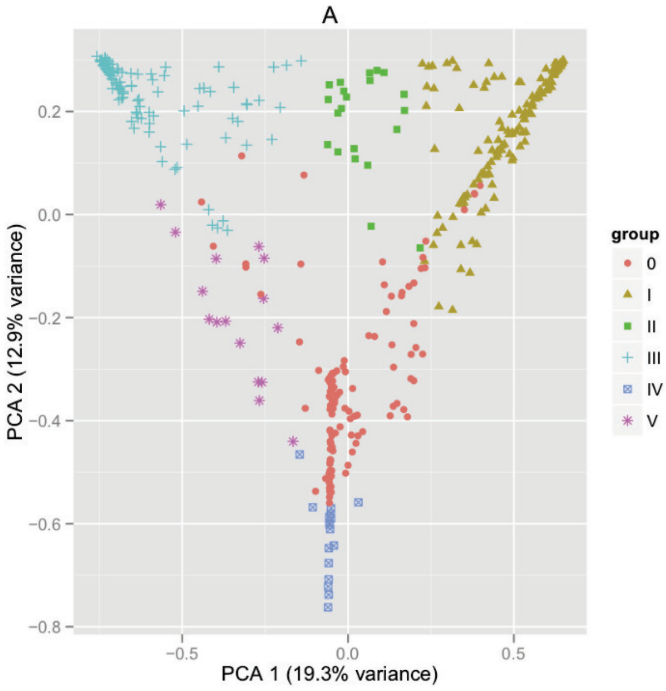

Figure 1. Two-dimensional (2D) plot is describing the principal component analysis (PCA) of OTU composition among analyzed samples. The plot represents a 2D projection of a multidimensional analysis where the relative abundance of each specific OTU defines a dimension. The plot visualizes the clustering and variability of studied vaginal bacterial communities. The variance described by the respective PCA components (Axis-1 and Axis-2) is written in brackets.

Each group harbored more than 15 samples. The remaining samples (n = 125; 28.9%) did not cluster in any of the five groups and were characterized by non-related heterogeneous bacterial communities (hereafter the non-classifiable heterogeneous 0-group). OTUs belonging to members of lactobacilli dominated four out of five groups. In group IV, the predominant bacterium was a member of the genus *Gardnerella* (Table 5).

*Table 5. Mean (±SD) values of selected parameters of the vaginal environment within bacterial community groups. Dominating OTUs are listed as top 3 with the mean relative abundance at least 3%. Linear regression models were used to test the differences between the groups (I-V) and 0-group: *** P<0.001; **P=0.01-0.001; * P=0.01-0.05.*

Community groups	**Group size**	**Vaginal pH**	**Shannon diversity index**	**Dominating OTUs in the bacterial community** (relative abundance/ prevalence)
I	**145**	4.58 (±0.9)***	0.8 (±0.6)***	*Lactobacillus iners_*1 (79.1% / 100%) *Gardnerella_ 1* (3.4% / 52.4%) *Lactobacillus crispatus_1* (3.0% / 59.3%)
II	**21**	4.54 (±0.6)*	1.4 (±0.5)**	*Lactobacillus iners_1* (47.7% / 100%) *Lactobacillus crispatus_1* (34.4% / 100%) *Lactobacillus jensenii* (4.0% / 81.0%)
III	**108**	4.19 (±0.5)***	0.7 (±0.4)***	*Lactobacillus crispatus_1* (83.7% / 100%) *Lactobacillus iners_1* (5.6% / 80.6%)
IV	**17**	4.88 (±0.6)	1.1 (±0.4)***	*Gardnerella_1* (68.5% / 100%) *Lactobacillus jensenii* (6.7%/ 35.3%) *Lactobacillus gasseri* (5.7%/ 47.1%)
V	**16**	4.64 (±0.7)	1.4 (±0.4)*	*Lactobacillus crispatus_1* (43.1% / 100%) *Gardnerella_1* (35.1%/ 100%) *Lactobacillus iners_1* (3.2% / 56.3%)
0	**125**	5.02 (±0.9)	1.8 (±0.6)	

In the 0-group, none of the OTUs had relative abundance over 0.2, but the prevalence of the following OTUs was relatively high in this group: *Atopobium vaginae* (52.4%), *Sneathia sanguinegens* (30.1%), and members of genera *Gardnerella* (74.1%), *Prevotella* (69.2%), *Ureaplasma* (33.6%), and *Clostridium* (16.8%).

Others have described similar clusters of vaginal community types (Zhou *et al.,* 2007, Zhou *et al.,* 2010, Ravel *et al.,* 2011). Typically, there are 5–8 vaginal community types that are differing based on their composition and also by dominating taxa. *L. crispatus* and *L. iners* are the most dominant species among women being dominant or one of the dominating species in at least one of the community types. (Table 5, Zhou *et al.,* 2007, Zhou *et al.,* 2010, Ravel *et al.,* 2011). Also, approximately 27–29% of women harbor heterogeneous vaginal microbiota that is characterized by low levels of lactobacilli (Table 5, Ravel *et al.,* 2011). Since the studies above (including ours), many newer studies often analyze the effect of differing vaginal microbiota, for instance to pregnancy outcomes, focusing on those community types (e.g. Romero *et al.,* 2014a, Avershina *et al.*, 2017, Freitas *et al.,* 2018).

Linear regression analysis showed that the Shannon diversity index increased with an increase in vaginal pH ($p = 0.003$) and with the presence of malodorous vaginal discharge ($p < 0.001$). Highest diversity and generally high vaginal pH were observed in women with heterogeneous vaginal microbiota (Table 5), which may be an indication that this community state can still imply to subclinical or upcoming disease as high pH values and

malodorous discharge are both symptoms of bacterial vaginosis (Amsel *et al.,* 1983). Moreover, studies such as by Srinivasan *et al.,* (2012) have shown that patients with bacterial vaginosis have higher bacterial diversity and richness than individuals without. Unfortunately, we did not determine the diagnosis for bacterial vaginosis based on Nugent nor Amsel criteria among women participating in our study. Thus, it is not known how many women would have been diagnosed with bacterial vaginosis based on widely used diagnostic criteria.

Nevertheless, heterogeneous vaginal microbiota itself cannot be considered as an indication for unhealthy vaginal microbiota as this type of community profile is relatively prevalent among women who do not have symptoms of bacterial vaginosis.

Moreover, studies have shown that during labor the bacteria associated with heterogeneous community type are significantly overrepresented when compared to 36th gestation week (Avershina *et al.*, 2017). Heterogeneous vaginal community type is also overrepresented among women during the postpartum period (MacIntyre *et al.*, 2015, DiGiulio *et al.,* 2015), which we also observed in our study in publication III, where only two out of seven women had high levels of lactobacilli colonizing vaginal microbiota onset of labor (Figure 2). In the postpartum period, vaginal microbial communities with high levels of lactobacilli were more common among these women than the onset of labor (Figure 2), but the results of Nugent scoring indicated that only two mothers had normal Nugent scores and two women (IDs 104 and 105) were diagnosed with bacterial vaginosis during that time. In turn, during the onset of labor, four out of seven women had normal Nugent scores, and none of the women had a score higher than 7 (indicative of bacterial vaginosis).

Although the sample size is too small to make definite conclusions, the vaginal community profile and Nugent scores together support the results of the aforementioned previous studies. These studies have hypothesized that the disruption of vaginal microbiota during pregnancy and postpartum period may be related to prenatal hormonal changes as the declining estrogen levels may reduce the estrogen-driven *Lactobacillus* spp. dominance (MacIntyre *et al.*, 2015, Avershina *et al.*, 2017).

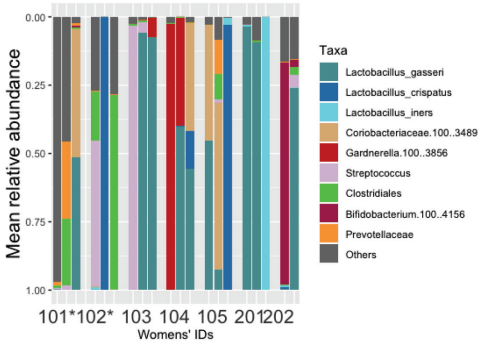

*Figure 2. The composition of vaginal microbiota of women onset of labor and during the postpartum period analyzed in the study described in publication III. Each column under specific ID represents each time point: before delivery, 6-8 weeks and six months after delivery. * indicates to ID's that tended to have higher bacterial diversity (on average [SD] 2.3 [1.5]) than other women (on average [SD] 0.64 [0.56]) especially at time points where lactobacilli were scarce in their vaginal microbiota although the difference was not statistically significant.*

4.1.2 The variation of vaginal fungal community

In addition to bacterial communities, we also analyzed fungal communities in 216 samples. The most prevalent and abundant fungal genus in analyzed samples was *Candida*. OTUs belonging to this genus were detected in 69.9% of samples and the mean relative abundance was 0.37. The prevalence is considerably higher than the prevalence of 20% found in asymptomatic healthy women in earlier studies (Sobel, 2000). The most prevalent and abundant OTU from genus *Candida* was *C. albicans* as 82% of the fungal OTUs identified as *Candida spp.* belonged to *C. albicans* Linear regression analysis incorporating different factors of medical history and lifestyle habits did not reveal any correlation between these factors and the relative abundance of *Candida spp*. Also, the relative abundance values of lactobacilli and *Gardnerella* were not correlated with the prevalence of *C. albicans*.

Our results indicate that colonization with *C. albicans* is widespread among healthy women indicating that the colonization with this species may not always lead to pathology. Nevertheless, asymptomatic colonization during pregnancy, especially the second trimester of pregnancy, has been associated with a higher risk of preterm birth (Farr *et al.,* 2015, Holzer *et al.,* 2017). The study by Roberts *et al.,* (2015) has observed a significant reduction in spontaneous preterm births in clotrimazole treated women when compared to untreated women.

4.2 The effect of preterm birth in combination with extremely low birth weight on the development of the gut microbiota (Publication II and III)

To analyze the effect of preterm birth on the development of the gut microbiota, we profiled gut microbial communities of 50 preterm infants with extremely low birth weight (< 1200 g) (ELBW infants) during the first two months of life using 16S rRNA gene sequencing. Only samples with higher than 500 trimmed sequences were included to analysis with a total of 101 samples being analyzed: 30 samples collected at the age of one week, 42 at one month and 29 at two months.

4.2.1 Development of the gut microbiota in preterm infants

Microbial diversity of the gut microbiota observed in ELBW infants increased significantly throughout the study period (mean [SD] Shannon diversity index was 0.38 [0.4] at one week, 0.71 [0.54] at one month, and 0.92 [0.44] at two months of age; all p-values were < 0.01). The gradual increase in diversity has been described both in preterm and full term infants (Schwiertz *et al.,* 2003, Stewart *et al.,* 2012, Abrahamsson *et al.,* 2014, Jakobsson *et al.,* 2014) indicating a normal process of diversification of the gut microbiota in early age. Nevertheless, the onset of the increase in microbial diversity is usually observed later in full term than in preterm infants (at 3-12 months of age) (Abrahamsson *et al.,* 2014, Jakobsson *et al.,* 2014). The earlier start of diversification in preterm infants may be related to lower microbial diversity observed up to 30 days of age when compared to full term infants (Arboleya *et al.*, 2012a, Arboleya *et al.*, 2012b).

We did not observe a correlation between most of the clinical factors and the parameters of the gut microbiota in ELBW infants. For instance, we did not observe a significant association between the composition of the gut microbiota in ELBW infants and the administration of either penicillin or ampicillin with gentamicin or the duration of broad-spectrum antibiotic treatment. The reason may be that in preterm infants, antibiotic treatment (e.g. early empiric antibiotic treatment with ampicillin and gentamicin) is often correlated with other factors such as lower birth weight and less

mature gestational age (Kuppala *et al.,* 2011, Greenwood *et al.,* 2014). Also, preterm infants are often administered a combination of different antibiotics that result in varying effects on the gut microbiota. For instance, Gibson *et al.*, (2016) observed an association between meropenem, cefotaxime and ticarcillin-clavulanate administration and reduced species richness while vancomycin and gentamicin, the antibiotics most commonly administered to preterm infants, had a non-uniform effect on species richness of the gut microbiota in preterm infants (Gibson *et al.*, 2016). We also did not observe significant differences between infants fed with breast milk versus formula and infants born via C-section versus born vaginally, which has been noted by others as well, who have analyzed the composition of the gut microbiota in preterm infants (LaRosa *et al.*, 2014, Stewart *et al.*, 2017a). The reason for not observing any differences between varying antibiotic usage, delivery modes and feeding regimen may be a cumulative or overshadowing effect of many of the coinciding factors, such as lower birth weight and less mature gestational age, but also NICU stays and other factors related to this care.

We were only able to observe a negative correlation between Shannon diversity index and the starting day of total enteral feeding (receiving breastmilk or formula) (p-value < 0.01), which may be due to a shorter period of nutrition deficiency in the gut resulting from the parenteral feeding. Shannon diversity index was also significantly higher among ELBW infants born to mothers suffering from chorioamnionitis (intra-amniotic infection that is a frequent causative factor for preterm birth) than in infants born to healthy mothers (mean Shannon diversity index value was [SD] 0.82 [0.59] vs 0.59 [0.45]; p-value < 0.01). The possible factors contributing to more diverse gut microbiota in these infants may be, among other things, the exposure to inflammation.

From three different time points, significant compositional changes in the gut microbiota of ELBW infants were observed between the first week and first month of life. Briefly, staphylococci dominated the gut microbiota of ELBW infants during the first week of life, but the mean [SD] relative abundance of this family had decreased significantly from 0.33 [0.43] to $5.1x10^{-4}$ [0.15] by the end of the first month of life (p-value < 0.001). At the same time, the relative abundance of members from families *Enterobacteriaceae* and *Veillionellaceae* increased significantly (mean [SD] relative abundance increased from 0.25 [0.42] to 0.66 [0.38] and from 0.003 [0.009] to 0.09 [0.19], respectively; p-values < 0.01). From the end of the first month to the second month of life, members from family *Enterobacteriaceae* dominated the gut microbiota of ELBW infants (Table 6). The rapid changes observed during the first month of life may be related to delayed development of the gut microbiota in preterm infants. The hypothesis is supported by previous studies that have observed different composition of the gut microbiota in pre- and full term infants shortly after birth, but in later time points, the gut microbiota observed in preterm infants is suggestive to a developmental trajectory leading from immature to mature microbial community state observed in full term infants (Claud *et al.,* 2013). Korpela *et al.*, (2018a) also observed progress in the gut microbiota of preterm infants from *Staphylococcus-Enterococcus* dominated composition to *Enterobacteriaceae* and finally *Bifdobacterium* dominated microbiota resembling the composition in full term infants and this progress correlated with the increase in postmenstrual age. Extremely premature infants (GA <28 weeks) tended to lag with this progress when compared to moderately or very premature infants (GA 28 – <32 weeks), which indicates that the delay in development of the gut microbiota is dependent on the maturity, indicated by postmenstrual age (Korpela *et al.*, 2018a). Chernikova *et al*. (2018) have observed similar results describing the negative correlation

between earlier gestational age and microbial diversity, suggesting that greater gestational immaturity at birth is related to a less diverse gut microbiome.

Table 6. Dominating taxa in the gut microbiota of preterm infants grouped by phylum. "Prev" indicates prevalence (%) and "Rel abund" to mean (SD) relative abundance.

Phylum	Lowest matching taxonomic level	1. week of age (n=30)		1. month of age (n=42)		2. months of age (n=29)	
		Prev	Rel abund	Prev	Rel abund	Prev	Rel abund
Proteobacteria	*Enterobacteriaceae*	67	0.19 (0.38)	95	0.38 (0.41)	93	0.28 (0.33)
	Escherichia.Shigella	17	0.02 (0.1)	40	0.16 (0.3)	59	0.13 (0.2)
	Serratia	17	0.03 (0.18)	19	0.04 (0.15)	31	0.04 (0.13)
Firmicutes	*Enterococcus*	50	0.11 (0.29)	81	0.1 (0.25)	62	0.02 (0.03)
	Staphylococcus	70	0.33 (0.43)	55	0.03 (0.15)	52	0.01 (0.03)
	Veillonella	23	0.002 (0.009)	50	0.06 (0.16)	76	0.14 (0.22)
	Clostridium	23	0.007 (0.04)	31	0.06 (0.17)	31	0.11 (0.29)
	Leuconostoc	30	0.08 (0.23)	19	0	24	0

Table 7. Dominating taxa in the gut microbiota of full term infants from publication III grouped by phylum. "Prev" indicates prevalence (%) and "Rel abund" to mean (SD) relative abundance.

Phylum	Lowest matching taxonomic level	48-72h after birth (n=9)		6-8 weeks after birth (n=9)		Six months after birth (n=9)	
		Prev	Rel abund	Prev	Rel abund	Prev	Rel abund
Proteobacteria	*Enterobacteriaceae*	100	0.4 (0.32)	100	0.35 (0.22)	78	0.35 (0.4)
Firmicutes	*Streptococcus*	100	0.02 (0.03)	100	0.04 (0.11)	100	0.16 (0.3)
	Veillonella	67	0.04 (0.06)	100	0.09 (0.12)	44	0.04 (0.07)
	Enterococcus	56	0.02 (0.04)	67	0.002 (0.007)	56	0.1 (0.28)
	Clostridium	22	0.03 (0.08)	44	0.06 (0.18)	44	6.16e-4 (0.001)
Bacteroidetes	*Bacteroides*	56	0.06 (0.18)	56	0.06 (0.15)	56	0.02 (0.04)
	Parabacteroides	56	0.07 (0.19)	44	0.005 (0.01)	56	1.7e-4 (2.2e-4)
Actinobacteria	*Rothia*	78	0.06 (0.19)	56	1.9e-4 (2.1e-4)	67	0.03 (0.1)

4.2.2 Differences between the development of the gut microbiota in pre- and full term infants

When compared to the results observed in publication III analyzing the composition of the gut microbiota in full term infants (Table 7 and Figure 3), the gut microbiota of ELBW infants seems to differ in two aspects. First, low abundance or absence of members from genera *Bifidobacterium* and *Bacteroides* appears to be characteristic specifically to preterm and not to full term infants (Table 7 and Figure 3; e.g. Butel *et al.,* 2007, LaTuga *et al.,* 2011, Arboleya *et al.,* 2015, Forsgren *et al.,* 2017). Bifidobacteria are essential members of the gut microbiota that are involved in carbohydrate metabolism especially in the breakdown of carbohydrates such as lactose and human milk oligosaccharides (HMO), which represent the majority of sugars present in the diet of a breastfed infant. The pan-genome of genus *Bifidobacterium* encompasses one of the largest predicted glycobiomes among known gut commensals (Milani *et al*., 2016). However, in the absence of bifidobacteria metabolic processing of carbohydrates (excl. HMOs) and energy metabolism are predominantly covered by other taxa, such as members from genus *Enterococcus* (Zwittink *et al.,* 2017). The prevalence of some of the bifidobacterial species has an antagonistic correlation with the presence of enterobacteria as well as enterococci (Nagpal *et al.*, 2017). Although this needs to be studied further, lack of bifidobacteria in the gut microbiota of ELBW infants may also contribute to a higher incidence of infections in preterm infants, such as necrotizing enterocolitis (NEC), which has been shown to be associated with increased levels of genus *Escherichia* (Wang *et al.*, 2009, Mai *et al.*, 2011, Stewart *et al.*, 2013).

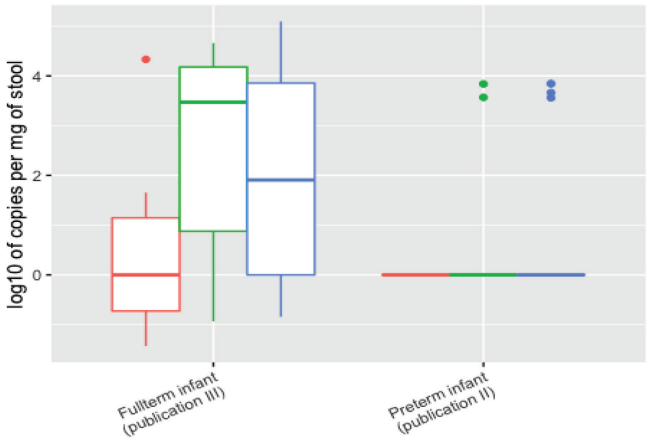

Figure 3. Log 10 of 16S rRNA gene copies of genus Bifidobacterium per mg of stool observed in full term and preterm infants. Three different boxes represent three different time points. For publication II, the first columns represent one week, second column one month and third column two months after birth. For publication III, the first columns represent 48-72h, second column 6-8 weeks (approx. two months) and third column six months after birth. It has to be noted that in publication II and III the quantification of gene copy number was carried out differently (described in the supplementary information of both publications).

Second, the dominance of members of family *Staphylococcaceae* in the gut microbiota at an early age seems to be characteristic to the most initial samples collected from preterm infants, which is later replaced by the dominance of members from family

Enterobacteriaceae (e.g. Aujoulat *et al.,* 2014, Gibson *et al.,* 2016). Full term infants are colonized with high levels of members from *Enterobacteriaceae* from an early age, whereas staphylococci are more minor constituents of the community (Palmer *et al.,* 2007). The high abundance of staphylococci in the gut microbiota of ELBW infants during the first weeks of life may be the reason why there is an increased risk of acquiring coagulase-negative staphylococcal infections in preterm neonates. For instance, a study by Soeorg *et al.,* (2013) has observed the genetical similarity between staphylococcal strains belonging to species *S. haemolyticus*, *S. epidermidis* and *S. hominis* causing late-onset sepsis (LOS) and the strains of these species found in the gastrointestinal tract of the infected infants. In our study 24 infants suffered from sepsis with 14 cases caused by staphylococci (mainly *S. epidermidis* and *S. haemolyticus* and *S.aureus*).

4.3 The influence of different maternal microbial communities on the development of infant gut and oral microbiota (Publication III, unpublished data)

In order to describe the impact of maternal microbial communities to the development of infant gut and oral microbiota during the first six months of infant's life, we analyzed the composition of bacterial communities in seven mothers and their nine infants using 16S rRNA gene sequencing. Analyzed maternal microbial communities included normal microbiota of mother's stomach, vagina, breast milk, oral cavity, and mammary areola. In total, 148 samples were analyzed after trimming and denoising the data. These samples were collected from mothers' gut (n=21), vagina (n=20), breast milk (n=17), oral cavity (n=20) and mammary areola (n=18) at four different time points (before delivery, 48-72 hours, 6-8 weeks and six months after delivery). Samples were also collected from infants' stool (n=27) and oral cavity (n=25) at three different time points (48-72 hours, 6-8 weeks and six months after delivery).

4.3.1 Diversity and similarity of different maternal microbial communities with infant gut and oral microbiota

The gut and oral microbiota of infants participating in this study had significantly lower diversity than their respective maternal community types (Shannon diversity index values 1.41 vs 3.3 and 1.68 vs 3.1, respectively for infant's gut vs mothers' gut and infant oral vs mothers' oral microbiota; p-values < 0.001). Infant community types were also less diverse than breast milk and mammary areola microbiota (Shannon diversity index values 2.45 and 2.16, respectively for breast milk and mammary areola microbiota; p-values < 0.03), although these maternal community types, in turn, were less diverse than maternal gut and oral microbiota.

Infant gut microbiota was not similar to any of the analyzed maternal community types (Cosine similarity index was < 0.07 in all cases). High similarity was observed between infant oral microbiota and the microbiota colonizing mothers' breast milk, mammary areola, and oral cavity (Cosine similarity index values 0.32, 0.48 and 0.42, respectively). When comparing individual mother-infant pairs, we did not observe infant gut and oral microbiota to exhibit significantly greater similarity to their own mother's microbial communities than to the microbial communities of unrelated mothers (Figure 4).

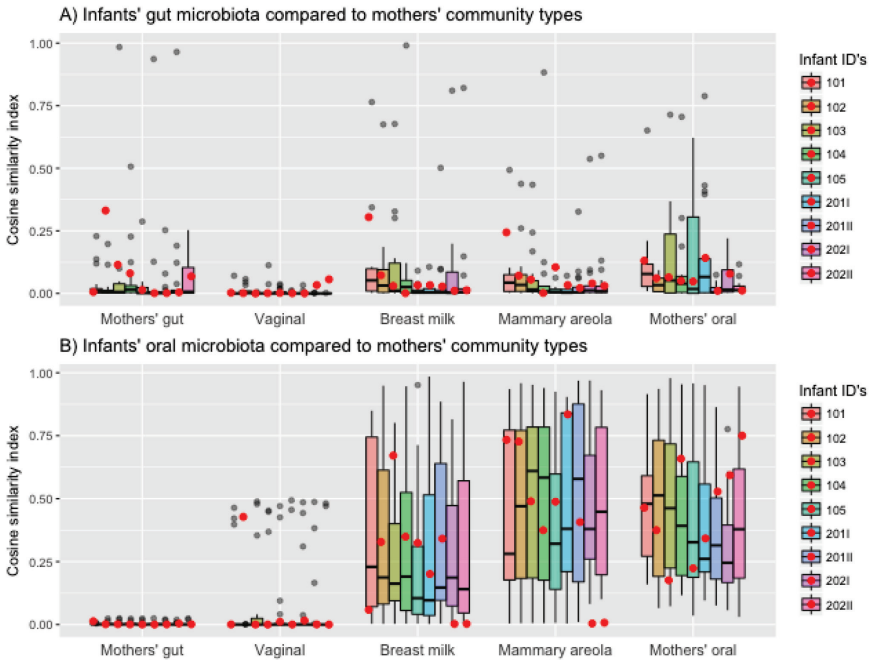

Figure 4. Average Cosine similarity index between infant's gut (A) and oral (B) microbiota and their own mother's microbial community types (red dots), and the community types observed in unrelated mothers (Tukey boxplot).

4.3.2 The influence of different maternal microbial communities on the development of infant gut microbiota

In our study, infants shared on average 67% of the OTUs in their gut microbiota with at least one of the mother's community types during 48-72 h after birth and this proportion did not change substantially over the study period (at 6-8 weeks it was on average [SD] 57% [13%] and at 6 months 57% [16%]). Nevertheless, the proportion of OTUs in infant gut microbiota, that were shared between the gut microbiota of individual mother-infant pairs, was not very high (mean [SD] proportion of OTUs in infant gut microbiota that was shared between mother-infant pairs: 32% [13%]). The proportion was comparable to the percentage of OTUs shared with the communities colonizing mother's oral cavity and mammary areola (Table 8). The low rate of shared OTUs indicates that when considering the possible transfer of microbes from a specific maternal microbial community, only a fraction of bacteria observed in infant gut microbiota has the likelihood of being vertically transferred. Also, it has to be taken into account that majority of the infants participating in this study were born via C-section, which may have a profound effect on the results as delivery mode has previously been shown to be a factor influencing the vertical transmission of microbes (Bäckhed *et al.*, 2015, Korpela *et al*., 2018b). This may explain why we observed a higher percentage of OTUs shared with mother's mammary areola and oral microbiota as the gut microbiota of C-section infants tend to be enriched with bacteria typically know to be of skin and mouth origin (Dominguez-Bello *et al*., 2010, Bäckhed *et al*., 2015).

OTUs shared between infant gut microbiota, and their mother's community types had different abundance in these community types, which may be due to lower bacterial diversity and therefore lower competition in the infant gut and an indication that the

fitness of bacteria reaching the infant's gut plays a more significant role than the quantitative contribution of microbial seeding as hypothesized by previous studies reporting similar results (Asnicar *et al.,* 2016, Ferretti *et al.,* 2018). Infant gut OTUs shared between individual mother-infant pairs (incl. different maternal community types) mapped to a variety of taxa with on average 54% of the OTUs being allocated to taxa typically observed in the gut microbiota (e.g. members of *Enterobacteriaceae, Clostridiales, Bacteroides*, and *Bifidobacterium*). Approximately 39% of the OTUs mapped to taxa typically related to skin and oral microbiota (mainly genus *Streptococcus*, but also families *Porphyromonadaceae,* and *Staphylococcaceae*). These proportions were observed at 48-72 h after birth and they did not change significantly over the study period (73% vs 27%; and 64% vs 36%, respectively for the gut- vs oral/skin-related taxa at 6-8 weeks and six months after birth). The variation of taxa changed over time as the proportion of members of order *Clostridiales* showed an increasing trend among gut-related taxa over the study period (at 48-72 h after birth only 18% of the OTUs mapped to taxa often related to the gut microbiota mapped to order *Clostridiales*. The proportion increased, although not significantly, to 40% and 50%, respectively at 6-8 weeks and six months after birth).

The proportion of OTUs shared between infant gut microbiota, and maternal communities of unrelated mothers were higher than the same ratio observed between individual mother-infant pairs regardless of the maternal community type (Table 8). Also, most of the OTUs observed in infant gut microbiota were simultaneously present in several of the maternal community types, either one's own or unrelated mothers, and this included OTUs dominating infant gut microbiota (mapped to family *Enterobacteriaceae,* and genera *Streptococcus* and *Veillonella*). There were only on average 9% of OTUs belonging to a variety of taxa that were exclusively shared with one's mother and not with unrelated mothers. All these observations above indicate to the shortcoming of 16S rRNA gene sequencing in the context of this analysis because it is unlikely that this method enables to identify taxa on the lower taxonomic level than genus (Konstantinidis & Tiedje, 2007), meaning that different species or, even more so, strains are left undetermined. Moreover, recent studies have shown that not all the strains of the species simultaneously observed in the gut microbiota of infants and their mothers are transmitted from mother to infant as only approximately 14–16.4% of strains from the species observed in mother-infant pairs are vertically transmissible (Asnicar *et al.,* 2016, Ferretti *et al.,* 2018). Thus, it is not possible to make conclusions about which specific OTUs observed in infant gut microbiota have been vertically transmitted from the mother and which maternal community type plays the most significant role in this transmission. It may also be the reason why we failed to observe the higher similarity between individual mother-infant pairs when compared to the similarity with unrelated mothers, as we cannot rule out that these differences may be visible when taking into account the variation observed at species or strain level.

Table 8. The proportion of OTUs observed in infant gut microbiota that was shared with mothers' community types; average (SD) proportion of OTUs shared between mother-infant pairs and with the communities of unrelated mothers.

