DISSERTATIONES
BIOLOGICAE
UNIVERSITATIS
TARTUENSIS
393

KREETE LÜLL

Investigating the relationships between human microbiome, host factors and female health

DISSERTATIONES BIOLOGICAE UNIVERSITATIS TARTUENSIS
393

DISSERTATIONES BIOLOGICAE UNIVERSITATIS TARTUENSIS
393

KREETE LÜLL

Investigating the relationships between human microbiome, host factors and female health

UNIVERSITY OF TARTU
Press
1632

Institute of Molecular and Cell Biology, Institute of Genomics, University of Tartu, Estonia

This dissertation is accepted for the commencement of the degree of Doctor of Philosophy in Gene Technology on March 30, 2022 by the Council of the Institute of Genomics and Institute of Molecular and Cell Biology, University of Tartu.

Supervisor: Elin Org, PhD
Estonian Genome Center, Institute of Genomics, University of Tartu, Tartu, Estonia.

Reviewer: Prof Angela Ivask, PhD
Chair of Genetics, Institute of Molecular and Cell Biology, University of Tartu, Tartu, Estonia

Opponent: Anne Salonen, PhD
Human Microbiome Research Program, Faculty of Medicine, University of Helsinki, Helsinki, Finland

Commencement: Room No. 105, 23B Riia St., Tartu, on May 13th, 2022, at 14:15.

The publication of this dissertation is granted by the Institute of Genomics and Institute of Molecular and Cell Biology at the University of Tartu, Estonia.

This research was funded by Estonian Research Council grants PUT 1371, PRG687 and PRG1076; EMBO Installation grant 3573; European Regional Development Fund project no.15-0012 GENTRANSMED; Estonian Center of Genomics/ Roadmap II project no.16-0125; Archimedes Foundation scolarships in smart specialization growth and Dora Plus scolarship for short-term study mobility; the Doctoral School of Biomedicine and Biotechnology scholarship. Data analyses were in part carried out at the High-Performance Computing Center of the University of Tartu, Estonia.

ISSN 1024-6479
ISBN 978-9949-03-862-6 (print)
ISBN 978-9949-03-863-3 (pdf)

University of Tartu Press
www.tyk.ee

"*It always seems impossible until it's done.*"

Nelson Mandela

TABLE OF CONTENTS

LIST OF ORIGINAL PUBLICATIONS

This thesis is based on the following original publications, referred to in the text by Roman numerals (Ref. I to Ref. III):

I Kurilshikov A, Medina-Gomez C, Bacigalupe R, Radjabzadeh D, Wang J, Demirkan A, Le Roy CI, Raygoza Garay JA, Finnicum CT, Liu X, Zhernakova DV, Bonder MJ, Hansen TH, Frost F, Rühlemann MC, Turpin W, Moon J-Y, Kim H-N, **Lüll K**, Barkan E, Shah SA, Fornage M, Szopinska-Tokov J, Wallen ZD, Borisevich D, Agreus L, Andreasson A, Bang C, Bedrani L, Bell JT, Bisgaard H, Boehnke M, Boomsma DI, Burk RD, Claringbould A, Croitoru K, Davies GE, van Duijn CM, Duijts L, Falony G, Fu J, van der Graaf A, Hansen T, Homuth G, Hughes DA, Ijzerman RG, Jackson MA, Jaddoe VWV, Joossens M, Jørgensen T, Keszthelyi D, Knight R, Laakso M, Laudes M, Launer LJ, Lieb W, Lusis AJ, Masclee AAM, Moll HA, Mujagic Z, Qibin Q, Rothschild D, Shin H, Sørensen SJ, Steves CJ, Thorsen J, Timpson NJ, Tito RY, Vieira-Silva S, Völker U, Völzke H, Võsa U, Wade KH, Walter S, Watanabe K, Weiss S, Weiss FU, Weissbrod O, Westra H-J, Willemsen G, Payami H, Jonkers DMAE, Vasquez AA, de Geus EJC, Meyer KA, Stokholm J, Segal E, Org E, Wijmenga C, Kim H-L, Kaplan RC, Spector TD, Uitterlinden AG, Rivadeneira F, Franke A, Lerch MM, Franke L, Sanna S, D'Amato M, Pedersen O, Paterson AD, Kraaij R, Raes J, Zhernakova A. 2021. **Large-scale association analyses identify host factors influencing human gut microbiome composition.** *Nat. Genet.* 53 (2), 156–165.

II **Lüll K**, Arffman RK, Sola-Leyva A, Molina NM, Aasmets O, Herzig K-H, Plaza-Díaz J, Franks S, Morin-Papunen L, Tapanainen JS, Salumets A, Altmäe S, Piltonen TT, Org E. 2021. **The gut microbiome in polycystic ovary syndrome and its association with metabolic traits.** *J Clin Endocrinol. Metab*. 106 (3), 858–871.

III **Lüll K**, Saare M, Peters M, Kakhiani E, Zhdanova A, Salumets A, Boyarsky K, Org, E. 2022. **Differences in microbial profile of endometrial fluid and tissue samples in women with IVF failure are driven by *Lactobacillus* abundance.** *AOGS* 101 (2), 212–220.

My contributions to the listed publications were as follows:

Ref. I Analyzed the Metabolic Syndrome in Men (METSIM) cohort microbiome data, wrote METSIM cohort description. Participated in consortium meetings, study design discussions and participated in the critical review of the paper.

Ref. II Participated in the study design, performed all of the data analysis, interpreted the results, prepared figures and tables, wrote the manuscript and participated in the critical review of the paper.

Ref. III Performed the DNA extraction of the samples, performed the data analysis, interpreted the results, prepared figures and tables, wrote the manuscript and participated in the critical review of the paper.

LIST OF ABBREVIATIONS

AUC	area under the curve
BA	bile acid
BMI	body mass index
DNA	deoxyribonucleic acid
DZ	dizygotic
EF	endometrial fluid
ET	endometrial tissue
FAI	free androgen index
FMT	fecal microbiota transplant
FPR	false positive rate
GWAS	genome-wide association study
HbA1c	glycated hemoglobin
IR	insulin resistance
IVF	*in vitro* fertilization
LPS	lipopolysaccharide
mbBTL	microbiome binary trait loci
mbQTL	microbiome quantitative trait loci
METSIM	METabolic Syndrome in Men
MWAS	microbiome-wide association study
MZ	monozygotic
ND	no data
NFBC1966	the Northern Finland Birth Cohort 1966
NGS	next-generation sequencing
NGT	normal glucose tolerance
OGTT	oral glucose tolerance test
PCOS	polycystic ovary syndrome
ROC	receiver operating characteristic
rRNA	ribosomal ribonucleic acid
SHBG	sex hormone binding globulin
SNP	single nucleotide polymorphism
TPR	true positive rate
T2D	type 2 diabetes
WD	Western diet
WGS	whole genome sequencing

INTRODUCTION

Over the last decade, with the becoming of new sequencing technologies and data analysis methods, the field of microbiomics has made an immense progress allowing to study microbiome in great detail. Studies of the microbiome have revealed that every human has a unique microbial fingerprint and the variability of it between individuals is enormous. It has become evident that the microbiome, especially the gut microbiome, does not affect only the health of gastrointestinal tract but has a profound impact on the overall health and wellbeing. Links between microbiome and numerous health conditions, including type 2 diabetes, types of cancer and even psychological disorders, are found. Studies have also made apparent that the human genetics has an impact on the microbiome and certain microbes seem to be heritable. By widening our knowledge on the role of microorganisms on different disease states and understanding the mechanisms through which microorganisms affect human health, it could be possible to incorporate that information into treatment regimen for more efficient outcomes.

The current thesis focuses on the factors influencing human gut microbiome composition as well as understanding the relationships of microbiome and female reproductive health. In the first part of the thesis, I will cover the basis of microbiome as an important cofounder to the human health based on the literature: I will give an overview of the intrinsic as well as extrinsic factors affecting the human microbiome and its variability; matters influencing the microbiome studies; relationships of genetics and microbiome and future perspectives of microbiome research. I will also go more into detail on the topics of female reproductive health and microbiome where I discuss on the importance of studying not only gut microbiome but also microbiome present in the female reproductive tract. In the experimental part, I will focus on characterizing the associations between gut microbiome and host genetics through which the value of large collaborations between study cohorts working as a big consortium becomes evident. Additionally, I will investigate the role of microbiome in female health by taking a closer look at the interplay between polycystic ovary syndrome (PCOS) and gut microbiome as well as investigating the microbiome of endometrium in women with fertility problems.

1. REVIEW OF THE LITERATURE

1.1. Human microbiome

Human organism is home for millions of different microorganisms living both inside and on the human body. These microorganisms are bacteria, archaea, viruses and fungi – referred together as microbiota. Therefore, the term "microbiota" stands for a range of microorganism that could be symbiotic, commensal, or pathogenic. In scientific papers another term is widely used – "the microbiome" which stands for the collection of genes and genomes of microbiota. Despite there being a difference between the two terms, a consensus definition for them among researchers remains debatable and the two terms are used interchangeably (Marchesi and Ravel, 2015). In the current thesis the term "microbiome" is used.

Our current knowledge on the microbes has deepened greatly with the coming of next-generation sequencing (NGS) allowing us to take a step forward from culture-based methods. This gives an opportunity to study anaerobic bacteria more in depth – something that was limited with culture-based approach. Broadly, human as a host provides two different types of living environment for the bacteria. To start with, the surface of human body, the skin, is colonized by numerous aerobic and facultative anaerobe bacteria (Byrd et al., 2018). Since skin is a an acidic, cool and rather nutrient poor, it is considered as a relatively harsh environment for the bacteria to live in (Byrd et al., 2018). Secondly, the bacteria inhabit the inside of human body. Mostly anaerobic bacteria inhabit different internal host environments such as the gut, oral cavity as well as vagina, and the microbiome in all of these differs from one another (Gilbert et al., 2018). To date, the most researched part of the microbes inhabits the human intestine. These microbes intermediate key immune, metabolic and physiological functions, and changes in their ecosystem can greatly influence human health and disease (Chow and Mazmanian, 2010; Fan and Pedersen, 2021; Ley et al., 2006). Despite gut being the most studied area of microbes in the human body, our knowledge of bacteria living in other ecosystems of the human organism is on the rise. There is progressively more research done with samples collected from specific sites of the human body, giving us the opportunity to search for links between diseases and microbes present at the disease site. Furthermore, with these approaches we could determine the microbial profile of the disease affected area in the body which could in turn provide insights into the pathophysiology of the disease.

At the same time, it is with the utmost importance to keep in mind that the microbiome has a wide interindividual variability and it is affected by various factors. These drivers of microbiome remain essential knowledge for microbiome research.

1.1.1. Variability of gut microbiome

The microbiome is assembled at birth and it is affected by various factors throughout the lifespan. One of the biggest resources in finding the variability affecting the gut microbiome have been different population studies. These have been a tremendous source providing crucial information on what could be the main factors influencing gut microbiome and providing a basis for further and more in-depth analysis on specific topics. Recent population studies from Estonia, Finland, United Kingdom, Israel, Belgium, and the Netherlands with sample sizes ranging from 1,000 to 8,000 have identified numerous factors that shape the microbiome such as biological gender, age, lifestyle factors, medication and diseases (Aasmets et al., 2022; Falony et al., 2016; Gacesa et al., 2020; Jackson et al., 2018; Rothschild et al., 2018; Salosensaari et al., 2021; Woo and Alenghat, 2017; Zhernakova et al., 2016). Interestingly, the found associations are able to explain only a rather small portion, approximately 15%, of the gut microbiome compositional variability (Gacesa et al., 2020).

As stated, the gut microbiome composition can be affected by a number of factors (Figure 1). Prominently, one of the major elements in gut microbiome variability is diet since it has a direct link with the gut (Albenberg and Wu, 2014). Diet is the main source of nutrients for the gut microbes and the nature of the diet determines the dominant types of gut bacteria. Whereas the digestive system can efficiently extract energy from fats and proteins, for a significant portion of carbohydrates, especially those of plant origin, the digestive system needs to work together with gut microbes in the digestive process. Through the digestive process, numerous important metabolites are being produced such as short-chained amino acids, bile acids (BAs), trimethylamine and more – all which directly or indirectly affect the human biology and are related to the development of different diseases including obesity, cardiovascular disease, and type 2 diabetes (T2D) (Dalile et al., 2019; Fan and Pedersen, 2021; Rowland et al., 2018). Furthermore, dietary changes modulate the composition, function and diversity of gut microbiome. It has been shown that an extreme short-term change in dietary habits can significantly change the gut microbiome consumption as well as make it more similar with people consuming the same diet. At the same time, long-term diet has been shown to have a large impact on gut microbiome composition (David et al., 2014; Wu et al., 2011). Over the few past decades, the dietary habits of mainly Western populations have changed immensely moving towards an increase in consuming simple carbohydrates and animal fats, leading to changes in the gut microbiome and resulting in upsurge in the incidence of metabolic diseases (Martinez et al., 2017).

Besides diet, one of the most acknowledged influencers of the gut microbiome are medications. Understanding the interaction between medications and microbiome is crucial for understanding drug mechanisms and development of possible side effects they may present (Doestzada et al., 2018; Vich Vila et al., 2020). Together with the discovery of penicillin began the era of antibiotics – drugs that without a doubt have an enormous impact on the medical field with their ability

to fight infections. Yet, as the time went by it became evident that antibiotic treatment, especially with broad-spectrum antibiotic, lowers the overall gut microbial diversity and changes the microbiome composition (Aasmets et al., 2022; Konstantinidis et al., 2020; Ramirez et al., 2020). It has been interesting to learn that the antibiotics taken early in life have a profound long-term effect on health which could eventually lead to development of several health complications such as allergies, obesity, irritable bowel syndrome and other diseases (Neuman et al., 2018). An *in vitro* study screening drug effects against human gut bacteria showed that 24% of human-targeted drugs have the effect of inhibiting the growth of gut bacteria and it was estimated that the actual percentage might be even higher (Maier et al., 2018). Interestingly, the same study discovered that the more prevalent gut microbes are impacted more by the pharmaceuticals than the less abundant bacteria (Maier et al., 2018). Another recent study revealed that not only the drug intake, but also medication combinations and dosage along with previous antibiotic history is needed to take into consideration to fully understand the medication-host-microbiome interactions in complex diseases (Forslund et al., 2021).

Additionally to aforementioned factors, it has been shown that living environment, having pets, physical activity, age, diseases and sleep all have a profound role on microbiome. Next to diet, exercise has been linked strongest to microbial composition (Walker et al., 2021). It has been suggested that exercising reduces the inflammation in the body which in turn changes the microbiome structure and this process works through the changes in cytokine profile: pro-inflammatory cytokines are reduced and anti-inflammatory cytokines elevated (Cook et al., 2016). When studying the role of exercise in microbiome health, it is crucial to keep in mind that normally people who exercise more tend to also eat healthier than those with more sedentary lifestyle which makes research on that topic more complex. Therefore, it can be challenging to pinpoint exactly which changes in the gut microbiome are derived by exercise and which by diet since they could be closely intertwined.

Analyses of population-based data have shown that populations themselves differ in their microbial composition. Similarly to genetics, it is important to note that associations seen on for example European population should not automatically be transferred to populations with different ethnical background. For that reason, it is necessary to perform large meta-analysis that would incorporate populations from all over the world. Increase in sample size would allow incorporate traits in the analysis which would possibly be underpowered in studies with 1,000 to 2,000 samples simply due to lacking sufficient number of cases to analyze. Doing this would help to identify universal factors influencing microbiome just like large consortia genome-wide association studies (GWASs) are doing in genetics. Additionally, longitudinal studies are needed to look at the microbiome variability among the same individuals during a longer period of time. As an example, a longitudinal study in 338 individuals spanning over the period of 4 years showed how genetic stability of gut microbes varies across species and that the gut microbial composition is more stable in individuals with

higher baseline diversity (Chen et al., 2021). In the sense of gut microbiome variability these are highly important findings since so far research monitoring the microbiome changes in the same person over time are lacking. On the brighter side, recently a large-scale longitudinal study by Israelis was introduced. The study is expected to recruit 10,000 participants for whom an extensive phenotype, microbiome, genetic, as well as environmental data are and will be collected with follow-ups performed for 25 years. This is a highly anticipated prospective which would without a doubt greatly benefit the microbiome field (Shilo et al., 2021).

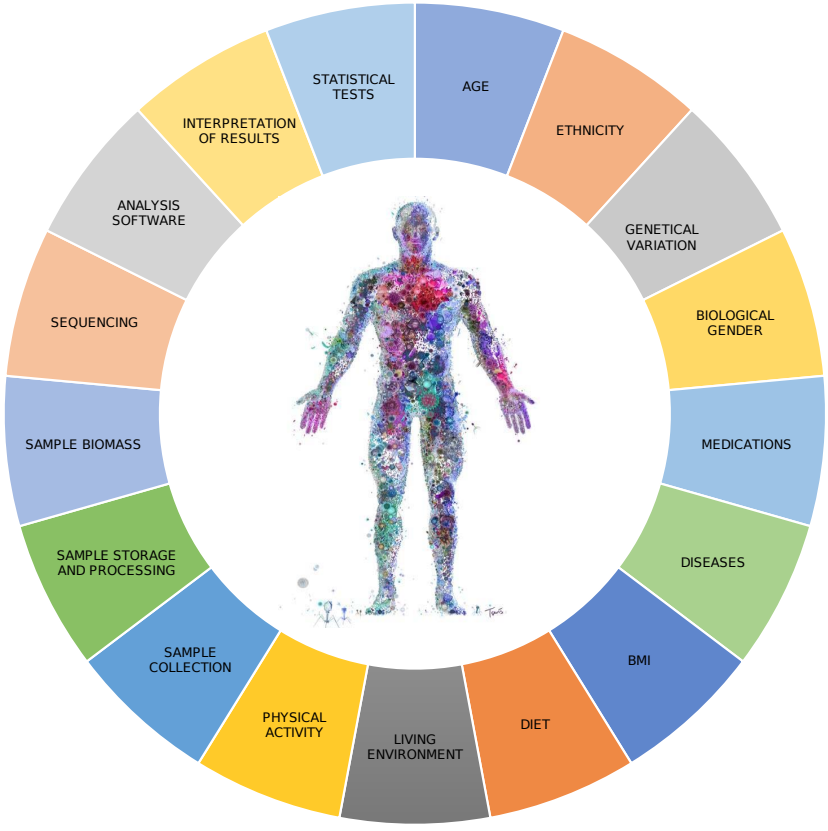


Figure 1. Factors influencing the microbiome variability and its analysis. Figure adapted from https://i0.wp.com/dirt.asla.org/wp-content/uploads/2018/06/microbiome.jpg?ssl=1.