Infant gut microbiota	Mothers' community type	Percentage of OTUs shared		
		48-72h after birth	6-8 weeks after birth	Six months after birth
Individual mother-infant pairs (average % [SD])	Mothers' gut	34% [5%]	34% [19%]	29% [11%]
	Vaginal[1]	19% [12%]	8% [4%][1]	17% [10%][1]
	Breast milk	25% [11%]	14% [6%]	31% [11%]
	Mammary areola	34% [18%]	19% [13%]	35% [18%]
	Mothers' oral	35% [12%]	28% [10%]	39% [15%]
Shared with unrelated mothers (average % [SD])	Mothers' gut	62% [13%]	60% [11%]	63% [8%]
	Vaginal[1]	42% [11%]	27% [8%][1]	41% [19%][1]
	Breast milk	50% [11%]	31% [13%]	36% [24%]
	Mammary areola	54% [11%]	39% [13%]	37% [19%]
	Mothers' oral	47% [12%]	40% [15%]	48% [18%]

[1] *Infant gut and oral microbiota at these time points were compared to vaginal microbiota observed in mothers before giving birth.*

4.3.2.1 The similarity of bifidobacterial communities observed in different maternal microbial communities and infant gut microbiota

When analyzing the abundance and prevalence of bifidobacterial, we observed that infant gut microbiota was colonized with high levels of members from genus *Bifidobacterium* throughout the study (Figure 3; changes in levels were not statistically significant). The levels of bifidobacteria in infant gut microbiota did not differ significantly from maternal community types (mean [SD] log10 of copies/uL of DNA per mg of stool was 3.3 [1], 0.2 [0.9], 0.2 [0.5], 0.5 [0.7] and -0.1 [0.5], respectively in mother's gut, vaginal, breast milk, mammary areola and oral microbiota; p – values > 0.05).

There were nine different species of *Bifidobacterium* observed in the gut microbiota of infants (Figure 5), with on average [SD] of 2.8 [1.7] species of *Bifidobacterium* present per infant. The highest number of *Bifidobacterium* species was observed in twin pairs, with on average 4.8 [1] species identified in one pair and on average 3.5 [1.4] in another. In maternal community types, the average number of *Bifidobacterium* species per sample tended to be lower than in infant gut microbiota (1.3 [1.5]; p-value = 0.05), except for mother's gut microbiota that harbored on average 3.1 [1] *Bifidobacterium* species per mother. The highest number of observations, where a specific bifidobacterial species were shared between individual mother-infant pairs, was observed for infant gut and mother's gut microbiota, although several species were also shared with mothers' vaginal and oral communities (Figure 5). None of the bifidobacterial species was exclusively observed in either infant gut or mother's community types, and all of them were shared in at least one mother-infant pair (simultaneously with several maternal community types) (Figure 5). *Bifidobacterium dentium* and *B. pseudocatenulatum* were the most frequent species shared between mother-infant pairs (Figure 5). There were no significant differences observed between infants born vaginally (n=2) vs via C-section or infants being mostly fed with formula (n=4) vs breastmilk.

The failure to observe differences and describe possible mother-to-infant transmission of bifidobacteria is probably due to the low number of subjects participating in our study, but it also may be because most of the infants were born via C-section. The frequency of mother-to-infant transmission of members from genus *Bifidobacterium* in C-section infants tends to be lower compared to vaginally born infants (Bäckhed *et al.*, 2015). Also, the strains of *Bifidobacterium adolescentis*, *B. bifidum, B. catenulatum, B. longum subsp. longum*, and *B. pseudocatenulatum* are monophyletic between the gut microbiota of mothers and their vaginally born infants, but these monophyletic strains are not present in mothers and their C-section infants (Makino et al., 2013).

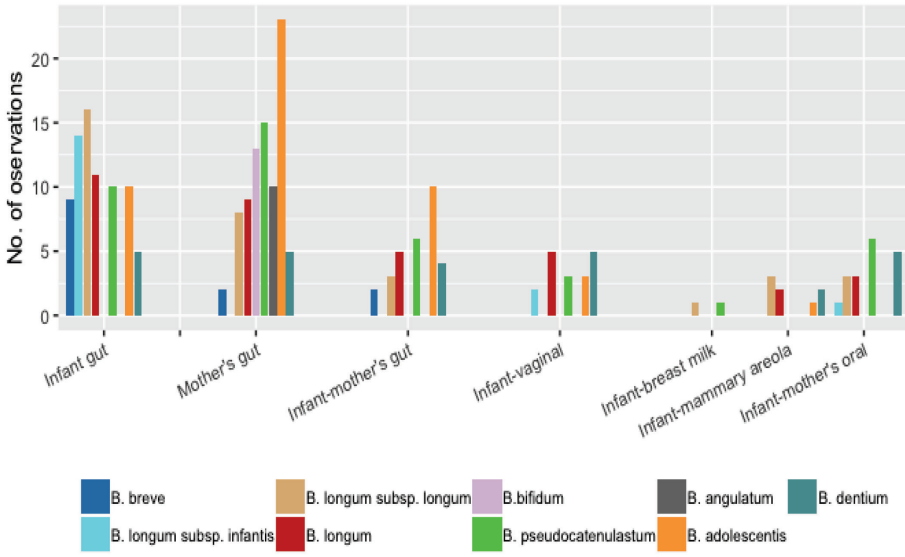

Figure 5. The prevalence of bifidobacterial species in infant and mother's gut and the incidence of bifidobacterial species shared between mother-infant pairs (mother's gut, vaginal, breast milk, mammary areola and oral microbial community).

4.3.3 The influence of different maternal microbial communities on the development of infant oral microbiota

In our study, infants shared on average 75.8% of the OTUs in their oral microbiota with at least one of the mother's community types during 48-72 h after birth and this proportion did not change substantially over the study period (at 6-8 weeks it was on average [SD] 71% [22%] and at 6 months 49% [30%]). The highest proportion of OTUs colonizing infant oral microbiota was shared with the mother's oral microbiota (on average [SD] 51% [11%]; Table 9). The similarity and the high numbers of shared OTUs with mothers' oral microbiota may be the result of the maternal habits of infant care (e.g., frequent use of the same spoon, licking the pacifier, kissing on the mouth). While not all mother-to-infant contacts involve direct interactions of the oral microbiota, a similar observation was made in a recent study where the salivary microbiota of romantically involved partners exhibited increasing similarity when partners kissed at relatively high frequencies (Kort *et al.,* 2014).

OTUs shared between individual mother-infant pairs (incl. different maternal community types) mapped to a variety of taxa with on average 45% of the OTUs being allocated to taxa typically observed in the oral microbiota (mainly members from genus *Streptococcus,* but also *Leptotrichia* and *Gemella*). 38% of the OTUs mapped to taxa usually related to gut and skin microbiota (e.g. members from genera *Veillonella, Prevotella, Bacteroides* and *Staphylococcus*) at 48-72 h after birth, and the proportion did not change significantly over the study period (50% vs 34%; and 50% vs 29%, respectively for oral- vs gut/skin-related taxa at 6-8 weeks and six months after birth).

As with the results observed for infant gut microbiota, the proportion of OTUs shared between infant oral microbiota and maternal communities of unrelated mothers was higher than the same ration found between individual mother-infant pairs regardless of the maternal community type (Table 9). Also, most of the OTUs observed in infant oral microbiota were simultaneously present in several of the maternal community types, either one's own or unrelated mothers, and this included OTUs dominating infant oral microbiota (mapped to genus *Streptococcus*). There were only on average 9% of infant oral OTUs belonging to a variety of taxa that were exclusively shared with one's mother and not with unrelated mothers.

Table 9. The proportion of OTUs observed in infant gut microbiota that was shared with mothers' community types; average (SD) proportion of OTUs shared between mother-infant pairs; and top 3 OTUs that were shared between these observations.

Infant oral microbiota	**Mothers' community type**	**Percentage of OTUs shared**		
		48-72h after birth	**6-8 weeks after birth**	**Six months after birth**
Individual mother-infant pairs (average % [SD])	Mothers' gut	20% [11%]	16% [9%]	18% [6%]
	Vaginal[1]	21% [11%]	11% [8%][1]	16% [13%][1]
	Breast milk	30% [13%]	33% [21%]	29% [14%]
	Mammary areola	43% [22%]	38% [24%]	33% [17%]
	Mothers' oral	50% [12%]	55% [9%]	46% [10%]
Shared with unrelated mothers (average % [SD])	Mothers' gut	47% [15%]	37% [8%]	45% [18%]
	Vaginal[1]	46% [14%]	28% [7%][1]	43% [17%][1]
	Breast milk	54% [13%]	70% [9%]	45% [20%]
	Mammary areola	68% [11%]	79% [6%]	45% [17%]
	Mothers' oral	67% [9%]	75% [12%]	60% [14%]

[1] Infant gut and oral microbiota at these time points were compared to vaginal microbiota observed in mothers before giving birth.

5 Conclusions

The current thesis reached to the following main conclusions:

1) Vaginal microbiota of women of reproductive age cluster into five different community types. Also, approximately a third of the samples (28.9%) do not cluster in any of the five groups and are characterized by heterogeneous bacterial communities. *Candida spp*., which can be detected in the vaginal microbiota of 69.9% of women, is more prevalent in this community than previously recognized. More precisely:

- The vaginal microbiota of almost all women (99%) is colonized with lactobacilli, which is consistent with the general knowledge that normal vaginal microbiota is dominated by members from genus *Lactobacillus*. Nevertheless, members of genera suggestive of a bacterial vaginosis-like microbiota (such as members from genera *Gardnerella* and *Atopobium*) are also relatively abundant and thus, their presence is not always indicative of a vaginal disease or symptoms. The vaginal mycobiome is dominated by *Candida spp.* (mean relative abundance 0.37). The most prevalent and abundant OTU from genus *Candida* colonizing vaginal microbiota is *C. albicans*. The abundance of this species among women indicates that the colonization with *C.albicans* is not always related to pathology.
- A group of women (28.9%) harbor heterogeneous vaginal microbiota that is characterized by lack of dominants (taxa with relative abundance higher than 0.2), yet the high prevalence of members from genera *Atopobium, Gardnerella,* and *Prevotella*. These women tend to have higher microbial diversity and high vaginal pH. Moreover, there is an association between higher microbial diversity and higher vaginal pH and additionally with the presence of malodorous vaginal discharge.
- During labor and the postpartum period, bacteria associated with heterogeneous community type seem to be overrepresented among women as only a few women have high levels of lactobacilli colonizing the vaginal microbiota onset of labor and Nugent score indicating a normal vaginal microbiota during the postpartum period (during six months after delivery), which may be related to prenatal hormonal changes (namely estrogen) that leads to a reduction in the abundance of lactobacilli.

2) The gut microbiota of extremely preterm infants differs considerably from full term infants especially during the first four weeks of life. More precisely:

- Extremely preterm infants harbor gut microbiota that differs from full term infants by having low abundance or absence of members from genera *Bifidobacterium* and *Bacteroides* throughout the first two months of life and by being dominated by members from family *Staphylococcaceae* during the first weeks of life. The high abundance of staphylococci in the gut microbiota of extremely preterm infants during the first month of life may be the reason why there is an increased risk of acquiring coagulase-negative staphylococcal infections in preterm neonates.
- Development of the gut microbiota in extremely preterm infants goes through rapid compositional changes (e.g. the dominance of members from family *Staphylococcaceae* is replaced by the dominance of members from family *Enterobacteriaceae*) and increase in microbial diversity during the first weeks of life. The community starts to resemble the gut microbiota of full term infants over time, which is an indication of delayed development of the gut microbiota in extremely preterm infants.

- Factors, such as antibiotic treatment, delivery mode and feeding regimen (formula vs breastmilk), do not seem to have an obvious effect on the development of the gut microbiota in extremely preterm infants. Nevertheless, the impact of these factors may be overshadowed by the combination of other factors, such as less mature gestational age, low birth weight, and neonatal intensive care unit stay. Microbial diversity, however, is significantly lower in infants who are fed parenterally for a more extended period, which may be related to the nutrition deficiency in the gut resulting from the parenteral feeding. Also, microbial diversity is significantly higher among infants born to mothers suffering from chorioamnionitis, which may be related to the exposure to inflammation and potential bacterial colonizers.

3) The gut microbiota of full term infants has low similarity with maternal community types during the first six months of life. At the same time, infant oral microbiota has a high resemblance to mothers' breast milk, mammary areola, and oral microbiota. More precisely:

- The gut and oral microbiota of full term infants do not exhibit significantly greater similarity to their own mother's microbial communities than to unrelated mothers. Moreover, the proportion of shared OTUs is higher with the communities of unrelated mothers than between individual mother-infant pairs. The biological significance of this observation is unknown, but it may indicate to the shortcoming of 16S rRNA gene sequencing in the context of this analysis, because it is unlikely that this method enables to identify taxa on the lower taxonomic level than genus and, thus, take into account variation observed at species or strain level.
- None of the maternal community types stands out as a potential primary source of bacteria to the infant gut microbiota. The highest proportion of OTUs shared between infant gut microbiota, and his or her mother's community types is shared with mother's gut microbiota (32%), but this proportion is similar with mother's oral and mammary areola microbiota, which indicates that when considering the possible transfer of microbes from a specific maternal microbial community, only a fraction of microbes has the likelihood of being vertically transferred. A relatively high proportion of OTUs shared with mother's oral and mammary areola microbiota may be related to the fact that majority of the participating infants were born via C-section as C-section infants tend to be enriched with bacteria typically know to be of skin and mouth origin. Also, C-section birth can be attributable for the failure to observe possible mother-to-infant transmission of members from *Bifidobacterium* as the frequency of transmission of these bacteria in C-section infants tends to be lower when compared to vaginally born infants.
- The highest proportion of OTUs shared between infant oral microbiota, and his or her mother's community types is shared with mother's oral microbiota (51%) which together with high similarity indicates that constant contact between microbial communities, probably resulting from the maternal habits of infant care, increases their similarity.

References

Abrahamsson, T. R., Jakobsson, H. E., Andresson, A. F., Björksten, B., Engstrand, L., and Jenmalm, M. C. (2011). Low diversity of the gut microbiota in infants with atopic eczema. J Allergy Clin Immunol 129, 434–440.

Abrahamsson, T. R., Jakobsson, H. E., Andersson, A. F., Björksten, B., L. Engstrand, L., and M. C. Jenmalm, M. C. (2014). Low gut microbiota diversity in early infancy precedes asthma at school age. Clin Exp Allergy 44, 842–850.

Adlerberth, I. & Wold A. E. (2009). Establishment of the gut microbiota in Western infants. Acta Paediatr 98, 229–238.

Aloisio, I., Mazzola, G., Corvaglia, L. T., Tonti, G., Faldella, G., Biavati, B., and Di Gioia, D. (2014). Influence of intrapartum antibiotic prophylaxis against group B *Streptococcus* on the early newborn gut composition and evaluation of the anti-*Streptococcus* activity of *Bifidobacterium* strains. Appl Microbiol Biotechnol 98, 6051–6060.

Amsel, R., Totten, P. A., Spiegel, C. A., Chen, K. C. S., Eschenbach, D. A., and Holmes, K. K. (1983). Nonspecific vaginitis: diagnostic criteria and epidemiologic associations. Am J Med 74, 14–22.

Asnicar, F., Manara, S., Zolfo, M., Truong, D. T., Scholz, M., Armanini, F., Ferretti, P., Gorfer, V., Pedrotti, A., Tett, A., and Segata, N. (2017). Studying vertical microbiome transmission from mothers to infants by strain-level metagenomic profiling. mSystems 2, e00164-16.

Arboleya, S., Binetti, A., Salazar, N., Fernández, N., Solís, G., Hernández-Barranco, A., Margolles, A., de los Reyes-Gavilán, C. G., and Gueimonde, M. (2012a). Establishment and development of intestinal microbiota in preterm neonates. FEMS Microbiol Ecol 79, 763–772.

Arboleya, S., Ang, L., Margolles, A., Yiyuan, L., Dongya, Z., Liang, X., Solís, G., Fernández, N., de los Reyes-Gavilán, C. G., and Gueimonde, M. (2012b). Deep 16S rRNA metagenomics and quantitative PCR analyses of the premature infant fecal microbiota. Anaerobe 18, 378–380.

Arboleya, S., Sánchez, B., Milani, S., Duranti, S., Solís, G., Fernández, N., de los Reyes-Gavilán, C. G., Ventura, M., Margolles, A., and Gueimonde, M. (2015). Intestinal Microbiota Development in Preterm Neonates and Effect of Perinatal Antibiotics. J Pediatr 166, 538–544.

Arboleya, S., Martinez-Camblor, P., Solís, G., Suárez, M., Fernández, N., de los Reyes-Gavilán, C. G., and Gueimonde, M. (2017). Intestinal microbiota and weight-gain in preterm neonates. Front Microbiol 8, 183.

Arike, L., Holmén-Larsson, J., and Hansson, G. C. (2017). Intestinal Muc2 mucin O-glycosylation is affected by microbiota and regulated by differential expression of glycosyltranferases. Glycobiology 27, 318–328.

Arumugam, M., Raes, J., Pelletier, E., Le Paslier, D., Yamada, T., Mende, D. R., Fernandes, G. R., Tap, J., Bruls, T., Batto, J.-M., Bertalan, M., Borruel, N., Casellas, F., Fernandez, L., Gautier, L., Hansen, T., Hattori, M., Hayashi, T., Kleerebezem, M., Kurokawa, K., Leclerc, M., Levenez, F., Manichanh, C., Nielsen, H. B., Nielsen, T., Pons, N., Poulain, J., Qin, J., Sicheritz-Ponten, T., Tims, S., Torrents, D., Ugarte, E., Zoetendal, E. G., Wang, J., Guarner, F., Pedersen, O., de Vos, W. M., Brunak, S., Doré, J., Weissenbach, J., Ehrlich, S. D., and Bork, P. (2011). Enterotypes of the human gut microbiome. Nature 473, 174–180.

Aujoulat, F., Roudière, L., Picaud, J.-C., Jacquot, A., Filleron, A., Neveu, D., Baum, T.-P., Marchandin, H., and Jumas-Bilak, E. (2014). Temporal dynamics of the very premature infant gut dominant microbiota. BMC Microbiol 14, 325.

Avershina, E., Slangsvold, S., Simpson, M. R., Storrø, O., Johnsen, R., Øien, T., and Rudi, K. (2017). Diversity of vaginal microbiota increases by the time of labor onset. Sci Rep 7, 17558.

Azad, M. B., Konya, T., Maughan, H., Guttman, D. S., Field, C. J., Chari, R. S., Sears, M. R., Becker, A. B., Scott, J. A, and Kozyrskyj, A. L. (2013). Gut microbiota of healthy Canadian infants: profiles by mode of delivery and infant diet at 4 months. CMAJ 185, 385–394.

Ballard, O., and Morrow, A. L. (2013). Human Milk Composition: Nutrients and Bioactive Factors. Pediatr Clin North Am 60, 49–74.

Barcenilla, A., Pryde, S. E., Martin, J. C., Duncan S. H., Stewart C. S., Henderson, C., and Flint, H. J. (2000). Phylogenetic relationships of butyrate-producing bacteria from the human gut. Appl Environ Microbiol 66, 1654–1661.

Barrett, E., Kerr, C., Murphy, K., O'Sullivan, O., Ryan, C. A., Dempsey, E. M., Murphy, B. P., O'Toole, P. W., Cotter, P. D., Fitzgerald, G. F., Ross, R. P., and Stanton, C. (2013). The individual-specific and diverse nature of the preterm infant microbiota. Arch Dis Child Fetal Neonatal Ed 98, F283–F283.

Bergström, A., Skov, T. H., Bahl, M. I., Roager, H. M., Christensen, L. B., Ejlerskov, K. T., Mølgaard, C., Michaelsen, K. F., and Licht, T. R. (2014). Establishment of Intestinal Microbiota during Early Life: a Longitudinal, Explorative Study of a Large Cohort of Danish Infants. Appl Environ Microbiol 80, 2889–2900.

Bezirtzoglou, E., Tsiotsias, A., and Welling, G. W. (2011). Microbiota profile in feces of breast- and formula-fed newborns by using fluorescence in situ hybridization (FISH). Anaerobe 17, 478–482.

Betrán, A. P., Ye, J., Moller, A.-B., Zhang, J., Gülmezoglu, A. M., and Torloni M. R. (2016). The Increasing Trend in Caesarean Section Rates: Global, Regional and National Estimates: 1990-2014. PLoS ONE 11, e0148343.

Biasucci, G., Rubini, M., Riboni, S., Morelli, L., Bessi, E., and Retetangos, C. (2010). Mode of delivery affects the bacterial community in the newborn gut. Early Hum Dev 86, 13–15.

Bisgaard, H., Li, N., Bonnelykke, K., Chawes, B. L. C., Skov, T., Paludan-Muller, G., Stokholm, J., Smith, B., and Krogfelt, K. A. (2011). Reduced diversity of the intestinal microbiota during infancy is associated with increased risk of allergic disease at school age. J Allergy Clin Immunol 128, 646–652.

Blekhman, R., Goodrich, J. K., Huang, K., Sun, Q., Bukowski, R., Bell, J. T., Spector, T. D., Keinan, A., Ley, R. E., Gevers, D., and Clark, A. G. (2015). Host genetic variation impacts microbiome composition across human body sites. Genome Biol 16, 191.

Blencowe, H., Cousens, S., Oestergaard, M. Z., Chou, D., Moller, A.-B., Narwal, R., Adler, A., Garcia, C. V., Rohde, S., Say, L., and Lawn, J. E. (2012). National, regional, and worldwide estimates of preterm birth rates in the year 2010 with time trends since 1990 for selected countries: a systematic analysis and implications. Lancet 379, 2162–2172.

Bokulich, N. A., Chung, J., Battaglia, T., Henderson, N., Jay, M., Li, H., Lieber, A., Wu, F., Perez-Perez, G. I., Chen, Y., Schweizer, W., Zheng, X., Contreras, M., Dominguez-Bello, M. G., and Blaser, M. J. (2016). Antibiotics, birth mode, and diet shape microbiome maturation during early life. Sci Transl Med 8, 343ra82.

Butel, M.-J., Suau, A., Campeotto, F., Magne, F., Aires, J., Ferraris, L., Kalach, N., Leroux, B., Dupont, C. (2007). Conditions of bifidobacterial colonization in preterm infants: a prospective analysis. J Pediatr Gastroenterol Nutr 44, 577–582.

Bäckhed, F., Roswall, J., Peng, Y., Feng, Q., Jia, H., Kovatcheva-Datchary, P., Li, Y., Xia, Y., Xie, H., Zhong, H., Khan, M. T., Zhang, J., Li, J., Xiao, L., Al-Aama, J., Zhang, D., Lee, Y. S., Kotowska, D., Colding, C., Tremaroli, V., Yin, Y., Bergman, S., Xu, X., Madsen, L., Kristiansen, K., Dahlgren, J., and Wang, J. (2015). Dynamics and stabilization of the human gut microbiome during the first year of life. Cell Host Microbe 17, 690–703.

Chang, P. V., Hao, L., Offermanns, S., and Medzhitov, R. (2014). The microbial metabolite butyrate regulates intestinal macrophage function via histone deacetylase inhibition. PNAS 111, 2247–2252.

Chelakkot, C., Choi, Y., Kim, D-K., Park, H. T., Ghim, J., Kwon, Y., Jeon, J., Kim, M-S., Jee, Y-K., Gho, Y. S., Park, H-S., Kim, Y-K., and Ryu, S. H. (2018). *Akkermansia muciniphila*-derived extracellular vesicles influence gut permeability through the regulation of tight junctions. Exp Mol Med 50, e450.

Cheng, J., Ringel-Kulka, T., Heikamp-de Jong, I., Ringel, Y., Carroll, I., de Vos, W. M., Salojärvi, J., and Satokari, R. (2016). Discordant temporal development of bacterial phyla and the emergence of core in the fecal microbiota of young children. ISME J 10, 1002–1014.

Chernikova, D. A., Madan, J. C., Housman, M. L., Zain-ul-abideen, M., Lundgren, S. N., Morrison, H. G., Sogin, M. L., Williams, S. M., Moore, J. H., Karagas, M. R., and Hoen, A. G. (2018). The premature infant gut microbiome during the first 6 weeks of life differs based on gestational maturity at birth. Pediatr Res 84, 71–79.

Chu, D. M., Ma, J., Prince, A. L., Antony, K. M., Seferovic, M. D., and Aagaard, K. M. (2017). Maturation of the infant microbiome community structure and function across multiple body sites and in relation to mode of delivery. Nat Med 23, 314–326.

Claud, E. C., Keegan, K. P., Brulc, J. M., Lu, L., Bartels, D., Glass, E., Chang, E. B., Meyer, F., and Antonopoulos, D. A. (2013). Bacterial community structure and functional contributions to emergence of health or necrotizing enterocolitis in preterm infants. Microbiome 1, 20.

Comelli, E. M., Simmering, R., Faure, M., Donnicola, D., Mansourian, R., Rochat, F., Corthesy-Theulaz, I., Cherbut, C. (2007). Multifaceted transcriptional regulation of the murine intestinal mucus layer by endogenous microbiota. Genomics 91, 70–77.

Corvaglia, L., Tonti, G., Martini, S., Aceti, A., Mazzola, G., Aloisio, I., Di Gioia, D., and Faldella, G. (2016). Influence of intrapartum antibiotic prophylaxis for group B *Streptococcus* on gut microbiota in the first month of life. JPGN 62, 304–308.

Coyte, K. Z., Schluter, J., and Foster K. R. (2015). The ecology of the microbiome: Networks, competition, and stability. Science 350, 663–666.

De Vadder, F., Kovatcheva-Datchary, P., Goncalves, D., Vinera, J., Zitoun, C., Duchampt, A., Bäckhed, F., and Mithieux, G. (2014). Microbiota-generated metabolites promote metabolic benefits via gut-brain neural circuits. Cell 156, 84–96.

DiGiulio, D. B., Callahan, B. J., McMurdie,P. J., Costello, E. K., Lyell, D. J., Robaczewska, A., Sun, C. L., Goltsman, D. S. A., Wong, R. J., Shaw, G., Stevenson, D. K., Holmes, S. P., and Relman, D. A. (2015). Temporal and spatial variation of the human microbiota during pregnancy. PNAS 112, 11060–11065.

Di Renzo, G. C., Melin, P., Berardi, A., Blennow, M., Carbonell-Estrany, X., Donzelli, G. P., Hakansson, S., Hod, M., Hughes, R., Kurtzer, M., Poyart, C., Shinwell, E., Stray-Pedersen, B., Wielgos, M., and El Helali, N. (2014). Intrapartum GBS screening and antibiotic prophylaxis: a European consensus conference. J Matern Fetal Neonatal Med 28, 766–782.

Dogra, S., Sakwinska, O., Soh, S.-E., Catherine Ngom-Bru, C., Wolfram M. Brück, W. M., Bernard Berger, B., Harald Brüssow, H., Yung Seng Lee, Y. S., Fabian Yap, F., Yap-Seng Chong, Y.-S., Keith M. Godfrey, K. M., and Joanna D. Holbrook, J. D. (2015). Dynamics of Infant Gut Microbiota Are Influenced by Delivery Mode and Gestational Duration and Are Associated with Subsequent Adiposity. mBio 6, e02419-14.

Dominguez-Bello, M. G., Costello, E. K., Contreras, M., Magris, M., Hidalgo, G., Fierer, N., and Knight, R. (2010). Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns. PNAS 107, 11971–11975.

Dominguez-Bello, M. G., De Jesus-Laboy, K. M., Shen, N., Cox, L. M., Amir, A., Gonzalez, A., Bokulich, N. A., Song, S. J., Hoashi, M., Rivera-Vinas, J. I., Mendez, K., Knight, R., and Clemente, J. C. (2016). Partial restoration of the microbiota of cesarean-born infants via vaginal microbial transfer. Nat Med 22, 250–253.

Donders, G. G. G. (2007). Definition and classification of abnormal vaginal flora. Best Pract Res Cl Ob 21, 355–373.

Donders, G. G. G., Van Calsteren, K., Bellen, G., Reybrouck, R., Van den Bosch, T., Riphagen, I., and Van Lierdee, S. (2009). Predictive value for preterm birth of abnormalvaginal flora, bacterial vaginosis and aerobic vaginitis during the first trimester of pregnancy. BJOG 116, 1315–1324.

Eggesbø, M., Moen, B., Peddada, S., Baird, D., Rugtveit, J., Midtvedt, T., Bushel, P. R., Sekelja, M., and Rudi, K. (2011). Development of gut microbiota in infants not exposed to medical interventions. APMIS 119, 17–35.

Fallani, M., Young, D., Scott, J., Norin, E., Amarri, S., Adam, R., Aguilera, M., Khanna, S., Gil, A., Edwards, C. A., and Dore, J. (2010). Intestinal microbiota of 6-week-old infants across Europe: geographic influence beyond delivery mode, breast-feeding, and antibiotics. JPGN 51, 77–84.

Fallani, M., Amarri, S., Uusijarvi, A., Adam, R., Khanna, S., Aguilera, M., Gil, A., Vieites, J. M., Norin, E., Young, D., Scott, J. A., Dore, J., and Edwards, C. A. (2011). Determinants of the human infant intestinal microbiota after the introduction of first complementary foods in infant samples from five European centres. Microbiology 157, 1385–1392.

Farr, A., Kiss, H., Holzer, I., Husslein, P., Hagmann, and Petricevic, L. (2015). Effect of asymptomatic vaginal colonization with Candida albicans on pregnancy outcome. Acta Obstet Gynecol Scand 94, 989–996.

Fernández, L., Langa, S., Martín, V., Maldonado, A., Jiménez, E., Martín, R., and Rodríguez, J. M. (2013). The human milk microbiota: Origin and potential roles in health and disease. Pharmacol Res 69, 1–10.

Ferretti, P., Pasolli, E., Tett, A., Ansicar, F., Gorfer, V., Fedi, S., Armanini, F., Truong, D. T., Manara, S., Zolfo, M., Beghini, F., Bertorelli, R., De Sanctis, V., Bariletti, I., Canto, R., Clementi, R., Cologna, M., Crifo, T., Cusumano, G., Gottardi, S., Innamorati, C., Mase, C., Postai, D., Savoi, D., Duranti, S., Lugli, G. A., Mancabelli, L., Turroni, F., Ferrario, C., Milani, C., Mangifesta, M., Anzalone, R., Viappiani, A., Yassour, M., Vlamakis, H., Xavier, R., Collado, C. M., Koren, O., Tateo, S., Soffiati, M., Pedrotti, A., Ventura, M., Huttenhower, C., Bork, P., and Segata, N. (2018). Mother-to-infant microbial transmission from different body sites shapes the developing infant gut microbiome. Cell Host Microbe 24, 133–145.e5.

Forsgren, M., Isolauri, E., S Salminen, S., and Rautava, S. (2017). Late preterm birth has direct and indirect effects on infant gut microbiota development during the first six months of life. Acta Paediatr 106, 1103–1109.

Fouhy, F., Guinane, C. M., Hussey, S., Wall, R., Ryan, C. A., Dempsey, E. M., Murphy, B., Ross, R. P., Fitzgerald, G. F., Stanton, C., and Cotter, P. D. (2012). High-throughput sequencing reveals the incomplete, short-term, recovery of the infant gut microbiota following parenteral antibiotic treatment with ampicillin and gentamycin. Antimicrob Agents Chemother 56, 5811–5820.

Frank, J. A., Reich, C. I., Sharma, S., Weisbaum, J. S., Wilson, B. A., and Olsen, G. J. (2008). Critical Evaluation of Two Primers Commonly Used for Amplification of Bacterial 16S rRNA Genes. Appl Environ Microbiol 74, 2461–2470.

Fredricks, D. N., Fiedler, T. L., and Marrazzo, J. M. (2005) Molecular identification of bacteria associated with bacterial vaginosis. N Engl J Med 353, 1899–1911.

Freitas, A. C., Bocking, A., Hill, J. E., Money, D. M., and the VOGUE Research Group. (2018). Increased richness and diversity of the vaginal microbiota and spontaneous preterm birth. Microbiome 6, 117.

Furusawa, Y., Obata1, Y., Fukuda, S., Endo, T. A., Nakato, G., Takahashi, D., Nakanishi, Y., Uetake, C., Kato, K., Kato, T., Takahashi, M., Fukuda, N. N., Murakami, S., Miyauchi, E., Hino, S., Atarashi, K., Onawa, S., Fujimura, Y., Lockett, T., Clarke, J. M., Topping, D. L., Tomita, M., Hori, S., Ohara, O., Morita, T., Koseki, H., Kikuchi, J., Honda, K., Hase, K., and Ohno, H. (2013). Commensal microbe-derived butyrate induces the differentiation of colonic regulatory T cells. Nature 504, 446–450.

Gajer, P., Brotman, R. M., Bai, G., Sakamoto, J., Schütte, U. M. E., Zhong, X., Koenig, S. S. K., Fu, L., Ma, Z. Zhou, X., Abdo, Z., Forney, L. J., and Ravel, J. (2012). Temporal Dynamics of the Human Vaginal Microbiota. Sci Transl Med 4, 132ra52.

Gibson, M. K., Wang, B., Ahmadi, S., Burnham, C.-A. D., Tarr, P. I., Warner, B. B., and Dantas, G. (2016). Developmental dynamics of the preterm infant gut microbiota and antibiotic resistome. Nat Microbiol 1, 16024.

Goodrich, J. K., Waters, J. L., Poole, A. C., Sutter, J. L., Koren, O., Blekhman, R., Beaumont, M., Van Treuren, W., Knight, R., Bell, J. T., Spector, T. D., Clark, A. G., and Ley, R. E. (2014). Human genetics shape the gut microbiome. Cell 159, 789–799.

Goodrich, J. K., Davenport, E. R., Beaumont, M., Jackson, M. A., Knight, R., Ober, C., Spector, T. D., Bell, J. T., Clark, A. G., and Ley, R. E. (2016). Genetic Determinants of the Gut Microbiome in UK Twins. Cell Host Microbe 19, 731–743.

Greenwood, C., Morrow, A. L., Lagomarcino, A. J., Altaye, M., Taft, D. H., Yu, Z., Newburg, D. S., Ward, D.V., and Schibler, K. R. (2014). Early empiric antibiotic use in preterm infants is associated with lower bacterial diversity and higher relative abundance of Enterobacter. J Pediatr 165, 23–29.

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

Haque, M. M., Merchant, M., Kumar, P. N., Dutta, A., and Mande, S. S. (2017). First-trimester vaginal microbiome diversity: A potential indicator of preterm delivery risk. Sci Rep 7, 16145.

Harmsen, H. J. M., Wildeboer–Veloo, A. C. M., Raangs, G. C., Wagendorp, A. A., Klijn, N., Bindels, J. G., and Welling, G. W. (2000). Analysis of intestinal flora development in breast-fed and formula-fed infants by using molecular identification and detection methods. J Pediatr Gastroenterol Nutr 30, 61–67.

Hsieh, C-Y., Osaka, T., Moriyama, E., Date, Y., Kikuchi, J., and Tsuneda, S. (2015). Strengthening of the intestinal epithelial tight junction by *Bifidobacterium bifidum.* Physiol Rep 3, e12327.

Holzer, I., Farr, A., Kiss, H., Hagmann, M., and Petricevic, L. (2017). The colonization with Candida species is more harmful in the second trimester of pregnancy. Arch Gynecol Obstet 295, 891–895.

Hunt K. M., Foster J. A., Forney L. J., Schutte U. M. E., Beck D. L., Abdo, Z., Fox, L. K., Williams, J. E., McGuire, M. K., and McGuire, M. A. (2011) Characterization of the Diversity and Temporal Stability of Bacterial Communities in Human Milk. PLoS ONE 6, e21313.

Ismail, I. H., Oppedisano, F., Joseph, S. J., Boyle, R. J., Licciardi, P. V., Robins-Browne, R. M., and Tang, M. L. K. (2012). Reduced gut microbial diversity in early life is associated with later development of eczema but not atopy in high-risk infants. Pediatr Allergy Immunol 23, 674–681.

Jakobsson, H. E., Abrahamsson, T. R., Jenmalm, M. C., Harris, K., Quince, C., Jernberg, C., Björkstén, B., Engstrand, L., and Andersson, A. F. (2014). Decreased gut microbiota diversity, delayed *Bacteroidetes* colonisation and reduced Th1 responses in infants delivered by Caesarean section. Gut 0, 1–8.

Jakobsson, H. E., Rodríguez-Piñeiro, A. M., Schütte, A., Ermund, A., Boysen, P., Bemark, M., Sommer, F., Bäckhed, F., Hansson, G. C., and Johansson, M. E. V. (2015). The composition of the gut microbiota shapes the colon mucus barrier. EMBO Rep 16, 164–177.

Jacquot, A., Neveu, D., Aujoulat, F., Mercier, G., Marchandin, H., Jumas-Bilak, E., and Picaud, J.-C. (2010). Dynamics and clinical evolution of bacterial gut microflora in extremely premature patients. J pediatr 158, 390–396.

Jost, T., Lacroix, C., Braegger, C. P., Rochat, F., and Chassard, C. (2014). Vertical mother–neonate transfer of maternal gut bacteria via breastfeeding. Environ Microbiol 16, 2891–2904.

Klaassens, E. S., Boesten, R. J., Haarman, M., Knol, J., Schuren, F. H., Vaughan, E. E., and de Vos, W. M. (2009). Mixed-species genomic microarray analysis of fecal samples reveals differential transcriptional responses of bifidobacteria in breast- and formula-fed infants. Appl Environ Microbiol 75, 2668–2676.

Koenig, J. E., Spor, A., Scalfone, N., Fricker, A. D., Stombaugh, J., Knight, R., Angenent, L. T., and Ley, R. E. (2011). Succession of microbial consortia in the developing infant gut microbiome. Proc Natl Acad Sci U S A, 108, 4578–85.

Konstantinidis, K. T. & Tiedje, J. M. (2007). Prokaryotic taxonomy and phylogeny in the genomic era: advancements and challenges ahead. Curr Opin Microbiol. 10, 504–509.

Korpela, K., Blakstad, E. W., Moltu, S. J., Strømmen, K., Nakstad, B., Rønnestad, A. E., Brække, K., Iversen, P. O., Drevon, C. A., and de Vos, W. (2018a). Intestinal microbiota development and gestational age in preterm neonates. Sci Rep 8, 2453.

Korpela, K., Costea, P., Coelho, L. P., Kandels-Lewis, S., Willemsen, G., Boomsma, D. I., Segata, N., and Bork, P. (2018b). Selective maternal seeding and environment shape the human gut microbiome. Genome Res 28, 561–568.

Kort, R., Caspers, M., van de Graaf, A., van Egmond, W., Keijser, B., Roeselers, G. (2014). Shaping the oral microbiota through intimate kissing. Microbiome 2, 41.

Kuppala, V. S., Meinzen-Derr, J., Morrow, A. L., and Schibler, K. R. (2011). Prolonged Initial Empirical Antibiotic Treatment is Associated with Adverse Outcomes in Premature Infants. J Pediatr 159, 720–725.

La Rosa, P. S., Warner, B. B., Zhou, Y., Weinstock, G. M., Sodergren, E., Hall-Moore, C. M., Stevens, H. J., Bennett, W. E., Shaikh, N., Linneman, L. A., Hoffmann, J. A., Hamvas, A., Deych, E., Shands, B. A., Shannon, W. D., and Tarr, P. I. (2014). Patterned progression of bacterial populations in the premature infant gut. PNAS 111, 12522–12527.

LaTuga, M. S., Ellis, J. C., Cotton, C. M., Goldberg, R. N., Wynn, J. L., Jackson, R. B., and Seed, P. C. (2011). Beyond bacteria: a study of the enteric microbial consortium in extremely low birth weight infants. PLoS ONE 6, e27858.

Leitich, H., Bodner-Adler, B., Brunbauer, M., Kaider, A., Egarter, C., and Husslein, P. (2003). Bacterial vaginosis as a risk factor for preterm delivery: a meta-analysis. Am J Obstet Gynecol 189, 139–147.

Leitich, H., and Kiss H. (2007). Asymptomatic bacterial vaginosis and intermediate flora as risk factors for adverse pregnancy outcome. Best Pract Res Clin Obstet Gynaecol 21, 375–390.

Li, M., Wang, B., Zhang, M., Rantalainen, M., Wang, S., Zhou, H., Zhang, Y., Shen, J., Pang, X., Zhang, M., Wei, H., Chen, Y., Lu, H., Zuo, J., Su, M., Qiu, Y., Jia, W., Xiao, C., Smith, L. M., Yang, S., Holmes, E., Tang, H., Zhao, G., Nicholson, J. K., Li, L., and Zhao, L. (2008). Symbiotic gut microbes modulate human metabolic phenotypes. PNAS 105, 2117–2122.

Liu, L., Johnson, H. L., Cousens, S., Perin, J., Scott, S., Lawn, J. E., Rudan, I., Campbell, H., Cibulskis, R., Li, M., Mathers, C., and Black, R. E. (2012). Global, regional, and national causes of child mortality: an updated systematic analysis for 2010 with time trends since 2000. Lancet 379, 2151–2161.

Lloyd-Price, J., Mahurkar, A., Rahnavard, G., Crabtree, J., Orvis, J., Hall, A. B., Brady, A., Creasy, H. H., McCracken, C., Giglio, M. G., McDonald, D., Franzosa, E. A., Knight, R., White, O., and Huttenhower, C. (2017). Strains, functions, and dynamics in the expanded Human Microbiome Project. Nature 550, 61–66.

Louis, P., Young, P., Holtrop, G., and Flint, H. J. (2010). Diversity of human colonic butyrate-producing bacteria revealed by analysis of the butyryl-CoA:acetate CoA-transferase gene. Environ Microbiol 12, 304–314.

Louis, P., and Flint, H. J. (2017). Formation of propionate and butyrate by the human colonic microbiota. Environ microbiol 19, 29–41.

MacIntyre, D. A., Chandiramani, M., Lee, Y. S., Kindinger, L., Smith, A., Angelopoulos, N., Lehne, B., Arulkumaran, S., Brown, R., Teoh, T. G., Holmes, E., Nicoholson, J. K., Marchesi, J. R., and Bennett, P. R. (2015). The vaginal microbiome during pregnancy and the postpartum period in a European population. Sci Rep 5, 8988.

Madan, J. C., Salari, R. C., Saxena, D., Davidson, L., O'Toole, G. A., Moore, J. H., Sogin, M. L., Foster, J. A., Edwards, W. H., Palumbo, P., and Hibberd, P. H. (2012). Gut microbial colonisation in premature neonates predicts neonatal sepsis. Arch Dis Child Fetal Neonatal Ed 97, F456–F462.

Magnúsdóttir, S., Ravcheev, D., de Crécy-Lagard, V., and Thiele, I. (2015). Systematic genome assessment of B-vitamin biosynthesis suggests co-operation among gut microbes. Front Genet 6, 148.

Mai, V., Torrazza, R. M., Ukhanova, M., Wang, X., Sun, Y., Li, N., Shuster, J., Sharma, R., Hudak, M. L., and Neu, J. (2013). Distortions in development of intestinal microbiota associated with late onset sepsis in preterm infants. PLoS ONE 8, e52876.

Makino, H., Kushiro, A., Ishikawa, E., Muylaert, D., Kubota, H., Sakai, T., Oishi, K., Martin, R., Ben-Amor, K., Oozeer, R., Knol, J., and Tanaka, R. (2011). Transmission of intestinal *Bifidobacterium longum subsp. longum* strains from mother to infant, determined by Multilocus Sequencing Typing and Amplified Fragment Length Polymorphism. Appl Environ Microbiol 77, 6788–6793.

Makino, H., Kushiro, A., Ishikawa, E., Kubota, H., Gawad, A., Sakai, T., Oishi, K., Martin, R., Ben-Amor, K., Knol, J., and Tanaka, R. (2013). Mother-to-infant transmission of intestinal bifidobacterial strains has an impact on the early development of vaginally delivered infant's microbiota. PLoS ONE 8, e78331.

March of Dimes, PMNCH, Save the Children, WHO. Born Too Soon: The Global Action Report on Preterm Birth. Eds Howson, C.P., Kinney, M. V., and Lawn, J. E. World Health Organization. Geneva, 2012.

Marcobal, A., Barboza, M., Froehlich, J. W., Block, D. E., German, J. B., Lebrilla, C. B., and Mills, D. A. (2010). Consumption of Human Milk Oligosaccharides by Gut-Related Microbes. J Agric Food Chem 58, 5334–5340.

Mazzola, M., Murphy, K., Ross, R. P., Di Gioia, D., Biavati, B., Corvaglia, L. T., Faldella, G., and Stanton, C. (2016). Early gut microbiota perturbations following intrapartum antibiotic prophylaxis to prevent broup B streptococcal disease. PLoS ONE 11, e0157527.

Mikamo, H., Sato, Y., Hayasaki, Y., Hua, Y. X., and Tamaya, T. (2000). Vaginal microflora in healthy women with Gardnerella vaginalis. J Infect Chemother 6, 173–177.

Milani, C., Turroni, F., Duranti, S., Lugli, G. A., Mancabelli, L., Ferrario, C., van Sinderen, D., and Ventura. M. (2016). Genomics of the genus *Bifidobacterium* reveals species-specific adaptation to the glycan-rich gut environment. Appl Environ Microbiol 82, 980–991.

Moles, L., Gómez, M., Heilig, H., Bustos, G., Fuentes, S., de Vos, W., Fernández, L., Rodríguez, J. M., and Jiménez, E. (2013). Bacterial diversity in meconium of preterm neonates and evolution of their fecal microbiota during the first month of life. PLoS ONE 8, e66986.

Morrow, A. L., Lagomarcino, A. J., Schibler, K. R., Taft, D. H., Yu, Z., Wang, B., Altaye, M., Wagner, M., Gevers, D., Ward, D. V., Kennedy, M. A., Huttenhower, C., and Newburg, D. S. (2013). Early microbial and metabolomic signatures predict later onset of necrotizing enterocolitis in preterm infants. Microbiome 1, 13.

Nagpal, R., Kurakawa, T., Tsuji, H., Takahashi, T., Kawashima, K., Nagata, S., Nomoto, K., and Yamashiro, Y. (2017). Evolution of gut *Bifidobacterium* population in healthy Japanese infants over the first three years of life: a quantitative assessment. Sci Rep 7, 10097.

Nicholson, J. K., Holmes, E., Kinross, J., Burcelin, R., Gibson, G., Jia, W., and Pettersson, S. (2012). Host-gut microbiota metabolic interactions. Science 336, 1262.

Olszak, T., An, D., Zeissig, S., Vera, M. P., Richter, J., Franke, A., Glickman, J. N., Siebert, R., Baron, R. M., Kasper, D. L., and Blumberg, R. S. (2012). Microbial exposure during early life has persistent effects on natural killer T cell function. Science 336, 489–493.

PAHO/ WHO. (2001). Guiding principles for Complementary Feeding of the Breastfed Child.

Palmer, C., Bik, E. M., DiGiulio, D. B., Relman, D. A., and Brown, P. O. (2007). Development of the human infant intestinal microbiota. PLoS Biol, 5:e177.

Penders, J., Thijs, C., Vink, C., Stelma, F. F., Snijders, B., Kummeling, I., van den Brandt, P. A., and Stobberingh, E. E. (2006). Factors influencing the composition of the intestinal microbiota in early infancy. Pediatrics 118, 511–521.

Petricevic, L., Domig, K. J., Nierscher, F. J., Sandhofer, M. J., Fidesser, M., Krondorfer, I., Husslein, P., Kneifel, W., and Kiss, H. (2014). Characterisation of the vaginal Lactobacillus microbiota associated with preterm delivery. Sci Rep 4, 5136.

Porta, A., Hsia, Y., Doerholt, K., Spyridis, N., Bielicki, J., Menson, E., Tsolia, M., Esposito, S., Wong, I. C. K., and Sharland, M. (2012). Comparing neonatal and paediatric antibiotic prescribing between hospitals: a new algorithm to help international benchmarking. J Antimicrob Chemother 67, 1278–1286.

Ravel, J., Gajer, P., Abdo, Z., Schneider, G. M., Koenig, S. S., McCulle, S. L., Karlebach, S., Gorle, R., Russell, J., Tacket, C. O., Brotman, R. M., Davis, C. C., Ault, K., Peralta, L., Forney, L. J. (2011). Vaginal microbiome of reproductive-age women. Proc Natl Acad Sci USA 108, 4680–4687.

Reichardt, N., Duncan, S. H., Young, P., Belenguer, A., McWilliam Leitch, C., Scott, K. S., Flint, H. J., and Louis, P. (2014). Phylogenetic distribution of three pathways for propionate production within the human gut microbiota. ISME J 8, 1323–1335.

Rinne, M. M., Gueimonde, M., Kalliomäki, M., Hoppu, U., Salminen, S. J., and Isolauri, E. (2005). Similar bifidogenic effects of prebiotic-supplemented partially hydrolyzed infant formula and breastfeeding on infant gut microbiota. FEMS Immunol Med Microbiol 43, 59–65.

Roberts, C. L., Algert, C. S., Rickard, K. L., and Morris, J. M. (2015). Treatment of vaginal candidiasis for the prevention of preterm birth: a systematic review and meta-analysis. Syst Rev 4, 31.

Roediger, W. E. W. (1982). Utilization of nutrients by isolated epithelial cells of the rat colon. Gastroenterol 83, 424-429.

Rogier, E. W., Frantz, A. L., Bruno, M. E. C., Wedlund, L., Cohen, D. A., Stromberg, A. J., and Kaetzel, C. S. (2014). Secretory antibodies in breast milk promote long-term intestinal homeostasis by regulating the gut microbiota and host gene expression. PNAS 111, 3074–3079.

Romero, R., Hassan, S. S., Gajer, P., Tarca, A. L., Fadrosh, D. W., Nikita, L., Galuppi, M., Lamont, R. F., Chaemsaithong, P., Miranda, J., Chaiworapongsa, T., and Ravel, J. (2014a). The composition and stability of the vaginal microbiota of normal pregnant women is different from that of non-pregnant women. Microbiome 2, 4.

Romero, R., Hassan, S. S., Gajer, P., Tarca, A. L., Fadrosh, D. W., Bieda, J., Chaemsaithong, P., Miranda, J., Chaiworapongsa, T., and Ravel, J. (2014b). The vaginal microbiota of pregnant women who subsequently have spontaneous preterm labor and delivery and those with a normal delivery at term. Microbiome 2, 18.

Schluter, J., and Foster, K. R. (2012). The evolution of mutualism in gut microbiota via host epithelial selection. PLoS Biol 10, e1001424.

Schrag, S. J., Farley, M. M., Petit, S., Reingold, A., Weston, E. J., Pondo, T., Hudson Jain, J., and Lynfield, R. (2016). Epidemiology of invasive early-onset neonatal sepsis, 2005 to 2014. Pediatrics 138, e20162013.

Schwiertz, A., Gruhl, B., Löbnitz, M., Michel, P., Radke, M., Blaut, M. (2003). Development of the intestinal bacterial composition in hospitalized preterm infants in comparison with breast-fed, full-term infants. Pediatr Res 54, 393–399.

Sender, R., Fuchs, S., and Milo, R. (2016). Revised Estimates for the Number of Human and Bacteria Cells in the Body. PLoS Biol 14, e1002533.

Singh, N., Thangaraju, M., Prasad, P. D., Martin, P. M., Lambert, N. A., Boettger, T., Offermanns, S., and Ganapathy, V. (2010). Blockade of dendritic cell development by bacterial fermentation products butyrate and propionate through a transporter (Slc5a8)-dependent inhibition of histone deacetylases. J Biol Chem 285, 27601–27608.

Sleator, R. D. (2010). The human superorganism – Of microbes and men. Med Hypotheses 74, 214–215.

Sobel, J. D. (2000). Practice guidelines for the treatment of fungal infections. For the Mycoses Study Group. Infectious Diseases Society of America. Clin Infect Dis 30, 652.

Soeorg, H., Huik, K., Parm, Ü., Ilmoja, M.-L., Metelskaja, N., Metsvaht, T., Lutsar, I. (2013). Genetic relatedness of coagulase-negative staphylococci from gastrointestinal tract and blood of preterm neonates with late-onset sepsis. Pediatr Infect Dis J 32, 389–393.

Srinivasan, S., Hoffman, N. G., Morgan, M. T., Matsen, F. A., Fiedler, T. L., Hall, R. W., Ross, F. J., McCoy, C. O., Bumgarner. R,, Marrazzo, J. M., and Fredricks, D. N. (2012). Bacterial communities in women with bacterial vaginosis: high resolution phylogenetic analyses reveal relationships of microbiota to clinical criteria. PLoS ONE 7, e37818.

Stam, J., van Stuijvenberg, M., Grüber, C., Mosca, F., Arslanoglu, S., Chirico, G., Braegger, C. P., Riedler, J., Boehm, G., and Sauer, P. J. J. (2012). Antibiotic use in infants in the first year of life in five European countries. Acta Pædiatr 101, 929–934.

Stewart, C. J., Marrs, E. C. L., Magorrian, S., Nelson, A., Lanyon, C., Perry, J. D., Embleton, N. D., Cummings, S. P., and Berrington, J. E. (2012). The preterm gut microbiota: changes associated with necrotizing enterocolitis and infection. Acta Paediatr 101, 1121–1127.

Stewart, C. J., Marrs, E. C. L., Nelson, A., Lanyon, C., Perry, J. D., Embleton, N. D., Cummings, S. P., and Berrington, J. E. (2013). Development of the preterm gut microbiome in twins at risk of necrotising enterocolitis and sepsis. PLoS ONE 8, e73465.

Stewart, C. J., Embleton, N. D., Clements, E., Luna P. N., Smith, D. P., Fofanova, T., Nelson, A., Taylor, G., Orr, C. H., Petrosino, J. F., Berrington, J. E., and Cummings, S. P. (2017a). Cesarean or vaginal birth does not impact the longitudinal development of the gut microbiome in a cohort of exclusively preterm infants. Front microbiol 8, 1008.

Stewart, C. J., Embleton, N. D., Marrs, E. C. L., Smith, D. P., Fofanova, T., Nelson, A., Skeath, T., Perry, J. D., Petrosino, J. F., Berrington, J. E., and Cummings, S. P. (2017b). Longitudinal development of the gut microbiome and metabolome in preterm neonates with late onset sepsis and healthy controls. Microbiome 5, 75.

Tanaka, S., Kobayashi, T., Songjinda, P., Tateyama, A., Tsubouchi, M., Kiyohara, C., Shirakawa, T., Sonomoto, K., and Nakayama, J. (2009). Influence of antibiotic exposure in the early postnatal period on the development of intestinal microbiota. FEMS Immunol Med Microbiol 56, 80–87.

The Human Microbiome Project Consortium. (2012). Structure, function and diversity of the healthy human microbiome. Nature 486, 207–214.

Trosvik, P., Stenseth, N. C., and Rudi, K. (2010). Convergent temporal dynamics of the human infant gut microbiota. ISME J 4, 151–158.

Turroni, F., Peano, C., Pass, D. A., Foroni, E., Severgnini, M., Claesson, M. J., Kerr, C., Hourihane, J., Murray, D., Fuligni, F., Gueimonde, M., Margolles, A., De Bellis, G., O'Toole, P. W., van Sinderen, D., Marchesi, J. R., and Ventura, M. (2012). Diversity of bifidobacteria within the infant gut microbiota. PLoS One 7, e36957.

Urbaniak, C., Angelini, M., Gloor, G. B., and Reid, G. (2016). Human milk microbiota profiles in relation to birthing method, gestation and infant gender. Microbiome 4, 1.

Victora, C. G., Bahl, R., Barros, A. J. D., França, G. V. A., Horton, S., Krasevec, J., Murch, S., Sankar, M. J., Walker, N., and Rollins, N. C. (2016). Breastfeeding in the 21st century: epidemiology, mechanisms, and lifelong effect. Lancet 387, 475–490.

Wang, Y., Hoenig, J. D., Malin, K. J., Qamar, S., Petrof, E. O., Sun, J., Antonopoulos, D. A., Chang, E. B., and Claud, E. C. (2009). 16S rRNA gene-based analysis of fecal microbiota from preterm infants with and without necrotizing enterocolitis. ISME J 3, 944–954.

Ward, R. E., Niæonuevo, M., Mills, D. A., Lebrilla, C. B., and German, J. B. (2007). In vitro fermentability of human milk oligosaccharides by several strains of bifidobacteria. Mol Nutr Food Res 51, 1398–1405.

Ward, D. V., Scholz, M., Zolfo, M., Taft, D. H., Schibler, K. R., Tett, A., Segata, N., and Morrow, A. L. (2016). Metagenomic sequencing with strain-level resolution implicates uropathogenic *E. coli* in necrotizing enterocolitis and mortality in preterm infants. Cell Rep 14, 2912–2924.

World Health Organization. (2013). Pocket book of hospital care for children: guidelines for the management of common childhood illnesses, 2nd ed.

World Health Organization. (2015). WHO recommendations for the prevention and treatment of maternal peripartum infections.

Yassour, M., Vatanen, T., Siljander, H., Hämäläinen, A.-M., Härkönen, T., Ryhänen, S. J., Franzosa, E. A., Vlamakis, H., Huttenhower, C., Gevers, D., Lander, E. S., Knip, M., and Xavier R. J. (2016). Natural history of the infant gut microbiome and impact of antibiotic treatment on bacterial strain diversity and stability. Sci Transl Med 343, 343ra81.

Yatsunenko, T., Rey, F. E., Manary, M. J., Trehan, I., Dominguez-Bello, M. G., Contreras, M., Magris, M., Hidalgo, G., Baldassano, R. N., Anokhin, A. P., Heath, A. C., Warner, B., Reeder, J., Kuczynski, J., Caporaso, J. J., Lozupone, C. A., Lauber, C., Clemente, J. C., Knights, D., Knight, R., and Gordon, J. I. (2012). Human gut microbiome viewed across age and geography. Nature 486, 222–227.

Zhou, X., Brown, C. G., Abdo, Z., Davis, C. C., Hansmann, M. A., Joyce, P., Foster, J. A., Forney, L. J. (2007). Difference in the composition of vaginal microbial communities found in healthy caucasian and black women. ISME 1, 121–133.

Zhou, X., Hansmann, M. A., Davis, C. C., Suzuki, H., Brown, C. J., Schütte, U., Pierson, J. D., Forney, L. J. (2010). The vaginal bacterial communities of Japanese women resemble those of women in other racial groups. FEMS Immunol Med Microbiol 58, 169–181.

Zozaya-Hinchliffe, M., Lillis, R., Martin, D. H., Ferris, M. J. (2010). Quantitative PCR assessments of bacterial species in women with and without bacterial vaginosis. J Clin Microbiol 48, 1812–1819.

Zwittink, R. D., van Zoeren-Grobben, D., Martin, R., van Lingen, R. A., Groot Jebbink, L. J. M., Boeren, S., Renes, I. B., van Elburg, R. M., Belzer, C, and Knol, J. (2017). Metaproteomics reveals functional differences in intestinal microbiota development of preterm infants. Mol Cell Proteomics 16, 1610–1620.

Zwittink, R. D., Renes, I. B., van Lingen, R. A., van Zoeren-Grobben, D., Konstanti, P., Norbruis, O. F., Martin, R., Groot Jebbink, L. J. M., Knol, J., and Belzer, C. (2018). Association between duration of intravenous antibiotic administration and early-life microbiota development in late-preterm infants. Eur J Clin Microbiol Infect Dis 37, 475–483.

Acknowledgements

I would like to thank my supervisors, Dr Epp Sepp and professor Madis Metsis for giving me the chance to carry out research in the interesting field of human microbiome studies. I am grateful to both of you for the guidance, support and enormous work you have done with the thesis. I am thankful to Dr Epp Sepp notably for being there for me throughout my PhD studies providing help and advice whenever needed, and to professor Madis Metsis especially for innovative ideas and unique perspective.

The publications used as the basis of the present thesis were a result of a wonderful collaboration with several researchers to whom I am grateful for their contribution. I would like to thank Dr Jaak Simm for all the extensive help with data analysis and biostatistics. Thanks to you I am truly fascinated by the endless possibilities of R Project for Statistical Computing. I would like to thank professor Andres Salumets for excellent guidance and advice. You helped to coordinate and deliver two of the publications. I would also like to thank Dr Lea Tummeleht, professor Irja Lutsar, Dr Jelena Štšepetova, Dr Anu Aaspõllu and Mrs Triin Lillsaar for their significant contribution in making this research possible.

The research projects described in the present thesis were supported and carried out under several private and academic institutions that made this work possible:

- Centre for Biology of Integrated Systems, Tallinn University of Technology;
- Competence Centre on Reproductive Medicine and Biology Ltd. and Competence Centre on Health Technologies Ltd.;
- Department of Microbiology, Institute of Biomedicine and Translational Medicine, University of Tartu.

This work has been partially supported by "TUT Institutional Development Program for 2016-2022" Graduate School in Biomedicine and Biotechnology receiving funding from the European Regional Development Fund under program ASTRA 2014-2020.4.01.16-0032 in Estonia.

I would like to thank all the amazing colleagues with whom I had the possibility to work together. The scientific environment we created in the Centre for Biology and Integrated Systems in Tallinn University of Technology and BiotaP LLC that was systematic, collaborative, and forward-looking was truly unique. I would also like to thank all the amazing colleagues from the Department of Microbiology in University of Tartu for their help, advice, and support.

Last but not least, I would like to thank my family and friends for their tireless support. I would specifically like to thank my partner Indrek for all the love and support.

Abstract

The Effect of Maternal Microbial Communities and Preterm Birth on the Development of Infant Gut Microbiota

The composition of the gut microbiota develops during the first years of an infant's life, with colonization beginning immediately after birth and the development lasting approximately three years. There are several factors that influence the development of the gut microbiota. These factors include mode of delivery, feeding regimen, antibiotic treatment and preterm birth. Although the role of these factors in the development of the gut microbiota has been studied relatively extensively, the results have been contradictory and thus the knowledge about the effect different factors have on this process is still inconclusive. Relatively little is also known about the mother-to-infant transmission of microbes, which has been hypothesized to be an essential source of microbes especially in case of vaginal birth. Studies focusing on mother-to-infant transmission of microbes, especially using large-scale molecular methods, have only lately been carried out and thus the knowledge is scarce.