1.1.2. Gut microbiome and host genetics

It is widely acknowledged that gut microbiome composition depends highly on environmental factors with one of the strongest effectors being diet and medication (Falony et al., 2016; Zhernakova et al., 2016). However, population-based studies reveal that a big proportion of microbial diversity continues to be unexplained after bearing in mind the environmental factors. This in turn raises the question on the role of host genetics on the gut microbiome (Wang et al., 2018).

Heritability is a measure that estimates the contribution of genes to phenotypic variance, it is a relative value that depends on the degree of variation a specific population has. The most traditional method to estimate heritability of a trait is by comparing the dissimilarities between monozygotic (MZ) and dizygotic (DZ) twins. Assuming that the twins are sharing the same environment and knowing that MZ twins share 100% of their genetics and DZ 50%, estimations on the heritability of a trait can be done directly. One of the first and largest twin studies on microbiome heritability was done on the TwinsUK cohort which estimated that the microbiome heritability lies between 1.9% and 8.1% (Goodrich et al., 2014, 2016; Rothschild et al., 2018). Several heritable taxa belonging to phylum Actinobacteria, Euryarchaeota, Firmicutes, and Tenericutes were identified, whereas Bacteroidetes which is one of the most dominant phylum in the gut showed very little heritability (Goodrich et al., 2016). For a subset of TwinsUK cohort a metagenomics sequencing was performed and the results showed not only heritable taxa but also high heritability of microbial gene ontology groups such as branched-chain amino acids biosynthesis and the module for sulfur reduction (Xie et al., 2016). On the whole, twin studies reveal that the role of host genetics in determining microbiome composition is small, and that genetic descent does not have a significant association with the gut microbiome (Rothschild et al., 2018). Inversely, significant similarities in their microbiome profile were detected in individuals sharing a household but being genetically unrelated, showing that environment has a far greater importance in shaping the gut microbiome than genetics (Rothschild et al., 2018). However, the observations seen in twin studies have been a driving force for population-based microbiome-wide association studies (MWASs) that identify genetic variants associated with microbiome composition.

The MWASs are essentially GWASs where instead of phenotype data microbiome data is used. In MWASs the abundance of the microbes can be used as a quantitative or binary trait. To this day several MWASs have examined the possible effect genetics might hold on microbiome and numerous microbial quantitative trait loci (mbQTLs) situated in genes associated with food and drug metabolism, immune response and irritable bowel syndrome have been identified (Table 1) (Blekhman et al., 2015; Bonder et al., 2016; Davenport et al., 2015; Goodrich et al., 2016; Hughes et al., 2020; Kolde et al., 2018; Kurilshikov et al., 2021; Lopera-Maya et al., 2022; Qin et al., 2022; Rothschild et al., 2018; Rühlemann et al., 2021; Turpin et al., 2016; Wang et al., 2016). During MWASs mostly three different types of association analysis are being done: a) quantitative that is searching associations between the abundance of taxa (usually common taxa with relative abundance of > 0.001) and genetic loci; b) binary analysis in which only absence or presence of a taxa is looked at, and c) associations with diversity. However, the reproducibility of these early MWASs has been limited and most reported associations tend to lose significance after correcting for multiple testing (Kurilshikov et al., 2017). The most common limitations to reproducibility are modest sample sizes, differences in the data processing methods as well as strong environmental effects. Additional limitation of these studies is the use of 16S

ribosomal ribonucleic acid (rRNA) sequencing which itself present a bias to the studies since it depends on which hypervariable region was used. Furthermore, 16S does not allow bacterial identification on species level or identification of bacterial pathway abundances (Lopera-Maya et al., 2022).

Recently, two large population-based studies in Finnish and Dutch cohorts using metagenomics have been published (Lopera-Maya et al., 2022; Qin et al., 2022). Similarly to previous studies with 16S rRNA sequencing methodology, the strongest host-microbiota associations found were with the *LCT* and *ABO* genes (Table 1) (Bonder et al., 2016; Goodrich et al., 2016; Lopera-Maya et al., 2022; Qin et al., 2022; Rothschild et al., 2018). However, if previously only genera were associated with these genes, then now with the usage of metagenomics approach specific species as well as pathways associated with *LCT* and *ABO* were identified. Additionally to the two well-known loci, *LCT* and *ABO*, other associated genetic regions harbor genes associated with different immune and metabolic phenotypes, hence providing interesting links with microbiome and diseases that can be studied deeper in future research. Unfortunately, similarly to the previous studies the replication of findings between the Finnish and Dutch studies was small. Besides the previously named two loci, only 3 out of 451 genome-wide significant single nucleotide polymorphisms (SNPs) from the Finnish study were replicated in the Dutch cohort (Lopera-Maya et al., 2022). This again indicates the need for larger studies and indicating that even studies with more than 7,000 samples suffer from power issues. Indeed, power studies added to the Dutch cohort research revealed that sample set with more than 50,000 individuals would be needed for studying bacteria present in at least 20% of the samples in order to identify associations with effect sizes similar to *LCT* and this number would increase even more when wanting to study rarer bacteria (Lopera-Maya et al., 2022).

Taken together, twin and population-based studies are a great source of information in studying the interplay between microbes and host genetics. However, similarly to genetic studies investigating traits with low heritability a great teamwork between research groups is needed in order to reach the needed power for new discoveries (Bonder et al., 2016; Lopera-Maya et al., 2022; Qin et al., 2022; Turpin et al., 2016; Wang et al., 2016).

Table 1. Summary list of main findings between genetic loci and bacterial taxa in microbiome-wide studies in humans.

Study	Population and sample size	Sequencing technology	Number of findings with bacterial taxa	Main findings
Blekhman et al., 2015	mixed (USA) n = 93	WGS	83 suggestive associations with bacterial taxa	Variants in *LCT* gene in a suggestive correlation with *Bifidobacterium*.
Davenport et al., 2015	Hutterites (North America) n = 127	16S rRNA	8 taxa associated with human genome	Variants in *PLD1* correlated with the genus *Akkermansia*.
Bonder et al., 2016	Dutch n = 1,514	WGS	9 loci associated with bacterial taxa	*LCT* associated with *Bifidobacterium*. Most significant associations were with *Blautia*, *Dialister invisus*, *Bacteroides xylanisolvens*, *Methnaobrevibacteri*. 33 loci were associated with bacterial pathways including pathways for bile acid metabolism, steroid degradation, sulfuric ester hydrolase activity.
Goodrich et al., 2016	United Kingdom n = 1,126	16S rRNA	28 loci associated with bacterial taxa	*LCT* associated with *Bifidobacterium*, as well as variants in *R3HDM1* which is in a strong linkage disequilibrium with *LCT*.
Turpin et al., 2016	Canadian n = 1,098; USA and Israel (replication) n = 463	16S rRNA	58 SNPs associated with the relative abundance of 33 taxa	4 taxa replicated in the replication cohort: variant in gene *UBR* associated with family *Rikenellaceae*, *CNTN6* associated with *Faecalibacterium*, SNP in *DMRTB1* associated with *Lachnospira* and variant in *SALL3* associated with *Eubacterium*.
Wang et al., 2016	German n = 1,812	16S rRNA	54 significant associations involving 40 loci and 22 bacterial traits	Of the 22 bacterial traits, the largest number belonged to Firmicutes, Proteobacteria, Bacteroidetes and Actinobacteria, at the phylum level.
Kolde et al., 2018	North America n = 298	WGS	5 species out of 118 were significantly correlated with the first genomic principal component	*FUT2* and *LCT* are associated with *Bifidobacterium longum*.

Study	Population and sample size	Sequencing technology	Number of findings with bacterial taxa	Main findings
Rothschild et al., 2018	mixed (Israel) n = 1,046	16S rRNA	7 associated SNPs found	2 out of the 7 found associations are near the *LCT* gene.
Hughes et al., 2020	Belgian n = 2,223; German n = 1,667	16S rRNA	13 loci associated with bacterial taxa	2 associations reached study-wide threshold: *Ruminococcus* with *RAPGEF1* and *Coprococcus* with *LINC01787*.
Kurilshikov et al., 2021	mixed (24 cohorts) n = 18,340	16S rRNA	31 loci associated with bacterial taxa	1 association passed the study-wide threshold: a variant in the *LCT* gene and *Bifidobacterium*. Other associations found remained suggestive.
Rühlemann et al., 2021	German n = 8,956	16S rRNA	17 loci associated with bacterial abundance	Associations of *ABO* histo-blood groups and *FUT2* with *Bacteroides* and *Faecalibacterium* spp.
Lopera-Maya et al., 2022	Dutch n = 7,738	WGS	24 loci associated with bacterial taxa and pathways	*LCT* and *ABO* associated with multiple microbial taxa as well as pathways. *LCT* was associated with decreased abundances of the species *Bifidobacterium adolescentis* and *Bifidobacterium longum*. Associations found at *ABO* locus include *Bifidobacterium bifidum* and *Collinsella aerofaciens*.
Qin et al., 2022	Finnish n = 5,959	WGS	471 taxa associated with at least one genetic variant	3 loci passed study wide significance threshold: *LCT* associated with *Bifidobacterium* and other taxa; *ABO* associated with *Faecalicatena lactaris*; *MED13L* associated with *Enterococcus faecalis*.

Abbreviations: WGS, whole genome sequencing; rRNA, ribosomal ribonucleic acid; *LCT*, Lactase; *PLD1*, Phospholipase D1; *R3HDM1*, R3H Domain Containing 1; *UBR*, Ubiquitin Protein Ligase E3 Component N-Recognin 1, *CNTN6I*, Contactin 6; *DMRTB1*, DMRT Like Family B With Proline Rich C-Terminal 1; *SALL3*, Spalt Like Transcription Factor 3; *FUT2*, Fucosyltransferase 2, *RAPGEF1*, Rap Guanine Nucleotide Exchange Factor 1; *LINC01787*, Long Intergenic Non-Protein Coding RNA 1787; *ABO*, Alpha 1-3-N-Acetylgalactosaminyltransferase And Alpha 1-3-Galactosyltransferase; *MED13L*, Mediator Complex Subunit 13L.

1.1.3. What matters in microbiome analysis

One of the biggest challenges in investigating the role of microbiome in human diseases is the low concordance between studies. This limits the ability to detect causal associations between microbes and pathology. The wide interindividual heterogeneity in microbiome composition compounds to the risk of finding false positives. Benchmarking process in microbiome studies is a difficult task to perform due to the large variability it has among individuals but also the usage of different study methods has an enormous impact on the results (Figure 1). To this day there is no golden mean to follow when performing microbiome analysis and new approaches are frequently being introduced to the field.

When talking on the topic of microbiome analysis, it is clear that various data pre-processing steps might not produce overlapping results (Hornung et al., 2019; Nearing et al., 2021). The first crucial steps in microbiome analysis is the sample collection and storage. At the time, the most used sampling measure for microbiome studies is stool sample which is considered as a gut proxy. However, the issue with stool samples is that while it describes well the microbial population of descending colon, it lacks the capability of fully describing the microbiome of gut epithelia and small intestine. In order to study the full microbial composition of the whole gut, various parts of the gut should be sampled using invasive procedures (Bajaj et al., 2012). Sample collection process highly depends on the research it aims. Large-scale population projects normally use remote sample collection where the sample is collected at home by the participant and sent to the research facility afterwards, whereas in a small research project focusing on a very specific study question sample collection process normally takes place in a clinic or study center (Vandeputte et al., 2017). In order to collect the stool sample, several commercially available collection tubes with easy collection protocol are available. Some of those allow for a short-term storage at room temperature, however it has been shown that this could potentially lead to changes in the microbiome composition in the sample (Penington et al., 2018). Many of these kits also contain a preservation liquid which are added in order to stabilize the sample and avoid compositional changes in the sample. Yet, studies have shown that there are significant differences in bacterial composition resulting from collecting stool samples in commercial tubes compared to instant freezing (Jones et al., 2021). The most suggested way in preserving a stool sample is freezing it immediately at –20 °C at the participants home and once the sample arrives to the study center it should be stored at –80 °C to avoid any growth of aerobic bacteria that would not be present in the gut (Jones et al., 2021).

Besides sample collection and storage techniques the next important step is DNA extraction for which there are numerous commercial kits available. It has become largely acknowledged that each extraction kit has its own so called kitome meaning that the kit itself might comprise some microbes. This kind of reagent contamination is especially important in low-biomass samples, since the kitome could have a larger impact on the observed community than the biological effect of interest (Debelius et al., 2016). DNA extraction kits by different

manufacturers use various protocols where the steps used for cell lysis vary greatly (e.g. usage of different enzymes such as lysozyme, mutanolysin or using mechanical bead beating) which in turn could result in detecting different overall microbial composition (Zhang et al., 2019a).

When it comes to the sequencing process, there are two that are widely used for sequencing in microbiome analysis. The first one being 16S rRNA gene sequencing which targets and amplifies portions of the hypervariable regions (V1-V9) present in all prokaryotes – bacteria and archaea. After sequencing, the reads are assigned to phylogenetic ranks based on 16S reference databases (Durazzi et al., 2021). The beneficial part of 16S rRNA sequencing is its relatively low price, but on the downside, it has limited resolution meaning that it cannot classify all the bacteria and identifying taxa down to species level may be impossible. Additionally, the 16S rRNA method cannot be used for detecting eukaryotes and viruses (Jovel et al., 2016; Kurilshikov et al., 2017). Also, in the case of 16S rRNA sequencing, the primer selection and hypervariable region greatly influence the observed microbial community. As an example, Kyono *et al.* studied the microbiome of cervicovaginal tract and showed how regions V4 and V3-V5 can detect *Bifidobacterium* and *Gardnerella*, known genera of this environment, while regions V1-V2 cannot (Kyono et al., 2018). The second method, shotgun metagenomics sequencing, does not only target the regions of 16S rRNA genes, but sequences all genomic DNA in the sample making it possible to identify not only bacteria but also viruses, fungi, and protozoa. Metagenomics allows taxa identification at the species level, has more power to identify less abundant taxa, makes it possible to annotate bacterial gene clusters as well as pathways and functional data (Durazzi et al., 2021). Metagenomics also eliminates the problem with hypervariable region selection present in 16S rRNA sequencing. The disadvantages of metagenomics sequencing include high sequencing cost and high bioinformatic load as a result of producing large number of reads. A downfall of both 16S rRNA and metagenomics sequencing is their dependence on reference databases, meaning that it is impossible to analyze genomes that are absent in the reference databases (Jovel et al., 2016; Kurilshikov et al., 2017).

Nevertheless, perhaps even a larger bias between studies arises from the bioinformatics and especially from the choice of methods for data processing and statistical analysis. The microbiome data are characterized by multiple distinct properties, which can significantly influence the results of the analysis. NGS-based microbiome studies are within the realm of compositional data, where the absolute number of microbes cannot be recovered from sequence data alone. The total number of reads that were sequenced varies between samples which confounds greatly to the results. One of the firsts methods to correct the problem of samples having different number of reads was to use rarefaction approach. Rarefaction in its essence is subsampling the read counts of each sample to a common size (Lozupone et al., 2011; Wong et al., 2016). However, there are issues regarding rarefaction including the omission of available valid data and arbitrary selection of the minimum number of reads. Due to the loss of information rarefaction causes, this approach has been questioned and alternative practices have

been proposed instead. In addition, with the total read count being uninformative for the analysis, microbiome data are known to contain only relative information. Thus log-ratio transformations that can alleviate the issues with the variability in read depth and relative nature of the data become more prevalent in microbiome studies (Greenacre et al., 2021). The log-transformed ratios are useful in the analysis since they are scale-invariant meaning that samples with low read counts are expected to have the same ratio as samples with many read counts (Gloor et al., 2017; Nearing et al., 2021). Moreover, microbiome data are sparse and zero-inflated, as many features (e.g. bacteria) are present in only few samples, this introduces a situation when in the case of log-transformation a constant arbitrary number, a so called pseudocount, needs to be added (Lin and Peddada, 2020). So far, adding the pseudocount is standard part of the compositional data analysis, however, there is an ongoing dispute against adding it to the data. Reasoning behind it is that we actually cannot be sure whether the bacteria is actually missing in the sample or its abundance is below the detectable threshold.

On the whole, methodological differences in the sample collection, processing and data analysis are noted to have strong impact on the microbial profile which can lead to the lack of reproducibility across studies. Due to which it is important for the microbiome research community to closely pay attention to the methods used in different studies and be critical with biological interpretation of microbiome research (Sun et al., 2021).

1.2. Female reproductive health in respect with microbiome

Recent advantages in DNA sequencing technology as well as computational resources have profoundly improved the microbiome research in ways that were impossible until fairly recently. This has made it possible to perform extensive studies on causes of different diseases and health conditions. Among other disease states the rapidly advancing field has now opened the door to study women's reproductive health from microbiome aspect. A lot of research has been done on the gut microbiome in regards to female health but it has become evident that when trying to understand the complex biological processes behind it there is a need to also look at the microbiome of the reproductive system itself. This is of course challenging since first, acquiring these samples is more difficult than obtaining the stool sample and secondly, these are mostly a low biomass samples which complicates the sample handling. However, more and more studies are being performed on this topic which as a result helps to broaden the understanding of female health.

1.2.1. Microbiome in polycystic ovary syndrome

PCOS is known as one of the most widespread endocrine and metabolic disorders in women at reproductive age worldwide. It has been estimated that PCOS affects approximately 8% to 18% of women, depending on the studied population and applied diagnostic criteria (Jobira et al., 2020; March et al., 2010; Teede et al., 2018; Zeng et al., 2020). The main diagnosis is based on the Rotterdam criteria which requires two symptoms out of the following three: excess of androgen (i.e. hyperandrogenism), persistent ovulatory dysfunction, and polycystic ovarian morphology (Rotterdam ESHRE/ASRM-Sponsored PCOS Consensus Workshop Group, 2004). PCOS is also noted as the most prevalent cause of anovulatory infertility and women having PCOS have a significantly increased risk of pregnancy-related complications compared to controls (Palomba et al., 2015). Though, PCOS is not only a reproductive disorder but also considered a syndrome with metabolic consequences affecting women throughout life. It is associated with a risk of developing a variety of metabolic derangements, including T2D, obesity, insulin resistance, hypertension, and nonalcoholic fatty liver disease (Azziz et al., 2016; Moran et al., 2010). It has also been shown that the metabolic rearranges are more predominant in women with PCOS who have hyperandrogenism (Barber et al., 2007; Moghetti et al., 2013).

The etiology of PCOS remains unknown but is believed to be multifactorial where genetics, intrauterine environment, lifestyle factors, and possibly alterations in the gut microbiome all have a role. So far gut microbiome dysbiosis has been associated with several metabolic diseases including obesity and T2D, both of which are also correlated with PCOS phenotype (Durack and Lynch, 2019). Regarding PCOS, research has shown that women with the disorder might present a lower bacterial richness in their gut compared to the women without PCOS (Chu et al., 2020; Insenser et al., 2018; Lindheim et al., 2017; Liu et al., 2017; Zhou et al., 2020). Studies have also been able to correlate the abundance of specific taxa or microbial diversity and androgen excess demonstrating that testosterone could be influencing the composition of the gut microbiome in women (Lindheim et al., 2017; Moreno-Indias et al., 2016). The overview of studies performed on gut microbiome and PCOS to date are summarized in Table 2.

Table 2. Summary of human gut microbiome studies concerning PCOS. Adapted from He and Li 2020 and He *et al*., 2021.

Study	**Country**	**Sample size**	**Sequencing technology (region)**	**Changes in diversity**	**Changes in composition**
Lindheim et al., 2017	Austria	PCOS = 20 control = 19	16S rRNA (V1-V2)	↓α diversity in PCOS β diversity distinguished	↓ Phylum Tenericutes ↓ Family *S24–7*
Liu et al., 2017	China	PCOS = 33 control = 15 *	16S rRNA (V3-V4)	α diversity unchanged β diversity: obese control group was more similar with non-obese PCOS and obese PCOS groups than with non-obese control group	↑ *Bacteroides*, *Escherichia/Shigella*, and *Streptococcus* ↓ *Ruminococcaceae*, *Akkermansia*
Torres et al., 2018	Poland	PCOS = 73 control = 48	16S rRNA (V4)	↓ α diversity in PCOS β diversity correlated with hyperandrogenism	↑ *Porphyromonas*, *Bacteroides coprophilus*, and *Faecalibacterium prausnitzii* ↓ *Blautia*, *Anaerococcus*, *Odoribacter*, *Roseburia*, and *Ruminococcus bromii*
Insenser et al., 2018	Spain	PCOS = 15 control n = 16 male control n = 15 *	16S rRNA (V4)	unchanged α diversity between PCOS and female controls; lower α diversity in women compared with men; ↓ β diversity in obese patients with PCOS compared with non-obese PCOS	↑ *Catenibacterium*, *Kandleria*, and *Oribacterium*
Zeng et al., 2019	China	PCOS = 8 IR PCOS = 9 control = 8	16S rRNA (V3-V4)	↓ α diversity in IR PCOS compared to healthy controls β diversity: controls clearly discriminated from the IR PCOS group	↑ *Prevotellaceae* in IR PCOS and non-IR PCOS compared to healthy controls ↑ *Bacteroides* in IR PCOS and non-IR PCOS compared to healthy controls

Study	Country	Sample size	Sequencing technology (region)	Changes in diversity	Changes in composition
Qi et al., 2019	China	PCOS = 50 control = 43	WGS	α diversity unchanged ↓ β diversity in PCOS	↑ *Bacteroidaceae*, *Bacteroides*, and *Bacteroides vulgatus*
Zhang et al., 2019b	China	PCOS = 38 control = 26	16S rRNA (V3-V4) and WGS	↓ α diversity in PCOS β diversity: clustering between PCOS and controls	↑ *Parabacteroides*, *Bacteroides*, *Lactobacillus*, *Oscillibacter*, *Escherichia/Shigella*, and *Clostridium* ↓ *Faecalibacterium*, *Bifidobacterium*, *Lachnospira* and *Blautia*
Zhou et al., 2020	China	PCOS = 60 control = 41 *	16S rRNA (V3-V4)	↑ α diversity in PCOS β diversity: no significant difference	↓ *Lactococcus* and *Alloprevotella* in non-obese PCOS compared with non-obese controls ↑ *Coprococcus* in obese PCOS compared to non-obese PCOS ↓ *Coprococcus_3*, *Lactobacillus* and *Prevotella_7* in obese PCOS compared to non-obese PCOS
Chu et al., 2020	China	PCOS = 14 control = 14 *	WGS	β diversity: no significant difference	↑ *Bacteroides* spp., *Escherichia* spp., *Shigella* spp., *Enterobacteria phage SfV*, *Parabacteroides merdae*, and *Comamonas kerstersii* ↓ *Bacteroides* spp., *Blautia hydrogenotrophica*, *Tannerella sp 6 1 58FAA CT1*, *Klebsiella pneumoniae*, *Faecalibacterium prausnitzii* and *Alistipes obesi*

Study	Country	Sample size	Sequencing technology (region)	Changes in diversity	Changes in composition
Jobira et al., 2020	USA	PCOS = 37 control = 21	16S rRNA (V3-V4)	↓ α diversity in PCOS altered β diversity	↑ *Actinobacteria*, *Streptococcaceae*, *Prevotella*, *Finegoldia* and *Lactobacillus* ↓ *Bacteroidetes*, *Bacteroidaceae*, *Porphyromonadaceae*, *Bacteroides* and *Parabacteroides*
Eyupoglu et al., 2020	Turkey	PCOS = 17 control = 15	16S rRNA (V3-V4)	α diversity unchanged β diversity similar between groups	↓ *Ruminococcaceae*
Mammadova et al., 2021	Turkey	PCOS = 22 control = 24	16S rRNA (V3-V4)	α diversity unchanged β diversity similar between groups	↑ *Clostridium cluster XVII* ↓ *Clostridium sensu stricto* and *Roseburia*

Arrows indicate the changes in PCOS compared to the control group. Abbreviations: PCOS, polycystic ovary syndrome; IR, insulin resistance; WGS, whole genome sequencing; rRNA, ribosomal ribonucleic acid. Asterisk marks that the study population included obese and non-obese participants.

Despite seeing correlations between PCOS and the gut microbiome in human studies as well as rodent models, the mechanisms responsible remain unclear. Because the diagnosis of PCOS is associated with metabolic disorders such as T2D and obesity, the role of intestinal microbes in PCOS may be related through metabolic processes. However, studies showing the mechanisms through which microbiome affects metabolic health is limited (Walter et al., 2020). Using tools such as fecal microbiota transplant (FMT) might provide an insight into the causal links between different diseases and gut microbiome. In the case of PCOS, Qi *et al.* performed an FMT of stool from women with and without PCOS into mice treated with antibiotics prior to the transplant. As a result they demonstrated that mice receiving FMT with PCOS stool resulted in a PCOS-like phenotype that included insulin resistance, infertility, disturbance of ovarian functions as well as altered BA metabolism and reduced interleukin-22 secretion (Qi et al., 2019). Notably, BAs are able to affect the growth of the gut microbes and the microbes can chemically modify the BAs (Wahlström et al., 2016). In their work Qi *et al.* reported that some secondary BAs (glycodeoxycholic acid and tauroursodeoxycholic acid) had lower levels in women with PCOS compared to unaffected women and when giving supplementation of these BAs to mice it had a protective effect against developing PCOS-like phenotype (Qi et al., 2019). This is suggestive that in the case of PCOS the effect of gut microbiome on the syndrome is mediated through BA metabolism, especially since traits such as obesity and T2D are closely related to both gut microbiome and PCOS. Another study done on mice, also indicated that the gut metabolites are more predictive of PCOS phenotype than the gut bacteria (Ho et al., 2021). Additionally, the role of gut microbes in PCOS has been suggested to be associated with endotoxemia – the presence of toxic compounds of bacterial origin found in the blood. The most well-known endotoxins today are lipopolysaccharides (LPS) which are a part of the cell wall of Gram-negative bacteria (Das et al., 2014). It has been suggested that an increase in intestinal permeability leads to more LPS in the blood circulation and the consequent activation of immune system induces insulin resistance and pro-inflammatory state (Duan et al., 2021).

While it seems to be evident that the effect of gut microbes on PCOS works through metabolism and metabolites, when aiming to look more into the reproductive part of the PCOS and bacteria, it would probably be necessary to study the microbial profile of the female reproductive tract. Studies on the reproductive system might give a better insight on how bacteria there might affect the etiology and reproductive part of PCOS.

In the future, larger studies as well as usage of PCOS-like animal models and reproductive tract sampling are beneficial in finding out which specific bacteria are responsible for the development of the syndrome and which are the precise mechanisms behind it.

1.2.2. Microbiome of female reproductive tract

For a long time, it was believed that healthy uterus is sterile and presence of microorganisms in the uterine cavity was considered to be pathological as well as an indicator of an ongoing infection. However, with the becoming of NGS the concept of "sterile womb" has been refuted and currently the existence of endometrial microbiome inside the uterus is undisputed, in fact, vagina, endometrium, ovaries, fallopian tubes, and vagina all harbor their own microbiome (Chen et al., 2017). Despite different parts of female reproductive system having differences in their microbial profile, one similarity between them is the predominance of genus *Lactobacillus* which is especially evident in the vagina (Chen et al., 2017; Punzón-Jiménez and Labarta, 2021). High abundance of *Lactobacillus* and low diversity in female reproductive system is associated with healthier reproductive health and better outcome of assisted reproduction techniques (Punzón-Jiménez and Labarta, 2021). Understanding the interactions between microbiome and female genital tract is important in the clinical point of view since it can be used as a tool to better the female health in infertility problems.

The female genital tract can be divided into two parts – the lower genital tract referring to vagina and the upper genital tract encompassing cervix, endometrium, Fallopian tubes, and ovaries (Figure 2). Probably the most investigated part of the female genital tract is the vagina. Several studies have come to a conclusion that the healthy vagina is largely dominated by *Lactobacillus* (Chen et al., 2017). Currently the vagina of healthy non-pregnant women is considered to have a high bacterial load with over 90% of the species belonging to *Lactobacillus* (Punzón-Jiménez and Labarta, 2021).

Moving up from the vagina, the bacterial diversity increases, cervical and endometrial microbiome is still highly *Lactobacillus* dominant, but its dominance has decreased compared to the other taxa present. In the Fallopian tubes and ovaries the microbial diversity increases and the abundance of *Lactobacillus* decreases compared to the vagina, cervix and endometrium. It is noteworthy that the pH levels in ovaries and Fallopian tubes change getting slightly alkaline which in itself could impact the microbial composition (Chen et al., 2017).

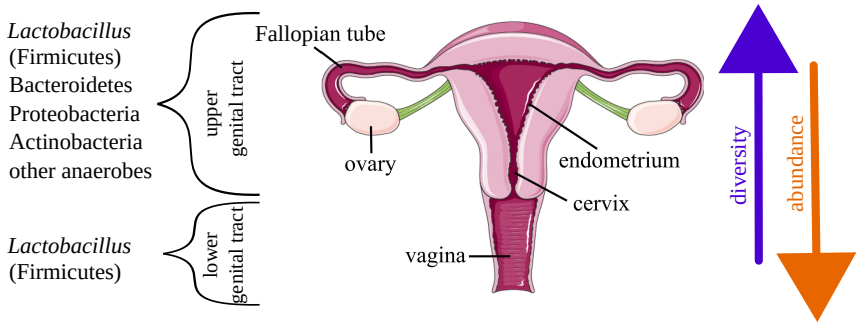


Figure 2. The female reproductive tract. Microbiome diversity increases in the upper genital tract whereas the microbiome abundance increases in the lower genital tract. (Adapted from Punzón-Jiménez and Labarta, 2021)

While it is acknowledged that the core of vaginal microbiome is *Lactobacillus*, to this day there is not a consensus found on what is the core of upper genital tract microbial composition. The explanation for this could be found when looking at how challenging the endometrial microbiome studies are. To start with, within one menstrual cycle female body goes through hormonal as well as physiological changes – all which can potentially affect the microbial community inside the uterus. Furthermore, receiving the endometrial sample is a complicated procedure in which a biopsy is taken via the cervix making it very prone to contamination from the vagina or cervix. Another method used for receiving the endometrial sample is during laparoscopy. Yet, usually women undergoing this surgical procedure have an existing medical condition. This on the other hand implies that it is highly unlikely for a healthy woman to undergo laparoscopy therefore, making it complicated to perform studies where endometrial samples for healthy controls are collected the same way as for the cases.

To this day, most of the studies involving endometrial microbiome are done on women suffering from different kinds of problems with their reproductive health (Franasiak et al., 2016; Hashimoto and Kyono, 2019; Kitaya et al., 2019; Kyono et al., 2018; Liu et al., 2018, 2019; Moreno et al., 2016, 2018). In the case of women undergoing assisted reproductive technology treatment, dysbiosis of endometrial microbiome has been associated with implantation failure and early spontaneous abortion. Based on early studies on the endometrial microbiome field, it is suggested that besides associations with reproductive outcome, endometrial microbiome is additionally associated with chronic endometritis (Fang et al., 2016; Liu et al., 2019; Moreno et al., 2018), endometriosis (Chen et al., 2017; Cregger et al., 2017; Hernandes et al., 2020; Khan et al., 2016; Wee et al., 2018), endometrial hyperplasia and cancer (Walsh et al., 2019; Walther-António et al., 2016) as well as endometrial polyps (Fang et al., 2016). Yet, it is still unknown what are the exact mechanism through which endometrial microbiome affects different disease states in the uterus nor is it clear whether the dysbiosis itself is a cause or a consequence of the pathology. Another unknown aspect in the endometrial microbiome is how the different microbes actually find their way into the uterus. So far there are a couple of ways suggested for bacteria entering the uterine cavity, among which are through gynaecological procedures, ascension through the cervix, retreating spread through fallopian tubes, during sexual intercourse as well as through hematogenous spread bacteria from the gut or oral cavities (Baker et al., 2018; Molina et al., 2020). Nevertheless, it is important to note that studies have suggested strong influence of vaginal microbes on the endometrial microbial composition thus implying that the most probable way for bacteria to enter uterus is through ascension from vagina (Kyono et al., 2018; Molina et al., 2020; Moreno et al., 2016).

Without a doubt, the microbial profile of the endometrium and the whole upper genital tract needs additional studies in the future to widen the knowledge on microbiome composition and its functions in the uterus in female health and disease.

1.3. Future perspectives of microbiome studies

The growth of microbiome studies has been rapid during the last 15 years and microbiome is on its way to the lead in scientific research. Preceding to 2005, there were approximately 500 gut microbiome publications in the Web of Science per year. By the beginning of 2020 that number had already rose up to more than 9,000 microbiome related articles per year (Li et al., 2020). When wondering what has caused the speedy rise in microbiome research, it is the fast development of research technologies (Gilbert et al., 2018). Improvements in the NGS technologies together with multi-omics approaches, including metagenomics, metabolomics and transcriptomics, have greatly enhanced the microbiome field and made it possible to understand the structure as well as function of the microbiome better (Gilbert et al., 2018).

In the beginning of microbiome research the main results were on correlations between taxa and health states. The gut microbiome has been associated with numerous diseases that go beyond being only gastrointestinal disorders, such as irritable bowel syndrome, Crohn's disease or colorectal cancer (Lloyd-Price et al., 2019; Wong and Yu, 2019). Links between the gut microbiome can be found with Alzheimer's disease (Vogt et al., 2017), obesity (Ridaura et al., 2013), T2D (Gurung et al., 2020; Qin et al., 2012) and many more. However, now the microbiome studies are altering from associations to finding causality. This is largely a normal progress in science since it is widely acknowledged that an association is not a proof of causation. One of the first landmark studies showing causality between the gut microbiome and host health was a research by Turnbaugh *et al*. published in *Nature* in 2006. In their work they used FMT technology on mice to demonstrate that obesity is transmissible by the gut microbiome (Turnbaugh et al., 2006). Different rodent studies with implementing FMT are progressively showing that the gut microbiome is at least partly causal for developing diseases. This is especially well described for metabolic diseases such as T2D and obesity (Arora and Bäckhed, 2016; Bäckhed et al., 2004; Fan and Pedersen, 2021; Gérard and Vidal, 2019; Qin et al., 2012; Turnbaugh et al., 2006).

Despite seeing that microbiome is causal in changes seen in these disease states, it is still uncertain what are the exact mechanisms through which microbiome contributes to the onset and progression of illnesses. Unraveling these questions behind disease development is certainly a question of utmost importance for the future research in the microbiome field. At the same time it is essential to keep in mind that the gut microbiome is only one aspect of the disease, and overemphasizing its importance should be avoided. The driving causes of different disease states are most likely a complex result of numerous driving factors. Which is why it is necessary to also include human genetics, lifestyle, sleeping patterns, environment, medication and other possible factors into the study structure.

Furthermore, thus far most of the published research on the topic of human health and gut microbiome have been cross-sectional studies looking at data at one specific time point which lacks the potential to identify the causality of the found associations. Future prospective studies looking at metagenomes from

pre-treatment or pre-diseases states are desperately needed to identify the causality of observed associations. As stated earlier, in order to study the microbial variability over time, longitudinal studies looking at the same individuals over a period of time are needed. A pioneer in this is a new cohort collected in Israel which aims to gather data of 10,000 individuals and follow them continuously over the time period of 25 years (Shilo et al., 2021). Recent studies suggest that repeated microbiome measurements are needed in microbiome field since while there are species whose abundance does not change to great extent over time, the so called microbiome fingerprint, the microbial profile does show temporal variation which could be a key determinant in understanding the roles of microbiome in health and disease (Chen et al., 2021; Vandeputte et al., 2021). For what is more, large existing biobanks have started collecting their own microbiome data enabling to add different omics datasets together with lifestyle and health data to the equation in explaining the microbiome variability.

Knowledge of microbiomes' role in human health is increasing expeditiously and every day researchers around the world are working with great effort to advance the field of microbiomics. Large population studies have a great importance in widening our understanding of microbiome and have already discovered many links between microbiome and wellbeing, inspiring new researches. The advancements will lay the first stone to rise novel therapeutic possibilities and move towards personalized medicine that would not only implement the human genetics but also human microbiome.