The aim of the present thesis was to study the composition of vaginal microbiota in reproductive-age women, to characterize the development pattern of the gut microbiota in extremely preterm infants, and to analyze the development of infant gut microbiota by identifying the similarities gut and oral microbiota of full term infants have with different maternal microbial communities (intestinal, vaginal, oral, breast milk, and mammary areola). For this reason, 16S rRNA sequencing was used to profile the vaginal microbiota of 494 reproductive age women, the gut microbiota of altogether 59 infants (9 were full term and 50 preterm), the oral microbiota of nine infants (all full term), and the microbiota of variety of body sites of seven mothers.

When analyzing the composition of vaginal microbiota in women of reproductive age, five different vaginal community types can be detected among women. In addition, approximately a third of the vaginal microbial profiles of women do not cluster in any of the five groups and are characterized by heterogeneous bacterial communities. A heterogeneous vaginal microbiota is characterized by lack of dominants, yet high prevalence of members from genera *Atopobium, Gardnerella,* and *Prevotella*. These women also tend to have high vaginal microbial diversity and high vaginal pH which have been also linked with a diseased state. Nevertheless, during labor and postpartum period, bacteria associated with heterogeneous vaginal community type seem to be overrepresented among women, which may be related to prenatal hormonal changes (namely estrogen) that lead to reduction in the abundance of lactobacilli. Additionally, *Candida spp*., which can be detected in the vaginal microbiota of approximately 70% of women, is more prevalent in this community than previously recognized.

The gut microbiota of extremely preterm infants differs considerably from full term infants especially during the first four weeks of life. Development of the gut microbiota in extremely preterm infants goes through rapid compositional changes (e.g. the dominance of members from family *Staphylococcaceae* is replaced by the dominance of members from *Enterobacteriaceae*) and increase in microbial diversity between first week and first month of life. The community starts to resemble the gut microbiota of full term infants over time, which indicates to a delayed development of the gut microbiota in extremely preterm infants. High abundance of staphylococci in the gut microbiota of extremely preterm infants during the first month of life may be the reason why there is

an increased risk of acquiring coagulase-negative staphylococcal infections in preterm neonates.

The results also showed that the gut microbiota of full term infants has low similarity with maternal community types during the first six months of life. None of the maternal community types stand out as a potential major source of bacteria to the infant gut microbiota as only approximately a third of the bacteria colonizing the infant gut is shared with specific mother's community types (more precisely with mother's gut, oral and mammary areola microbiota). Thus, when considering the possible transfer of microbes from a specific maternal microbial community, only a fraction of microbes has the likelihood of being vertically transferred. Relatively high proportion of bacteria shared with mother's oral and mammary areola microbiota may be related to the fact that majority of the participating infants were born via C-section and the gut microbiota of C-section infants tend to be enriched with bacteria typically know to be of skin and mouth origin. Infant oral microbiota, on the other hand, has high similarity with several of the mother's community types (breast milk, mammary areola, and oral microbiota) and over half of the bacteria colonizing infant oral microbiota are shared with mother's oral microbiota (51%), which indicates that constant contact between microbial communities, in this case probably resulting from the maternal habits of infant care, increases their similarity.

Lühikokkuvõte

Ema mikroobikoosluste ja enneaegse sünni mõju imiku seedetrakti mikrobioota kujunemisele

Seedetrakti mikrobioota kujunemine algab vahetult pärast lapse sündi ja kestab umbes kolm esimest eluaastat. Seda protsessi mõjutavad mitmed erinevad tegurid, millest olulisimad on sünniviis, dieet, antibakteriaalne ravi ja enneaegsus. Kuigi nende tegurite mõju seedetrakti mikrobioota kujunemisele on uuritud suhteliselt palju, on tulemused olnud vasturääkivad ja seetõttu on arusaam nende tegurite rollist seedetrakti mikrobioota kujunemisprotsessis endiselt puudulik. Lisaks on senimaani ainult vähesel määral uuritud mikroobide ülekandumist emalt lapsele, mida peetakse üheks peamiseks mikroobiallikaks seedetrakti mikrobioota kujunemisprotsessis, eriti vaginaalse sünniviisi puhul. Selle protsessi uurimisele on hakatud rohkem keskenduma alles viimasel ajal ja see kehtib eriti uuringute kohta, kus on kasutatud molekulaarseid metoodikaid, mis põhinevad nukleiinhappe analüüsidel. Sellest tulenevalt saab teadmisi mikroobide ülekande kohta praegu veel pidada ebapiisavaks.

Käesoleva dissertatsiooni eesmärgiks oli uurida reproduktiiveas naiste vaginaalset mikrobioobikooslust, analüüsida seedetrakti mikrobioota kujunemist sügavalt enneaegsetel imikutel ja hinnata ajaliste imikute soole ning suu mikrobioota sarnasust ema erinevate kehapiirkondade mikrobiootaga (suu, sool, rinnapiim, nahk rinnanibu ümber, tupp). Kokku uuriti vaginaalse mikrobioota kooslust 494-l reproduktiiveas naisel, seedetrakti mikrobioota kooslust 59 imikul (neist üheksa olid ajalised ja 50 sügavalt enneaegsed), suu mikrobiootat üheksal ajalisel imikul ja erinevate kehapiirkondade mikrobioota kooslust seitsmel emal. Uuringu läbiviimisel kasutati 16S rRNA sekveneerimise meetodit.

Analüüsides reproduktiiveas naiste tupe mikrobiootat, selgus, et reproduktiiveas naiste vaginaalne mikrobioota jaguneb valdavalt viieks erinevaks koosluse tüübiks. Sellele lisaks esineb umbes kolmandikul naistest vaginaalne mikrobioota, mida iseloomustab heterogeenne kooslus ja mis ei kuulu mitte ühtegi viiest eelpool mainitud koosluse tüübist. Heterogeenset vaginaalset mikrobiootat iseloomustab dominantsete liikide puudumine ja perekondade *Atopobium, Gardnerella* ja *Prevotella* suur esinemissagedus. Heterogeenne vaginaalne mikrobioota on sageli ka suure bakteriaalse mitmekesisusega ja nendel naistel on sageli kõrge vaginaalne pH, mida on varasemalt seostatud ka vaginaalse keskkonna haigusliku seisundiga. Siiski on varasemad uuringud näidanud, et heterogeense vaginaalse mikrobioota esinemissagedus suureneb sünnituse ajal ja vahetult pärast sünnitust. See võib olla seotud prenataalse perioodi hormonaalsete muutustega (peamiselt östrogeeni taseme muutustega), mis põhjustab laktobatsillide arvukuse languse ning sellest tulenevalt heterogeense koosluse väljakujunemise. Lisaks vaginaalse mikrobioota bakteriaalse komponendi kirjeldamisele tuvastasime 70%-l naistest vaginaalse mikrobioota koosluses ka seene perekonna *Candida spp*. esinemise, mida esineb meie uuringu põhjal suuremal hulgal tervetel naistel kui varasemalt arvatud.

Sügavalt enneaegsete imikute seedetrakti mikrobioota erineb esimesel neljal elunädalal märkimisväärselt ajaliste imikute seedetrakti mikrobiootast ning sellel perioodil toimuvad sügavalt enneaegsete imikute seedetrakti mikrobioota kujunemise raames järsud koosluse muutused (nt stafülokokkide domineerimine asendub sugukonna *Enterobacteriaceae* esindajate domineerimisega), millega kaasneb ka bakteriaalse mitmekesisuse suurenemine. Muudatuste tulemusena hakkab kooslus

järjest enam sarnanema ajaliste laste seedetrakti mikrobiootaga, mis viitab, et sügavalt enneaegsete imikute seedetrakti mikrobioota kujunemine toimub võrreldes ajaliste imikutega tõenäoliselt hilinemisega. Stafülokokkide suur arvukus elu esimestel nädalatel võib olla põhjuseks, miks sügavalt enneaegsete imikute puhul on suurenenud risk haigestuda koagulaas-negatiivsete stafülokokkide põhjustatud infektsioonidesse.

Meie uuringu tulemused näitasid samuti, et ajaliste imikute seedetrakti mikrobioota ei sarnane esimesel elupoolaastal ema erinevate kehapiirkondade mikroobikooslustega. Ema seedetrakti, vaginaalse, rinnapiima, rinnanibu ümbritseva naha ja suu mikroobikooslused ei tõuse esile kui potentsiaalsed mikroobide allikad imku seedetrakti mikrobioota kujunemise protsessis ja ainult umbes kolmandik imiku seedetrakti koloniseerivatest bakteritest on tuvastatavad ka ema seedetrakti, suu või rinnanibu ümbritseva naha mikrobioota kooslustes (teiste ema koosluste puhul veelgi vähem). See viitab sellele, et ainult väike osa imikute seedetraktis esinevatest mikroobidest on saanud emalt vertikaalselt üle kanduda. Suhteliselt suur bakterite osakaal, mis on jagatud imiku seedetrakti ja ema suu ning rinnanibu ümbritseva naha mikrobioota vahel, võib olla seotud nende imikute sünniviisiga (enamik neist sündis keisrilõikega). Varasemad uuringud on näidanud, et keisrilõikega sündinud imikute seedetrakti mikrobiootas on arvukalt baktereid, mis kuuluvad suu ja naha mikroobikooslusesse. Erinevalt seedetrakti mikrobiootast on imiku suu mikrobioota esimesel kuuel elukuul väga sarnane ema rinnapiima, rinnanibu ümbritseva naha ja suu mikrobiootaga Üle poole imiku suud koloniseerivatest bakteritest on jagatud ema suu mikrobioota kooslusega. See näitab suure tõenäosusega, et pidev kokkupuude (antud juhul imiku eest hoolitsemise aegne kontakt) kahe erineva koosluse vahel suurendab nende koosluste sarnasust.

Appendix

Publication I

Drell, T., Lillsaar, T., Tummeleht, L., Simm, J., Aaspõllu, A., Väin, E., Saarma, I., Salumets, A., Donders, G. G. G., and Metsis, M. (2013). Characterization of the vaginal micro- and mycobiome in asymptomatic reproductive-age Estonian women. PLoS One 8, e54379.

OPEN ACCESS Freely available online

PLOS ONE

Characterization of the Vaginal Micro- and Mycobiome in Asymptomatic Reproductive-Age Estonian Women

Tiina Drell[1,2]*, Triin Lillsaar[1,2], Lea Tummeleht[1¤], Jaak Simm[1], Anu Aaspõllu[1], Edda Väin[3], Ivo Saarma[3], Andres Salumets[2,4,5], Gilbert G. G. Donders[6,7], Madis Metsis[1,2]

1 Centre for Biology of Integrated Systems, Tallinn University of Technology, Tallinn, Estonia, **2** Competence Centre on Reproductive Medicine and Biology, Tartu, Estonia, **3** Fertilitas Hospital, Haabneeme, Estonia, **4** Department of Obstetrics and Gynaecology, University of Tartu, Tartu, Estonia, **5** Institute of Biomedicine, University of Tartu, Tartu, Estonia, **6** Department Obstetrics and Gynecology, Regional Hospital H Hart, Tienen, Belgium, **7** Femicare vzw, Clinical Research for Women, Tienen, Belgium

Abstract

The application of high-throughput sequencing methods has raised doubt in the concept of the uniform healthy vaginal microbiota consisting predominantly of lactobacilli by revealing the existence of more variable bacterial community composition. As this needs to be analyzed more extensively and there is little straightforward data regarding the vaginal mycobiome of asymptomatic women we aimed to define bacterial and fungal communities in vaginal samples from 494 asymptomatic, reproductive-age Estonian women. The composition of the vaginal microbiota was determined by amplifying bacterial 16S rRNA and fungal internal transcribed spacer-1 (ITS-1) regions and subsequently sequencing them using 454 Life Sciences pyrosequencing. We delineated five major bacterial community groups with distinctive diversity and species composition. Lactobacilli were among the most abundant bacteria in all groups, but also members of genus *Gardnerella* had high relative abundance in some of the groups. Microbial diversity increased with higher vaginal pH values, and was also higher when a malodorous discharge was present, indicating that some of the women who consider themselves healthy may potentially have asymptomatic bacterial vaginosis (BV). Our study is the first of its kind to analyze the mycobiome that colonizes the healthy vaginal environment using barcoded pyrosequencing technology. We observed 196 fungal operational taxonomic units (OTUs), including 16 OTUs of *Candida spp.*, which is more diverse than previously recognized. However, assessing true fungal diversity was complicated because of the problems regarding the possible airborne contamination and bioinformatics used for identification of fungal taxons as significant proportion of fungal sequences were assigned to unspecified OTUs.

Citation: Drell T, Lillsaar T, Tummeleht L, Simm J, Aaspõllu A, et al. (2013) Characterization of the Vaginal Micro- and Mycobiome in Asymptomatic Reproductive-Age Estonian Women. PLoS ONE 8(1): e54379. doi:10.1371/journal.pone.0054379

Editor: David N. Fredricks, Fred Hutchinson Cancer Center, United States of America

Received August 27, 2012; **Accepted** December 11, 2012; **Published** January 23, 2013

Funding: This research was funded by BiotaP LCC Laboratories, Tallinn University of Technology PhD program and Estonian Ministry of Education and Research (SF0180044s09) and Enterprise Estonia (EU30200). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The study received funding from a commercial source, BiotaP LCC. The company partly financed and provided sequencing service for the project as part of their workflow development for metagenomic sequencing service. This does not alter the authors' adherence to all the PLOS ONE policies on sharing data and materials.

* E-mail: tiinadrell@gmail.com

¤ Current address: Department of Infectious Diseases, Estonian University of Life Sciences, Tartu, Estonia

Introduction

Similar to most cavities and surfaces on the human body, the health and functioning of the vagina are closely linked to its microbial inhabitants. These microbes maintain adequate pH, help to prevent the acquisition of pathogens, stimulate the local innate immune system, and decrease symptoms and complications during pregnancies [1–3]. According to the ecological theory applied to all natural habitats, including microbial econiches within the human body, communities diverge from each other by two critical characteristics: taxon composition and physiological potential [4]. When these characteristics are altered, the functioning of the vaginal environment becomes disturbed and disease may emerge. The biggest enigma in recent years is determining how to define a normal state: how much variation within the prevailing vaginal microbial communities can still be considered within the normal boundaries. Originally, healthy vaginal microbiota was considered to be relatively non-diverse and consisted mainly of the *Lactobacillus* species, which produce lactic acid to maintain low vaginal pH and bacteriocins to kill potential pathogenic invaders [5]. Recent studies using cultivation-independent molecular methods, however, have shown that the composition of healthy vaginal microbiota may vary to a greater extent than was initially thought [6–8]. In particular, a significant proportion of asymptomatic, reproductive-age women are colonized by glucose-consuming species other than lactobacilli, such as *Megasphaera sp.*, *Atopobium vaginae*, *Gardnerella vaginalis*, *Prevotella sp.*, or *Streptococcus sp.*, even though these species have been linked to abnormal vaginal microbiota that are present in certain pathologic conditions, such as BV. Moreover, because BV affects up to 30% of all reproductive-age women in Western countries [9], any unusual shift in microbial communities should be addressed through further analysis. This also applies to some of the opportunistic fungal pathogens, such as *Candida albicans*, which can colonize approximately 20% of asymptomatic women [10]. Furthermore, even a slight modification in host defense and the composition of

resident bacterial community can encourage the emergence of opportunistic infections caused by *C. albicans* [11]. Still, there is surprisingly little straightforward data regarding the presence of the *Candida* species and other fungi in the vaginal environment of asymptomatic women as well as the interactions and possible coexistence patterns with bacterial communities. The application of advanced molecular technologies and computational biology promises to increase our understanding regarding both the normal vaginal ecosystem as well as the different microbial diseases, especially those considered to be polymicrobial in origin [12].

In this study, using parallel pyrosequencing technology followed by taxonomic identification, we were able to define in detail the bacterial and fungal components of microbial communities present in the vaginal samples of asymptomatic reproductive-age women. In addition, we estimated the variation of normal vaginal microbial communities and linked those variations to medical history or lifestyle habits of the women.

Materials and Methods

Participants and sample collection

For this study, 494 healthy pre-menopausal Caucasian women, age 15–44 y (mean[±SD] 31.1 [±6.4] y), having a routine gynecological check-up visit at the Fertilitas private hospital, Estonia, from June to October 2010 were enrolled. Informed written consent was obtained from all participants prior to enrollment. In case of under-aged participants, verbal consent was obtained from one of their guardians by the doctor carrying out the sample collection and recruitment. We confined to verbal consent of the guardians because we expected the under-aged group of people to be old and aware enough to be sufficiently competent to make a decision about this study and thus no extra measures to document the approval of the guardians was taken. The study design (including the participation of under-aged women based on aforementioned terms) was approved by the Ethics Committee of Medical Research in Tallinn, Estonia. Exclusion criteria included pregnancy, menstruation, or any signs or symptoms of urogenital disorders. The presence of urogenital infections was excluded by inspection and *in speculo* examination before a sample was taken from the vaginal fornix and cervix using Rovers® Cervix-Brush® Combi (Rovers Medical Devices B.V. Oss, The Netherlands). Amsel criteria or Gram staining were not specifically applied for the exclusion of BV. The vaginal pH was measured during sampling with pH indicator strips (range 2.0–9.0; 0.5 unit increments; Merck & Co, Inc., USA). Collected samples were then stored at −20°C until batch analysis was performed.

Participants were asked to complete a questionnaire on their medical history and lifestyle habits, which specifically addressed their sexual behavior and gynecologic hygiene. The opinion of a gynecologist regarding physical appearance of the vaginal environment (e.g. redness, characterization of vaginal discharge etc.) was included with the questionnaire.

DNA extraction

DNA was extracted from samples using a modified BioSprint 96 DNA Blood Kit (Qiagen, USA) protocol and Kingfisher Flex Magnetic Particle Processor (Thermo Scientific, Bremen, Germany). The BioSprint 96 DNA Blood Kit protocol was used with a pre-initial homogenization step using ceramic beads (Ø 0.4–0.6 mm, 0.5 g per sample; Saint-Gobain ZIRPRO, Le Pontet Cedex, France) and TissueLyser II (6 min at 30 Hz; Qiagen). In addition to samples collected from patients, negative controls were included to every separate extraction batch of 95 samples. Extracted DNA samples were stored at −20°C for further analysis.

PCR amplification

Before pyrosequencing, the amplification of the desired sequence was carried out with two sequential PCR reactions. In the first PCR reaction, the V1-V2 hypervariable region of 16S rRNA genes (16S rDNA) was amplified with 8F and 357R broad range primers to assess the bacterial component of the vaginal microbiota [13]. The complete sequences of the primers were as follows: 8F-5′ TTGGCAGTCTCAGNNNNNNNN**AGTTTGATCCTGGCTCAG** 3′ and 357R-5′ GTCTCCGACTCAGNNNNNNNN**CTGCTGCCTYCCGTA** 3′. For assessment of the fungal component of the vaginal microbiota, the Internal transcribed spacer-1 (ITS-1) region of eukaryotic ribosomal DNA was amplified using ITS1F and ITS2 primers [14]. The complete sequence for the forward primer (ITS1F) was 5′ GTCTCCGACTCAGNNNNNNNN**CTTGGTCATTTAGAGGAAGTAA** 3′ and reverse primer (ITS2) was 5′ TTGGCAGTCTCAGNNNNNNNN**GCTGCGTTCTTCATCGATGC** 3′. Underlined letters indicate partial 454-specific sequencing adapters, bold letters denote amplicon-specific primer sequences (8F/ITS2 or 357R/ITS1F), and the 8 bp barcode is marked as 8 Ns, which refers to a unique sequence tag to barcode each sample [15]. Cycling parameters were 15 min at 95°C, followed by 5 cycles of 30 s at 95°C, 30 s at 50°C, and 60 s at 72°C, and then 35 cycles of 30 s at 95°C, 30 s at 65°C, and 60 s at 72°C, with a final extension at 72°C for 10 min.

To extend the partial 454-specific adapter sequences at both ends of each amplicon to full-length sequences, a second PCR was performed with full sequencing adapters (Primers A 5′-CCATCTCATCCCTGCGTGTCTCCGACTCAG-3′ and B 5′-CCTATCCCCTGTGTGCCTTGGCAGTCTCAG-3′) using the amplicon derived from the previous PCR step of the 16S rRNA gene or ITS-1 region amplification diluted 50-fold. Cycling parameters for the second PCR were 15 min at 95°C, followed by 5 cycles of 30 s at 95°C, 30 s at 62°C, and 60 s at 72°C, and then 20 cycles of 30 s at 95°C and 60 s at 72°C, with a final extension at 72°C for 10 min.

All PCR reactions were carried out in total reaction volume of 10 µL consisting of 5 µL Maxima Hot Start PCR Master Mix (Fermentas, Germany), 1 µL DNA template or water in case of negative controls added to every PCR reaction, and 0.2 µM each primer, with water added to reach the final volume. Reactions were performed on a Thermal cycler 2720 (Applied Biosystems, California, US).

454 pyrosequencing

Barcoded amplicons with full adapter sequences were pooled together by extraction of amplicons from a 1.5% agarose gel using the QIAquick Gel Extraction kit (Qiagen) and sequenced with the GS Junior Sequencing System (454 Life Sciences, Roche, Germany) according to the manufacturer's protocol.

Analysis of bacterial sequences

Pyrosequencing noise was removed from the initial pre-trimmed dataset of the 16S rRNA gene V1-V2 region (based on ".qual" file output of 454 sequencing system), and only sequences longer than 170 bp were included for further processing. Reference sequences of aligned 16S rDNA were obtained from the SILVA ribosomal RNA database [16]. Operational taxonomic units (OTUs) with a 97% identity threshold were defined by the average neighbor hierarchical clustering algorithm using mothur 1.24.1 software [17]. To filter chimeric sequences, UChime was used by applying the "chimera.uchime" procedure from the mothur software in *de novo* mode, which first splits sequences into groups and then checks each sequence within a group using the more abundant groups as

reference. For additional denoising, OTUs that had less than 5 sequences were removed. Taxonomic assignments were performed against the SILVA bacterial database using Ñaive Bayesian classifier with a confidence cutoff of 75% [18]. To represent sequences of the 10 most relatively abundant OTUs, an additional taxonomic assignment was carried out using BLASTN against the NCBI nt database (updated in May, 2012). Different OTUs assigned to the same species were numbered starting from the most relatively abundant OTU. For clustering of similar samples, we used a similarity tree based on hierarchical Yue-Clayton theta value [19] with a group cutoff of 0.30. Only clusters $\geq$15 members were included.

Analysis of fungal sequences

Pyrosequencing noise was removed from the initial pre-trimmed dataset of ITS1 region (based on ".qual" file output of 454 sequencing system) and only sequences longer than 170 bp were included for further processing. The fungal sequences were first clustered with cd-hit-est software [20] with 97% cluster identity. Cd-hit-est chooses one representative sequence for each OTU (cluster), which can then be used for further analysis. In this current study, the cd-hit-est results were compared against the UNITE fungal database [21] using a BLAST search with 97% identity and an e-value below 10^{-20}. In case of more than one match, the one that had the lowest e-value was kept for further analysis. OTUs that did not match the UNITE database were discarded. Similarly to the bacterial analysis, for additional denoising, only OTUs with more than 5 sequences were kept for further analysis. In addition, samples with less than 75 sequences were excluded from the dataset. The taxonomy for fungal OTUs was taken from the UNITE database.

Statistical analysis

Statistics was performed using R 2.13.1 software and its packages: gplots [22], ggplot2 [23] and vegan [24]. The level of a significant difference was set at 5%. To describe general bacterial diversity, the Shannon diversity index was calculated. The Shannon index is a mathematical measure of species diversity in a community that accounts for both abundance and evenness of the species present. Taxonomic richness was also expressed through the number of observed OTUs. For reviewing the proportions of different bacteria, the relative abundances of OTUs identified within the same genus were summed and referred to as a summarized genus.

To analyze the correlation of medical history or lifestyle habits of the participants with Shannon diversity index values or summarized abundance of *Candida spp.*, a linear regression analysis was used. Before applying linear regression models, a correlation matrix was generated for all of the factors that were included into the models. If the Pearson's correlation coefficient was higher than 0.7 between the covariates, then one of them was removed. Both linear regression models included factors of gynecological diseases experienced recurrently and within the last 6 months prior to the sampling [BV, vulvovaginal candidosis (VVC), myco- and ureaplasmosis and no gynecological disease within the last 6 months], contraceptives (not used, condom, vaginal ring, contraceptive patch, hormonal pills, copper-containing intrauterine device, or levonorgestrel-releasing device), age of the participants, time from the last intercourse, characteristics of vaginal discharge (physiological, mucopurulent, homogeneous, white color or malodorous), and redness of the vaginal canal. In addition, models testing the correlation with the Shannon diversity index included factors of sexually transmitted diseases (STDs) experienced within the last 6 months prior to the sampling (gonorrhea, chlamydia, genital herpes, or no STDs) and characteristics of cervix uteri (conical or cylindrical shape, deformed, ovula Nabothi or with redness). The model testing the correlation with the summarized abundance of *Candida spp.* included only vaginal pH as an additional factor.

Linear regression analysis was also used for analyzing the correlation between the clustering of vaginal bacterial communities and vaginal pH, Shannon diversity index, and the number of observed bacterial OTUs. The prevalence of *C. albicans* among analyzed samples in relation to the relative abundance of the genera *Lactobacillus* and *Gardnerella* (comprised of the summarized relative abundance of OTUs belonging to aforementioned genera) was tested by a logistic regression model. All regression analyses were carried out with Holm-Bonferroni correction.

Results

Medical history and lifestyle habits of the participants

Among the 494 recruited asymptomatic Estonian women, the vaginal pH value varied in the range of 3–9, it was greater than 7 in 11 women. The mean value of vaginal pH was 4.7 ($\pm$0.8). Of the women assessed, 28.5% (n = 141) had experienced VVC and 10.5% (n = 52) had experienced BV repeatedly during their adult lives. During the last 6 months prior to the sampling, VVC was detected in 13.6% (n = 67) of the patients and BV in 5.7% (n = 28) patients respectively. Data regarding aerobic vaginitis are lacking, as this was not recognized as a separate entity by our clinicians in the current study.

Because there was no restriction on intercourse prior to the sampling, the last intercourse event was recorded as follows: 60 individuals (12.2%) had the last intercourse within 24 h prior to the sampling, 108 (21.9%) had sexual intercourse 24–48 h before sampling, 148 (30%) during the previous 7 d and 167 (33.8%) more than a week prior to the sampling. Twelve participants did not provide this information. A majority of participants (92.5%) reported themselves as heterosexual, while 13 (2.6%) reported themselves as homosexual and 25 participants (5.1%) did not answer this question.

Vaginal bacterial communities

In total, we obtained 828,551 16S rRNA gene sequences with an average sequence read length of 300 bp. The number of sequences was higher than 400 in 432 samples, which were included for further analysis. Four hundred sequences per sample was the level whereby the rarefaction curves for 95% of the samples reached a 5% plateau. This means that increasing the number of sequences per sample did not increase the number of OTUs obtained in 95% of the samples (Fig. 1). After exclusion of these samples, the mean number of sequences per sample was 1330 ($\pm$818).

For further analysis, we only included OTUs (n = 208; FASTA file S1) that were present in at least one sample with a relative abundance of $\geq$1%. The additional cut-off was applied to normalize the filtering across samples with different sequences assigned to them and also to filter out artifacts generated by the limitations of laboratory workflow. The mean taxonomic richness (number of OTUs) per sample was 26.9 ($\pm$24). A comparison of these OTUs against the SILVA and NCBI databases showed that three of the OTUs were unclassified bacteria with an average summarized relative abundance of 0.01%. Lactobacilli accounted for the greatest proportion of the recovered OTUs (69.4%), followed by OTUs belonging to genera of summarized *Gardnerella* (11.2%), *Prevotella* (3.8%), *Atopobium* (2.2%), *Streptococcus* (1.5%),

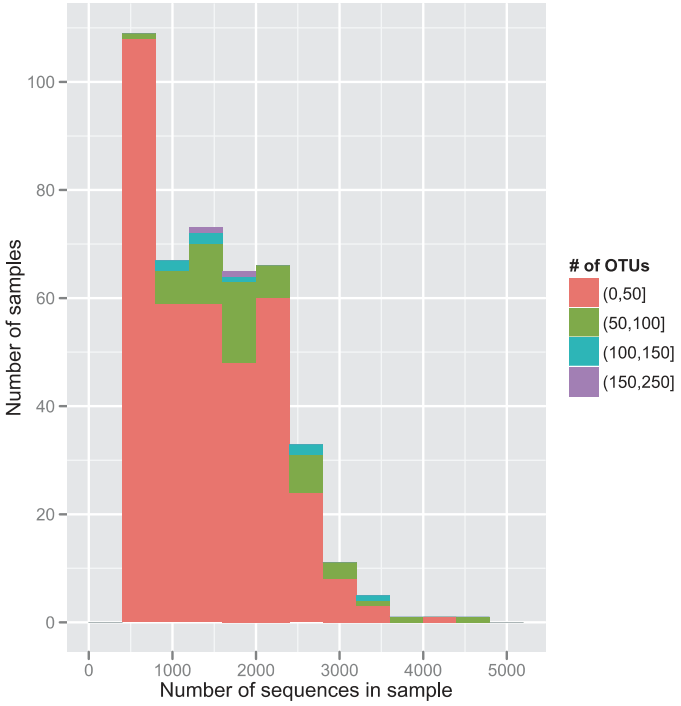

Figure 1. Distribution of 432 samples according to number of sequences and number of OTUs. The cutoff value is set to 400 sequences. Round and square brackets in the figure indicate to the exclusion and inclusion of adjacent value, respectively.
doi:10.1371/journal.pone.0054379.g001

Ureaplasma (0.9%), *Escherichia coli* (0.5%), *Mycoplasma* (0.2%) and *Staphylococcus* (0.1%).

Lactobacilli were recovered from 98.8% of the samples, members of *Gardnerella* from 70.6%, *Prevotella* from 55.8%, *Ureaplasma* from 41%, *Atopobium* from 38%, *Streptococcus* from 30.6%, *Staphylococcus* from 13.9%, *Escherichia coli* from 6.5% and *Mycoplasma* from 4.6% of the samples. The relative abundance and prevalence of the most abundant and meaningful OTUs, which predominantly belonged to the aforementioned genera, is shown in Figure 2. Special emphasis should be laid on the OTU-s with identical names but different numbers. By visually analyzing the alignment of representative sequences of these OTU-s and ignoring the potential pyrosequencing errors, considerable similarities can be noted in all, except for *Prevotella* 1 and 2 (See Fig. S1).

Bacterial community parameters in relation to the medical history and lifestyle habits

There was no correlation between the number of bacterial sequences of 16S rRNA gene retrieved per sample from the trimmed dataset and the Shannon diversity index of bacterial communities [correlation coefficient (CC) = −0.05], nor was there a correlation between the number of sequences per sample and the number of observed OTUs (CC = 0.36). These results confirmed that taxonomic richness was not created by the number of sequences by itself.

Linear regression analysis showed that the Shannon diversity index of vaginal bacterial communities increased with an increase in vaginal pH ($P = 0.003$) and with the presence of malodorous vaginal discharge ($P<0.0001$) and also that the Shannon diversity index was not influenced by other factors included to the analysis ($P>0.5$).