2. AIMS OF THE STUDY

The aim of this thesis is to investigate the variation of the human microbiome and its associations with health. The specific objectives of the thesis are as follows:

1. to study the effects of host genetics on gut microbiome composition in a large consortium study with 24 different cohorts

2. to explore the associations between microbiome and female health, focusing on polycystic ovary syndrome and endometrial microbiome analysis of *in vitro* fertilization (IVF) patients

3. RESULTS AND DISCUSSION

3.1. The potential of large-scale association analyses in microbiome studies (Ref. I)

The work of the last decade has clearly shown that the gut microbiota composition is highly influenced by multiple environmental factors, such as diet and medication (Aasmets et al., 2022; Falony et al., 2016; Jackson et al., 2018; Rothschild et al., 2018; Salosensaari et al., 2021; Zhernakova et al., 2016). However, the role of genetics on the gut microbiome has been less studied. Evaluating the associations between the gut microbiome and genetic variations is vital in understanding the role of microbiome in human organism. First studies with twins and later population-based studies with unrelated individuals have shown the heritability of microbiome and detected associations between the gut microbiome and genetic variations (Bonder et al., 2016; Goodrich et al., 2016; Lopera-Maya et al., 2022; Qin et al., 2022). Based on twin studies, the effect of genetics on microbiome is relatively small, especially compared to the environmental effects. Due to large variability in microbiome composition between individuals and genetics having a relatively small effect on microbiome, large sample sizes are needed which would help to detect the associations between genetic loci and microbiome. In order to overcome the hurdle of small sample sets, the MiBioGen consortium with more than 18,000 individuals belonging to 24 different cohorts from all over the world was established.

3.1.1. Description of cohort and methods

For this study we used a METabolic Syndrome In Men (METSIM) cohort, which is a longitudinal population-based cohort consisting a total of 10,197 randomly selected non-diabetic Finnish men (Laakso et al., 2017). In the MiBioGen consortium meta-study, a subset of the METSIM cohort was used. A total of 522 men with an average age of 61.91 (standard deviation of 5.41) and average BMI of 27.92 (standard deviation of 3.61) were comprised in the study. All of the participants had genotyping as well as microbial 16S rRNA sequencing data. Genotyping was done using the Illumina Omni ExpressExome microarray. Microbial DNA was extracted from frozen fecal samples using the PowerSoil DNA Isolation Kit (MO BIO Laboratories, Carlsbad, CA, USA) following the manufacturer's instructions. Amplification of the V4 hypervariable region of the 16S rRNA gene was done with the 515F and 806R primer and sequenced with the Illumina MiSeq platform using Illumina OmniExpressExome microarray at the University of California, Los Angeles.

The whole meta-analysis study comprised of 18,340 individuals from 24 different cohorts from Belgium, Canada, Denmark, Finland, Germany, Israel, the Netherlands, the United Kingdom, South Korea, the United States and Sweden.

The analysis performed by each of the participating cohort in-house were 1) 16S rRNA data processing; 2) host SNP microarray data processing; 3) running the association study benchmark focusing on 10 most abundant bacterial genera of the cohort; 4) running MWAS for all taxa (microbiome quantitative trait loci, mbQTL); 5) running MWAS on alpha diversity; 6) running binary trait (bacterial presence/absence (microbiome binary trait loci, mbBTL) MWAS. The taxonomy binning based on SILVA 128 database done using standard pipeline created by the consortium for all the cohorts (Quast et al., 2013).

The microbiome data was corrected for age and principal components (for other cohorts also for sex but since METSIM is a male-only cohort, this step was not performed). Cut-offs and data transformations for taxonomies in each cohort were following: taxonomy present in more than 10% of the samples were kept for the mbQTL mapping whereas for mbBTL mapping taxa present in more than 10% of the samples but less than 90% of the samples were included. To account for differences in sequencing depth, datasets from every cohort were rarefied to 10,000 reads per sample. Study-wide cut-offs for bacterial taxa in mbQTL included a minimum effective sample size of 3,000 samples and presence in at least three cohorts. For mbBTL analysis, a mean abundance higher than 1% in the samples was required. In total, 211 taxa were included in the mbQTL analysis and 177 taxa in mbBTL analysis.

For SNPs the data processing steps were: minor allele frequency > 5%; imputation quality of more than 0.4; SNP call rate was set to 0.95 and higher and ambiguous SNPs were removed. For binary traits logistic regression with Chi square-based p-value estimation was used and for non-zero samples linear regression model on log-transformed counts with Fisher test-based p-value estimation was used. A P-value of $P < 5 \times 10^{-8}$ was considered to reach the nominal genome-wide significance level and $P < 1.95 \times 10^{-10}$ was considered to pass the strict correction for the number of taxa tested and reach study-wide significance.

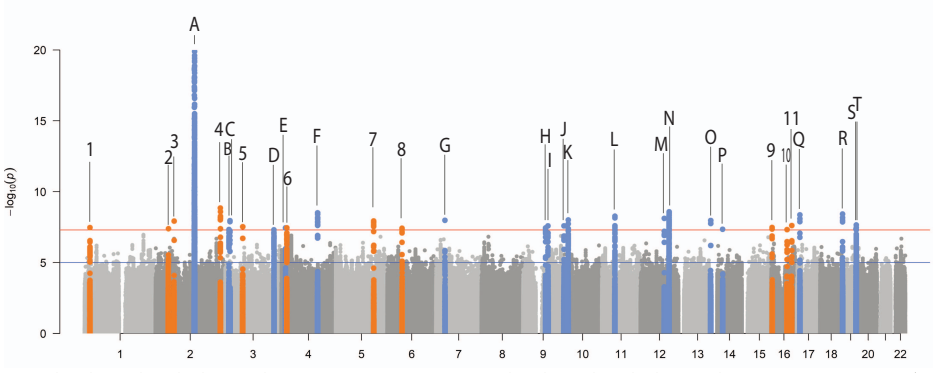
3.1.2. Gut microbiome is associated with 31 genetic loci

To explore the effect of host genetics two different types of MWAS meta-analysis were performed 1) on the microbial abundance levels (mbQTL) and 2) on the presence or absence (mbBTL) of bacterial taxa in the gut microbiome. The mbQTL analysis identified 20 loci associated with the abundance of 27 taxa and the mbBTL analysis 10 loci associated with presence or absence of bacterial taxa that reached the nominal genome-wide significance level ($P < 5 \times 10^{-8}$). For one bacterium, two independent loci were identified, leading the overall number of loci associated with gut microbiome to 31 (Figure 3).

Out of the 31 associated loci, only one locus passed the study-wide multiple testing correction ($P < 1.95 \times 10^{-10}$). This was the *LCT* locus on chromosome 2 associated with the genus *Bifidobacterium* ($P = 8.63 \times 10^{-21}$). The *LCT* gene is responsible for lactase persistence in adult European population. The associations at the *LCT* locus have previously been described and to this day this is the strongest association reported between human genetics and microbiome composition

(Blekhman et al., 2015; Bonder et al., 2016; Goodrich et al., 2016; Rühlemann et al., 2021). What makes the association between *LCT* and *Bifidobacterium* noteworthy is that this association has been consistently described in studies with different ethnicities, a wide range of sample sizes as well as studies implementing different pipelines (Blekhman et al., 2015; Bonder et al., 2016; Goodrich et al., 2016; Lopera-Maya et al., 2022; Rühlemann et al., 2021). A recent study done on Dutch cohort was able to describe the association with *LCT* locus not only on genus level, but more specifically, showed association between *LCT* and *Bifidobacterium longum* and *Bifidobacterium adolescentis* (Lopera-Maya et al., 2022).

Despite not reaching the strict correction for the number of taxa tested, there were several loci identified that are enriched for genes related to metabolism. One such example is the *FUT2* gene. The *FUT2* locus was associated with the abundance of the *Ruminococcus torques* genus group. *Ruminococcus sp.* are known to degrade complex carbohydrates while as *FUT2* encodes an enzyme responsible for the secretion of fucosylated mucus glycans in the gastrointestinal mucosa (Crost et al., 2018; Kashyap et al., 2013). These functions are supportive of the link between the *Ruminococcus* and *FUT2*. Associations with *FUT2* have been described also in the Dutch and German cohorts (Lopera-Maya et al., 2022; Rühlemann et al., 2021).



Locus	Pvalue	Type	Chr	rsID	Bacteria
A	1.28E-20	Quant	2	rs182549	genus Bifidobacterium
B	4.66E-08	Quant	3	rs9864379	order Gastranaerophilales
C	1.10E-08	Quant	3	rs75754569	genus Peptococcus
D	4.88E-08	Quant	3	rs4428215	family Oxalobacteraceae
E	3.55E-08	Quant	4	rs10805326	genus Intestinibacter
F	3.06E-09	Quant	4	rs11098863	genus Enterorhabdus
G	1.04E-08	Quant	7	rs17159861	genus group Eubacterium coprostanoligenes
H	3.57E-08	Quant	9	rs602075	genus Allisonella
I	2.57E-08	Quant	9	rs736744	genus Oxalobacter
J	2.55E-08	Quant	10	rs12781711	genus Ruminococcaceae UCG013
K	9.80E-09	Quant	10	rs61841503	family Peptostreptococcaceae
L	5.29E-09	Quant	11	rs10769159	genus Ruminococcus1
M	7.57E-09	Quant	12	rs12320842	genus Faecalibacterium
N	2.58E-09	Quant	12	rs11110281	genus Streptococcus
O	1.08E-08	Quant	13	rs7322849	genus Bifidobacterium
P	4.42E-08	Quant	14	rs8009993	genus Ruminococcaceae UCG009
Q	4.31E-09	Quant	17	rs7221249	genus Erysipelatoclostridium
R	3.74E-09	Quant	19	rs67476743	genus Tyzzerella3
S	4.26E-08	Quant	19	rs830151	genus Candidatus Soleaferrea
T	2.21E-08	Quant	19	rs35866622	genus group Ruminococcus torques
1	3.34E-08	Binary	1	rs555115	genus Turicibacter
2	4.07E-08	Binary	2	rs57440956	genus Ruminiclostridium6
3	1.17E-08	Binary	2	rs56081657	genus Odoribacter
4	1.42E-09	Binary	2	rs7574352	family Peptococcaceae
5	2.98E-08	Binary	3	rs17066404	family Peptococcaceae
6	3.53E-08	Binary	4	rs207281	genus Enterorhabdus
7	1.16E-08	Binary	5	rs62373497	genus Coprococcus1
8	3.79E-08	Binary	6	rs9474033	family Pseudomonadaceae
9	3.34E-08	Binary	16	rs437867	genus Lachnospiraceae UCG010
10	4.67E-08	Binary	16	rs17319026	genus Anaerostipes
11	2.42E-08	Binary	16	rs11647069	family Bacteroidales S24.7

Figure 3. Manhattan plot of the mbTL mapping meta-analysis results. MbQTLs are indicated by letters. MbBTLs are indicated by numbers. Horizontal lines define nominal genome-wide significance ($P = 5 \times 10^{-8}$; red) and suggestive genome-wide ($P = 1 \times 10^{-5}$; blue) thresholds.

The analysis on the host genetics and bacterial alpha diversity revealed no associations as well as diversity did not present any heritability in the heritability estimation analysis based on the TwinsUK and the Netherlands Twin Registry cohorts.

This part of the thesis stresses the value that collaborative work and large sample sizes have in research specifically in microbiome studies. Large studies like this provide an exceptional source of data for human microbiome studies. However, multicohort studies also come with obstacles which could overshadow the chances of discovering true associations. The biggest challenges include a) large heterogeneity due to multi-ethnicity; b) different sample collection and DNA extraction methods between cohorts; c) usage of various sequencing and genotyping technologies. This all is combined with challenges present in all MWASs not only in multi-cohort studies such as the large environmental effect on microbiome which could conceal the effect of genetics; the burden of multiple testing correction; and challenges in microbiome analysis due to its complex structure. For what is more, the expected effect size of host genetics on the microbiome seems to be modest meaning that in order to identify the genetic variants associated with microbiome a vast increase in sample size is needed for adequate power (Kurilshikov et al., 2017). Based on recent power calculations more than 50,000 study participants are needed to find new associations between human genetics and the gut microbiome with effect sizes close to *LCT* (Lopera-Maya et al., 2022).

3.2. The gut microbiome in Finnish female cohort and its associations with polycystic ovary syndrome (Ref. II)

Evaluating the human gut microbiome profile of different cohorts across the world has been one of the first aspects described in microbiome studies. It is a crucial step in understanding the overall composition of the gut ecosystem of a certain cohort before moving on to investigate the deeper associations between phenotypical traits and microbes.

The gut microbiome is widely studied in different metabolic diseases, however its probable role in PCOS which is also categorized as a metabolic and endocrine syndrome is rather understudied to this day. Given the fact that PCOS is a complex disorder characterized by a variety of traits that have been linked to the gut health it is reasonable to study its associations with gut microbiome. For what is more, the etiology of PCOS even to this day remains partially unknown. It is known to be a multifactorial syndrome where lifestyle as well as genetics have their role in the developmental process and now more and more studies are being published that are investigating the possible link between PCOS and the gut microbiome.

In this study our goal is to describe the gut microbiome profile of late fertile age women in a homogenous Finnish female cohort and to compare the gut microbiome in women with and without PCOS. We also correlated the gut microbiome with PCOS-associated metabolic markers to find possible interaction between the two of them.

3.2.1. Description of cohort and methods

For this study we used a subset of the Northern Finland Birth Cohort 1966 (NFBC1966) that includes the 2 northernmost provinces of Finland (Rantakallio, 1988; University of Oulu). Participants donated blood and fecal samples, performed an oral glucose tolerance test (OGTT) and their anthropometric measurements were taken. Also, the Finnish register for drug imbursements was used in order to identify study participants who had been prescribed antibiotics, letrozole, antimycotics or tamoxifen within three months prior to sample collection, these individuals were then excluded. The total study population was 303 women including 102 women with PCOS and 201 age- and BMI-matched women without PCOS (hereinafter referred as non-PCOS).

The bacterial DNA extraction from the stool samples was done with the QIAamp Stool Mini Kit (Qiagen, Venlo, The Netherlands) and the sequencing of V3-V4 regions of the 16S rRNA gene were done on an Illumina MiSeq instrument. The open-source software QIIME 2 2019.7 (Bolyen et al., 2019) was used for the raw sequencing data analysis and statistical data analysis was performed using the statistical software R v.3.6.1 (under RStudio v.1.2.1335).

3.2.2. The gut microbiome profile of Finnish women corresponds to the Western population

In this part of the study, we characterized the gut microbial profile of our study participants of Finnish women belonging to the NFBC1966 cohort. The microbial profiling revealed that the study population is representative of a typical Western diet (WD) diversity profile of gut microbiota with Firmicutes (54.0%) and Bacteroidetes (31.9%) being the most prevalent phyla followed by Proteobacteria (6.7%), Actinobacteria (3.4%) and Verrucomicrobia (2.4%) (Figure 4A) (Senghora et al., 2018).

In paragraph 3.1 in the thesis (Ref. I) we used another population-based cohort from Finland, the METSIM cohort of 522 Finnish men collected from Kuopio in Eastern Finland. Comparing gut microbiome profiles in two Finnish population-based cohorts, one of which being female-only and the other male-only cohort, we observed that the overall composition on phylum level for both Finnish cohorts are similar to one another, with only the Proteobacterium being more prevalent in the female cohort than in the male-only METSIM cohort, 6.7% and 2.4%, respectively (Figure 4A, C).

We also assessed the core microbiome of both Finnish cohorts. The core was defined as the genera shared by > 95% of the samples. In the female NFBC1966 cohort we identified eight genera as core: *Bacteroides* (19.9%), followed by *Alistipes* (7.5%), *Faecalibacterium* (4.9%), *Roseburia* (2.5%), *Blautia* (2.5%), *Lachnoclostridium* (1.5%), *Ruminococcaceae* uncultured (1.2%), and *Oscillibacter* (1.1%) (Figure 4B). Six of the eight genera detected in our female cohort

overlapped with a large international study by MiBioGen consortium incorporating 24 different populations (Ref. I). The additional two genera found in our study were *Ruminococcaceae* and *Oscillibacter*. Figure 4D illustrates the core found for METSIM cohort. Similarly to NFBC1966 all core genera in METSIM belonged to the two most abundant phyla Firmicutes and Bacteroidetes. However, the core of METSIM cohort consisted of more genera than in NFBC1966 cohort (Figure 4B, D).

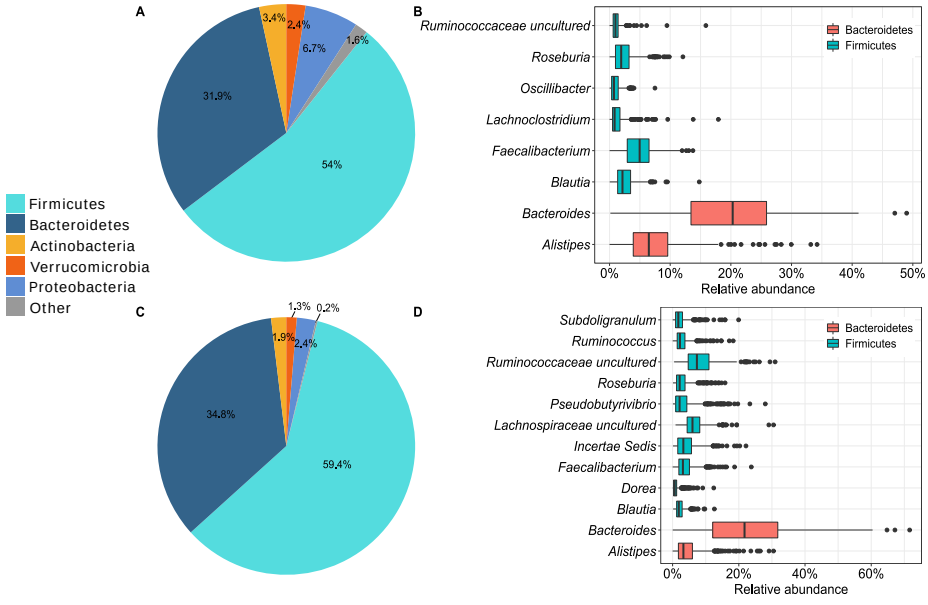


Figure 4. Landscape of microbiome composition of the NFBC1966 (panels A and B) and METSIM cohort (panels C and D). (A and C) Pie chart indicate the average relative abundances of the top major phyla of the cohorts. (B and D) Mean relative abundance of the core microbiome (taxa present in over 95% of individuals) of the cohorts at genus level.

3.2.3. Two time point clinical data helps to evaluate the clinical phenotype of PCOS

The clinical characteristics of the study participants were assessed at two time points (at the age of 31 and 46), however, the stool samples were taken only during the second time point. Analyzing the clinical markers revealed that compared to controls women with PCOS had higher testosterone ($P = 0.01$) levels at 31 years as well as higher free androgen index (FAI) ($P = 0.01$) measures at 46 years (Table 3). In some women with PCOS, testosterone levels are normal, however a large proportion of PCOS women suffer from androgen excess, including testosterone (Sheehan, 2003). The reasoning behind elevated testosterone in PCOS is that the levels of insulin in the blood and the levels of luteinizing hormone produced by pituitary gland are too high. This in turn causes

the ovaries to produce an excessive amount of testosterone. To this day there are different options in PCOS to decrease testosterone production. Since the excess testosterone production happens predominately in ovaries, the reduction of gonadotropin production together with increasing sex hormone binding globulin (SHBG), oral contraceptives are used to decrease the testosterone levels (Wiegratz et al., 2003). Another possible way to target testosterone levels, is by lowering insulin levels which can be accomplished through weight loss or by the usage of metformin (Harborne et al., 2003; Sheehan, 2003).

At age 31, women with PCOS reported having fewer menses per year and having problems with infertility, which was also seen at the age 46. One widely prevalent symptoms of PCOS estimated to occur up to 85% of women with PCOS is oligomenorrhea which by general definition means that a women experiences fewer than eight menstrual cycles per year (Harris et al., 2018). Problems with ovulation, usually caused by hormonal imbalance, are normally the primary cause of infertility in women with PCOS making it comprehensible for us to see rise in infertility problems as well as less menses per year among the PCOS cases in our study.

As anticipated, there were no differences in BMI between the groups since the controls were BMI-matched to the PCOS cases. Although not statistically significant, compared to the non-PCOS SHBG tended to be lower in PCOS group (Table 3).

Table 3. Clinical, metabolic, and hormonal characteristics of the NFBC1966 cohort.

	31 years			46 years		
	Control	**PCOS**	**P value[a]**	**Control**	**PCOS**	**P value[a]**
BMI, kg/m^2	25.09±4.81 (n = 201)	25.35±5.41 (n = 100)	0.92	27.44±4.98 (n = 202)	26.96±5.08 (n = 102)	0.82
SHBG, nmol/L	50.60±15.09 (n = 36)	40.71±13.90 (n = 14)	0.21	58.76±30.99 (n = 200)	55.76±27.07 (n = 102)	0.64
testosterone, nmol/L	1.14±0.61 (n = 152)	1.33±0.62 (n = 75)	0.01	0.86±0.33 (n = 201)	0.93±0.32 (n = 102)	0.08
FAI	3.51±3.82 (n = 10)	5.70±3.07 (n = 3)	0.29	1.77±1.11 (n = 200)	2.04±1.10 (n = 102)	0.01
menses per year	12.15±1.25 (n = 190)	10.43±2.73 (n = 93)	1.10×10^{-12}	ND	ND	ND
infertility problems ever in life, *n*	28 (16%) (n = 179)	38 (42%) (n = 91)	4.92×10^{-06}	24 (13%) (n = 187)	37 (39%) (n = 95)	1.21×10^{-06}
parity	1.51±1.26 (n = 197)	1.5±1.19 (n = 98)	0.97	2.38±1.41 (n = 188)	2.25±1.56 (n = 95)	0.13
miscarriages, n	0.21±0.51 (n = 198)	0.24±0.63 (n = 98)	0.90	0.46±1.05 (n =170)	0.47±0.97 (n = 83)	0.71
fasting glucose, mmol/L	4.96±0.49 (n = 156)	5.15±1.24 (n = 78)	0.19	5.05±0.92 (n = 195)	5.38±0.58 (n = 95)	0.55
2 h glucose, mmol/L	ND	ND	ND	5.88±1.63 (n = 193)	5.69±1.45 (n = 92)	0.24

	31 years			46 years		
	Control	**PCOS**	**P value[a]**	**Control**	**PCOS**	**P value[a]**
fasting insulin, mU/L	8.52±3.79 (n = 155)	8.52±3.79 (n = 155)	0.70	9.82±5.65 (n = 195)	10.2±7.63 (n = 95)	0.79
2 h insulin, mU/L	ND	ND	ND	63.03±50.61 (n = 193)	57.34±41.81 (n = 93)	0.62
Matsuda index	ND	ND	ND	4.84±2.55 (n = 190)	5.24±3.35 (n = 89)	0.77
disposition index	ND	ND	ND	186.15±87.32 (n = 190)	189.47±89.59 (n = 89)	0.74

Data are presented as mean ± standard deviation for continuous traits and as absolute proportions and prevalence (%). [a] Wilcoxon sign rank test (continuous variable) or Fisher's Exact test (categorical variable). The number of women in separate analyses varies due to non-response to some items. Abbreviations: BMI, body mass index; FAI, Free Androgen Index; *n*, number of individuals; ND, no data; SHBG, sex hormone binding globulin; y, year.

3.2.4. Differences in gut microbiome in PCOS women compared to women without PCOS

In the gut microbiome analysis, we first compared the diversity measures between the PCOS and non-PCOS women. Both, alpha (Shannon diversity metric, inverse Simpson diversity and a number of observed taxa) and beta diversity analyses indicated no differences between the two study groups (Figure 2A in Ref II). Similarly, we saw no differences in beta diversity (Figure 2B and 2C in Ref II). Previous studies have reported contradictory results where some of the studies report changes in diversity measures while others do not (Table 2) (Eyupoglu et al., 2020; Insenser et al., 2018; Mammadova et al., 2021; Qi et al., 2019; Zeng et al., 2019; Zhang et al., 2019b; Zhou et al., 2020). For example, a metagenomics study performed on 93 individuals of Chinese ancestry reported no differences in alpha diversity while beta diversity of women with PCOS was significantly decreased compared with non-PCOS individuals (Qi et al., 2019). The most probable justification for not detecting differences between the groups in diversity analysis and for the results being contradictory could be the fact that PCOS as a whole does not change the gut microbiome to such an extent which would reflect in the overall microbial community.

The differential abundance analysis performed in order to detect bacteria distinguishing microbial profile of PCOS from non-PCOS ended with no statistically significant results found (Supporting Information Table 4 in Ref. II). However, we also performed a so called Selbal analysis which allows to identify groups of microbial taxa differentiating between the study groups. During this analysis we identified a microbial balance consisting of four genera whose balance is predictive of PCOS (AUC = 0.64). The equation is following:

$$\frac{\sqrt{(Eubacterium\ ventriosus\ \text{group} \times Bifidobacterium)}}{\sqrt{(Prevotella \times Streptococcus)}},$$

where higher solution notes for women with PCOS and lower for non-PCOS (Figure 5). Interestingly, despite not detecting any bacteria with statistically significant differences in their relative abundance between the PCOS and non-PCOS, all bacteria found in the Selbal analysis but the *Bifidobacterium* belonged to the top 10 bacteria identified in the differential abundance analysis (Supporting Information Table 4 in Ref. II). These findings add proof that the named four bacteria may have a role in the pathogenesis of PCOS and the relationship could possibly be detected on a statistically significant level in a larger study.

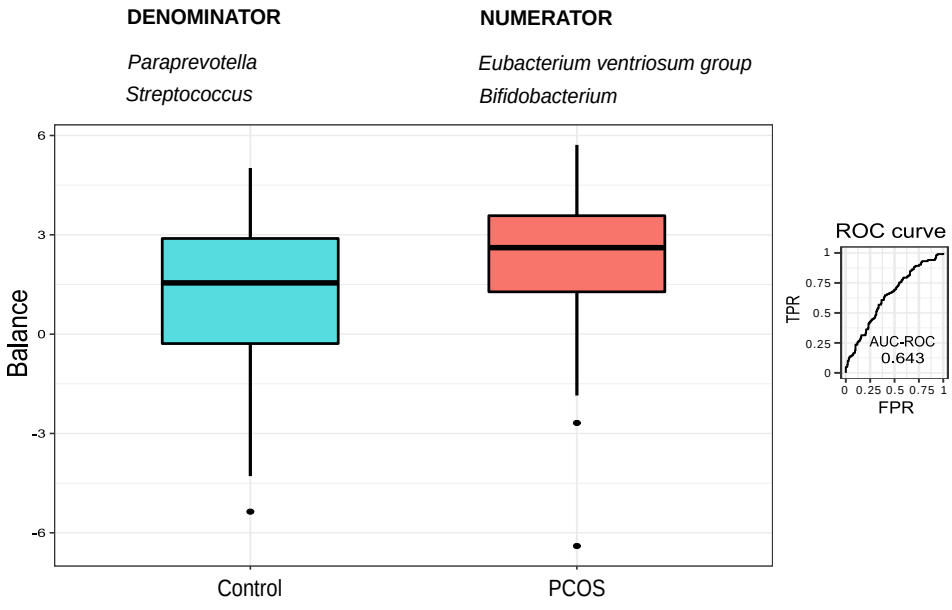


Figure 5. Selbal analysis results. The balance is made out of two groups of taxa: *Paraprevotella-Streptococcus* and *Eubacterium ventriosum* group *Bifidobacterium*. The boxplots characterize the distribution of the balance scores for PCOS women and healthy controls. The right part holds the ROC curve with its AUC value (0.643). Abbreviations: AUC, area under the curve; FPR, false positive rate; PCOS, polycystic ovary syndrome; ROC, receiver operating characteristic; TPR, true positive rate.

Finally, we performed correlation analysis of the top 10 bacteria found in the differential abundance analysis with traits connected to PCOS, such as hormonal markers, BMI as well as features of insulin and glucose metabolism. As a result, two genera – *Ruminococcaceae* UCD-002 and *Clostridiales* Family XIII AD 3001 group – presented correlations with various PCOS-related markers (Figure 3 in Ref. II). Particularly, the abundance of genus *Ruminococcaceae* UCG-002 was positively correlated with disposition index (P = 0.001), SHBG (P = 0.001) and Matsuda index (P = 0.009). Additionally, the abundance of *Clostridiales* Family XIII AD3011 group was positively correlated with SHBG (P = 0.006) and Matsuda index (P = 0.010) and negatively correlated with glycated hemoglobin (HbA1c) (P = 0.003), 2h glucose level (P = 0.006) as well as BMI (P = 0.010). Interestingly, some genera belonging to the family of *Ruminococcaceae* have been associated with PCOS in previous studies (Liu et al., 2017; Torres et al., 2018). *Ruminococcaceae* UCG-002 could potentially possess metabolically similar functions to the genera found in earlier studies making it an important genus to be further studied when learning the development of PCOS.

3.2.5. Changes in microbial diversity in PCOS patients are dependent on their metabolic health

The WD is defined as an unhealthy diet with high fat consumption characterized by frequent snacking, binge and overeating, as well as prolonged postprandial state (Malesza et al., 2021). WD has numerous metabolic aspects such as insulin resistance, dyslipidemia, hyperinsulinemia, and oxidative stress, for what is more, the consequences of high-fat diet include gut microbiota dysbiosis, gut barrier dysfunction, dyslipidemia, and increased intestinal permeability – all of which strongly promote the development of chronic low-grade systematic inflammation which is considered as one of the key contributors to the pathogenesis of PCOS (Malesza et al., 2021; Rudnicka et al., 2021).

Since we identified significant associations between PCOS and different metabolic traits, we next hypothesized that the effect of PCOS on the gut microbiome may be related to metabolic processes. In order to test this hypothesis, we categorized women with PCOS into subgroups based on their OGTT results. In total there were 76 individuals with normal glucose tolerance (NGT) and 14 with prediabetes which is a state where blood sugar levels are higher than normal, but not as high as in the case of type 2 diabetes. Comparing to NGT group, we reported decreased lower alpha diversity ($P = 0.018$) for individuals with prediabetes and significant differences in beta diversity ($P = 0.003$) (Figure 6A, B) which is consisted with the results of previous studies (Gurung et al., 2020; Menni et al., 2020). Additionally, the relative abundance of genus *Dorea* was significantly lower in the NGT group compared to preT2D suggesting a probable role of *Dorea* in metabolic diseases and their etiology (Figure 6C). In the literature, we found supporting evidence for this hypothesis. Namely, *Dorea* has previously been reported to be positively correlated with fasting blood glucose, glutamate, branched chain amino acids and BMI, all of which are important components of metabolic health (Naderpoor et al., 2019; Ottosson et al., 2018). Moreover, in our cohort *Dorea* had statistically significantly correlations with multiple metabolic features, including positive correlations with fasting glucose and fasting insulin levels and borderline positive correlations with BMI and HbA1c. This adds new evidence to the effect of *Dorea* in metabolic health and through it also to PCOS.

Taken together, our study on women with PCOS in their late reproductive years indicates that there might be no direct links between PCOS and gut microbiome, rather the mechanisms could work through microbial metabolites. Indeed, a recent research on gut microbiome and PCOS suggests that the link between the PCOS and gut microbiome most likely works through metabolism and more specifically, through metabolites produced by the gut microbiota. Namely, a recent study using untargeted metabolomics and metagenomics study on PCOS-like mouse saw that the microbial metabolites, specifically the primary and secondary BAs, are able to predict the PCOS more precisely than the microbes themselves (Ho et al., 2021). Given that supplementation with BAs such as glycodeoxycholic acid and tauroursodeoxycholic acid have shown to be lower in women with PCOS than in non-PCOS women (Qi et al., 2019), additional studies

are needed to determine what are the protective mechanisms through which these BAs act in PCOS. These discoveries can be useful for development of novel therapies implemented in PCOS treatment.

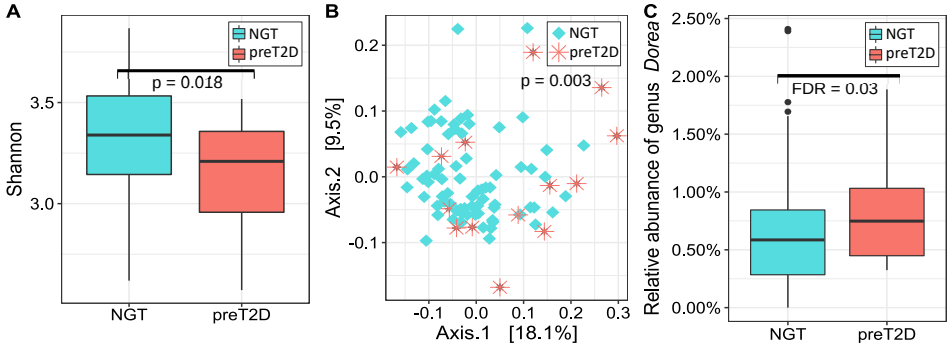


Figure 6. Comparison of the gut microbiome diversity between NGT (n = 76) and preT2D samples (n = 14) among women with PCOS. (A) Box plots of the Shannon diversity index (alpha diversity), median values and interquartile ranges have been indicated in the plot (P = 0.018). (B) Beta diversity is represented by principal coordinate analysis based on unweighted UniFrac distance (P = 0.003). (C) Genus *Dorea* with statistically significant differences (FDR = 0.03). Abbreviations: NGT, normal glucose tolerance; preT2D, pre-type 2 diabetes. Each plot point represents a single individual, the shapes and color indicate study groups (rhombus and blue: NGT, star and red: preT2D)

Our research also demonstrates the complexity of PCOS and indicates that it is highly metabolic as well as hormonal disorder where glucose tolerance and prediabetic state have an important part – both of which are strongly associated with the gut microbiome. In our study, the women had already surpassed their reproductive peak which could make it difficult to discover the associations with fertility. Therefore, in order to look at the associations between microbiome and fertility in PCOS it would be beneficial to perform a study in young adults with PCOS. As a future prospective it would also be interesting to combine polygenic risk scores calculated for PCOS with microbiome to predict the risk of developing the disorder.

3.3 Differences in microbial profile of endometrial fluid and tissue samples in women with *in vitro* fertilization failure are driven by *Lactobacillus* abundance (Ref. III)

The gut microbiome has been in the spotlight of microbiome studies for over a decade. However, with the improvement of methods used in the microbiome research it has become possible to start studying the microbial composition of more complex ecosystems with low biomass. This in turn makes it achievable to so to say get closer to the place of interest in the body. When so far most of the microbiome studies were done on using stool samples then now it is becoming more and more prevalent to study the sites closer to a so-called active illness location.

The microbes residing inside the female reproductive tract have been under the interest of gynecologists for a long period of time and a number of microbes have been associated to several gynecological disorders. This rises the idea that the microbiome could potentially be used as an additional biomarker when identifying possible causes behind infertility problems.

Using the 16S rRNA sequencing data of endometrial microbiome samples taken from endometrial fluid and tissue, we here aim to assess the microbial community of these samples in women who have had at least one *in vitro* fertilization treatment cycle failure.

3.3.1. Description of cohort and methods

The current study was carried out in collaboration with the Center of Human Reproduction Genesis clinic in St Petersburg, Russia where 25 women with previous IVF procedure failure were recruited. All of the women had undergone at least one unsuccessful IVF treatment cycle and had longstanding problems with infertility (Table I in Ref. III).

Two different types of endometrial samples were collected from each participant – an endometrial tissue (ET) and an endometrial fluid (EF) sample. The sample collection process was performed in the middle of the secretory phase of the menstrual cycle. A 16S rRNA sequencing of the V3-V4 hypervariable region was carried out and the statistical analyses between the ET and EF samples were carried out with R software (v4.0.2) using packages phyloseq (v1.32.0), vegan (v2.5-6), microbiome (v1.10.0) and ALDEx2 (v1.20.0).

Additionally, according to their *Lactobacillus* abundance, the samples were categorized as *Lactobacillus* dominant or non-*Lactobacillus* dominant. To be specific, samples with relative abundance of *Lactobacillus* $\geq$ 50% were considered as *Lactobacillus* dominant and samples with *Lactobacillus* $<$ 50% were considered as non-*Lactobacillus* dominant.

3.3.2. Microbial profile of endometrial tissue and fluid samples

Analysis of the microbial community composition of the ET and EF samples revealed that the most prevalent phyla in both of the sample types were Firmicutes, Proteobacteria and Actinobacteria (Supporting Information Figure 1 in Ref. III). In total there were 9 different phyla in ET samples and 10 phyla in EF samples, out of which 9 were present in both sample types (Supporting Information Figure 1 in Ref. III). Similarly to the human gut microbiome, a large variation in microbiome composition was present across samples. For example, the abundance of the most prevalent genus *Lactobacillus* fluctuated from 0 to 99.8% (mean 50.2%) in ET samples and from 0 to 99.7% (mean 71.8%) in EF samples. Previous work on endometrial microbiota have also shown that the genus *Lactobacillus* is the most prevalent habitant in this ecosystem (Liu et al., 2018). The fact that the abundance of *Lactobacillus* is higher in fluid samples than in tissue samples could potentially reflect the vaginal cavity microbial habitat.