Clustered bacterial communities

Three hundred and seven (71.1%) vaginal bacterial communities clustered into 5 major groups based on the similarities in relative abundance values of bacterial OTUs (OTU composition) (Fig. 3). Each group harbored ≥15 members. The remaining vaginal communities (N = 125; 28.9%) could not be classified in any of the 5 groups and were rather erratic, non-related, heterogeneous communities that did not cluster into any other entity (hereafter referred to as the non-classifiable or 0-group). OTUs belonging to members of lactobacilli dominated four out of five groups. In two of the groups (I and II), the most abundant OTU was *Lactobacillus iners*_1, and in two other groups (III and V) *L. crispatus_1* was the most abundant. In group IV, lactobacilli did not have the most abundant OTUs, while in this group, the predominant bacterium was *Gardnerella*_1 (Table 1; Fig. 4).

Among the non-classifiable group, there were no dominating bacteria that had relative abundance over 20%, but the prevalence of the following OTUs was relatively high: *Gardnerella_1* (74.1%) and members belonging to genera *Prevotella* (69.2%), *Atopobium vaginae*_1 (52.4%), *Ureaplasma sp.* (33.6%), *Sneathia sanguinegens*

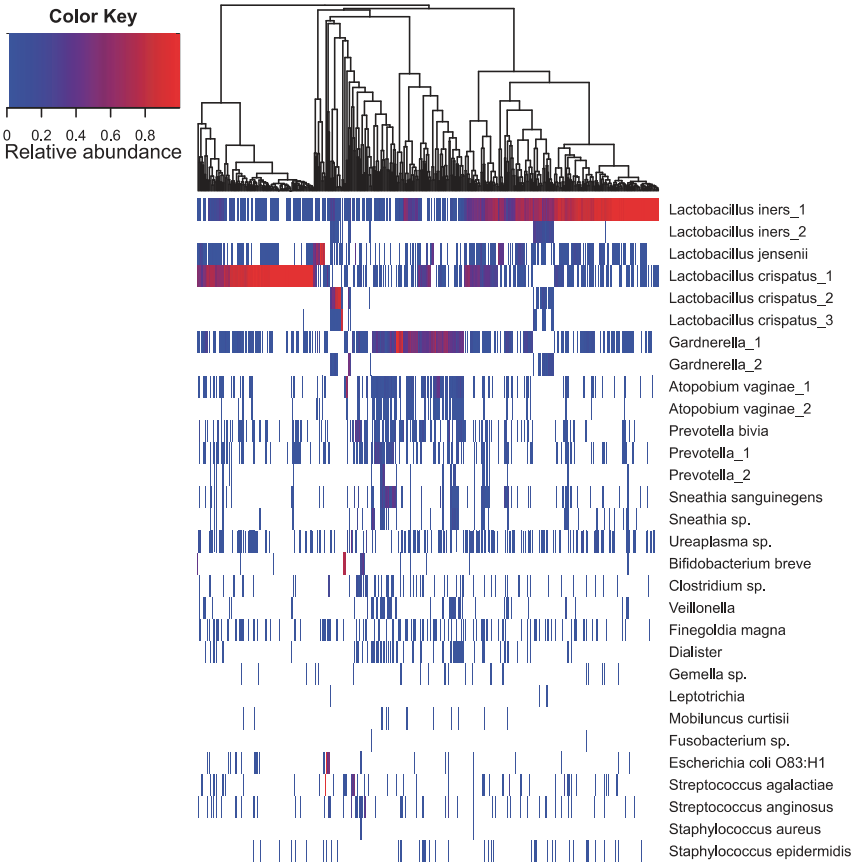

Figure 2. Relative abundance of bacterial OTUs found in the vaginal communities of healthy Estonian women (n = 432). Bacterial taxonomic assignments are indicated on the right of the heatmap at the Genus and Species level. The relative abundance is color coded and indicated by the color key on the left top of the map. The tree on the top of the heatmap characterizes the similarity of analyzed samples based on OTU composition of vaginal microbiota.
doi:10.1371/journal.pone.0054379.g002

(30.1%) and *Clostridium sp.* (16.8%). A comparison of the community groups (I–V) with 0-group in linear regression analysis revealed significant differences in the mean bacterial diversity expressed by the Shannon diversity index, in the mean number of observed OTUs, and in the mean values of vaginal pH (Table 1).

Vaginal fungal communities

From 494 collected samples, we analyzed fungal communities in 251 samples. In total, we obtained 276,336 ITS-1 region specific sequences and the average sequence read length was 220 bp. The count of sequences was higher than 75 in 216 samples, which were included for further analysis. Seventy-five sequences per sample was the level whereby the rarefaction curves for 95% of the samples reached a 5% plateau. After this exclusion of the samples, the mean number of sequences per sample was 1076.7 (±1282.2). Only OTUs (n = 196; FASTA file S2) that were present in at least one sample with a relative abundance ≥1% were included for further analysis. The additional cut-off was applied to normalize the filtering across samples with different sequences assigned to them and also to filter out artifacts generated by the limitations of laboratory workflow. The mean taxonomic richness per sample was 7.8 (±4). A comparison of these OTUs against the UNITE database resulted in 161 unique hits at the species taxonomic level.

The mean relative abundance of *Candida spp.* was 36.9%. We were able to detect 16 different OTUs belonging to genus *Candida*: *C. parapsilosis* (n = 1), *C. dubliniensis* (n = 1), *Candida sp.* VI04616 (n = 1), *Pichia kudriavzevii* (*Candida krusei*) (n = 3) and *C. albicans* (n = 10). The relative abundance of summarized *C. albicans* was 34.1%, *P. kudriavzevii* 2.3%, *C. parapsilosis* 0.3%, *Candida sp.* VI04616 0.3% and *C. dubliniensis* 0.04%. OTUs belonging to genus *Candida* were detected in 69.9% of samples. The prevalence of *C. albicans* was 67.6%, *P. kudriavzevii* 9.7%, *C. dubliniensis* 0.5% and both *C. parapsilosis* and *Candida sp.* VI04616 1.9% of 216 samples analyzed. The prevalence and summarized relative abundance of all predominant fungal OTUs are shown in Table 2. A considerable number of these OTUs were unspecified

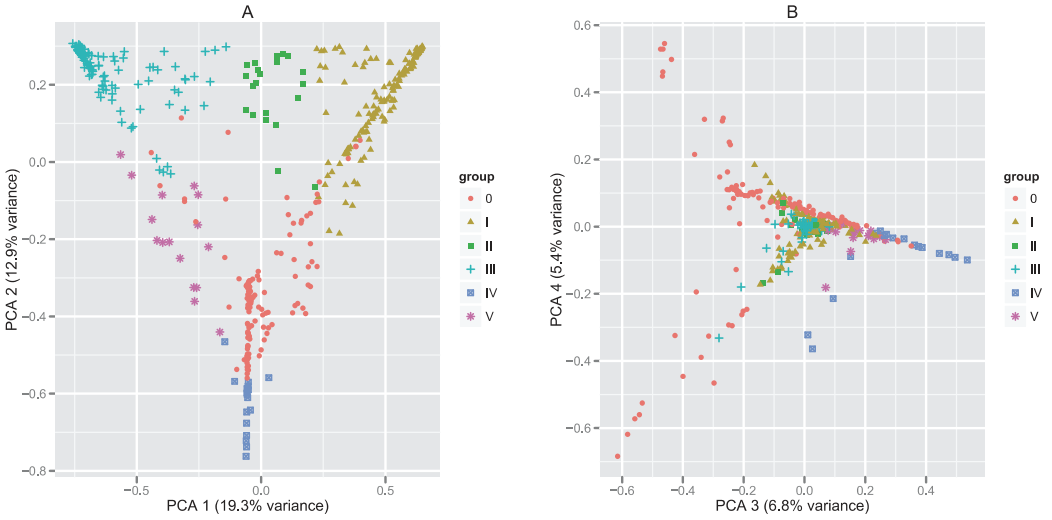

Figure 3. Two-dimensional (2D) plots describing the principal component analysis (PCA) of OTU composition among analyzed samples. The plots represent 2D projections of a multidimensional analysis where the relative abundance of each specific OTU defines a dimension. Both plots are the projection of the same analysis viewed at a different angle. The plots visualize the clustering and variability of studied vaginal bacterial communities. First (a) 2D plot of the first two PCA components describes the clustering of groups I–V. Second (b) 2D plot of third and fourth PCA component confirms that the samples belonging to non-classifiable group (0) are not clustering into separate or any other entity. The variance described by the respective PCA components (Axis-1 and Axis-2) is written in brackets.
doi:10.1371/journal.pone.0054379.g003

Table 1. Mean (±SD) values of selected parameters of vaginal environment within bacterial community groups.

Community groups	*Group size*	*Vaginal pH*	*Shannon diversity index*	*Number of observed OTUs*	*Dominating OTUs in bacterial community (relative abundance/prevalence)*
I	**145**	4.58(±0.9)***	0.8(±0.6)***	26.1(±27.4)***	*Lactobacillus iners_1* (79.1%/100%)
					Gardnerella_ 1 (3.4%/52.4%)
					Lactobacillus crispatus_1 (3.0%/59.3%)
II	**21**	4.54(±0.6)*	1.4(±0.5)**	34.2(±24.7)	*Lactobacillus iners_1* (47.7%/100%)
					Lactobacillus crispatus_1 (34.4%/100%)
					Lactobacillus jensenii (4.0%/81.0%)
III	**108**	4.19(±0.5)***	0.7(±0.4)***	20.4(±14.8)***	*Lactobacillus crispatus_1* (83.7%/100%)
					Lactobacillus iners_1 (5.6%/80.6%)
IV	**17**	4.88(±0.6)	1.1(±0.4)***	24.9(±11.5)*	*Gardnerella_1* (68.5%/100%)
					Lactobacillus jensenii (6.7%/35.3%)
					Lactobacillus gasseri (5.7%/47.1%)
V	**16**	4.64(±0.7)	1.4(±0.4)*	20.6(±11.1)**	*Lactobacillus crispatus_1* (43.1%/100%)
					Gardnerella_1 (35.1%/100%)
					Lactobacillus iners_1 (3.2%/56.3%)
0	**125**	5.02(±0.9)	1.8(±0.6)	39.4(±27.8)	

Dominating OTUs are listed as top 3 with the mean relative abundance at least 3%. Linear regression models were used to test the differences between the groups (I–V) and 0-group:
***P<0.001;
**P = 0.01–0.001;
*P = 0.01–0.05.
doi:10.1371/journal.pone.0054379.t001

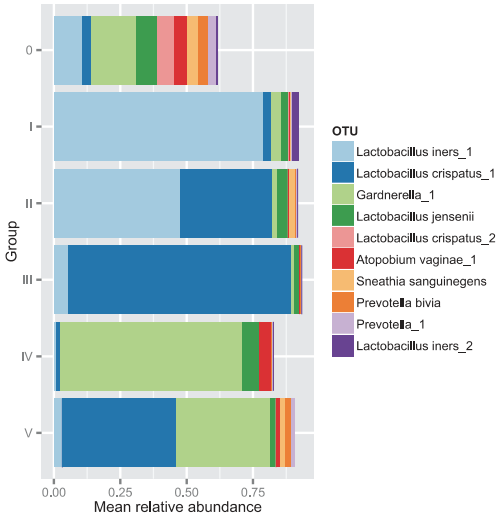

Figure 4. The distribution of 10 most relatively abundant OTUs in determined vaginal bacterial community groups. The distribution is presented based on mean relative abundance values (x-axis) of these OTU-s among bacterial community groups (y-axis).
doi:10.1371/journal.pone.0054379.g004

(uncultured OTUs that did not have a taxonomic assignment lower than kingdom), and their summarized relative abundance was 38.6%. The prevalence of these OTUs was 85.2% of samples.

Fungal community parameters in relation to the medical history, lifestyle habits, and bacterial community parameters

There was no correlation (CC = −0.05) between the number of sequences per sample and the calculated Shannon diversity index for fungal communities, indicating that taxonomic richness was not created by the number of sequences by itself. Linear regression analysis incorporating different factors of medical history and lifestyle habits did not reveal any correlation between these factors and summarized abundance of *Candida spp.* (P>0.05). Also, the relative abundance values of lactobacilli and *Gardnerella* were not correlated with the prevalence of *C. albicans* (P>0.05) when tested by logistic regression analysis. Figure 5 shows the relative abundance of the most abundant microbes (10 bacteria and 6 fungi) among 181 samples that had both the 16S rRNA gene and ITS-1 sequencing data available. When analyzing the heatmap patterns, three pairs of fungal OTU-s were almost identically distributed between the 181 samples. These OTU pairs were *C. albicans* and uncultured *Saccharomycetales*; uncultured fungus_1 and uncultured fungus_2; and *Pichia kudriavzevii* and uncultured fungus_3. By visually analyzing the alignment of representative sequences of these OTU-s, there were considerable similarities between uncultured fungus_1 and uncultured fungus_2; and *Pichia kudriavzevii* and uncultured fungus_3 but not between *C. albicans* and uncultured *Saccharomycetales* (Fig. S2).

Discussion

This study is the first to analyze vaginal microbial communities of Estonian women by using parallel pyrosequencing technology. We have shown that the composition of vaginal microbiota varies widely among healthy women with homogeneous ethnic origin. According to their taxonomic composition, the majority of

Table 2. The prevalence and summarized relative abundance of predominant fungal taxa.

Division	*Order*	*OTU*	*Prevalence*	*Summarized relative abundance*
Ascomycota			0.87	0.58
	Saccharomycetales		0.74	0.37
		Candida albicans	0.68	0.34
		uncultured *Saccharomycetales*	0.57	0.06
		Pichia kudriavzevii (C. krusei)	0.10	0.02
	Capnodiales		0.32	0.05
		Davidiellaceae sp PIMO_97	0.21	0.03
		Cladosporium perangustum	0.13	0.01
		Cladosporium sp BMP2897	0.02	0.01
		uncultured *Epicoccum*	0.06	0.01
	Eurotiales		0.21	0.03
		Eurotium amstelodami	0.06	0.01
	Pleosporales		0.16	0.03
		Alternaria alternata	0.06	0.01
	Helotiales		0.04	0.01
Basidiomycota			0.27	0.03
	Sporidiobolales		0.03	0.01
		Rhodotorula sp LH51	0.02	0.01

Only taxons with ≥1% relative abundance and occurring in more than 1% of the tested samples were included. Unspecified OTUs are excluded from the table.
doi:10.1371/journal.pone.0054379.t002

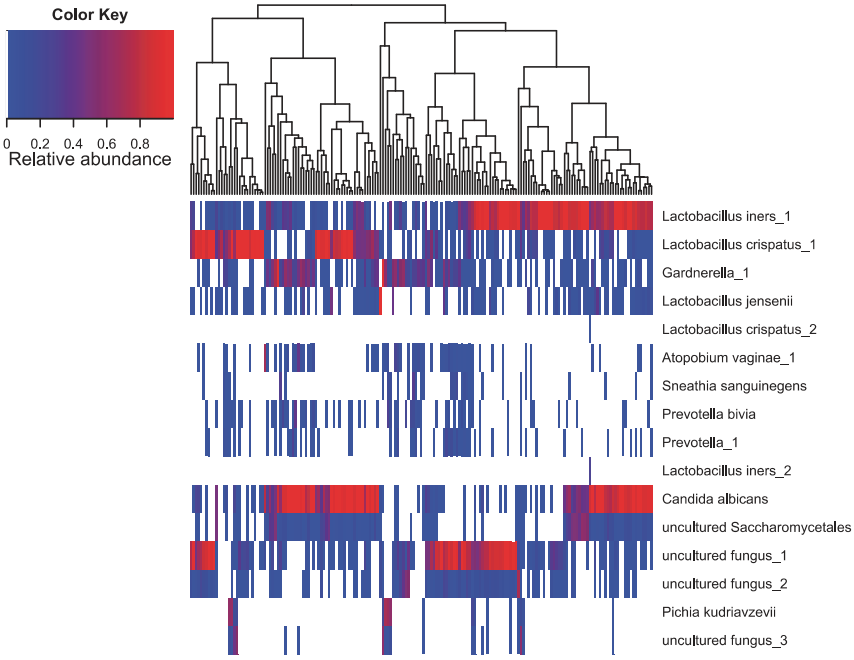

Figure 5. Relative abundance of the most abundant bacterial and fungal OTUs found in the vaginal communities of 181 women. Taxonomic assignments are indicated on the right of the heatmap at the Genus and Species level. The relative abundance is color coded and indicated by the color key on the left top of the map. The tree on the top of the heatmap characterizes the similarity of analyzed samples.
doi:10.1371/journal.pone.0054379.g005

bacterial communities clustered into five major community groups that were characterized by significantly different bacterial diversity, whereas a considerable number of communities (28.9%) did not form differentiated groups. As the women participating in the study were asymptomatic, it is not surprising that four major vaginal microbial community groups were dominated by the *Lactobacillus* species (*Lactobacillus iners* and *L. crispatus*). In general, 99% of the women were colonized with lactobacilli, and the mean relative abundance of this genus was 70%. This is consistent with the results of Srinivasan et al., who also found that women without BV have vaginal microbiotas dominated by *Lactobacillus* species, notably either *L. crispatus* or *L. iners* [25]. Nevertheless, members of genera suggestive of a "BV-like" microbiota, such as *Gardnerella* and *Atopobium*, were also relatively abundant in clustered community groups of our results.

Detection of *Gardnerella vaginalis* and *Atopobium vaginae* in vaginal samples has been historically related to vaginal microbiota that is shifted towards a diseased state [26–28]. The study by Srinivasan et al. associated these species with several clinical criteria used in the diagnosis of BV and they also described relatively strong co-occurrence of these species [25]. However, previous studies have also shown that *G. vaginalis* could be detected in approximately 50% of vaginal microbial communities belonging to healthy women [29,30]. In our study, 70.6% of the women were colonized with OTUs belonging to genus *Gardnerella*, which supports these recent results. Similar to the genus *Gardnerella*, OTUs belonging to genera *Prevotella* and *Atopobium* were widespread among healthy women participating in our study (prevalence 55.8% and 38% of samples, respectively), which is in concordance with recent studies showing *A. vaginae* and *Prevotella spp.* to be prevalent in the healthy vaginal environment [6,7].

Additionally, aerobic bacteria, such as members of the genera *Streptococcus* and *Staphylococcus*, as well as *Escherichia coli*, were detected in healthy vaginal microbiota. The abundance of these microbes was low (<2%), but the prevalence among participating women was relatively high, reaching 30.6% for streptococci. All of these bacteria have been linked to aerobic vaginitis [31].

Also, of note, *Ureaplasma sp.* was present in 41% of the studied vaginal microbial communities, but with very low abundance (0.9%), indicating its affiliation with normal vaginal microbiota. However, this does not rule out the possibility that *Ureaplasma sp.* can cause disease when present in high numbers.

Broad taxonomic variation and clustering of vaginal communities has been performed before in large scale studies that have analyzed the composition of normal vaginal microbiota in ethnically diverse participants, including Asian, Hispanic, African American, and Caucasian women from the United States and Japanese women from Japan. In those studies, most of the bacterial community groups were dominated by *Lactobacillus* species (*L. crispatus*, *L. gasseri*, *L. iners*, and *L. jensenii*), but the studies also found that a number of communities were dominated by other genera, such as *Prevotella*, *Atopobium*, and *Gardnerella* [6,8,32]. Interestingly, the number of vaginal microbial communities with miscellaneous community profiles accounted for as much as 29% of our study group, which is very similar to the

results of Ravel *et al.*, who found heterogeneous community profiles in 27% of healthy women studied in the United States [6].

Different community groups may represent different stages in the menstrual cycle, as the shifts in vaginal microbiota in relation to the menstrual cycle have been previously reported [33,34]. Srinivasan *et al.*, who used qPCR to evaluate levels of key vaginal bacteria, demonstrated elevated levels of *G. vaginalis* and *L. iners* in the vaginal environment during menstruation, whereas levels of *L. jensenii* and *L. crispatus* decreased during that time [33]. The shifts in the composition of vaginal microbiota during the menstrual cycle were observed by Gajer *et al.*, who showed that the fluctuations in the community composition were mainly affected by time in the menstrual cycle and community class (classification was based on clustering done by Ravel *et al.* [6]) [34]. However, the fluctuation of community composition related to the menstrual cycle did not occur in all cases.

Among asymptomatic women participating in our study, microbial diversity increased with higher vaginal pH values and was also higher when a malodorous discharge was present. High pH values and malodorous discharge have both been related to BV before [35]. Moreover, studies such as by Srinivasan et al. have shown that patients suffering from BV have higher bacterial diversity and richness than individuals without BV [25]. Asymptomatic BV is relatively common, as approximately half of the women who are diagnosed with BV are asymptomatic [35] and could have been excluded from the cohort of current study by applying Amsel diagnostic criteria and Nugent scoring. As this was not done and the exclusion factors included only symptomatic urogenital disorders, women having asymptomatic BV were probably not excluded from our study. The abundant presence of asymptomatic BV and other abnormal vaginal microbiota types has been recognized for decades and is a continuing concern to clinicians, as these microbiota types are associated with diverse adverse outcomes during pregnancy (such as preterm delivery and chorioamnionitis [3,36]) and increased risk of HIV and other STD transmission [2,37]. Moreover, the simple treatment of BV with metronidazole does not prevent these complications [3]. On the other hand, because these women with more diverse vaginal microbiota in combination with higher vaginal pH or malodorous discharge did not seem to have a reduced quality of life, one can question the current diagnostic strategies and the application of therapy in the absence of symptoms.

Although yeast infections are frequent and can disturb the vaginal microbiota dramatically, studies addressing vaginal microbiota have mainly focused on bacterial inhabitants, while discarding information about the vaginal mycobiome. Furthermore, studies analyzing vaginal fungal colonization have mostly focused on the *Candida* species only in the context of vaginal pathology [38]. The variation of fungal taxa in the vaginal environment may not be limited to a few known yeasts, such as *Candida* and *Saccharomyces* species. Ghannoum *et al.* used barcoded pyrosequencing techniques to analyze the oral mycobiome and found a remarkably higher prevalence of fungal taxons during mouth rinsing than shown before with culture-based methods [39]. Our study is the first of its kind to analyze the mycobiome that colonizes the healthy vaginal environment using barcoded pyrosequencing technology. Our results showed vaginal colonization with *Candida spp.* in 64.5% of patients, which is considerably higher than the prevalence of 20% found in asymptomatic healthy women in earlier studies [40]. 82% of the fungal OTUs identified as *Candida spp.* belonged to *C. albicans*, which is also in full accordance with earlier studies.

C. parapsilosis, *C. dubliniensis*, *Candida sp.* VI04616, and *Pichia kudriavzevii* (*Candida krusei*) were the only non-*albicans Candida* OTUs detected in our study. All aforementioned species, with the exception of *Candida sp.* VI04616, have been linked to vaginal pathology in previous studies. *C. parapsilosis* can cause VVC and leads to symptomatic disease in two thirds of colonized women [41]. *C. dubliniensis* is mostly an opportunistic pathogen in the oral cavity of immunocompromised individuals, but it has recently been isolated from vaginal samples collected from women suffering from VVC in Turkey [42]. *Pichia kudriavzevii* (*Candida krusei*) has been reported to cause chronic vaginitis in older women, which is notably resistant to fluconazole [43]. Nevertheless, women participating in our study were all asymptomatic, indicating that *C. parapsilosis*, *C. dubliniensis*, and *P. kudriavzevii* can be affiliated with normal vaginal microbiota without causing disease.

When analyzing the fungal non-*Candida* OTUs we came across several complications. First, most of the abundant fungal non-*Candida* OTUs were identified as potential air-borne contaminants (e.g. OTUs belonging to genus *Cladosporium*) as seen in Table 2. This may be an indication of contamination problem facing the studies analyzing fungal diversity with highly sensitive methodology, although the workflow in current study was carried out under relatively stringent conditions (DNA was extracted under laminar flow cabinet and PCR reactions were prepared under UV PCR cabinets) and the negative controls included to the process in DNA extraction and PCR reaction stage did not give a positive signal with ITS-1 specific PCR. Thus, stricter conditions for sample collection and/or lowering the number of PCR cycles used may solve or reduce this issue. Second, the summarized relative abundance of unspecified OTUs was 38.6%. The failure to specify the strains is due to two problems; the information about vaginal fungi is under-represented in sequence databases used in our study (the NCBI-based UNITE database). Also, the clustering method of cd-hit appears to be relatively imprecise for metagenomic identification of fungal OTUs. As shown in Figure 5, several fungal OTUs shared almost identical heatmap patterns, suggesting an almost identical prevalence of these OTUs among the studied samples. The probability for this kind of distribution to happen by chance or have a biological meaning is low, and therefore these OTUs are most likely bioinformatic or sequencing artifacts of a single OTU (this is supported by the alignment analysis revealing substantial sequence similarities in majority of these kinds of OTU pairs). In that case, it suggests that there is an over-estimation of the actual fungal OTU number, but the over-estimation is not considerable, because the UNITE database search resulted in 161 unique hits. Importantly, this number is not considerably smaller than the overall number of detected fungal OTUs (n = 196). We defined fungal OTUs by widely used similarity level of 97%, but lowering this value to 96% or lower may offer a solution to this bias.

The limitation of primer efficiency can be another factor that may be responsible for biases in the results. This may be problematic when comparing results with other studies. The primer pair chosen by us to amplify V1-V2 region of 16S rRNA gene has been suggested to be suitable for community clustering and taxonomic assignments in a wide range of datasets based on bioinformatic analysis [44]. Nevertheless, it has been shown that the primer-binding sites of 27F (8F) primer have several sequence variations and mismatches against several bacteria, including *Gardnerella vaginalis* [45]. Also, ITS region is widely used for studies on fungal diversity, but most of the primers used to amplify different parts of ITS have several disadvantages that may bias the results of these diversity studies [46]. For instance, the primer pair used in current study is shown to be fungal specific but on the other hand has relatively low coverage of fungal sequences

belonging to taxonomic groups of *Ascomycete* and *Basidiomycetes* under strict PCR conditions [46].

In conclusion, we have presented the first large-scale study addressing the vaginal microbial community composition of healthy Estonian women while concomitantly considering the mycobiome. We were able to confirm the findings of previous analogous studies assessing the composition of healthy vaginal microbiota by observing the clustering of microbial profiles into five distinctive community groups. Finally, the fungal component of these communities is more diverse than expected, although the bioinformatics for identification of these taxons is still currently incomplete.

Supporting Information

Figure S1 Alignment of representative sequences of *Lactobacillus iners_1* and *L.iners_2* (A); *L.crispatus_1*, *L.crispatus_2* and *L.crispatus_3* (B); *Gardnerella_1* and *Gardnerella_2* (C); *Atopobium vaginae_1* and *A.vaginae_2* (D); and *Prevotella_1* and *Prevotella_2* (E), respectively. Alignment was carried out with Muscle 3.8.31 using default parameters except for the gap value, which was −4 in current study.
(EPS)

Figure S2 Alignment of representative sequences of *C. albicans* and uncultured *Saccharomycetales* (A); uncultured fungus_1 and uncultured fungus_2 (B); and *Pichia kudriavzevii* and uncultured fungus_3 (C), respectively. Alignment was carried out with Muscle 3.8.31 using default parameters except for the gap value, which was −4 in current study.
(EPS)

FASTA file S1 List of all bacterial OTU-s and their representative sequences present in the dataset of current study. Taxonomy assignments are based on SILVA database.
(FNA)

FASTA file S2 List of all fungal OTU-s and their representative sequences present in the dataset of current study. Taxonomy assignments are based on UNITE database.
(FNA)

Acknowledgments

We thank gynecologists of Fertilitas Hospital for the recruitment of participants, co-operation and consulting.

Author Contributions

Conceived and designed the experiments: TD TL MM. Performed the experiments: TD TL AA. Analyzed the data: LT JS. Contributed reagents/materials/analysis tools: EV IS. Wrote the paper: TD LT GD AS.