We used differential abundance analysis to search for possible community differences between the sample types. Here, the only difference found between the sample types was genus *Lactobacillus* ($p = 0.01$), yet, this difference did not remain statistically significant after correcting for multiple testing (Table II in Ref. III). Likewise, we detected no statistically significant differences between the EF and ET in alpha and beta diversity analysis (Figure 2A, 2B, 2C and 2D in Ref. III) The fact that no differences were found between the ET and EF samples could be explained by our relatively small sample size as well as by high variability in microbiome composition between individuals.

Since our dataset comprises individuals with several different reasons behind their infertility, we decided to compare the microbiome of the individuals based on their infertility reasons (Supporting Information Table 3 in Ref. III). As a result, we detected no differences in the microbiome between the infertility groups which illustrates that the endometrial microbiome is probably not affected by infertility causes.

3.3.3. Genus *Lactobacillus* drives the differences between endometrial tissue and fluid samples

To look more in depth into the ET-EF pairs of samples we decided to perform analysis with mismatching *Lactobacillus* dominance. A reasoning behind it was also the fact that the relative abundance of *Lactobacillus* had a statistically significant effect on the variability in beta diversity ($r^2 = 0.34$, FDR $< 9.9\times10^{-5}$) suggesting that named genus has a big effect on the sample composition.

In total there were 8 individuals with mismatching *Lactobacillus* dominance samples [with *Lactobacillus* abundance for *Lactobacillus* dominated and non-*Lactobacillus* dominated groups being 83.7% and 13.1%, respectively ($P = 8.6\times10^{-4}$)]. Out of these 8 individuals 6 had *Lactobacillus* dominated fluid

samples and 2 individuals had *Lactobacillus* dominated tissue samples. Here we saw differences in both alpha and beta diversity analyses. The Shannon diversity was higher in ET samples in comparison to EF samples (P = 0.06, Supporting Information Figure 2 in Ref. III). This is also concordant with a previous study showing higher alpha diversity in ET samples (Liu et al., 2018). Furthermore, there was a clear clustering between the sample types in beta diversity (Figure 2E in Ref. III). Interestingly, we saw now clustering based on individual which is usually common in human gut samples (Figure 2D and 2E in Ref. III).

Our results conclude that the main driver between the differences in endometrial samples is the genus *Lactobacillus*. Furthermore, other research has shown that *Lactobacillus* is an important genus in terms of fertility problems and its role in female reproduction needs further investigation. Additional work using metagenomics would be beneficial in understanding the importance of *Lactobacillus* on species level as well as search for pathways which could be associated with the female reproduction. Deeper understanding of the role of *Lactobacillus* in fertility could lead into discovery of microbiome-based biomarker helping to determine the causes of infertility and potentially other problems related to female reproductive health.

CONCLUSIONS

Understanding the effects of microbes on human health and disease has a great potential in improving the field of personalized medicine by adding microbial component to the equation. Microbiome data is a valuable asset in understanding the mechanisms behind the etiology and mechanisms behind different disease states in human body. Comprehensive microbiome studies elucidate the roles of bacteria in humans and provide bases for future explorative research to deepen the understanding of the interplay between human as a host and microorganisms inhabiting the human body.

The main conclusions drawn from this thesis are as follows:

- Large-scale meta-analysis incorporating multiple cohorts across the world are needed to unravel the associations between the gut microbiome and genetics. A large MWAS with 18,340 individuals was able to detect associations between 31 genetic loci and microbiome as well as confirm previously described association between the *LCT* locus and the *Bifidobacterium* genus. The huge interindividual variability in microbiome as well as the heavy burden of multiple testing present in microbiome studies makes it complicated to uncover associations between rarer taxa and genetic variants. This is a clear indication that in the future, studies with additional individuals are needed to fully understand the interrelation between microbiome and genetics.

- The role of gut microbiome in PCOS most likely works through the metabolic traits. PCOS is highly linked with metabolic health and this thesis shows that the differences in the gut microbiome profile become evident when considering the prediabetic state of PCOS women. The results shown here indicate lower alpha diversity and significantly higher abundance of genus *Dorea* in prediabetic PCOS women compared to women with PCOS but normal glucose tolerance.

- The differences seen in the microbiome of endometrial tissue and fluid samples in women undergoing IVF treatment are driven by the *Lactobacillus* genus. This finding confirms the importance of *Lactobacillus* in the female reproductive health. The study as a whole illustrates the need in microbiome studies to move further from the gut microbiome and closer to the active site of the disease, in this case the endometrium. This knowledge provides an opportunity to find new microbiome-based biomarkers for early prediction, better diagnosis and more accurate treatment.

SUMMARY IN ESTONIAN

Inimese mikrobioomi mõjutavad faktorid ning seosed naiste tervisega

Viimaste paari kümnendi edasiarengud sekveneerimistehnoloogiates on hüppeliselt kaasa aidanud mikrobioomi kui teadusharu uuringutele. DNA analüüsimeetodite täiustamine on teinud võimalikuks oma eluks anaeroobset keskkonda vajavate bakterite uurimise, mis on mikrobioomi uuringute jaoks avanud uue maailma. Nüüd on võimalik ulatuslikumalt uurida soolestikus ning mujal organismis elavaid baktereid ning otsida nende seoseid tervisega.

Inimese mikrobioomikooslus omab suurt olulisust meie tervisele, sellel on mõju paljudele protsessidele organismis nagu näiteks ainevahetusele, immuun- ja närvisüsteemi toimimisele. Tulenevalt mikrobioomi olulisusest organismi heaolule, on oluline välja selgitada, mis on need faktorid, mis mikrobioomikooslust ning seeläbi meie tervist mõjutavad.

Käesolev doktoriväitekiri uurib seoseid inimese geneetilise varieeruvuse ning mikrobioomi koosseisu vahel ja analüüsib lähemalt, kuidas mikrobioom võib mõjutada naiste tervist. Teaduskirjandusele toetudes annab töö esimene osa ülevaate sellest, millised faktorid mõjutavad enim soolestiku mikrobioomikooslust ja selle varieeruvust inimeste vahel ning kuidas analüüsi läbiviimine võib teadustulemusi mõjutada. Arutluse all on lisaks mikrobioomi ja geneetika omavahelised seosed ning lähemalt vaadatakse seoseid nii soolestiku kui ka reproduktiivsüsteemi mikrobioomi ja naiste tervise vahel. Kirjanduse ülevaate viimases osas arutletakse mikrobioomi teaduse tulevikust, mõtiskledes selle kasutamise võimaluste üle tervise hindamisel. Doktoritöö eksperimentaalosas analüüsitakse kõigepealt soolestiku mikroobide ning inimese geneetilise varieeruvuse vahelisi seoseid suuremahulises koostööprojektis. Lisaks uuritakse soolestiku mikroobide mõju polütsüstiliste munasarjade sündroomile (PCOS) ning analüüsitakse endomeetriumis leiduvaid baktereid ning nende potentsiaalset efekti naiste reproduktiivtervisele.

Geneetika ning mikrobioomivaheliste seoste uurimise suurimaks takistuseks võib lugeda mikrobioomi suurt varieeruvust inimeste vahel ning geneetika küllaltki väikest mõju mikrobioomi varieeruvusele. Seetõttu jääb väikese valimiga teadustöödel puudu vajalikust uuringu statistilisest võimusest, mis võimaldaks tuvastada geneetika ja mikrobioomi vahelisi seoseid. Valimi suuruse probleemidest ülesaamiseks kasutatakse geneetiliste seoste tuvastamiseks inimese genoomist konsortsiumipõhiseid lähenemisi, kus toimub erinevate uurimisgruppide andmete ühine analüüs. Käesolevas töö uuriti inimese geneetilise varieeruvuse ja mikrobioomi vahelisi seoseid suures koostööprojektis (MiBioGen konsortsium), kus analüüsiti rohkem kui 18 000 inimese soolestiku mikrobioomi andmeid. Töö tulemusena tuvastati 31 geenilookuse mõju soole mikrobioomile ja kinnitati varem näidatud seos *LCT* geenilookuse ning *Bifidobacterium* perekonna vahel. Lisaks saadud tulemustele näitas antud töö erinevatest faktoritest mõjutatud mikrobioomi suurt varieeruvust (sh. uuringu disain ja erinevused populatsioonide vahel), mis raskendab erinevate uuringute andmete kombineerimist. Töö illustreerib, et

harvaesinevate bakterite seose tuvastamiseks geneetikaga, on vajalik kordades suurema valimi kaasamine.

Mikrobioomi mõju inimese tervisele on näidatud paljude erinevate haiguste puhul, ulatudes kaugemale soolestiku tervisest, kus elab suurim mikroobikooslus inimese organismis. Viimastel aastatel on üha enam avaldatud uuringuid teemadel, mis näitavad naiste reproduktiivtervise ja mikrobioomi vahelisi seoseid. PCOS on üheks enimlevinumaks endokriinhaiguseks viljakas eas olevatel naistel, mõjutades nii naise ainevahetust kui ka reproduktiivtervist. Leidsime oma töös, et soolestiku mikrobioomis olev bakterikooslus, millesse kuuluvad bakteriperekonnad *Paraprevotella*, *Streptococcus*, *Eubacterium ventriosum* ning *Bifidobacterium*, on võimeline eristama PCOS diagnoosiga naisi tervetest naistest. Kuna meie tulemused näitasid, et PCOS'ga seoses olnud soolestiku mikroobid on samuti seotud erinevate ainevahetust mõjutavate tunnustega nagu kehamassiindeks, glükoosi ja insuliini tasemed, püstitasime hüpoteesi, et soolestiku mikrobioomi mõju PCOS'i puhul võib toimida läbi ainevahetuslike protsesside. Selle testimiseks võrdlesime soolestiku mikrobioomi erinevusi normaalse glükoositaluvusega ning eeldiabeediga PCOS naistel. Meie andmed näitasid, et eeldiabeedi tunnustega PCOS diagnoosiga naistel oli märkimisväärselt vähenenud soolestiku mikroobikoosluse mitmekesisus ning kõrgem *Dorea* perekonna osakaal võrreldes normaalse glükoositaluvusega PCOS naistega. *Dorea* oli lisaks nii meie kui varasemates publikatsioonides seotud erinevate PCOS'ga seotud metaboolsete tunnustega nagu kehamassiindeks ning glükoosi- ja insuliinitasemed.

Soolestiku mikrobioomi roll on väga kompleksne ning mõjutab kogu organismi tervikuna, kuid on oluline uurida ka selliseid mikroorganisme, kes elavad väljaspool soolestikku ja konkreetselt uurimise all oleva haigusega seotud keskkondades. Üheks osaks inimese tervisest on viljakus ning oluline on analüüsida viljakusega seotud mikrobioomi muutusi reporduktiivorganitest võetud proovidest. On teada mitmeid patogeene, kelle esinemine reproduktiivteedes toob kaasa põletiku. Oma töös võrdlesime viljakusprobleemidega naiste mikroobikooslusi nii endomeetriumi koest kui ka emakasisesest vedelikust võetud proovidest. Tulemustest järeldus, et suurim endomeetriumi mikrobioomi mõjutaja on perekond *Lactobacillus*. Saadud tulemused on kooskõlas varasemate teadustöödega, kus on kahanenud *Lactobacillus* perekonna sagedust reproduktiivteedes on seostatud vähenenud viljakusega. Teadmised endomeetriumi mikrobioomi koosseisust ja mõjust reproduktiivsusele on olulise tähtsusega, kuna see võimaldab tulevikus leida mikrobioomipõhiseid biomarkereid, mis aitaksid ennustada võimalikke probleeme viljakusega ja seeläbi aitaksid kaasa uute ravivõimaluste arendamisele.

Kokkuvõttes andis käesolev doktoritöö uusi teadmisi mikrobioomi ja geneetika vahelistest seostest ning laiendas oluliselt meie teadmisi mikrobioomi ja naiste tervise valdkonnas.

REFERENCES

Aasmets, O., Krigul, K.L., Lüll, K., Metspalu, A., and Org, E. (2022). Gut metagenome associations with extensive digital health data in a volunteer-based Estonian microbiome cohort. Nat. Commun. *13*, 869.

Albenberg, L.G., and Wu, G.D. (2014). Diet and the intestinal microbiome: associations, functions, and implications for health and disease. Gastroenterology *146*, 1564–1572.

Arora, T., and Bäckhed, F. (2016). The gut microbiota and metabolic disease: current understanding and future perspectives. J. Intern. Med. *280*, 339–349.

Azziz, R., Carmina, E., Chen, Z., Dunaif, A., Laven, J.S.E., Legro, R.S., Lizneva, D., Natterson-Horowtiz, B., Teede, H.J., and Yildiz, B.O. (2016). Polycystic ovary syndrome. Nat. Rev. Dis. Prim. *2*, 16057.

Bäckhed, F., Ding, H., Wang, T., Hooper, L. V, Koh, G.Y., Nagy, A., Semenkovich, C.F., and Gordon, J.I. (2004). The gut microbiota as an environmental factor that regulates fat storage. Proc. Natl. Acad. Sci. U. S. A. *101*, 15718–15723.

Bajaj, J.S., Hylemon, P.B., Ridlon, J.M., Heuman, D.M., Daita, K., White, M.B., Monteith, P., Noble, N.A., Sikaroodi, M., and Gillevet, P.M. (2012). Colonic mucosal microbiome differs from stool microbiome in cirrhosis and hepatic encephalopathy and is linked to cognition and inflammation. Am. J. Physiol. Liver Physiol. *303*, G675–G685.

Baker, J.M., Chase, D.M., and Herbst-Kralovetz, M.M. (2018). Uterine Microbiota: Residents, Tourists, or Invaders? Front. Immunol. *9*, 208.

Barber, T.M., Wass, J.A.H., McCarthy, M.I., and Franks, S. (2007). Metabolic characteristics of women with polycystic ovaries and oligo-amenorrhoea but normal androgen levels: implications for the management of polycystic ovary syndrome. Clin. Endocrinol. (Oxf). *66*, 513–513.

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

Bolyen, E., Rideout, J.R., Dillon, M.R., Bokulich, N.A., Abnet, C.C., Al-Ghalith, G.A., Alexander, H., Alm, E.J., Arumugam, M., Asnicar, F., et al. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat. Biotechnol. *37*, 852–857.

Bonder, M.J., Kurilshikov, A., Tigchelaar, E.F., Mujagic, Z., Imhann, F., Vila, A.V., Deelen, P., Vatanen, T., Schirmer, M., Smeekens, S.P., et al. (2016). The effect of host genetics on the gut microbiome. Nat. Genet. 1–9.

Byrd, A.L., Belkaid, Y., and Segre, J.A. (2018). The human skin microbiome. Nat. Rev. Microbiol. *16*, 143–155.

Chen, C., Song, X., Wei, W., Zhong, H., Dai, J., Lan, Z., Li, F., Yu, X., Feng, Q., Wang, Z., et al. (2017). The microbiota continuum along the female reproductive tract and its relation to uterine-related diseases. Nat. Commun. *8*, 875.

Chen, L., Wang, D., Garmaeva, S., Kurilshikov, A., Vich Vila, A., Gacesa, R., Sinha, T., Segal, E., Weersma, R.K., Wijmenga, C., et al. (2021). The long-term genetic stability and individual specificity of the human gut microbiome. Cell *184*, 2302–2315.e12.

Chow, J., and Mazmanian, S.K. (2010). A Pathobiont of the Microbiota Balances Host Colonization and Intestinal Inflammation. Cell Host Microbe *7*, 265–276.

Chu, W., Han, Q., Xu, J., Wang, J., Sun, Y., Li, W., Chen, Z.J., and Du, Y. (2020). Metagenomic analysis identified microbiome alterations and pathological association between intestinal microbiota and polycystic ovary syndrome. Fertil. Steril. *113*, 1286-1298.e4.

Cook, M.D., Allen, J.M., Pence, B.D., Wallig, M.A., Gaskins, H.R., White, B.A., and Woods, J.A. (2016). Exercise and gut immune function: evidence of alterations in colon immune cell homeostasis and microbiome characteristics with exercise training. Immunol. Cell Biol. *94*, 158–163.

Cregger, M.A., Lenz, K., Leary, E., Leach, R., Fazleabas, A., White, B., and Braundmeier, A. (2017). Reproductive Microbiomes: Using the Microbiome as a Novel Diagnostic Tool for Endometriosis. Reprod. Immunol. Open Access *02*, 1–7.

Crost, E.H., Le Gall, G., Laverde-Gomez, J.A., Mukhopadhya, I., Flint, H.J., and Juge, N. (2018). Mechanistic Insights Into the Cross-Feeding of Ruminococcus gnavus and Ruminococcus bromii on Host and Dietary Carbohydrates. Front. Microbiol. *9*, 2558.

Dalile, B., Van Oudenhove, L., Vervliet, B., and Verbeke, K. (2019). The role of short-chain fatty acids in microbiota–gut–brain communication. Nat. Rev. Gastroenterol. Hepatol. *16*, 461–478.

Das, A.P., Kumar, P.S., and Swain, S. (2014). Recent advances in biosensor based endotoxin detection. Biosens. Bioelectron. *51*, 62–75.

Davenport, E.R., Cusanovich, D.A., Michelini, K., Barreiro, L.B., Ober, C., and Gilad, Y. (2015). Genome-Wide Association Studies of the Human Gut Microbiota. PLoS One *10*, e0140301.

David, L.A., Maurice, C.F., Carmody, R.N., Gootenberg, D.B., Button, J.E., Wolfe, B.E., Ling, A. V, Devlin, A.S., Varma, Y., Fischbach, M.A., et al. (2014). Diet rapidly and reproducibly alters the human gut microbiome. Nature *505*, 559–563.

Debelius, J., Song, S.J., Vazquez-Baeza, Y., Xu, Z.Z., Gonzalez, A., and Knight, R. (2016). Tiny microbes, enormous impacts: What matters in gut microbiome studies? Genome Biol. *17*, 1–12.

Doestzada, M., Vila, A.V., Zhernakova, A., Koonen, D.P.Y., Weersma, R.K., Touw, D.J., Kuipers, F., Wijmenga, C., and Fu, J. (2018). Pharmacomicrobiomics: a novel route towards personalized medicine? Protein Cell *9*, 432–445.

Duan, L., An, X., Zhang, Y., Jin, D., Zhao, S., Zhou, R., Duan, Y., Zhang, Y., Liu, X., and Lian, F. (2021). Gut microbiota as the critical correlation of polycystic ovary syndrome and type 2 diabetes mellitus. Biomed. Pharmacother. *142*, 112094.

Durack, J., and Lynch, S. V. (2019). The gut microbiome: Relationships with disease and opportunities for therapy. J. Exp. Med. *216*, 20–40.

Durazzi, F., Sala, C., Castellani, G., Manfreda, G., Remondini, D., and De Cesare, A. (2021). Comparison between 16S rRNA and shotgun sequencing data for the taxonomic characterization of the gut microbiota. Sci. Rep. *11*, 3030.

Eyupoglu, N.D., Ergunay, K., Acikgoz, A., Akyon, Y., Yilmaz, E., and Yildiz, B.O. (2020). Gut microbiota and oral contraceptive use in overweight and obese patients with polycystic ovary syndrome. J. Clin. Endocrinol. Metab.

Falony, G., Joossens, M., Vieira-Silva, S., Wang, J., Darzi, Y., Faust, K., Kurilshikov, A., Bonder, M.J., Valles-Colomer, M., Vandeputte, D., et al. (2016). Population-level analysis of gut microbiome variation. Science (80–.). *352*, 560–564.

Fan, Y., and Pedersen, O. (2021). Gut microbiota in human metabolic health and disease. Nat. Rev. Microbiol. *19*, 55–71.

Fang, R.L., Chen, L.X., Shu, W.S., Yao, S.Z., Wang, S.W., and Chen, Y.Q. (2016). Barcoded sequencing reveals diverse intrauterine microbiomes in patients suffering with endometrial polyps. Am. J. Transl. Res. *8*, 1581–1592.

Forslund, S.K., Chakaroun, R., Zimmermann-Kogadeeva, M., Markó, L., Aron-Wisnewsky, J., Nielsen, T., Moitinho-Silva, L., B Schmidt, T.S., Falony, G., Vieira-Silva, S., et al. (2021). Combinatorial, additive and dose-dependent drug–microbiome associations. Nataliya Sokolovska *600*, 500.

Franasiak, J.M., Werner, M.D., Juneau, C.R., Tao, X., Landis, J., Zhan, Y., Treff, N.R., and Scott, R.T. (2016). Endometrial microbiome at the time of embryo transfer: next-generation sequencing of the 16S ribosomal subunit. J. Assist. Reprod. Genet. *33*, 129–136.

Gacesa, R., Kurilshikov, A., Vila, A.V., Sinha, T., Klaassen, M.A.Y., Bolte, L.A., Andreu-Sánchez, S., Chen, L., Collij, V., Hu, S., et al. (2020). The Dutch Microbiome Project defines factors that shape the healthy gut microbiome. BioRxiv *2020.11.27*.

Gérard, C., and Vidal, H. (2019). Impact of Gut Microbiota on Host Glycemic Control. Front. Endocrinol. (Lausanne). *10*, 1–13.

Gilbert, J.A., Blaser, M.J., Caporaso, J.G., Jansson, J.K., Lynch, S. V, and Knight, R. (2018). Current understanding of the human microbiome. Nat. Med. *24*, 392–400.

Gloor, G.B., Macklaim, J.M., Pawlowsky-Glahn, V., and Egozcue, J.J. (2017). Microbiome Datasets Are Compositional: And This Is Not Optional. Front. Microbiol. *8*, 1–6.

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., et al. (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.

Greenacre, M., Martínez-Álvaro, M., and Blasco, A. (2021). Compositional Data Analysis of Microbiome and Any-Omics Datasets: A Validation of the Additive Logratio Transformation. Front. Microbiol. *12*, 1–11.

Gurung, M., Li, Z., You, H., Rodrigues, R., Jump, D.B., Morgun, A., and Shulzhenko, N. (2020). Role of gut microbiota in type 2 diabetes pathophysiology. EBioMedicine *51*, 102590.

Harborne, L., Fleming, R., Lyall, H., Norman, J., and Sattar, N. (2003). Descriptive review of the evidence for the use of metformin in polycystic ovary syndrome. Lancet *361*, 1894–1901.

Harris, H.R., Babic, A., Webb, P.M., Nagle, C.M., Jordan, S.J., Risch, H.A., Rossing, M.A., Doherty, J.A., Goodman, M.T., Modugno, F., et al. (2018). Polycystic Ovary Syndrome, Oligomenorrhea, and Risk of Ovarian Cancer Histotypes: Evidence from the Ovarian Cancer Association Consortium. Cancer Epidemiol. Biomarkers Prev. *27*, 174–182.

Hashimoto, T., and Kyono, K. (2019). Does dysbiotic endometrium affect blastocyst implantation in IVF patients? J. Assist. Reprod. Genet. *36*, 2471–2479.

He, F.F., and Li, Y.M. (2020). Role of gut microbiota in the development of insulin resistance and the mechanism underlying polycystic ovary syndrome: A review. J. Ovarian Res. *13*, 1–13.

He, Y., Jin, X., Wang, H., Dai, H., Lu, X., Zhao, J., Zhang, H., Chen, W., and Wang, G. (2021). The emerging role of the gut microbiome in polycystic ovary syndrome. F&S Rev. *2*, 214–226.

Hernandes, C., Silveira, P., Rodrigues Sereia, A.F., Christoff, A.P., Mendes, H., Valter de Oliveira, L.F., and Podgaec, S. (2020). Microbiome Profile of Deep Endometriosis Patients: Comparison of Vaginal Fluid, Endometrium and Lesion. Diagnostics *10*, 163.

Ho, B., Ryback, D., Benson, B., Mason, C.N., Torres, P.J., Quinn, R.A., Thackray, V.G., and Kelley, S.T. (2021). Gut Metabolites Are More Predictive of Disease and Cohoused States than Gut Bacterial Features in a Polycystic Ovary Syndrome-Like Mouse Model. MSystems *6*, e01149-20.

Hornung, B.V.H., Zwittink, R.D., and Kuijper, E.J. (2019). Issues and current standards of controls in microbiome research. FEMS Microbiol. Ecol. *95*.

Hughes, D.A., Bacigalupe, R., Wang, J., Rühlemann, M.C., Tito, R.Y., Falony, G., Joossens, M., Vieira-Silva, S., Henckaerts, L., Rymenans, L., et al. (2020). Genome-wide associations of human gut microbiome variation and implications for causal inference analyses. Nat. Microbiol. *5*, 1079–1087.

Insenser, M., Murri, M., Del Campo, R., Martínez-García, M.Á., Fernández-Durán, E., and Escobar-Morreale, H.F. (2018). Gut Microbiota and the Polycystic Ovary Syndrome: Influence of Sex, Sex Hormones, and Obesity. J. Clin. Endocrinol. Metab. *103*, 2552–2562.

Jackson, M.A., Verdi, S., Maxan, M.-E., Shin, C.M., Zierer, J., Bowyer, R.C.E., Martin, T., Williams, F.M.K., Menni, C., Bell, J.T., et al. (2018). Gut microbiota associations with common diseases and prescription medications in a population-based cohort. Nat. Commun. *9*, 2655.

Jobira, B., Frank, D.N., Pyle, L., Silveira, L.J., Kelsey, M.M., Garcia-Reyes, Y., Robertson, C.E., Ir, D., Nadeau, K.J., and Cree-Green, M. (2020). Obese Adolescents With PCOS Have Altered Biodiversity and Relative Abundance in Gastrointestinal Microbiota. J. Clin. Endocrinol. Metab. *105*, e2134–e2144.

Jones, J., Reinke, S.N., Ali, A., Palmer, D.J., and Christophersen, C.T. (2021). Fecal sample collection methods and time of day impact microbiome composition and short chain fatty acid concentrations. Sci. Rep. *11*, 13964.

Jovel, J., Patterson, J., Wang, W., Hotte, N., O'Keefe, S., Mitchel, T., Perry, T., Kao, D., Mason, A.L., Madsen, K.L., et al. (2016). Characterization of the gut microbiome using 16S or shotgun metagenomics. Front. Microbiol. *7*, 1–17.

Kashyap, P.C., Marcobal, A., Ursell, L.K., Smits, S.A., Sonnenburg, E.D., Costello, E.K., Higginbottom, S.K., Domino, S.E., Holmes, S.P., Relman, D.A., et al. (2013). Genetically dictated change in host mucus carbohydrate landscape exerts a diet-dependent effect on the gut microbiota. Proc. Natl. Acad. Sci. U. S. A. *110*, 17059–17064.

Khan, K.N., Fujishita, A., Masumoto, H., Muto, H., Kitajima, M., Masuzaki, H., and Kitawaki, J. (2016). Molecular detection of intrauterine microbial colonization in women with endometriosis. Eur. J. Obstet. Gynecol. Reprod. Biol. *199*, 69–75.

Kitaya, K., Nagai, Y., Arai, W., Sakuraba, Y., and Ishikawa, T. (2019). Characterization of Microbiota in Endometrial Fluid and Vaginal Secretions in Infertile Women with Repeated Implantation Failure. Mediators Inflamm. *2019*, 4893437.

Kolde, R., Franzosa, E.A., Rahnavard, G., Hall, A.B., Vlamakis, H., Stevens, C., Daly, M.J., Xavier, R.J., and Huttenhower, C. (2018). Host genetic variation and its microbiome interactions within the Human Microbiome Project. Genome Med. *10*, 1–13.

Konstantinidis, T., Tsigalou, C., Karvelas, A., Stavropoulou, E., Voidarou, C., and Bezirtzoglou, E. (2020). Effects of Antibiotics upon the Gut Microbiome: A Review of the Literature. Biomedicines *8*, 502.

Kurilshikov, A., Wijmenga, C., Fu, J., and Zhernakova, A. (2017). Host Genetics and Gut Microbiome: Challenges and Perspectives. Trends Immunol. *38*, 633–647.

Kurilshikov, A., Medina-Gomez, C., Bacigalupe, R., Radjabzadeh, D., Wang, J., Demirkan, A., Le Roy, C.I., Raygoza Garay, J.A., Finnicum, C.T., Liu, X., et al. (2021). Large-scale association analyses identify host factors influencing human gut microbiome composition. Nat. Genet. *53*, 156–165.

Kyono, K., Hashimoto, T., Nagai, Y., and Sakuraba, Y. (2018). Analysis of endometrial microbiota by 16S ribosomal RNA gene sequencing among infertile patients: a single-center pilot study. Reprod. Med. Biol. *17*, 297–306.

Laakso, M., Kuusisto, J., Stančáková, A., Kuulasmaa, T., Pajukanta, P., Lusis, A.J., Collins, F.S., Mohlke, K.L., and Boehnke, M. (2017). The Metabolic Syndrome in Men study: a resource for studies of metabolic and cardiovascular diseases. J. Lipid Res. *58*, 481–493.

Ley, R.E., Peterson, D.A., and Gordon, J.I. (2006). Ecological and evolutionary forces shaping microbial diversity in the human intestine. Cell *124*, 837–848.

Li, D., Gao, C., Zhang, F., Yang, R., Lan, C., Ma, Y., and Wang, J. (2020). Seven facts and five initiatives for gut microbiome research. Protein Cell *11*, 391–400.

Lin, H., and Peddada, S. Das (2020). Analysis of microbial compositions: a review of normalization and differential abundance analysis. Npj Biofilms Microbiomes *6*, 60.

Lindheim, L., Bashir, M., Münzker, J., Trummer, C., Zachhuber, V., Leber, B., Horvath, A., Pieber, T.R., Gorkiewicz, G., Stadlbauer, V., et al. (2017). Alterations in Gut Microbiome Composition and Barrier Function Are Associated with Reproductive and Metabolic Defects in Women with Polycystic Ovary Syndrome (PCOS): A Pilot Study. PLoS One *12*, e0168390.

Liu, R., Zhang, C., Shi, Y., Zhang, F., Li, L., Wang, X., Ling, Y., Fu, H., Dong, W., Shen, J., et al. (2017). Dysbiosis of Gut Microbiota Associated with Clinical Parameters in Polycystic Ovary Syndrome. Front. Microbiol. *8*, 324.

Liu, Y., Wong, K.K.W., Ko, E.Y.L., Chen, X., Huang, J., Tsui, S.K.W., Li, T.C., and Chim, S.S.C. (2018). Systematic comparison of bacterial colonization of endometrial tissue and fluid samples in recurrent miscarriage patients: Implications for future endometrial microbiome studies. Clin. Chem. *64*, 1743–1752.

Liu, Y., Ko, E.Y.-L., Wong, K.K.-W., Chen, X., Cheung, W.-C., Law, T.S.-M., Chung, J.P.-W., Tsui, S.K.-W., Li, T.-C., and Chim, S.S.-C. (2019). Endometrial microbiota in infertile women with and without chronic endometritis as diagnosed using a quantitative and reference range-based method. Fertil. Steril. *112*, 707-717.e1.

Lloyd-Price, J., Arze, C., Ananthakrishnan, A.N., Schirmer, M., Avila-Pacheco, J., Poon, T.W., Andrews, E., Ajami, N.J., Bonham, K.S., Brislawn, C.J., et al. (2019). Multi-omics of the gut microbial ecosystem in inflammatory bowel diseases. Nature *569*, 655–662.

Lopera-Maya, E.A., Kurilshikov, A., van der Graaf, A., Hu, S., Andreu-Sánchez, S., Chen, L., Vila, A.V., Gacesa, R., Sinha, T., Collij, V., et al. (2022). Effect of host genetics on the gut microbiome in 7,738 participants of the Dutch Microbiome Project. Nat. Genet.

Lozupone, C., Lladser, M.E., Knights, D., Stombaugh, J., and Knight, R. (2011). UniFrac: an effective distance metric for microbial community comparison. ISME J. *5*, 169–172.

Maier, L., Pruteanu, M., Kuhn, M., Zeller, G., Telzerow, A., Anderson, E.E., Brochado, A.R., Fernandez, K.C., Dose, H., Mori, H., et al. (2018). Extensive impact of non-antibiotic drugs on human gut bacteria. Nature *555*, 623–628.

Malesza, I.J., Malesza, M., Walkowiak, J., Mussin, N., Walkowiak, D., Aringazina, R., Bartkowiak-Wieczorek, J., and Mądry, E. (2021). High-Fat, Western-Style Diet, Systemic Inflammation, and Gut Microbiota: A Narrative Review. Cells *10*, 3164.

Mammadova, G., Ozkul, C., Yilmaz Isikhan, S., Acikgoz, A., and Yildiz, B.O. (2021). Characterization of gut microbiota in polycystic ovary syndrome: Findings from a lean population. Eur. J. Clin. Invest. *51*, e13417.

March, W.A., Moore, V.M., Willson, K.J., Phillips, D.I.W., Norman, R.J., and Davies, M.J. (2010). The prevalence of polycystic ovary syndrome in a community sample assessed under contrasting diagnostic criteria. Hum. Reprod. *25*, 544–551.

Marchesi, J.R., and Ravel, J. (2015). The vocabulary of microbiome research: a proposal. Microbiome *3*, 31.

Martinez, K.B., Leone, V., and Chang, E.B. (2017). Western diets, gut dysbiosis, and metabolic diseases: Are they linked? Gut Microbes *8*, 130–142.

Menni, C., Zhu, J., Le Roy, C.I., Mompeo, O., Young, K., Rebholz, C.M., Selvin, E., North, K.E., Mohney, R.P., Bell, J.T., et al. (2020). Serum metabolites reflecting gut microbiome alpha diversity predict type 2 diabetes. Gut Microbes *11*, 1632–1642.

Moghetti, P., Tosi, F., Bonin, C., Di Sarra, D., Fiers, T., Kaufman, J.-M., Giagulli, V.A., Signori, C., Zambotti, F., Dall'Alda, M., et al. (2013). Divergences in Insulin Resistance Between the Different Phenotypes of the Polycystic Ovary Syndrome. J. Clin. Endocrinol. Metab. *98*, E628–E637.

Molina, N.M., Sola-Leyva, A., Jose Saez-Lara, M., Plaza-Diaz, J., Tubic-Pavlovic, A., Romero, B., Clavero, A., Mozas-Moreno, J., Fontes, J., and Altmäe, S. (2020). New opportunities for endometrial health by modifying uterine microbial composition: Present or future? Biomolecules *10*, 593.

Moran, L.J., Misso, M.L., Wild, R.A., and Norman, R.J. (2010). Impaired glucose tolerance, type 2 diabetes and metabolic syndrome in polycystic ovary syndrome: a systematic review and meta-analysis. Hum. Reprod. Update *16*, 347–363.

Moreno-Indias, I., Sánchez-Alcoholado, L., Sánchez-Garrido, M.Á., Martín-Núñez, G.M., Pérez-Jiménez, F., Tena-Sempere, M., Tinahones, F.J., and Queipo-Ortuño, M.I. (2016). Neonatal Androgen Exposure Causes Persistent Gut Microbiota Dysbiosis Related to Metabolic Disease in Adult Female Rats. Endocrinology *157*, 4888–4898.

Moreno, I., Codoñer, F.M., Vilella, F., Valbuena, D., Martinez-Blanch, J.F., Jimenez-Almazán, J., Alonso, R., Alamá, P., Remohí, J., Pellicer, A., et al. (2016). Evidence that the endometrial microbiota has an effect on implantation success or failure. Am. J. Obstet. Gynecol. *215*, 684–703.

Moreno, I., Cicinelli, E., Garcia-Grau, I., Gonzalez-Monfort, M., Bau, D., Vilella, F., De Ziegler, D., Resta, L., Valbuena, D., and Simon, C. (2018). The diagnosis of chronic endometritis in infertile asymptomatic women: a comparative study of histology, microbial cultures, hysteroscopy, and molecular microbiology. Am. J. Obstet. Gynecol. *218*, 602.e1-602.e16.

Naderpoor, N., Mousa, A., Gomez-Arango, L., Barrett, H., Nitert, M.D., and de Courten, B. (2019). Faecal Microbiota Are Related to Insulin Sensitivity and Secretion in Overweight or Obese Adults. J. Clin. Med. *8*, 452.

Nearing, J.T., Douglas, G.M., Hayes, M., Macdonald, J., Desai, D., Allward, N., Jones, C.M.A., Wright, R., Dhanani, A., Comeau, A.M., et al. (2021). Microbiome differential abundance methods produce disturbingly different results across 38 datasets. BioRxiv 2021.05.10.443486.