References

1. Klebanoff SJ, Hillier SL, Eschenbach DA, Waltersdorph AM (1991) Control of the microbial flora of the vagina by H2O2-generating lactobacilli. J Infect Dis 164: 94–100.
2. Schwebke JR (2001) Role of vaginal flora as a barrier to HIV acquisition. Curr Infect Dis Rep 3: 152–155.
3. Donders GGG, van Calsteren K, Bellen G, Reybrouck R, van den Bosch T, et al. (2009) Predictive value for preterm birth of abnormal vaginal flora, bacterial vaginosis and aerobic vaginitis during the first trimester of pregnancy. BJOG 116: 1315–1324.
4. Foster JA, Krone SM, Forney LJ (2008) Application of ecological network theory to the human microbiome. Interdiscip Perspect Infect Dis, ID829501.
5. Antonio MA, Hawes SE, Hillier SL (1999) The identification of vaginal *Lactobacillus* species and the demographic and microbiologic characteristics of women colonized by these species. J Infect Dis 180: 1950–1956.
6. Ravel J, Gajer P, Abdo Z, Schneider GM, Koenig SS, et al. (2011) Vaginal microbiome of reproductive-age women. Proc Natl Acad Sci U S A 108 (1): 4680–4687.
7. Zhou X, Bent SJ, Schneider MG, Davis CC, Islam MR, et al. (2004) Characterization of vaginal microbial communities in adult healthy women using cultivation-independent methods. Microbiology 150: 2565–73.
8. Zhou X, Hansmann MA, Davis CC, Suzuki H, Brown CJ, et al. (2010) The vaginal bacterial communities of Japanese women resemble those of women in other racial groups. FEMS Immunol Med Microbiol 58: 169–181.
9. Danielsson D, Teigen PK, Moi H (2011) The genital econiche: focus on microbiota and bacterial vaginosis. Annals of the New York Academy of Sciences 1230: 48–58.
10. Sobel JD (2007) Vulvovaginal candidosis. Lancet 369(9577): 1961–1971.
11. Jenkinson HF, Douglas LJ (2002) Chapter 18 Interactions between *Candida* Species and Bacteria in Mixed Infections. pp. 357–373. *In* Brogden KA and Guthmiller JM (eds), Polymicrobial infections and disease, ASM Press, Washington, DC.
12. White BA, Creedon DJ, Nelson KE, Wilson BA (2011) The vaginal microbiome in health and disease. TEM 22: 389–93.
13. McKenna P, Hoffmann C, Minkah N, Aye PP, Lackner A, et al. (2008) The macaque gut microbiome in health, lentiviral infection, and chronic enterocolitis. PLoS Pathog 4(2): e20.
14. Buée M, Reich M, Murat C, Morin E, Nilsson RH, et al. (2009) 454 Pyrosequencing analyses of forest soils reveal an unexpectedly high fungal diversity. New Phytologist 184: 449–456.
15. Hamady M, Walker JJ, Harris JK, Gold NJ, Knight R (2008) Error-correcting barcoded primers for pyrosequencing hundreds of samples in multiplex. Nature Methods 5: 235–237.
16. Pruesse E, Quast C, Knittel K, Fuchs BM, Ludwig W, et al. (2007) SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. Nucleic Acids Res 35:7188–7196.
17. Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, et al. (2009) Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. Appl Environ Microbiol 75(23): 7537–41.
18. Wang Q, Garrity GM, Tiedje J M, Cole JR (2007) Naïve Bayesian Classifier for Rapid Assignment of rRNA Sequences into the New Bacterial Taxonomy. Appl Environ Microbiol 73:5261–7.
19. Yue JC, Clayton MK (2005) A similarity measure based on species proportions. Commun Stat Theor Methods 34: 2123–2131.
20. Li W, Godzik A (2006) Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. Bioinformatics 22: 1658–9.
21. Kõljalg U, Larsson KH, Abarenkov K, Nilsson RH, Alexander IJ, et al. (2005) UNITE: a database providing web-based methods for the molecular identification of ectomycorrhizal fungi. New Phytol 166: 1063–8.
22. Warnes GR, Bolker B, Bonebakker L, Gentleman R, Liaw WHA, et al. (2011) Available: http://CRAN.R-project.org/package=gplots.
23. Wickham H (2009) ggplot2: elegant graphics for data analysis. Springer New York.
24. Dixon P (2003) VEGAN, a package of R functions for community ecology. Journal of Vegetation Science, 14: 927–930.
25. Srinivasan S, Hoffman NG, Morgan MT, Matsen FA, Fiedler TL, et al. (2012) Bacterial communities in women with bacterial vaginosis: high resolution phylogenetic analyses reveal relationships of microbiota to clinical criteria. PLoS ONE 7: e37818.
26. Donders GGG (2007) Definition and classification of abnormal vaginal flora. Best Pract Res Cl Ob 21: 355–73.
27. Haggerty CL, Totten PA, Ferris M, Martin DH, Hoferka S, et al. (2009) Clinical characteristics of bacterial vaginosis among women testing positive for fastidious bacteria. Sex Transm Infect 85: 242–8.
28. Zozaya-Hinchliffe M, Lillis R, Martin DH, Ferris MJ (2010) Quantitative PCR assessments of bacterial species in women with and without bacterial vaginosis. J Clin Microbiol 48: 1812–1819.
29. Mikamo H, Sato Y, Hayasaki Y, Hua YX, Tamaya T (2000) Vaginal microflora in healthy women with *Gardnerella vaginalis*. J Infect Chemother 6: 173–177.
30. Fredricks DN, Fiedler TL, Marrazzo JM (2005) Molecular identification of bacteria associated with bacterial vaginosis. N Engl J Med 353: 1899–911.
31. Donders GGG, Vereecken A, Bosmans E, Dekeersmaecker A, Salembier G, et al. (2002) Definition of a type of abnormal vaginal flora that is distinct from bacterial vaginosis:aerobic vaginitis. BJOG 109: 34–43.
32. Zhou X, Brown CG, Abdo Z, Davis CC, Hansmann MA, et al. (2007) Difference in the composition of vaginal microbial communities found in healthy caucasian and black women. ISME 1: 121–133.
33. Srinivasan S, Liu C, Mitchell CM, Fiedler TL, Thomas KK, et al. (2010) Temporal variability of human vaginal bacteria and relationship with bacterial vaginosis. PloS one 5, e10197.

34. Gajer P, Brotman RM, Bai G, Sakamoto J, Schütte UME, et al. (2012) Temporal Dynamics of the Human Vaginal Microbiota. Sci Transl Med 4: 132ra52.
35. Amsel R, Totten PA, Spiegel CA, Chen KCS, Eschenbach DA, et al. (1983) Nonspecific vaginitis: diagnostic criteria and epidemiologic associations. Am J Med 74: 14–22.
36. Rezeberga D, Lazdane G, Kroica J, Sokolova L, Donders GGG (2008) Placental histological inflammation and reproductive tract infections in a low risk pregnant population in Latvia. Acta Obstet Gynecol Scand 87: 360–365.
37. Van de Wijgert JHHM, Morrison CS, Cornelisse PGA, Munjoma M, Moncada J, et al. (2008) Bacterial Vaginosis and Vaginal Yeast, but not Vaginal Cleansing, Increase HIV-1 Acquisition in African Women. J Acquir Immune Defic Syndr 48: 203–210.
38. Trama JP, Mordechai E, Adelson ME (2005) Detection and identification of Candida species associated with Candida vaginitis by Real-Time PCR and pyrosequencing. Mol Cell Probes 19:145–152.
39. Ghannoum MA, Jurevic RJ, Mukherjee PK, Cui F, Sikaroodi M, et al. (2010) Characterization of the oral fungal microbiome (mycobiome) in healthy individuals. PLoS pathogens 6, e1000713.
40. Sobel JD (2000) Practice guidelines for the treatment of fungal infections. Clin Infect Dis 30: 000–652.
41. Kennedy MA, Sobel JD (2010) Vulvovaginal candidiasis caused by non-*albicans Candida* species: new insights. Curr Infect Dis Rep 12(6): 465–470.
42. Us E, Cengiz S (2007) Prevalence and phenotypic evaluation of *Candida dubliniensis* in pregnant women with vulvovaginal candidosis in a university hospital in Ankara. Mycoses 50: 13–20.
43. Singh S, Sobel JD, Bhargava P, Boikov D, Vazquez JA (2002) Vaginitis due to Candida krusei: Epidemiology, Clinical Aspects and Therapy. Clin Infect Dis 35: 1066–1070.
44. Liu Z, DaSantis TZ, Andersen GL, Knight R (2008) Accurate taxonomy assignments from 16S rRNA sequences produced by highly parallel pyrosequencers. Nucleic Acids Res 36: e120.
45. Srinivasan S, Fredricks DN (2008) The human vaginal bacterial biota and bacterial vaginosis. Interdiscip Perspect Infect Dis.
46. Bellemain E, Carlsen T, Brochmann C, Coissac E, Taberlet P, et al. (2010) ITS as an environmental DNA barcode for fungi: an in silico approach reveals potential PCR biases. BMC Microbiol 10: 189.

Publication II

Drell, T., Lutsar, I., Štšepetova, J., Parm, Ü., Metsvaht, T., Ilmoja, M., Simm, J., and Sepp, E. (2014). The development of gut microbiota in critically ill extremely low birth weight infants assessed with 16S rRNA gene based sequencing. Gut Microbes 5, 304–503.

RESEARCH PAPER

Gut Microbes 5:3, 304–312; May/June 2014; © 2014 Landes Bioscience

The development of gut microbiota in critically ill extremely low birth weight infants assessed with 16S rRNA gene based sequencing

Tiina Drell[1,2,3,*], Irja Lutsar[1], Jelena Štšepetova[1], Ülle Parm[1], Tuuli Metsvaht[1,4], Mari-Liis Ilmoja[5], Jaak Simm[2], and Epp Sepp[1]

[1]Institute of Microbiology; University of Tartu; Tartu, Estonia; [2]Centre for Biology of Integrated Systems; Tallinn University of Technology; Tallinn, Estonia; [3]Competence Centre on Reproductive Medicine and Biology; Tartu, Estonia; [4]Anaesthesiology and Intensive Care Clinic; Tartu University Hospital; Tartu, Estonia; [5]Department of Anaesthesiology and Intensive Care; Tallinn Children's Hospital; Tallinn, Estonia

Keywords: gut microbiota, microbiome profiling, preterm neonates, 16S rRNA gene sequencing, extremely low birth weight

Abbreviations: ELBW, Extremely low birth weight; BW, birth weight; NICU, Neonatal intensive care unit; OTU, Operational taxonomic unit; GA, Gestational age; LOS, Late onset sepsis; NEC, Necrotizing enterocolitis; qPCR, Quantitative Real-Time PCR; PROM, Premature rupture of membranes; TPN, Total parenteral nutrition

Objective: An increasing number of studies that are using high-throughput molecular methods are rapidly extending our knowledge of gut microbial colonization in preterm infants whose immaturity and requirement for extensive treatment may result in altered colonization process. We aimed to describe the profile of gut microbiota in 50 extremely low birth weight (<1200 g) critically ill infants at three different time points during the first two months of life by using 16S rRNA gene specific sequencing.

Patients and Methods: Stool samples were collected at the age of one week, one month and two months. Bacterial community profiling was done using universal amplification of 16S rRNA gene and 454 pyrosequencing.

Results: The diversity of gut microbiota in preterm neonates in the first week of life was low but increased significantly over two months. The gut microbiota was dominated by facultative anaerobic bacteria (*Staphylococcus spp.* and *Enterobacteriaceae*) and lacked colonization with bacteria known to provide resistance against pathogens (*Bacteroides, Bifidobacterium*, and *Lactobacillus)* throughout the study. Colonization of *Escherichia coli* and uncultured *Veillionella* was positively correlated with maturity. Infants born to mothers with chorioamnionitis had significantly higher bacterial diversity than those without.

Conclusions: High prevalence and abundance of potentially pathogenic *Enterobacteriaceae* and *Staphylococcaceae* with low prevalence and abundance of colonization resistance providing taxa bifidobacteria, *Bacteroides* and lactobacilli may lead to high infection risk via microbial translocation from the gut. Additionally, our data suggest that maternal chorioamnionitis may have an effect on the diversity of infants' gut microbiota; however, the mechanisms involved remain to be elucidated.

Introduction

Normal gut colonization begins straight after birth with maternal vaginal and faecal microbiota being the key source of colonizing microbes.[1] In term neonates during the first days of life the most abundant colonizers are staphylococci, γ-proteobacteria (e.g., *Enterobacteriaceae*), and bifidobacteria,[2-4] but the composition changes over time and gradually becomes more adult-like with the dominance of *Bacteroidetes* and *Firmicutes*.[3]

Within the past two decades, the survival rate of preterm infants has greatly increased due to improved standards of neonatal intensive care. For this group of infants, many parameters of normal development of gut bacterial colonization must be redefined as, in addition to immaturity, preterm infants often require intensive care and antibiotic treatment, all having a considerable effect on the development of the gut microbiota.[5] Previous studies have shown that the gut microbiota in preterm infants is dominated by members of family *Enterobacteriaceae* and has decreased numbers of lactobacilli and bifidobacteria.[6,8] Also, during the first days of life preterm infants are colonized with only a few microbial species.[6-8] Increasing number of studies using 16S rRNA gene specific sequencing are in general confirming but on the other hand expanding the previous knowledge.[9-14] We aimed to broaden the understanding of the composition of gut microbiota in preterm infants by describing the profile of gut microbiota in 50 extremely low birth weight (ELBW; birth weight [BW] <1200 g) neonates, all requiring treatment in 3rd level neonatal intensive care unit (NICU) and

*Correspondence to: Tiina Drell; Email; tiinadrell@gmail.com
Submitted: 10/18/2013; Revised: 03/28/2014; Accepted: 04/10/2014; Published Online: 04/10/2014
http://dx.doi.org/10.4161/gmic.28849

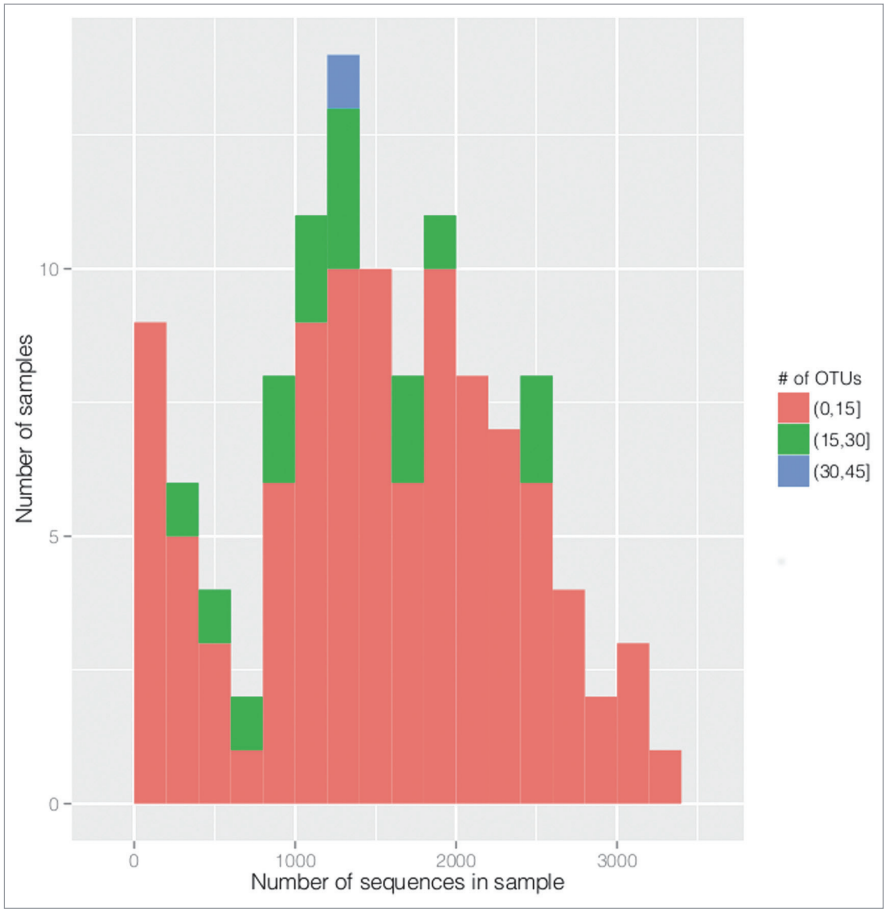

Figure 1. The distribution of 118 samples according to number of sequences and number of OTUs. Round and square brackets in the figure indicate to the exclusion and inclusion of adjacent value, respectively.

receiving antibiotic treatment, with the focus of composition dynamics between three time points collected during the first two months of life by using 16S rRNA gene specific sequencing.

Results

Sequencing data

We retrieved 188 222 high-quality sequence reads with an average length of 280 bp. The number of trimmed sequences was higher than 500 in 101 out of 118 samples (30 samples collected at the age of one week, 42 at one month, and 29 at two months). The cut-off was set to 500 sequences per sample because this was the level whereby the rarefaction curves for 95% of the samples reached a 1% plateau. This means that the increasing number of sequences per sample does not increase the number of OTUs obtained in 95% of the samples (**Fig. 1**). Eighty-five operational taxonomic units (OTUs) met the criteria of having at least 5 sequences assigned to them and thus were included into further analysis (Accession numbers for representative sequences of OTUs in GenBank are KJ527501-KJ527585). The mean (SD) taxonomic richness per sample was 8.15 (4.19). This did not change between the age of one week and one month being 7.13 (5.15) and 7.83 (3.77), respectively, but it increased significantly by the age of two months (9.64 [3.34]; $P = 0.04$). The mean (SD) Shannon diversity index was 0.67 (0.51) and it increased over time being 0.38 (0.4) at one week, 0.71(0.54) at one month, and 0.92 (0.44) at two months of age (all $P < 0.01$).

Prevalence and mean relative abundance of bacterial phyla, families, and genera

Retrieved sequences were distributed between seven bacterial phyla, 24 families and 35 genera. Major phyla (prevalence > 10%

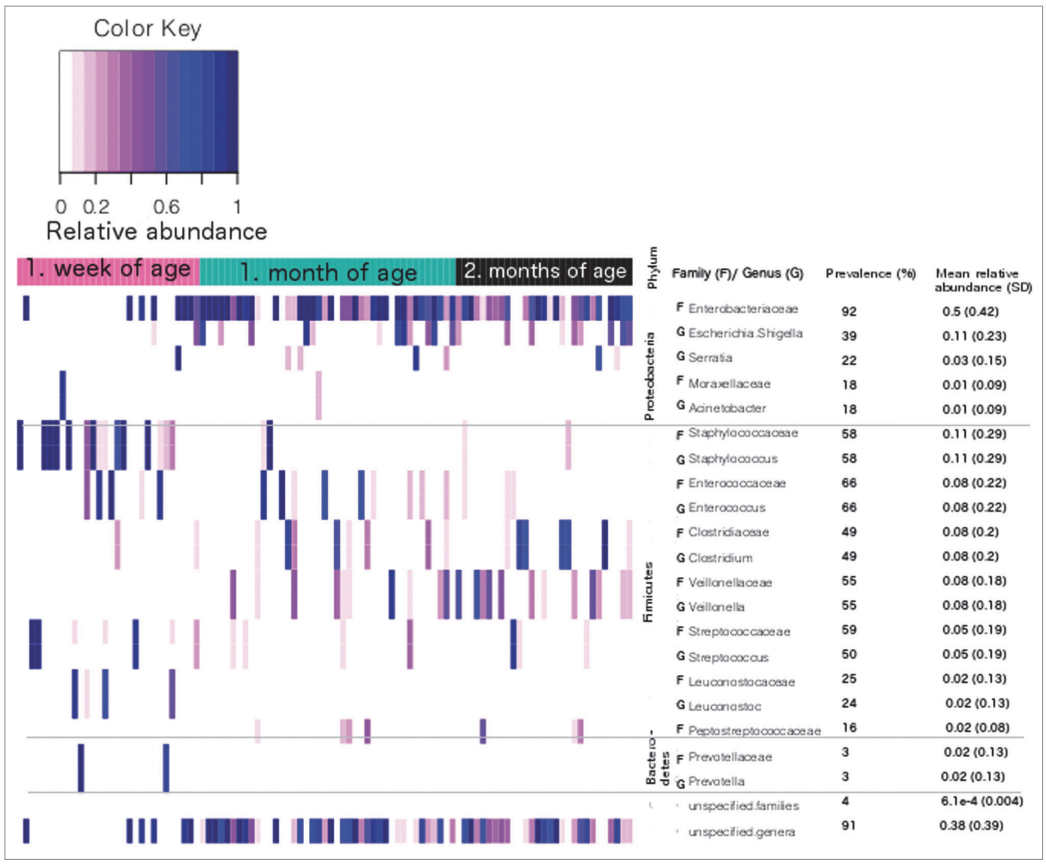

Figure 2. The distribution of 10 most abundant families (f) and genera (g) between the ages of one week, one month, and two months.

and relative abundance > 0.01) were *Proteobacteria* (prevalence 94%; mean relative abundance [SD] 0.51 [0.42]), whose abundance increased significantly between the age of one week and one month (from 0.28 [0.44] to 0.66 [0.39]; $P < 0.01$); *Firmicutes* (100%; 0.45 [0.41]), whose abundance decreased significantly between the age of one week and one month (from 0.65 [0.45] to 0.32 [0.37]; $P < 0.01$); and *Bacteroidetes* (13%; 0.02 [0.14]). Minor phyla were *Actinobacteria* (13%; 0.002 [0.007]) and *Verrucomicrobia*, which was present in one sample with relative abundance of 0.14.

Ten most abundant families and genera are presented in **Figure 2**. Among these families, the prevalence of *Enterobacteriaceae* and *Enterococcaceae* increased significantly between one week and one month of age (from 77% to 98% and from 53% to 81%; $P < 0.01$ and $P = 0.02$, respectively). This increase was also observed for mean relative abundance values of *Enterobacteriaceae* and *Veillionellaceae* that increased from 0.25 (0.42) to 0.66 (0.38) ($P < 0.001$) and from 0.003 (0.009) to 0.09 (0.19) ($P < 0.01$), respectively. At the same time the prevalence of *Prevotellaceae* and the abundance of *Staphylococcaceae* decreased significantly from 13% to 0 ($P = 0.03$) and 0.33 (0.43) to 5.1e-4 (0.15) ($P < 0.001$), respectively. The only change observed between the age of one month and two months was the decrease in abundance of *Enterococcaceae* from 0.1 (0.25) to 0.02 (0.03) ($P = 0.03$). The distribution of 10 most abundant families in individual samples can be seen in **Figure S1**.

The majority of these changes were also observed in approximately the same proportions among the most abundant genera belonging to aforementioned families. These genera included *Staphylococcus, Enterococcus*, *Veillonella*, and *Prevotella* (**Fig. 2**). In family *Enterobacteriaceae*, most abundant genus was *Escherichia/Shigella*, which similarly to family *Enterobacteriaceae*, increased significantly between the ages of one week and one month with the values of 17% vs. 41% for prevalence and 0.02 (0.1) vs. 0.16 (0.3) for mean (SD) relative abundance ($P = 0.04$ and $P < 0.01$, respectively). It has to be taken into consideration, that significant proportion of sequences (prevalence 91% and mean relative abundance [SD] 0.38 [0.39]) was unspecified on

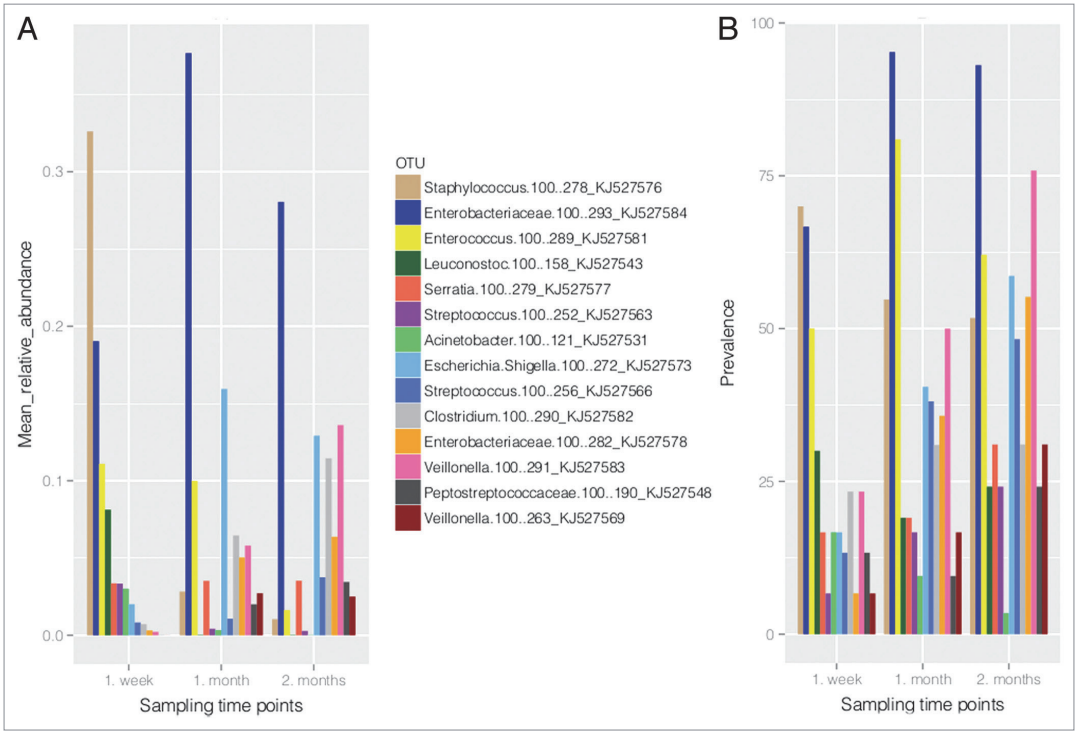

Figure 3. The distribution of the most dominant OTUs (prevalence > 10% and mean relative abundance > 0.01) between the ages of one week, one month and two months based on their mean relative abundance (**A**) and prevalence (**B**).

genus level (**Fig. 2**). The majority of these sequences belonged to family *Enterobacteriaceae* (88%, 0.35 [0.39]), in particular to OTU named *Enterobacteriaceae.100.293* (86%, 0.29 [0.38]). Based on additional identification against NCBI nt database this OTU matched to genus *Klebsiella*.

Prevalence and relative abundance of bacterial OTUs

Fourteen OTUs out of 85 dominated the gut microbiota of ELBW infants (prevalence and mean relative abundance higher than 10% and 0.01, respectively). The distribution of these OTUs in studied time points is shown in **Figure 3**. At the age of one week, three of the most prevalent and abundant OTUs in gut microbiota of ELBW infants were *Staphylococcus.100.278* (GenBank accession number KJ527576), *Enterobacteriaceae.100.293* (*Klebsiella sp.*; KJ527584), and *Enterococcus.100.289* (KJ527581). By the age of one month, OTUs corresponding to *Escherichia/Shigella.100.272* (KJ527573) together with *Enterobacteriaceae.100.293* and *Enterococcus.100.289* predominated. At the age of two months the most prevalent and abundant OTUs were *Enterobacteriaceae.100.293* (*Klebsiella sp.*), *Veillonella.100.291* (KJ527583), *Enterococcus.100.289*, and *Escherichia/Shigella.100.272* (**Fig. 3**).

None of the OTUs assigned to genus *Bifidobacterium* crossed the detection limit. Also, only one OTU was assigned to genus *Bacteroides* and two to *Lactobacillus*. These three OTUs had prevalence lower than 4%, although relative abundance was high in one or two samples (>0.01, reaching up to 0.6).

Correlation between gut microbiota and clinical parameters

We observed positive correlation between Shannon diversity index and corrected gestational age (GA+ postnatal age; $P < 0.001$) and negative correlation between Shannon diversity index and the starting day of total enteral feeding ($P < 0.01$). Also, Shannon diversity index was significantly higher among ELBW infants born to mothers with chorioamnionitis than in infants born to mothers without (mean values [SD] 0.82 [0.59] vs. 0.59 [0.45]; $P < 0.01$).

The prevalence and relative abundance of dominating OTUs correlated only with corrected GA: relative abundance of *Staphylococcus.100.272* decreased ($P = 0.02$) and the prevalence of *Escherichia/Shigella.100.272* and *Veillonella.100.291* increased with corrected GA ($P < 0.01$ and $P = 0.03$, respectively).

Discussion

Current study is a thorough addition to several recent 16S rRNA gene sequencing based studies analyzing the composition

of gut microbiota in preterm infants. We observed significant changes in the gut microbiota of ELBW infants during the first weeks of life when extensive interventions (NICU stay, artificial lung ventilation, total parenteral nutrition [TPN], and empiric and broad-spectrum antibiotic treatment) were applied. By the age of one month, however, the composition of gut microbiota had shifted toward more stable complex with the changes occurring between the age of one and two months being not as extensive as in the first weeks of life regardless of significant increase in diversity. Nevertheless, key colonization resistance providing bacteria (*Bifidobacterium*, *Bacteroides*, and *Lactobacillus*) were lacking during the whole study period. These results are generally confirmatory to previous knowledge on this subject. Interestingly, ELBW infants born to mothers with chorioamnionitis had significantly higher bacterial diversity in their gut. This raises the question if diversity may be affected by more extensive exposure to microbes, inflammatory reactions and antibiotic treatment in case of chorioamnionitis or is the presence of chorioamnionitis correlated with some unknown factor affecting the diversity in these infants. In clinical data available for our study, chorioamnionitis did not correlate with any factor.

Similarly to previous studies assessing the composition of gut microbiota in preterm infants, we observed gradual increase of diversity in ELBW infants throughout the study.[7,15] As the gradual increase of diversity has also been reported in vaginally born fullterm infants,[16] it seems to be a normal process of diversification of gut microbiota in early age regardless of maturity. Nevertheless, Jacquot et al.[17] have shown that bacterial diversity increases more slowly in extremely preterm (GA < 28 wk) than in moderately preterm infants (GA ⩾ 28 wk),[17] supported by the increase of Shannon diversity index with increasing corrected GA seen in our study. Additionally, many studies have concluded that preterm infants have generally low bacterial diversity,[8,10,15] which can be seen by us in ELBW infants as well.

Shannon diversity index was higher among ELBW infants who started to receive enteral feeds earlier. This is most probably due to a shorter period of nutrition deficiency in the gut resulting from parenteral feeding. We also found that Shannon diversity index was significantly higher among ELBW infants born to mothers with chorioamnionitis (frequent causative factor for preterm birth) as opposed to those without, which is contradicting the results of Madan et al. [18] Unfortunately, we did not have the additional information about the type or duration of antibiotics used to treat these women. Thus, we can only hypothesize that the β-lactams in amnionic fluid, exposure to inflammation and potential colonizers (chorionamnionitis is a polymicrobial infection; over 65% of positive amniotic fluid cultures involve two or more organisms[19]) may, among other factors, contribute to more diverse gut microbiota in these infants.

Members of family *Enterobacteriaceae* from phylum *Proteobacteria* were the dominating organisms of the gut microbiota of ELBW infants except immediately after birth when the slight predominance of *Staphylococcaceae* from phylum *Firmicutes* was observed. Increased levels of *Proteobacteria* have been reported extensively by studies assessing the composition of gut microbiota of preterm infants[14,20-22] and interestingly it has lately been linked to necrotizing enterocolitis.[14,22,23]

The dominance of *Staphylococcaceae* in the gut microbiota of newborns has been hypothesized to be the result of improved hygienic condition during delivery.[24] High abundance of staphylococci may increase the risk for acquiring coagulase-negative staphylococcal infections, for which preterm neonates are especially predisposed to.[25] A correlation between higher abundance of most dominant OTU from genus *Staphylococcus* and smaller corrected GA of ELBW infants indicate that the risk of acquiring these infections through staphylococcal colonization is at its highest during the first week of life.

A representative from genus *Escherichia/Shigella* was one of the dominating OTUs, which higher prevalence was correlated with increasing corrected GA. The most widely reported member of genus *Escherichia* in gut microbiota of preterm infants has historically been *Escherichia coli*,[6,8,17] with the prevalence of this microbe rising during the first month of life.[8] If we consider our results confirmatory to this knowledge, we verify that *E. coli* is a habitant of more mature gut microbita. Nevertheless, it has to be noted that although it is very likely, we cannot confirm with full certainty if this OTU is corresponding to *E. coli* or not due to the limited capacity of resolving different species based on short lengths of sequences.

Similarly to OTU corresponding to genus *Escherichia/Shigella*, more mature ELBW infants were more likely to harbor OTU corresponding to genus *Veillonella* in their gut microbiota. The increase of *Veillonella* in gut microbiota over time has also been described in healthy infants.[2]

Although similar overall dynamics of bacterial families has been observed in fullterm infants[3] there is a much wider inter-individual variability of the composition of gut microbiota in fullterm than in preterm infants.[7] Also, high abundance of bifidobacteria and *Bacteroides* in fullterm[1,2,26] but low abundance or absence in preterm infants has been frequently observed.[6,8-10,26,27] Moreover, Butel et al.[28] have shown that when the GA at birth is less than 33 wk, the gut colonization by bifidobacteria is decreased.[28] This may be an explanation why bifidobacteria did not cross the detection limit in our study in which GA for all ELBW infants was less than 32 wk. In addition to lack of bifidobacteria, low prevalence and abundance of *Bacteroides* and *Lactobacillus* are all indicators of impaired colonization process. It has been widely hypothesized that gut colonization with bifidobacteria and lactic acid bacteria such as lactobacilli, is supported by breast-feeding although there are increasing evidence that this may not be the case.[29] Nevertheless, our results may seem to favor this hypothesis as most of the ELBW infants participating in current study did not receive breast-feeding and the levels of lactic acid bacteria were very low to statistically argue against it, but based on other similar studies it is unlikely that feeding regimen is affecting this colonization in preterm infants. For instance, Butel et al.[28] failed to see a statistical correlation between breast-feeding and bifidobacterial colonization in preterm infants.[28] Also, studies using 16S rRNA gene based sequencing are supporting these

findings by reporting low levels of lactobacilli, bifidobacteria, and *Bacteroides* despite of studying mainly breast-fed preterm infants.[9,10]

One of the limitations in our study is a possible primer bias especially toward bifidobacteria as 27F primer has been shown to have low affinity for this group of bacteria.[30] Nevertheless, by analyzing the levels of bifidobacteria with quantitative Real-Time PCR and genus *Bifidobacterium* specific primers, we were able to validate that the prevalence and levels of bifidobacteria are indeed low in ELBW infants. We also ruled out primer bias against genera *Bacteroides* and *Lactobacillus*. Thus, in the context of current study the primer bias does not have a considerable effect on the results.

In conclusion, the detailed description of gut microbiota in critically ill ELBW infants during the first two months of life confirmed low prevalence and abundance of bifidobacteria, *Bacteroides*, and lactobacilli throughout the study period, suggesting prevailing impaired colonization resistance, and showed the shift in dominance from *Staphylococcaceae* to *Enterobacteriaceae* during first weeks of life. Our data suggest that maternal chorioamnionitis may have an effect on the diversity of ELBW infants' gut microbiota; however, the mechanisms involved remain to be elucidated.

Patients and Methods

Participants and sample collection

A total of 118 stool samples collected from 50 preterm infants met the inclusion criteria. Briefly, approximately 200 mg of stool was available from infants with BW <1200 g participating in a cluster-randomized study comparing the efficacy of ampicillin and penicillin, both combined with gentamicin, in risk-factor based empiric treatment of suspected early onset neonatal sepsis. The details of the study are described elsewhere.[31] Thus, the patients were all receiving antibiotic treatment. Stool samples were collected in a clean screw-top container at the age of one week (n = 38), one month (n = 45), and two months (n = 35). Samples were initially stored at 4 °C for a maximum of 4 h and then transferred to -80 °C until analyzed.

Additionally, extensive demographic and clinical data were obtained for the participants (**Table 1**). The study was conducted in the NICUs of Tallinn Children's Hospital and Tartu University Hospital. Altogether 50 mothers (median age [min;max] was 30.5 [16;44] years) participated in this study. Twelve mothers received antenatal antibiotics and 23 received antibiotics during delivery. Eighteen mothers were suffering from chorioamnionitis and 12 had premature rupture of membranes >18 h (PROM). The diagnosis of chorioamnionitis was confirmed based on clinical criteria: fever and/or high C reactive protein levels, odorous amniotic fluid, and placental inflammation.

Of note, two outbreaks of bloodstream infections were observed during the study. In hospital ward B an outbreak of methicillin-resistant *Staphylococcus aureus* infection involving three patients and in hospital ward A five patients had a blood stream infection caused by *K. pneumoniae*.[4]

454 pyrosequencing of 16S rRNA gene

DNA was extracted from stool samples using QiaAmp Stool DNA Mini kit (Qiagen) according to manufacturer's instructions and stored at -20 °C. The amplification of 16S rRNA V1-V4 hypervariable region was performed with primers that included 454 specific adaptor sequences at 5′ end following the 8-bp barcode marked as Ns (unique sequence tag to barcode each sample) and universal 27F and 685R primers.[32,33] Full primer sequences were as follows: 27F with B adaptor 5′CCTATCCCCT GTGTGCCTTG GCAGTCTCAG NNNNNNNNAG AGTTTGATCC TGGCTCAG3′ and 685R with A adaptor 5′CCATCTCATC CCTGCGTGTC TCCGACTCAG NNNNNNNNTC TACGCATTTC ACCGCTAC3′.

Cycling parameters were 3 min at 95 °C, followed by 5 cycles of 30 s at 95 °C, 30 s at 47 °C and 60 s at 72 °C, then 30 cycles of 30 s at 95 °C, 30 s at 71 °C and 60 s at 72 °C with a final extension at 72 °C for 10 min. PCR reactions were performed in total volume of 25 μL including 3 μL of DNA template and primers at concentration 0.2 μM. PCR products were purified using Agencourt AMPure XP (Beckman Coulter) and sequenced with 454 FLX+ systems at GATC Biotech AG.

Additional PCR and qPCR were performed for genera *Lactobacillus, Bacteroides*, and *Bifidobacterium* to analyze primer 27F bias toward these genera in current study (**Fig. S2**).

Data analysis

The initial pre-trimmed data set was denoised using PyroNoise and UChime ("chimera.uchime" task in de novo mode) implemented in MOTHUR software 1.27.0. Only sequences longer than 150 bp were included for further processing. OTUs were generated with the average neighbor hierarchical clustering algorithm with identity threshold of 97%. For additional denoising, OTUs with less than 5 sequences were removed. Reference sequences of aligned 16S rDNA were obtained from the SILVA rRNA database[34] against what the taxonomic assignments were performed using Naive Bayesian classifier with a confidence cutoff of 90%.[35] OTU naming was based on the lowest taxonomic level identified for OTUs based on SILVA database match followed by the value of this match (100 equals 100% match) and serial number in the OTU list generated for this study (the list was generated in random order). Additional taxonomic assignment for identifying species to which OTUs potentially corresponded to, was performed using BLASTN against the NCBI nt database (last accessed in February, 2014). The relative abundance values and sequence counts were all normalized.

Statistics

Statistical analysis was performed using R 2.13.2 software. For analyzing general bacterial diversity Shannon diversity index was calculated.[36] Categorical values were compared with Fisher exact test and the continuous variables by Welch Two sample T-test. Mixed effect models were used to analyze the correlation of clinical parameters with microbial characteristics (Shannon diversity index; prevalence and relative abundance of five most

Table 1. Demographic and clinical factors for participating ELBW infants (n = 50)

Factor	Value
Duration of NICU days; median (IQR)	25.7 (12.5–41.8)
Number of patients/ Duration of artificial lung ventilation (days); median (IQR)	48/ 5.8 (3.4–17.5)
Number of patients in hospital ward A	28
Number of patients born with caesarean section	22
GA (weeks); mean (SD)	26.64 (0.32)
BW (g) mean (SD)	886.64 (28.27)
Number of male patients	26
Number of patients receiving empiric antibiotic treatment with penicillin	25
Duration of empiric antibiotic treatment in days; median (IQR)	3 (0.5–14.5)
Number of patients receiving broad-spectrum antibiotic treatment:	37
β-lactam + betalactamase inhibitors	23
III ja IV generation cephalosporines	10
carbapenems	6
Starting day of broad-spectrum antibiotic treatment; median (IQR)	7 (0–25)
Duration of broad-spectrum antibiotic treatment in days; median (IQR)	11.5 (0–38.8)
Number of patients on following feeding regimen at day 7*:	
TPN	12
Breast milk containing regimen	17
Formula	21
Mean % (range) of breast milk in breast milk containing regimen	81.4 (25–100)
Starting day of enteral feeding regimen; median (IQR)	3 (1–7)
Starting day of total enteral feeding; median (IQR)	17 (14–22)
Number of participants suffering from sepsis:	24
Early onset sepsis	2
Late onset sepsis	22
Number of NEC II, III patients	7
Number of patients who died in hospital	2

*Feeding regimen of ELBW infants was documented on Day 7 and categorized into three groups based on the route and character of the feeds as follows: (1) TPN, only parenteral feeding or including enteral feeds providing calories less than 10% of daily total; (2) breast milk containing regimen, breast milk constituting at least 11% of enteral feeds (the actual proportion of breast milk ranged from 12 to 100%); and (3) formula feeding, formula constituting more than 89% of enteral feeds. In case of breast milk containing regimen only biological mother's fresh or frozen breast milk was used; pasteurization or donor milk were not available. For formula feeding a ready-made liquid preterm formula Nenatal by Nutricia providing 82 Kcal/100 ml was used. Fortification was started only when enteral volume of 100 ml/kg was reached. TPN was started with glucose (4–6 g/kg/d) and amino acids (1 g/kg/d) within the first hours of life and increased by 2 g/kg/d for glucose and 1 g/kg/d for aminoacids as tolerated. Lipids (0.5–1 g/kg/d) were started on the second day of life and advanced by 0.5–1 g/kg/d as tolerated.

dominating OTUs). These models included the following fixed effects: corrected GA (GA+ postnatal age in weeks); PROM; mode of delivery; duration of artificial lung ventilation; duration of empiric antibiotic treatment; the starting day and duration of broad spectrum antibiotic treatment (carbapenems, third or fourth generation cephalosporins, and β-lactamase resistant penicillins); feeding regimen; presence of late onset sepsis; maternal usage of antibiotics prior to delivery and during delivery; and presence of maternal chorioamnionitis. Feeding regimen included three separate variables: the volume of received breast milk documented on Day 7; the starting day of enteral feeding, which consisted of breast milk containing regimen (at least 11% of breast milk in enteral feeds) and formula feeding (more than 89% of formula in enteral feeds); and the starting day of total enteral feeding (marks the end of TPN). The random effects in mixed effect models were hospital ward (A and B) and the type of empiric antibiotic treatment used (penicillin or ampicillin plus gentamicin).

All analyses were performed with Holm-Bonferroni correction and the overall level of a significant difference was set at 5%.

Ethics

The study was approved by the Ethics Committee of University of Tartu and informed consent was signed by parents or guardians.

Disclosure of Potential Conflicts of Interest

No potential conflict of interest was disclosed.

Acknowledgments

This study was supported by grants of Estonian Ministry of Education and Research (target financing No SF0180132s08) and the Estonian Science Foundation (grant No 6984 and grant No 9059).

Supplemental Materials

Supplemental materials may be found here: www.landesbioscience.com/journals/gutmicrobes/article/28849

References

1. Vaishampayan PA, Kuehl JV, Froula JL, Morgan JL, Ochman H, Francino MP. Comparative metagenomics and population dynamics of the gut microbiota in mother and infant. Genome Biol Evol 2010; 2:53-66; PMID:20333224; http://dx.doi.org/10.1093/gbe/evp057
2. Eggesbø M, Moen B, Peddada S, Baird D, Rugtveit J, Midtvedt T, Bushel PR, Sekelja M, Rudi K. Development of gut microbiota in infants not exposed to medical interventions. APMIS 2011; 119:17-35; PMID:21143523; http://dx.doi.org/10.1111/j.1600-0463.2010.02688.x
3. Palmer C, Bik EM, DiGiulio DB, Relman DA, Brown PO. Development of the human infant intestinal microbiota. PLoS Biol 2007; 5:e177; PMID:17594176; http://dx.doi.org/10.1371/journal.pbio.0050177
4. Parm Ü, Metsvaht T, Sepp E, Ilmoja M-L, Pisarev H, Pauskar M, Lutsar I. Mucosal surveillance cultures in predicting Gram-negative late-onset sepsis in neonatal intensive care units. J Hosp Infect 2011; 78:327-32; PMID:21684633; http://dx.doi.org/10.1016/j.jhin.2011.03.025
5. Mshvildadze M, Neu J, Mai V. Intestinal microbiota development in the premature neonate: establishment of a lasting commensal relationship? Nutr Rev 2008; 66:658-63; PMID:19019028; http://dx.doi.org/10.1111/j.1753-4887.2008.00119.x
6. Magne F, Abély M, Boyer F, Morville P, Pochart P, Suau A. Low species diversity and high interindividual variability in faeces of preterm infants as revealed by sequences of 16S rRNA genes and PCR-temporal temperature gradient gel electrophoresis profiles. FEMS Microbiol Ecol 2006; 57:128-38; PMID:16819956; http://dx.doi.org/10.1111/j.1574-6941.2006.00097.x
7. Schwiertz A, Gruhl B, Löbnitz M, Michel P, Radke M, Blaut M. Development of the intestinal bacterial composition in hospitalized preterm infants in comparison with breast-fed, full-term infants. Pediatr Res 2003; 54:393-9; PMID:12788986; http://dx.doi.org/10.1203/01.PDR.0000078274.74607.7A
8. Gewolb IH, Schwalbe RS, Taciak VL, Harrison TS, Panigrahi P. Stool microflora in extremely low birthweight infants. Arch Dis Child Fetal Neonatal Ed 1999; 80:F167-73; PMID:10212075; http://dx.doi.org/10.1136/fn.80.3.F167
9. Mshvildadze M, Neu J, Shuster J, Theriaque D, Li N, Mai V. Intestinal microbial ecology in premature infants assessed with non-culture-based techniques. J Pediatr 2010; 156:20-5; PMID:19783002; http://dx.doi.org/10.1016/j.jpeds.2009.06.063
10. LaTuga MS, Ellis JC, Cotton CM, Goldberg RN, Wynn JL, Jackson RB, Seed PC. Beyond bacteria: a study of the enteric microbial consortium in extremely low birth weight infants. PLoS One 2011; 6:e27858; PMID:22174751; http://dx.doi.org/10.1371/journal.pone.0027858
11. Mai V, Torrazza RM, Ukhanova M, Wang X, Sun Y, Li N, Shuster J, Sharma R, Hudak ML, Neu J. Distortions in development of intestinal microbiota associated with late onset sepsis in preterm infants. PLoS One 2013; 8:e52876; PMID:23341915; http://dx.doi.org/10.1371/journal.pone.0052876
12. Stewart CJ, Marrs EC, Nelson A, Lanyon C, Perry JD, Embleton ND, Cummings SP, Berrington JE. Development of the preterm gut microbiome in twins at risk of necrotising enterocolitis and sepsis. PLoS One 2013; 8:e73465; PMID:24023682; http://dx.doi.org/10.1371/journal.pone.0073465
13. Torrazza RM, Ukhanova M, Wang X, Sharma R, Hudak ML, Neu J, Mai V. Intestinal microbial ecology and environmental factors affecting necrotizing enterocolitis. PLoS One 2013; 8:e83304; PMID:24386174; http://dx.doi.org/10.1371/journal.pone.0083304
14. Claud EC, Keegan KP, Brulc JM, Lu L, Bartels D, Glass E, Chang EB, Meyer F, Antonopoulos DA. Bacterial community structure and functional contributions to emergence of health or necrotizing enterocolitis in preterm infants. Microbiome 2013; 1:20; PMID:24450928; http://dx.doi.org/10.1186/2049-2618-1-20
15. Stewart CJ, Marrs ECL, Magorrian S, Nelson A, Lanyon C, Perry JD, Embleton ND, Cummings SP, Berrington JE. The preterm gut microbiota: changes associated with necrotizing enterocolitis and infection. Acta Paediatr 2012; 101:1121-7; PMID:22845166; http://dx.doi.org/10.1111/j.1651-2227.2012.02801.x
16. Koenig JE, Spor A, Scalfone N, Fricker AD, Stombaugh J, Knight R, Angenent LT, Ley RE. Succession of microbial consortia in the developing infant gut microbiome. Proc Natl Acad Sci U S A 2011; 108(Suppl 1):4578-85; PMID:20668239; http://dx.doi.org/10.1073/pnas.1000081107
17. Jacquot A, Neveu D, Aujoulat F, Mercier G, Marchandin H, Jumas-Bilak E, Picaud J-C. Dynamics and clinical evolution of bacterial gut microflora in extremely premature patients. J Pediatr 2011; 158:390-6; PMID:20961563; http://dx.doi.org/10.1016/j.jpeds.2010.09.007
18. Madan JC, Salari RC, Saxena D, Davidson L, O'Toole GA, Moore JH, Sogin ML, Foster JA, Edwards WH, Palumbo P, et al. Gut microbial colonisation in premature neonates predicts neonatal sepsis. Arch Dis Child Fetal Neonatal Ed 2012; 97:F456-62; PMID:22562869
19. Tita ATN, Andrews WW. Diagnosis and management of clinical chorioamnionitis. Clin Perinatol 2010; 37:339-54; PMID:20569811; http://dx.doi.org/10.1016/j.clp.2010.02.003
20. Barrett E, Kerr C, Murphy K, O'Sullivan O, Ryan CA, Dempsey EM, Murphy BP, O'Toole PW, Cotter PD, Fitzgerald GF, et al. The individual-specific and diverse nature of the preterm infant microbiota. Arch Dis Child Fetal Neonatal Ed 2013; 98:F334-40; PMID:23303303; http://dx.doi.org/10.1136/archdischild-2012-303035
21. Chang JY, Shin SM, Chun J, Lee J-H, Seo J-K. Pyrosequencing-based molecular monitoring of the intestinal bacterial colonization in preterm infants. J Pediatr Gastroenterol Nutr 2011; 53:512-9; PMID:21734604
22. Wang Y, Hoenig JD, Malin KJ, Qamar S, Petrof EO, Sun J, Antonopoulos DA, Chang EB, Claud EC. 16S rRNA gene-based analysis of fecal microbiota from preterm infants with and without necrotizing enterocolitis. ISME J 2009; 3:944-54; PMID:19369970; http://dx.doi.org/10.1038/ismej.2009.37
23. Mai V, Young CM, Ukhanova M, Wang X, Sun Y, Casella G, Theriaque D, Li N, Sharma R, Hudak M, et al. Fecal microbiota in premature infants prior to necrotizing enterocolitis. PLoS One 2011; 6:e20647; PMID:21674011; http://dx.doi.org/10.1371/journal.pone.0020647
24. Adlerberth I, Lindberg E, Aberg N, Hesselmar B, Saalman R, Strannegård I-L, Wold AE. Reduced enterobacterial and increased staphylococcal colonization of the infantile bowel: an effect of hygienic lifestyle? Pediatr Res 2006; 59:96-101; PMID:16380405; http://dx.doi.org/10.1203/01.pdr.0000191137.12774.b2
25. Stoll BJ, Hansen N. Infections in VLBW infants: studies from the NICHD Neonatal Research Network. Semin Perinatol 2003; 27:293-301; PMID:14510320; http://dx.doi.org/10.1016/S0146-0005(03)00046-6
26. Arboleya S, Binetti A, Salazar N, Fernández N, Solís G, Hernández-Barranco A, Margolles A, de Los Reyes-Gavilán CG, Gueimonde M. Establishment and development of intestinal microbiota in preterm neonates. FEMS Microbiol Ecol 2012; 79:763-72; PMID:22126419; http://dx.doi.org/10.1111/j.1574-6941.2011.01261.x
27. Morrow AL, Lagomarcino AJ, Schibler KR, Taft DH, Yu Z, Wang B, Altaye M, Wagner M, Gevers D, Ward DV, et al. Early microbial and metabolomic signatures predict later onset of necrotizing enterocolitis in preterm infants. Microbiome 2013; 1:13; PMID:24450576; http://dx.doi.org/10.1186/2049-2618-1-13
28. Butel M-J, Suau A, Campeotto F, Magne F, Aires J, Ferraris L, Kalach N, Leroux B, Dupont C. Conditions of bifidobacterial colonization in preterm infants: a prospective analysis. J Pediatr Gastroenterol Nutr 2007; 44:577-82; PMID:17460489; http://dx.doi.org/10.1097/MPG.0b013e3180406b20
29. Adlerberth I, Wold AE. Establishment of the gut microbiota in Western infants. Acta Paediatr 2009; 98:229-38; PMID:19143664; http://dx.doi.org/10.1111/j.1651-2227.2008.01060.x
30. Sim K, Cox MJ, Wopereis H, Martin R, Knol J, Li M-S, Cookson WO, Moffatt MF, Kroll JS. Improved detection of bifidobacteria with optimised 16S rRNA-gene based pyrosequencing. PLoS One 2012; 7:e32543; PMID:22470420; http://dx.doi.org/10.1371/journal.pone.0032543
31. Metsvaht T, Ilmoja M-L, Parm U, Maipuu L, Merila M, Lutsar I. Comparison of ampicillin plus gentamicin vs. penicillin plus gentamicin in empiric treatment of neonates at risk of early onset sepsis. Acta Paediatr 2010; 99:665-72; PMID:20096030
32. Keinänen-Toivola MM, Revetta RP, Santo Domingo JW. Identification of active bacterial communities in a model drinking water biofilm system using 16S rRNA-based clone libraries. FEMS Microbiol Lett 2006; 257:182-8; PMID:16553851; http://dx.doi.org/10.1111/j.1574-6968.2006.00167.x
33. Hamady M, Walker JJ, Harris JK, Gold NJ, Knight R. Error-correcting barcoded primers for pyrosequencing hundreds of samples in multiplex. Nat Methods 2008; 5:235-7; PMID:18264105; http://dx.doi.org/10.1038/nmeth.1184

34. Pruesse E, Quast C, Knittel K, Fuchs BM, Ludwig W, Peplies J, Glöckner FO. SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. Nucleic Acids Res 2007; 35:7188-96; PMID:17947321; http://dx.doi.org/10.1093/nar/gkm864

35. 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; 73:5261-7; PMID:17586664; http://dx.doi.org/10.1128/AEM.00062-07

36. Magurran AE. Chapter four: An index of diversity... In: Measuring biological diversity. Malden, MA: Blackwell Publishing; 2004: 100-130.

Publication III

Drell, T., Štšepetova, J., Simm, J., Rull, K., Aleksejeva, A., Antson, A., Tillmann, V., Metsis, M., Sepp, E., Salumets, A., and Mändar, R. (2017). The infuence of different maternal microbial communities on the development of infant gut and oral microbiota. Sci Rep 7, 9940.

SCIENTIFIC REPORTS

OPEN

The Influence of Different Maternal Microbial Communities on the Development of Infant Gut and Oral Microbiota

Received: 10 March 2017
Accepted: 25 July 2017
Published online: 30 August 2017

Tiina Drell[1,2,3], Jelena Štšepetova[2], Jaak Simm[3,10,11], Kristiina Rull[6,7], Aira Aleksejeva[6], Anne Antson[4,5], Vallo Tillmann[4,5], Madis Metsis[12], Epp Sepp[2], Andres Salumets[1,7,8,9] & Reet Mändar[1,2]

Very few studies have analyzed how the composition of mother's microbiota affects the development of infant's gut and oral microbiota during the first months of life. Here, microbiota present in the mothers' gut, vagina, breast milk, oral cavity, and mammary areola were compared with the gut and oral microbiota of their infants over the first six months following birth. Samples were collected from the aforementioned body sites from seven mothers and nine infants at three different time points over a 6-month period. Each sample was analyzed with 16S rRNA gene sequencing. The gut microbiota of the infants harbored distinct microbial communities that had low similarity with the various maternal microbiota communities. In contrast, the oral microbiota of the infants exhibited high similarity with the microbiota of the mothers' breast milk, mammary areola and mouth. These results demonstrate that constant contact between microbial communities increases their similarity. A majority of the operational taxonomic units in infant gut and oral microbiota were also shared with the mothers' gut and oral communities, respectively. The disparity between the similarity and the proportion of the OTUs shared between infants' and mothers' gut microbiota might be related to lower diversity and therefore competition in infants' gut microbiota.

The composition of gut and oral microbiota develops during the first years of an infant's life, with colonization of the gastrointestinal tract and oral cavity beginning immediately after birth. The most abundant colonizers in infant gut microbiota have been reported to include staphylococci, gammaproteobacteria (e.g., *Enterobacteriaceae*), and bifidobacteria[1,2]. In contrast, *Streptococcus* is dominant in infant oral microbiota[3]. Accumulating evidence has shown that several maternal factors (e.g. type of delivery and feeding regimen) influence the development of the infants' gut and oral microbiota[4,5]. For example, immediately after birth, yet prior to removal of the *vernix caseosa*, vaginally born infants acquire bacterial communities both in their gut and oral cavity that resemble their mothers' vaginal microbiota[6]. Similarly, infants born via caesarean section harbor communities that are similar to those found on their mothers' skin[6]. Thus, maternal microbial communities appear to be a key source of microbes during the initial colonization process of infant gut and oral microbiota[6]. However, it has not been sufficiently characterized how different types of maternal microbial communities affect the development of infant gut and oral microbiota during the first months of life after the initial colonization process has started.

[1]Competence Centre on Health Technologies, Tartu, Estonia. [2]Department of Microbiology, Institute of Biomedicine and Translational Medicine, University of Tartu, Tartu, Estonia. [3]Department of Gene Technology, Tallinn University of Technology, Tallinn, Estonia. [4]Department of Pediatrics, Institute of Clinical Medicine, University of Tartu, Tartu, Estonia. [5]Children's Clinic of Tartu University Hospital, Tartu, Estonia. [6]Women's Clinic of Tartu University Hospital, Tartu, Estonia. [7]Department of Obstetrics and Gynaecology, Institute of Clinical Medicine, University of Tartu, Tartu, Estonia. [8]Department of Obstetrics and Gynecology, University of Helsinki and Helsinki University Hospital, Helsinki, Finland. [9]Institute of Biomedicine and Translational Medicine, University of Tartu, Tartu, Estonia. [10]Department of Electrical Engineering (ESAT), STADIUS Center for Dynamical Systems, Signal Processing, and Data Analytics, KU Leuven, Leuven, Belgium. [11]iMinds Medical IT, Leuven, Belgium. [12]School of Natural Sciences and Health, Tallinn University, Tallinn, Estonia. Correspondence and requests for materials should be addressed to T.D. (email: tiinadrell@gmail.com)

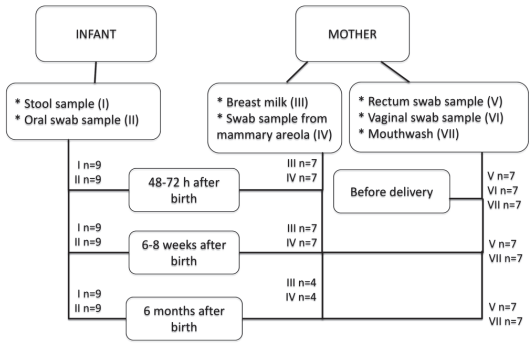

Figure 1. The number and type of samples that were collected from the mother-infant pairs at each of the time points indicated.

Therefore, the aim of this study was to compare the effect of different maternal microbial communities–intestinal, vaginal, oral, breast milk, and mammary areola–on the development of the infants' gut and oral microbiota during the first six months of an infant's life.

Material and Methods

Study group and sampling. Pregnant women were enrolled in this study during the first period of spontaneous onset of labor or up to 24 h before elective caesarean section with intact amniotic membranes. They were enrolled at the delivery ward of the Women's Clinic of Tartu University Hospital between May 2012 and September 2013. All the recruited women had uncomplicated term pregnancies, they did not have any infectious diseases that required antibiotic treatment, and they had no history of diabetes or hypertensive disorders during the second half of pregnancy. A total of seven mothers (mean age: 33.1 ± 4.5 y) were recruited with their newborn babies (mother-infant pairs). Two pairs included twins, and all four infants were enrolled in this study. Only one woman was not multiparous (mean parity, 1.4 ± 0.8). Five women gave birth via caesarean section and received prophylactic cefuroxime treatment prior to the incision in the skin. All infants were term (mean gestational age, 38.9 ± 1.6 weeks) and had normal birth weights (mean weight, $3{,}379.3 \pm 564.1$ g). Clinical factors describing the mother-infant pairs are provided in Supplementary Tables S1.

Six types of samples were collected from each mother-infant pair before birth or 48–72 h after birth, at 6–8 weeks and 6 months after birth. In addition, a vaginal swab sample was collected from each mother before giving birth. Altogether 139 samples were collected (Fig. 1). Prior to sampling, vaginal or skin disinfectant had not been used and the mammary areola had not been cleaned. All the collected samples were stored at −20 °C immediately after sampling.

Bacterioscopic smears were made from a separate vaginal swab sample that was collected from each mother before giving birth. The samples were Gram stained and scored according to the Nugent method[7]: 0–3 (normal), 4–6 (intermediate), and 7–10 (bacterial vaginosis (BV)).

Bacterial community profiling. DNA was extracted using MoBio PowerFecal DNA Isolation Kit (Mo Bio, Carlsbad, California, USA) according to the manufacturer's instructions. However, additional steps were included for the different types of samples. Briefly, all of the swab samples were initially vortexed for approximately 10 min in 750 μL sterile phosphate-buffered saline (PBS). 400 μL of the swab samples in PBS and 800 μL of mother's mouthwash were added into individual dry bead tubes. 1 ml of mother's breast milk was initially centrifuged at 6,000 rpm for 20 min. The supernatants were removed and each pellet was resuspended in 400 μL of PBS before being transferred to individual dry bead tubes. Tubes were then centrifuged at 13,200 rpm for 30 min. The supernatants were removed and the protocol for MoBio PowerFecal DNA Isolation Kit was followed from step 2. DNA from stool samples was extracted according to the manufacturer's protocol. Samples representing the same microbial community type formed a single DNA extraction batch. Extracted DNA was stored at −20 °C prior to analysis.

Amplification of the V1–V2 hypervariable region of 16S rRNA was performed by using barcoded universal 27F-YM and 357 R primers[8]. The primer sequences were as follows: 5′-CAA GCA GAA GAC GGC ATA CGA GAT NNNNNNNN AGA GTT TGA TYM TGG CTC AG-3′ (Illumina TruSeq adapter, sample-specific 8-bp index marked as Ns followed by 27F-YM at the 3′ end) and 5′-AAT GAT ACG GCG ACC ACC GAG ATC TAC ACC TGC TGC CTY CCG TA-3′ (TruSeq universal adapter sequence and 357 R at the 3′ end). The conditions for amplification included: 10 min at 98 °C, followed by 5 cycles of 30 s at 98 °C, 30 s at 55 °C, and 45 s at 72 °C, 30 cycles of 30 s at 98 °C and 60 s at 72 °C, and a final extension step at 72 °C for 10 min.

PCR reactions had a total volume of 20 μL, with 10 μL of Phusion High-Fidelity PCR Master Mix (Thermo Scientific, Waltham, Massachusetts, USA), 5 μL of DNA template, and each primer at the concentration of 0.2 μM. PCR products were purified using Agencourt AMPure XP (Beckman Coulter, Brea, California, USA) and were sequenced with an Illumina MiSeq system at the Genome Centre in University of Tartu, Estonia (single end sequencing using MiSeq v2 kit and 300 cycles).

Analysis of bifidobacteria with DGGE and real-time PCR was performed by using the methodology described in Supplementary methods.

Data and statistical analysis. MOTHUR software 1.27.0 was used to trim, denoise and align the sequences obtained to generate operational taxonomic units (OTUs) and to assign taxonomy. Sequences were trimmed and discarded based on a quality score <25 and length >225 bp, respectively. OTUs were generated by using an average neighbor hierarchical clustering algorithm with the identity threshold of 97%. Reference sequences for aligned 16S rRNA gene sequences were obtained from the SILVA ribosomal RNA database and comparisons with taxonomic assignments were performed with a Ňaive Bayesian classifier with a confidence cutoff of 90%. OTUs with less than two sequences and OTUs present in fewer than two samples were also discarded. To determine the most likely species name for pivotal OTUs mapped to genus *Lactobacillus*, an additional taxonomic assignment against the NCBI 16S ribosomal RNA sequences database was performed with BLASTN.

Statistical analyses were performed with R 3.4.0 software. The cut-off values of the minimal number of trimmed sequences assigned to the samples were set to individual community types after which the sequence counts were normalized for the whole dataset. To analyze general bacterial diversity and similarity between the samples, Shannon diversity index, Cosine similarity index (CSI) and Jaccard distance values were calculated, respectively. CSI measures the similarity between samples taking into account both the abundance and the prevalence of the OTUs (on the scale of 0 to 1 with 1 being the most similar) when Jaccard distance, which is a measure of dissimilarity, indicates to the differences between samples based only on the prevalence of the OTUs (on the scale of 0 to 1 with 1 being the most dissimilar). Continuous variables were compared with the Wilcoxon rank-sum test. The probability of an OTU present in infants' gut or oral microbiota being shared with a specific maternal community type was analyzed using logistic regression analysis. A factor representing the presence (or absence) of an OTU in both infant's and mother's community type was a dependent variable whereas the identifier for the pairs of samples (infant's stool or oral sample paired with mother's rectum, vaginal, breast milk, mammary areola or mouthwash sample) was a predictor. Only the dominant OTUs colonizing infants' gut and oral microbiota (relative abundance >0.005) were analyzed with separate logistic regression analyses.

All analyses were carried out with the Holm-Bonferroni correction. P-values less than 0.05 were considered to be statistically significant.

Ethical considerations. This study received approval from the Research Ethics Committee of the University of Tartu (no. 210T-7) and written informed consent was obtained from each participant upon admission to the delivery ward. All methods were performed in accordance with the relevant guidelines and regulations.

Results

Bacterial community profiling using 16S rRNA sequencing. A total of 1,106,448 high quality sequence reads were generated in this study. The cut-off set for the minimal number of trimmed sequences assigned to the samples varied from 200 to 1,400 depending on the community type (Supplementary Table S2). At these cut-off values the rarefaction curves for at least 90% of the samples reached a 5% plateau. In total, 135 samples and 1,530 OTUs exceeded these cut-off values and were further analyzed (GenBank accession numbers for representative sequences of the OTUs: KP117311-KP118840). One breast milk, one mother's oral and two infants' oral samples did not cross these cut-off values. The retrieved OTUs were distributed among 12 phyla, 89 families, and 170 genera.

The highest Shannon diversity index values were observed in the mothers' gut and oral microbiota and these communities differed significantly from the infants' gut and oral microbiota that harbored significantly lower microbial diversity ($p < 0.001$). The diversity of the microbial communities did not change significantly during the analyzed time period (Supplementary Figure S1).

Similarity between infants' gut and oral microbiota versus mothers' microbial communities from various sites. Based on the Cosine similarity index and Jaccard distance values, distinct patterns of similarity and dissimilarity were observed among the samples examined (Figs 2 and 3). These patterns were also concordant with the observed distribution of dominant OTUs between the community types (Fig. 4).

The gut microbiota of the infants did not share similarities with any of the analyzed maternal community types throughout the study, yet high similarity was observed between the infants' oral microbiota and the microbiota colonizing the mothers' breast milk, the mammary areola, and the oral cavity (Fig. 2). These four community types harbored a combination of dominant OTUs belonging mostly to genus *Streptococcus* that were present in each population with similarly high relative abundance (Fig. 4). Exceptions were the OTUs that mapped to the genera *Staphylococcus* and *Propionibacterium*. These OTUs had low relative abundance (<0.01) in both the infants' and mothers' oral microbiota.

Most of the OTUs observed in the infants' gut and oral microbiota were simultaneously shared with several of the maternal community types and we did not observe significantly higher probability of the OTUs dominating infants' gut and oral microbiota being shared with a specific maternal community type over other maternal community types. Only 20%, 26% and 28% of the OTUs colonizing the infants' gut, and 22%, 19% and 27% colonizing infants' oral microbiota at 48–72 h, 6–8 weeks and 6 months after birth, respectively, were observed exclusively in these community types. All of these OTUs had very low relative abundance (<0.01). Highest proportion of the OTUs observed in the pooled data of infants' gut microbiota (55%, 60% and 63%, respectively at 48–72 h, 6–8 weeks and 6 months after birth) were observed in the pooled data of mothers' gut microbiota (mainly members of *Clostridiales* and *Bacteroidales*, which represented more than 68% of the OTUs observed in both communities) (Fig. 5), but when analyzing individual mother-infant pairs, the proportion of the OTUs shared between an infant's and his or her own mother's gut microbiota was not that high (mean [SD] proportion of the OTUs shared

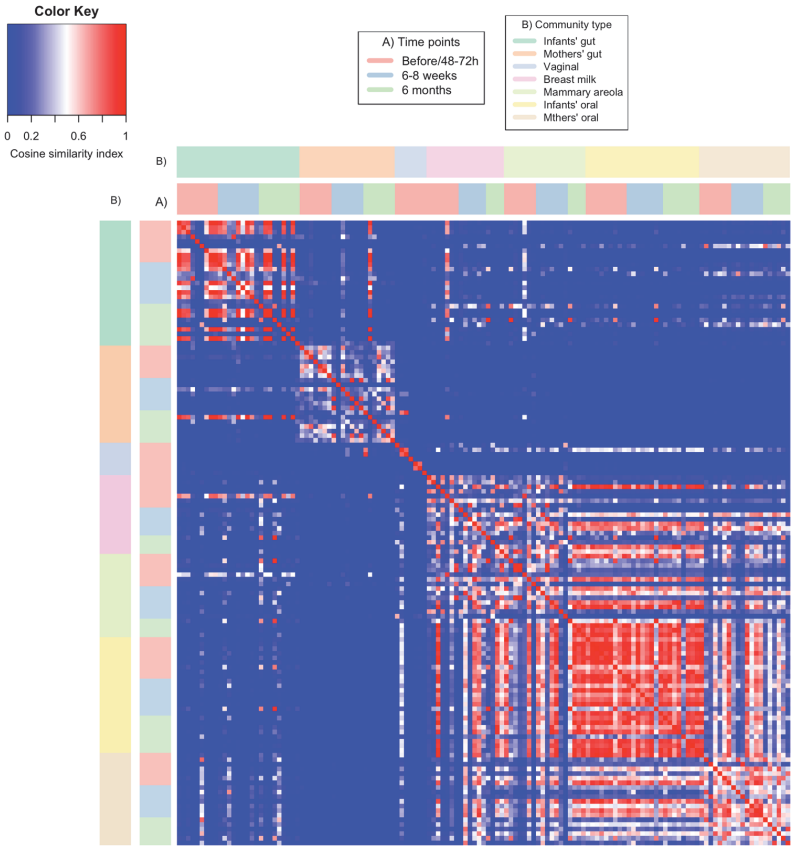

Figure 2. Correlation plot representing the Cosine similarity index values between the analyzed microbial community types. The time points represent the sampling times in relation to birth of the infants.

between mother-infant pairs: 32% [13%], 34% [19%] and 29% [11%], respectively at 48–72 h, 6–8 weeks and 6 months after birth). This proportion was similar to the proportion shared between infant's gut microbiota and the communities colonizing his or her mother's oral cavity (35% [12%], 28% [10%] and 39% [15%], respectively at 48–72 h, 6–8 weeks and 6 months after birth) and mammary areola (34% [18%], 19% [13%] and 35% [18%], respectively at 48–72 h, 6–8 weeks and 6 months after birth) (Fig. 6). In the infants' oral microbiota, the highest proportion of OTUs was shared with the mothers' oral microbiota (members of various taxa) when analyzed both between two community types (i.e. all samples pooled together) (51%, 61% and 48%, respectively at 48–72 h, 6–8 weeks and 6 months after birth; Fig. 5) and between individual mother-infant pairs (mean [SD] proportion of the OTUs 50% [12%], 55% [9%] and 46% [10%], respectively at 48–72 h, 6–8 weeks and 6 months after birth; Fig. 6).

The infants' gut and oral microbiota did not exhibit significantly greater similarity to their own mother's microbial communities than to the other mothers' microbial communities and the similarity did not change significantly during the study (Fig. 7 and Supplementary Figure S2).

The composition of the infants' gut and oral microbiota. *Firmicutes* and *Proteobacteria* were equally dominant in the infants' gut microbiota throughout the study (mean relative abundance [SD] was 0.42 [0.27] vs. 0.39 [0.3], respectively) (Supplementary Figure S3). On lower taxonomic levels, *Enterobacteriaceae* from phylum *Proteobacteria* dominated (0.37 [0.31]), while *Clostridiales* (mostly *Lachnospiraceae*), *Streptococcaceae*, and *Veillonellaceae* from phylum *Firmicutes* were also abundant (0.13 [0.2], 0.12 [0.24], and 0.1 [0.15], respectively). In contrast, *Bacteroidaceae* and *Bifidobacteriaceae* were relatively scarce (0.06 [0.14] and 0.05 [0.08], respectively). Real-time PCR detected high levels of bifidobacteria in the infants' gut microbiota throughout the study (average [SD]: 5,974 [15,208] copies per mg of feces; Supplementary Figure S4), and no significant fluctuations were observed. There were six bifidobacterial taxa observed in the infants' gut microbiota (*B. longum, B. longum subsp. infantis*, *B. longum subsp. longum, B. breve, B. pseudocatenulatum* and *B. adolescentis*), with an average of 3.6[2] bifidobacterial taxa colonized per infant. The highest number of bifidobacterial taxa was observed in the twin pairs, with six taxa identified in ID-s 201I and 201II and five taxa identified in ID-s 202I and 202II. Gut

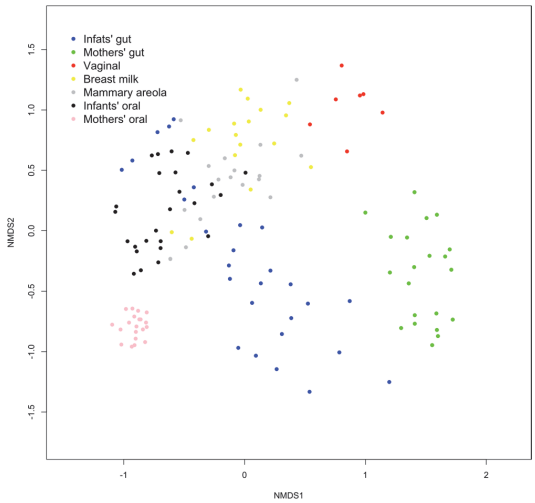

Figure 3. Non-metric multidimensional scaling (NMDS) plot visualizing the Jaccard distance between the samples of analyzed microbial community types.

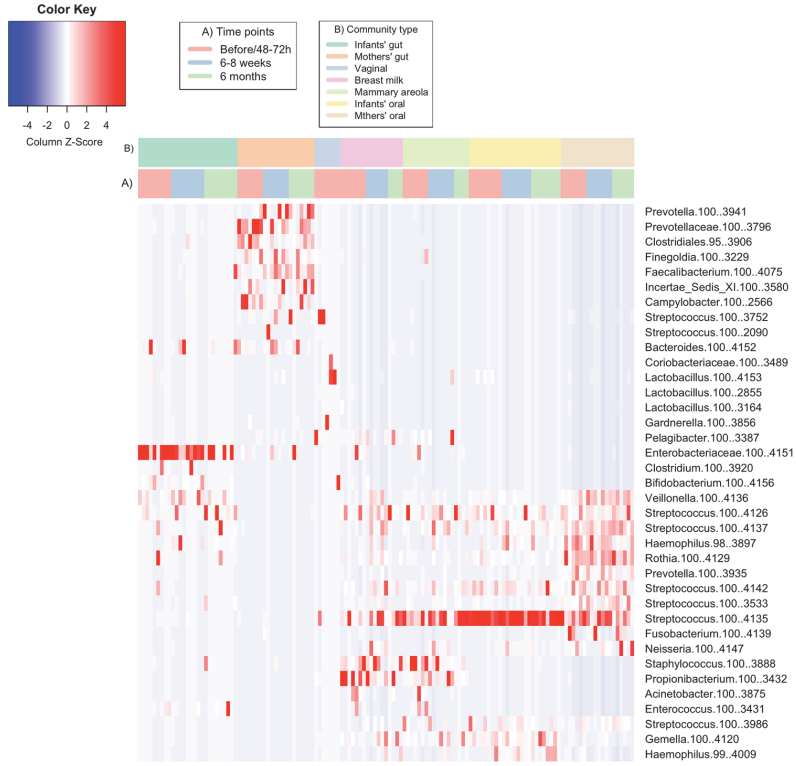

Figure 4. Relative abundance of the most dominant OTUs that colonized the analyzed community types.

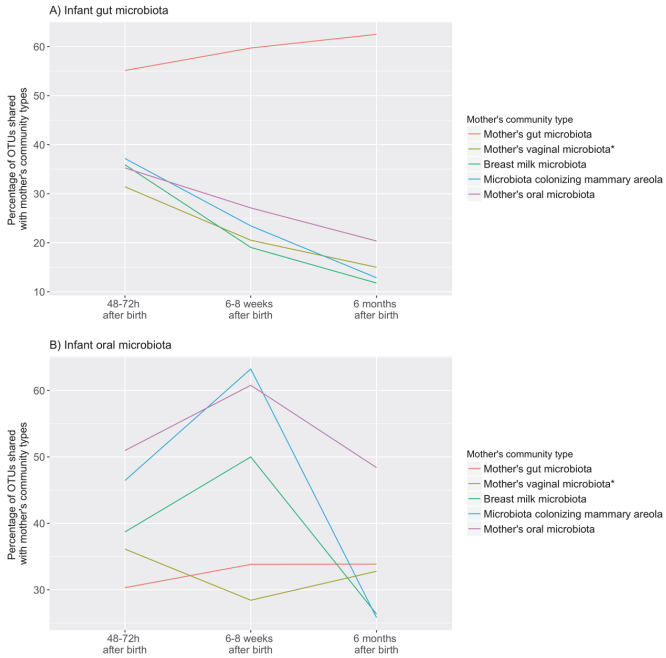

Figure 5. Proportion of OTUs observed in the pooled data of infants' gut (**A**) and oral (**B**) microbiota that were shared with mothers' community types. *Infants' gut and oral microbiota at every analyzed time point were compared to vaginal microbiota observed in mothers before giving birth.

microbiota was dynamic in most of the infants over the time period studied (Supplementary Figure S5-A). The composition was moderately similar among infants and it did not change significantly during the study period when considering both the CSI and Jaccard distance.

Firmicutes was also prominent in the infants' oral microbiota throughout the study (mean [SD] relative abundance, 0.82 [0.16]) (Supplementary Figure S3). On lower taxonomic levels, *Streptococcaceae* dominated (0.7 [0.18]), followed by *Pasteurellaceae* (0.08 [0.09]). The infants' oral microbiota was relatively stable during the time period studied (Supplementary Figure S6-C) and was very similar among the infants.

The gut and oral microbiota of the twins in this study did not exhibit a greater similarity among these paired siblings than was observed among the other infants (mean [SD] CSI: 0.35 [0.38] vs. 0.32 [0.34] and 0.76 [0.24] vs. 0.75 [0.22] for the gut and oral microbiota among twins vs. the rest of the infants, respectively). Also, gut and oral microbiota of infants born vaginally *versus* via caesarean section did not exhibit greater similarity (0.22 [0.16] vs. 0.35 [0.38] and 0.8 [0.12] vs. 0.73 [0.24] for the gut and oral microbiota among infants born vaginally vs. via caesarean section, respectively).

The composition of microbial communities colonizing the mothers. *Firmicutes* dominated all the maternal community types analysed, but *Bacteroidetes* and *Actinobacteria* were also very abundant in the gut and vaginal microbiota, respectively (Supplementary Figure S3). On lower taxonomic levels, the communities differed substantially with gut microbiota being dominated by order *Clostridiales* (0.4 [0.14]) along with family *Prevotellaceae* (0.2 [0.14]); oral microbiota by *Streptococcaceae* (0.4 [0.18]), followed by *Prevotellaceae, Micrococcaceae, Fusobacteriaceae*, *Pasteurellaceae*, and *Veillonellaceae* (0.09 [0.05], 0.09 [0.07], 0.08 [0.12], 0.08 [0.06], and 0.06 [0.03], respectively); breast milk microbiota and the microbiota colonizing mammary areola by *Streptococcus, Staphylococcus*, *Propionibacterium, Gemella, Acinetobacter*, and *Enterococcus* in both instances (0.32 [0.28], 0.12 [0.17], 0.09 [0.1], 0.05 [0.08], 0.03 [0.06] and 0.01 [0.04]; 0.48 [0.33], 0.13 [0.17], 0.05 [0.09], 0.05 [0.08], 0.03 [0.09] and 0.02 [0.05], respectively in each case). Gut and oral microbiota was relatively stable and generally similar between different women over the time period studied (Supplementary Figures S5-B and S6-D). In most of the women, the composition of breast milk microbiota and the microbiota colonizing mammary areola was dynamic during the time period studied (Supplementary Figures S6-A and S6-B). All the women had normal to intermediate Nugent scores (Supplementary Table S3) and the composition of vaginal microbiota varied among women (Supplementary Figure S7).

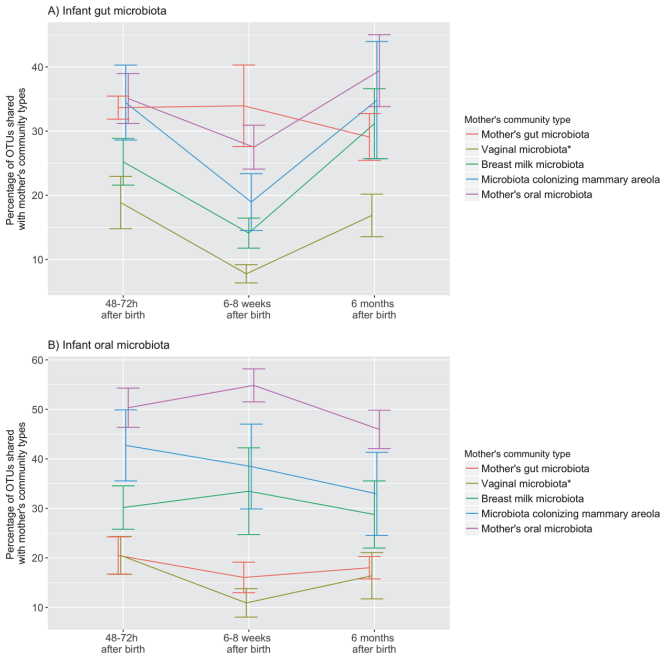

Figure 6. Average proportion with 95% confidence interval of OTUs observed in infants' gut (**A**) and oral (**B**) microbiota that were shared with his or her mother's community types (the proportion shared between specific mother-infant pairs). *Infants' gut and oral microbiota at every analyzed time point was compared to vaginal microbiota observed in mothers before giving birth.

Discussion

To our knowledge this is one of the first studies to analyze the effect of mothers' microbiota of various body sites to infants' gut and oral microbiota. Based on the mother-infant pairs that we examined, infant gut microbiota appears to harbor a distinctive microbial community that exhibits low similarity with the microbiota that colonize the mother's gut, vaginal, skin, breast milk, and oral cavity during the first six months of the infant's life. In contrast, the infants' oral microbiota, as well as the mothers' breast milk microbiota, mammary areola microbiota, and oral microbiota exhibited high similarity to each other.

Although approximately 63% of the OTUs observed in an infant's gut microbiota were also observed in the mother's gut microbiota (Fig. 5), only on average 32% were shared between individual mother-infant pairs, which was not considerably higher than the proportion of OTUs shared with the mother's oral and skin communities (Fig. 6). High numbers of OTUs observed in both infants' and mothers' gut microbiota that had considerably different relative abundances (Fig. 4) may be related to the lower species diversity (Supplementary Figure S1) and therefore competition in infants' gut microbiota as hypothesized by Asnicar *et al.*[9]. Relatively low levels of OTUs shared between infant's and one's own mother's gut microbiota and considerable numbers of OTUs shared between the infant's gut microbiota and microbiota colonizing one's own mother's oral cavity and mammary areola may on the other hand be related to the delivery mode of the infants participating in this study as most of them were born via caesarean section (Supplementary Table S1-B). Bäckhed *et al.*[10] have shown that 72% of the early colonizers of the vaginally delivered infants' gut matches species found in the stool of their own mother, whereas only 41% of these species are detected in infants born via caesarean section. They also observed enriched presence of bacteria typically know to be of skin and mouth origin colonizing the gut microbiota of C-section infants[10].

The infants' oral microbiota shared a high similarity with communities colonizing oral cavity, breast milk and mammary areola of the mothers (Fig. 2). Although the same OTUs dominated in all four aforementioned community types (Fig. 4), dominance of the same OTUs mapped to genus *Streptococcus* in infants' and mothers' oral microbiota indicates that mother's oral microbiota has the biggest influence on the development of infants' oral microbiota during the first six months of life, because members of *Streptococcus* are the predominant habitants of oral microbiota in both infants and adults[3]. Additionally, the highest number of OTUs colonizing the infants' oral microbiota was observed in the mothers' oral microbiota and this was also the case when analyzing the proportion of shared OTUs between individual mother-infant pairs (Figs 5 and 6). The similarity with oral microbiota may be the result of maternal habits of infant care (e.g., frequent use of the same spoon, licking the pacifier, kissing on the mouth). While not all mother-to-infant contacts involve direct interactions of oral microbiota, a

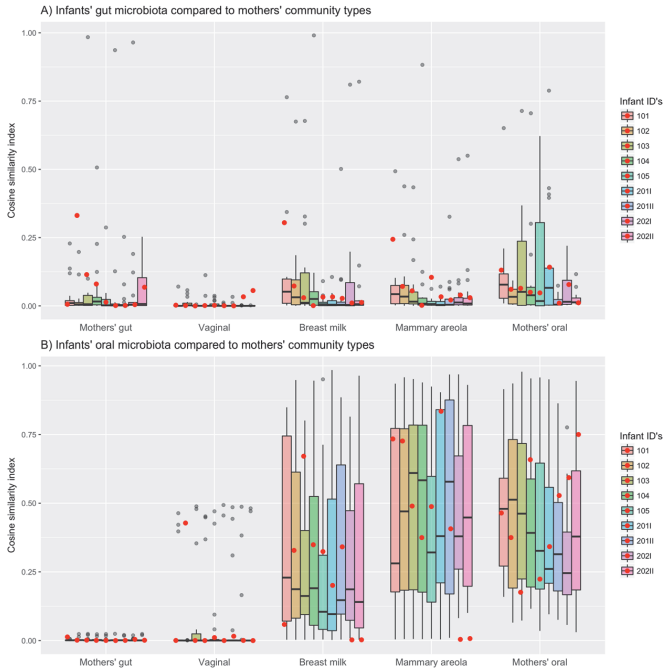

Figure 7. Average Cosine similarity index between infant's gut (**A**) and oral (**B**) microbiota and their own mother's microbial community types (red dots), and the community types observed in the rest of the mothers (Tukey boxplot).

similar observation was made in a recent study where the salivary microbiota of romantically involved partners exhibited increasing similarity when partners kissed at relatively high frequencies[11]. Mother's oral microbiota may also influence the development of an infant's oral microbiota via the placenta. For example, in a recent study by Aagaard *et al.*[12] mothers' oral microbes appeared to be present in the placenta[12].

High abundance of *Streptococcus* OTUs observed in communities colonizing breast milk and mammary areola and the abundant presence of *Staphylococcus* and *Propionibacterium*, which are typical colonizers of human skin[13], in breast milk microbiota (Supplementary Figure S6) indicates that there may be a retrograde reflux taking place during breastfeeding as have been hypothesized by Ansicar *et al.*[9]. Because mammary areolae were not cleaned before sampling the study may be overestimating the similarity between infants' oral microbiota and microbiota colonizing mammary areola.

We did not observe higher similarity of microbial communities between infants and their own mothers when compared to the other mothers (Fig. 7). This may seem surprising considering that several previous studies have observed mother-to-infant transmission of bacterial strains[9,10], but most probably our results indicate to the shortcoming of 16S rRNA gene sequencing in identifying taxa on lower than genus level[14]. Thus, our results do not rule out significant similarities between individual mother-infant pairs (and distinctively different community pattern from other mothers) when analysing the composition of the communities on strain level.

Another limitation of this study was the small sample size, which is most probably the reason why we did not observe the effect of type of delivery and feeding regimen to the community composition of infants' gut and oral microbiota. These effects are well described by many of the previous studies [e.g. refs 6, 10]. Nevertheless, the highest number of bifidobacterial taxa was observed among twins' gut microbiota who were breastfed for a shorter period of time and received prebiotic formula instead. This result is consistent with a study conducted by Barrett *et al.*[15] where the greatest number of bifidobacterial strains and diversity were observed in the infants who received formula containing prebiotics (e.g. galacto-oligosaccharides and polyfructose)[15].

In conclusion, the infants' gut microbiota was found to be dissimilar from all of the maternal community types which were analyzed in this study while the infants' oral microbiota exhibited a high similarity with the mothers' oral, breast milk, and mammary areola microbiota. These results emphasize the significant effect of constant contact between these microbial communities. However, both the infant gut and oral microbiota were found to share the highest proportion of the OTUs with the corresponding maternal community type. The disparity between the similarity and the proportion of the OTUs observed both in the infants' and mothers' gut microbiota might be related to lower species diversity and therefore lower competition in infants' gut microbiota, which leads to difference in the abundance of the shared OTUs.

References

1. Eggesbø, M. *et al.* Development of gut microbiota in infants not exposed to medical interventions. *APMIS.* **119**, 17–35 (2011).
2. Palmer, C., Bik, E. M., DiGiulio, D. B., Relman, D. A. & Brown, P. O. Development of the human infant intestinal microbiota. *PLoS Biol.* **5**, e177 (2007).
3. Cephas, K. D. *et al.* Comparative analysis of salivary bacterial microbiome diversity in edentulous infants and their mothers or primary care givers using pyrosequencing. *PLoS ONE.* **6**, e23503 (2011).
4. Adlerberth, I. & Wold, A. E. Establishment of the gut microbiota in Western infants. *Acta Paediatr.* **98**, 229–238 (2009).
5. Zaura, E., Nicu, E. A., Krom, B. P. & Keijser, B. Acquiring and maintaining a normal oral microbiome: current perspective. *Front Cell Infect Microbiol.* **4**, 85 (2014).
6. Dominguez-Bello, M. G. *et al.* Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns. *PNAS.* **107**, 11971–11975 (2010).
7. Nugent, R. P., Krohn, M. A. & Hillier, S. H. Reliability of diagnosing bacterial vaginosis is improved by a standardized method of gram stain interpretation. *J. Clin Microbiol.* **29**, 297–301 (1991).
8. McKenna, P. *et al.* The macaque gut microbiome in health, lentiviral infection, and chronic enterocolitis. *PLoS Pathog.* **4**, e20 (2008).
9. Asnicar, F. *et al.* Studying vertical microbiome transmission from mothers to infants by strain-Level metagenomic profiling. *mSystems.* **2**, e00164–16 (2016).
10. Bäckhed, F. *et al.* Dynamics and stabilization of the human gut microbiome during the first year of life. *Cell Host Microbe.* **17**, 690–703 (2015).
11. Kort, R. *et al.* Shaping the oral microbiota through intimate kissing. *Microbiome.* **2**, 41 (2014).
12. Aagaard, K. *et al.* The placenta harbors a unique microbiome. *Sci Transl Med.* **6**, 237ra65 (2014).
13. Kong, H. H. & Segre, J. A. Skin microbiome: looking back to move forward. *J. Invest Dermatol.* **132**, 933–939 (2012).
14. Konstantinidis, K. T. & Tiedje, J. M. Prokaryotic taxonomy and phylogeny in the genomic era: advancements and challenges ahead. *Curr Opin Microbiol.* **10**, 504–509 (2007).
15. Barrett, E. *et al.* The neonatal gut harbours distinct bifidobacterial strains. *Arch Dis Child Fetal Neonatal Ed.* **100**, F405–F410 (2015).

Acknowledgements

We thank Heili Varendi and Anneli Aus from Tartu University Children's Clinic for their technical assistance. This study was supported by the Estonian Research Council (grants No. IUT34-19 and IUT34-16), the Estonian Ministry of Education and Research (grant No. KOGU-HUMB), the European Union through the European Regional Development Fund (SFOS reg. No. 3.2.0701.11-0023), Enterprise Estonia (grant No. EU48695), and the Horizon 2020 innovation programme (WIDENLIFE, EU692065). Jaak Simm is funded by the KU Leuven Research Council (GOA/10/09 MaNet, CoE PFV/10/016 SymBioSys) and PhD/Postdoc grants from the Flemish Government (IOF, Hercules Stitching, iMinds Medical Information Technologies SBO 2014). Funding was also provided by Intel, Janssen Pharmaceutica, and the Institute for the Promotion of Innovation through Science and Technology in Flanders (IWT).

Author Contributions

Tiina Drell wrote the main manuscript text, prepared the figures and carried out majority of the experiments. Jelena Štšepetova carried out real-time PCR and DGGE experiments. Jaak Simm was responsible for bioinformatic analysis of the data. Kristiina Rull, Aira Aleksejeva and Anne Antson were responsible for the sample collection. Vallo Tillmann was a member of the team who designed the project and was responsible for the application of the ethics approval from the Research Ethics Committee of the University of Tartu. Madis Metsis was a member of the team who designed the project. Epp Sepp was a member of the team who designed the project and supervised the work. Andres Salumets and Reet Mändar were responsible senior scientists who supervised the work and also were members of the team who designed the project. All authors reviewed the manuscript.

Additional Information

Supplementary information accompanies this paper at doi:10.1038/s41598-017-09278-y

Competing Interests: The authors declare that they have no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Curriculum vitae

Personal data

Name: Tiina Drell
Date of birth: 22.11.1985
Place of birth: Tallinn, Estonia
Citizenship: Estonia

Contact data

E-mail: tiinadrell@gmail.com

Education

2010–2019 Tallinn University of Technology—PhD
2007–2009 MSC in Biomedicine, University of Tartu
2004–2007 BSC in Gene technology, University of Tartu
1992–2004 Viimsi High school

Language competence

English Fluent

Professional employment

2014–… Ministry of Social Affairs of Estonia, Adviser
2011–2014 BiotaTec LLC, Microbiologist
2010–2014 Competence Centre on Reproductive Medicine and Biology, Researcher
2009–2012 Tallinn University of Technology, Centre for Biology of Integrated Systems, Specialist

Elulookirjeldus

Isikuandmed

Nimi: Tiina Drell
Sünniaeg: 22.11.1985
Sünnikoht: Tallinn, Eesti
Kodakondsus: Eesti

Kontaktandmed

E-post: tiinadrell@gmail.com

Hariduskäik

2010–2019 Tallinna Tehnikaülikool – PhD
2007–2009 MSC biomeditsiini erialal, Tartu Ülikool
2004–2007 BSC geenitehnoloogia erialal, Tartu Ülikool
1992–2004 Viimsi Keskkool

Keelteoskus

Inglise keel – Kõrgtase

Teenistuskäik

2014–... Sotsiaalministeerium, nõunik
2011–2014 BiotaTec OÜ, Mikrobioloog
2010–2014 Reproduktiivmeditsiini TAK AS, Teadur
2009–2012 Tallinna Tehnikaülikool, Integreeritud süsteemide bioloogia keskus, Spetsialist