Neuman, H., Forsythe, P., Uzan, A., Avni, O., and Koren, O. (2018). Antibiotics in early life: dysbiosis and the damage done. FEMS Microbiol. Rev. *018*, 489–499.

Ottosson, F., Brunkwall, L., Ericson, U., Nilsson, P.M., Almgren, P., Fernandez, C., Melander, O., and Orho-Melander, M. (2018). Connection Between BMI-Related Plasma Metabolite Profile and Gut Microbiota. J. Clin. Endocrinol. Metab. *103*, 1491–1501.

Palomba, S., de Wilde, M.A., Falbo, A., Koster, M.P.H., La Sala, G.B., and Fauser, B.C.J.M. (2015). Pregnancy complications in women with polycystic ovary syndrome. Hum. Reprod. Update *21*, 575–592.

Penington, J.S., Penno, M.A.S., Ngui, K.M., Ajami, N.J., Roth-Schulze, A.J., Wilcox, S.A., Bandala-Sanchez, E., Wentworth, J.M., Barry, S.C., Brown, C.Y., et al. (2018). Influence of fecal collection conditions and 16S rRNA gene sequencing at two centers on human gut microbiota analysis. Sci. Rep. *8*, 4386.

Punzón-Jiménez, P., and Labarta, E. (2021). The impact of the female genital tract microbiome in women health and reproduction: a review. J. Assist. Reprod. Genet. *38*, 2519–2541.

Qi, X., Yun, C., Sun, L., Xia, J., Wu, Q., Wang, Y., Wang, L., Zhang, Y., Liang, X., Wang, L., et al. (2019). Gut microbiota–bile acid–interleukin-22 axis orchestrates polycystic ovary syndrome. Nat. Med. *25*, 1225–1233.

Qin, J., Li, Y., Cai, Z., Li, S., Zhu, J., Zhang, F., Liang, S., Zhang, W., Guan, Y., Shen, D., et al. (2012). A metagenome-wide association study of gut microbiota in type 2 diabetes. Nature *490*, 55–60.

Qin, Y., Havulinna, A.S., Liu, Y., Jousilahti, P., Ritchie, S.C., Tokolyi, A., Sanders, J.G., Valsta, L., Brożyńska, M., Zhu, Q., et al. (2022). Combined effects of host genetics and diet on human gut microbiota and incident disease in a single population cohort. Nat. Genet. *54*, 134–142.

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., and Glöckner, F.O. (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res. *41*, D590-6.

Ramirez, J., Guarner, F., Bustos Fernandez, L., Maruy, A., Sdepanian, V.L., and Cohen, H. (2020). Antibiotics as Major Disruptors of Gut Microbiota. Front. Cell. Infect. Microbiol. *10*.

Rantakallio, P. (1988). The longitudinal study of the northern Finland birth cohort of 1966. Paediatr. Perinat. Epidemiol. *2*, 59–88.

Ridaura, V.K., Faith, J.J., Rey, F.E., Cheng, J., Duncan, A.E., Kau, A.L., Griffin, N.W., Lombard, V., Henrissat, B., Bain, J.R., et al. (2013). Gut Microbiota from Twins Discordant for Obesity Modulate Metabolism in Mice. Science (80-.). *341*, 1241214.

Rothschild, D., Weissbrod, O., Barkan, E., Kurilshikov, A., Korem, T., Zeevi, D., Costea, P.I., Godneva, A., Kalka, iris N., Bar, N., et al. (2018). Environment dominates over host genetics in shaping human gut microbiota. Nature *555*, 210–215.

Rotterdam ESHRE/ASRM-Sponsored PCOS Consensus Workshop Group (2004). Revised 2003 consensus on diagnostic criteria and long-term health risks related to polycystic ovary syndrome. Fertil. Steril. *81*, 19–25.

Rowland, I., Gibson, G., Heinken, A., Scott, K., Swann, J., Thiele, I., and Tuohy, K. (2018). Gut microbiota functions: metabolism of nutrients and other food components. Eur. J. Nutr. *57*, 1–24.

Rudnicka, E., Suchta, K., Grymowicz, M., Calik-Ksepka, A., Smolarczyk, K., Duszewska, A.M., Smolarczyk, R., and Meczekalski, B. (2021). Chronic Low Grade Inflammation in Pathogenesis of PCOS. Int. J. Mol. Sci. *22*, 3789.

Rühlemann, M.C., Hermes, B.M., Bang, C., Doms, S., Moitinho-Silva, L., Thingholm, L.B., Frost, F., Degenhardt, F., Wittig, M., Kässens, J., et al. (2021). Genome-wide association study in 8,956 German individuals identifies influence of ABO histo-blood groups on gut microbiome. Nat. Genet. *53*, 147–155.

Salosensaari, A., Laitinen, V., Havulinna, A.S., Meric, G., Cheng, S., Perola, M., Valsta, L., Alfthan, G., Inouye, M., Watrous, J.D., et al. (2021). Taxonomic signatures of cause-specific mortality risk in human gut microbiome. Nat. Commun. *12*, 2671.

Senghora, B., Sokhnab, C., Ruimycde, R., and Lagiera, J.-C. (2018). Gut microbiota diversity according to dietary habits and geographical provenance. Hum. Micro-biome J. *7–8*, 1–9.

Sheehan, M.T. (2003). Polycystic Ovarian Syndrome: Diagnosis and Management.

Shilo, S., Bar, N., Keshet, A., Talmor-Barkan, Y., Rossman, H., Godneva, A., Aviv, Y., Edlitz, Y., Reicher, L., Kolobkov, D., et al. (2021). 10 K: a large-scale prospective longitudinal study in Israel. Eur. J. Epidemiol. *36*, 1187–1194.

Sun, Z., Huang, S., Zhang, M., Zhu, Q., Haiminen, N., Carrieri, A.P., Vázquez-Baeza, Y., Parida, L., Kim, H.C., Knight, R., et al. (2021). Challenges in benchmarking metagenomic profilers. Nat. Methods *18*, 618–626.

Teede, H.J., Misso, M.L., Costello, M.F., Dokras, A., Laven, J., Moran, L., Piltonen, T., Norman, R.J., and Network, I.P. (2018). Recommendations from the international evidence-based guideline for the assessment and management of polycystic ovary syndrome. Hum. Reprod. *33*, 1602–1618.

Torres, P.J., Siakowska, M., Banaszewska, B., Pawelczyk, L., Duleba, A.J., Kelley, S.T., and Thackray, V.G. (2018). Gut Microbial Diversity in Women With Polycystic Ovary Syndrome Correlates With Hyperandrogenism. J. Clin. Endocrinol. Metab. *103*, 1502–1511.

Turnbaugh, P.J., Ley, R.E., Mahowald, M.A., Magrini, V., Mardis, E.R., and Gordon, J.I. (2006). An obesity-associated gut microbiome with increased capacity for energy harvest. Nature *444*, 1027–1031.

Turpin, W., Espin-Garcia, O., Xu, W., Silverberg, M.S., Kevans, D., Smith, M.I., Gutt-man, D.S., Griffiths, A., Panaccione, R., Otley, A., et al. (2016). Association of host genome with intestinal microbial composition in a large healthy cohort. Nat. Genet. *48*, 1413–1417.

University of Oulu University of Oulu: Northern Finland Birth Cohort 1966.

Vandeputte, D., Tito, R.Y., Vanleeuwen, R., Falony, G., and Raes, J. (2017). Practical considerations for large-scale gut microbiome studies. FEMS Microbiol. Rev. *41*, S154–S167.

Vandeputte, D., De Commer, L., Tito, R.Y., Kathagen, G., Sabino, J., Vermeire, S., Faust, K., and Raes, J. (2021). Temporal variability in quantitative human gut micro-biome profiles and implications for clinical research. Nat. Commun. *12*, 6740.

Vich Vila, A., Collij, V., Sanna, S., Sinha, T., Imhann, F., Bourgonje, A.R., Mujagic, Z., Jonkers, D.M.A.E., Masclee, A.A.M., Fu, J., et al. (2020). Impact of commonly used drugs on the composition and metabolic function of the gut microbiota. Nat. Commun. *11*, 1–11.

Vogt, N.M., Kerby, R.L., Dill-McFarland, K.A., Harding, S.J., Merluzzi, A.P., Johnson, S.C., Carlsson, C.M., Asthana, S., Zetterberg, H., Blennow, K., et al. (2017). Gut micro-biome alterations in Alzheimer's disease. Sci. Rep. *7*, 13537.

Wahlström, A., Sayin, S.I., Marschall, H.U., and Bäckhed, F. (2016). Intestinal Crosstalk between Bile Acids and Microbiota and Its Impact on Host Metabolism. Cell Metab. *24*, 41–50.

Walker, R.L., Vlamakis, H., Lee, J.W.J., Besse, L.A., Xanthakis, V., Vasan, R.S., Shaw, S.Y., and Xavier, R.J. (2021). Population study of the gut microbiome: associations with diet, lifestyle, and cardiometabolic disease. Genome Med. *13*, 188.

Walsh, D.M., Hokenstad, A.N., Chen, J., Sung, J., Jenkins, G.D., Chia, N., Nelson, H., Mariani, A., and Walther-Antonio, M.R.S. (2019). Postmenopause as a key factor in the composition of the Endometrial Cancer Microbiome (ECbiome). Sci. Rep. *9*, 19213.

Walter, J., Armet, A.M., Finlay, B.B., and Shanahan, F. (2020). Establishing or Exaggerating Causality for the Gut Microbiome: Lessons from Human Microbiota-Associated Rodents. Cell *180*, 221–232.

Walther-António, M.R.S., Chen, J., Multinu, F., Hokenstad, A., Distad, T.J., Cheek, E.H., Keeney, G.L., Creedon, D.J., Nelson, H., Mariani, A., et al. (2016). Potential contribution of the uterine microbiome in the development of endometrial cancer. Genome Med. *8*, 1–15.

Wang, J., Thingholm, L.B., Skiecevičie, J., Rausch, P., Kummen, M., Hov, J.R., Degenhardt, F., Heinsen, F.A., Rühlemann, M.C., Szymczak, S., et al. (2016). Genome-wide association analysis identifies variation in Vitamin D receptor and other host factors influencing the gut microbiota. Nat. Genet. *48*, 1396–1406.

Wang, J., Kurilshikov, A., Radjabzadeh, D., Turpin, W., Croitoru, K., Bonder, M.J., Jackson, M.A., Medina-Gomez, C., Frost, F., Homuth, G., et al. (2018). Meta-analysis of human genome-microbiome association studies: the MiBioGen consortium initiative. Microbiome *6*, 101.

Wee, B.A., Thomas, M., Sweeney, E.L., Frentiu, F.D., Samios, M., Ravel, J., Gajer, P., Myers, G., Timms, P., Allan, J.A., et al. (2018). A retrospective pilot study to determine whether the reproductive tract microbiota differs between women with a history of infertility and fertile women. Aust. New Zeal. J. Obstet. Gynaecol. *58*, 341–348.

Wiegratz, I., Kutschera, E., Lee, J.H., Moore, C., Mellinger, U., Winkler, U.H., and Kuhl, H. (2003). Effect of four different oral contraceptives on various sex hormones and serum-binding globulins. Contraception *67*, 25–32.

Wong, S.H., and Yu, J. (2019). Gut microbiota in colorectal cancer: mechanisms of action and clinical applications. Nat. Rev. Gastroenterol. Hepatol. *16*, 690–704.

Wong, R.G., Wu, J.R., and Gloor, G.B. (2016). Expanding the UniFrac Toolbox. PLoS One *11*, e0161196.

Woo, V., and Alenghat, T. (2017). Host–microbiota interactions: epigenomic regulation. Curr. Opin. Immunol. *44*, 52–60.

Wu, G.D., Chen, J., Hoffmann, C., Bittinger, K., Chen, Y.-Y., Keilbaugh, S.A., Bewtra, M., Knights, D., Walters, W.A., Knight, R., et al. (2011). Linking long-term dietary patterns with gut microbial enterotypes. Science *334*, 105–108.

Xie, H., Guo, R., Zhong, H., Feng, Q., Lan, Z., Qin, B., Ward, K.J., Jackson, M.A., Xia, Y., Chen, X., et al. (2016). Shotgun Metagenomics of 250 Adult Twins Reveals Genetic and Environmental Impacts on the Gut Microbiome. Cell Syst. *3*, 572-584.e3.

Zeng, B., Lai, Z., Sun, L., Zhang, Z., Yang, J., Li, Z., Lin, J., and Zhang, Z. (2019). Structural and functional profiles of the gut microbial community in polycystic ovary syndrome with insulin resistance (IR-PCOS): a pilot study. Res. Microbiol. *170*, 43–52.

Zeng, X., Xie, Y.-J., Liu, Y.-T., Long, S.-L., and Mo, Z.-C. (2020). Polycystic ovarian syndrome: Correlation between hyperandrogenism, insulin resistance and obesity. Clin. Chim. Acta *502*, 214–221.

Zhang, C., Thakkar, P. V., Powell, S.E., Sharma, P., Vennelaganti, S., Betel, D., and Shah, M.A. (2019a). A Comparison of Homogenization vs. Enzymatic Lysis for Microbiome Profiling in Clinical Endoscopic Biopsy Tissue Samples. Front. Microbiol. *9*, 1–9.

Zhang, J., Sun, Z., Jiang, S., Bai, X., Ma, C., Peng, Q., Chen, K., Chang, H., Fang, T., and Zhang, H. (2019b). Probiotic Bifidobacterium lactis V9 Regulates the Secretion of Sex Hormones in Polycystic Ovary Syndrome Patients through the Gut-Brain Axis. MSystems *4*, e00017-19.

Zhernakova, A., Kurilshikov, A., Bonder, M.J., Tigchelaar, E.F., Schirmer, M., Vatanen, T., Mujagic, Z., Vila, A.V., Falony, G., Vieira-Silva, S., et al. (2016). Population-based metagenomics analysis reveals markers for gut microbiome composition and diversity. Science (80-.). *352*, 565–569.

Zhou, L., Ni, Z., Cheng, W., Yu, J., Sun, S., Zhai, D., Yu, C., and Cai, Z. (2020). Characteristic gut microbiota and predicted metabolic functions in women with PCOS. Endocr Connect *9*, 63–73.

ACKNOWLEDGMENTS

To start with, I cannot believe that I have actually made it to the part where I am writing the acknowledgments of my PhD thesis. For sure, this came to be the reality only thanks to many wonderful people who I have crossed paths with during this rollercoaster called PhD studies.

First and foremost, my deepest gratitude goes to my amazing supervisor Elin Org who made the bold move of taking me in as her first PhD student and introducing me to the incredible world of microbiome. No words can express the true gratitude I have to you for guiding me through all these years sharing wisdom, giving advice and being so patient with me. Elin, you definitely deserve the title of world's best supervisor! Plus, you have the amazing ability not only being great in science but also knowing that life is about so much more. Therefore, I thank you for all the fun activities outside the office you have put together for our team. Not many are lucky to have the opportunity to go to mushroom picking and snowboarding with their supervisor. Me and I think the whole microbiome team is so incredibly blessed to have you (well, except when it is April Fool's day).

My deepest gratitude goes to Prof Andres Metspalu thanks to whom Estonian Genome Center is the place it is today. I am internally grateful that you welcomed me here when I was still doing my Bachelor's studies and thank you for always encouraging PhD students to dream big and never give up. Talking about my Bachelor's, I want to thank Prof Reedik Mägi, who supervised me before PhD and introduced me to the possibility of doing PhD in the first place. God knows that I had no intentions to be in science but I am grateful that you sent me to hear about the possible projects available that eventually lead me to Elin and microbiome. I sincerely want to thank Prof Andres Salumets for sending two interesting projects towards me that eventually ended up being a big part of my thesis.

I wish to thank all the co-authors I have had the privilege to work together with. This thesis would not have become a reality without you.

And of course none of it would not be possible without the most amazing fellow microbiome PhD students Kertu and Oliver. A big curtsy to the both of you. Sharing the office with you has been amazing with equal parts of science and fun. Kertu, you are always so full of life and positivity that it will lift everybody's spirits and helps to get through even the hardest of days. It is incredible how passionate you are about everything you do. Oliver, the best microbiome statistician in the world without the slightest doubt. Thank you for enlightening me on microbiome analysis and making even equations understandable, I can assure that no-one has managed that for a very long time. And thank you for the bringing in the additional stress relief force Nuru.

A huge thank you to my dearest fellow non-microbiome doctoral students Marili, Maarja, Anette, Nele and Ivana with whom I have had the honor to walk

through the difficult PhD journey. I am forever grateful for your support, motivation and friendships. Ivana, I want to thank you for all the "PhD depression cure sessions" we have had over the years, they were so needed and 100% made the world a better place to live in. Marili, your optimism is contagious. Thank you for keeping up the good spirits even at times when it seemed impossible. Anette, thank you for sharing the office with me for years, I loved how you popped up behind the wall with fancy candies. Maarja, I really love your no fuss straight to business attitude and honesty as well as the way you always stand up for yourself, I really need to learn that from you. And of course big regards to the original office doggo Käbi. Nele, how did you manage to clone yourself? I mean, you have to have done it how else is it possible to manage to do all that you do!? I need a recipe ASAP. A huge "thank you" goes to Kristi who did not make the cut for fellow doctoral students list since she is a badass who already graduated and was the biggest role model showing that it is actually doable! Tusen tack ska du ha! And thank you all who have found time to do sports with me! I truly think sport is essential in life and I am so thankful that I always have someone to go for a jog or have a little muscle pumping session with.

With all my heart I wish to praise and thank the amazing group of friends who go by the union name "homaarid" – Mona, Niko, Kiiri, Freddy, Kristel, Ahto, Anette, Mikk, Marili and Oliwer. I really have loved all the get-togethers and sauna sessions where so many world problems have been solved. I am so happy that within the last few years there have also been so many wonderful news from all of you guys. Cannot wait for the future with the extended version of homaarid!

Dear Matthias, you get your own little paragraph since bff's do not fit anywhere else. I just want to thank you for being you. The most remarkable friend who definitely knows a bit too much to let go of. I cherish your spontaneous being which always leads to crazy adventures whether it is a 1-hour trip that should take 10 minutes in the city center of Tartu or on an unknown (but totally safe) mountain path in Montenegro.

My dear fellow Aussie moms Merle and Marta, thank you for finding your way into my life. I am so grateful for all the hikes and walks together with you and your wonderful fur babies. One thing is that thanks to Maya[2] and Rio Teemu has as crazy friends as himself to play with. But even more importantly, I value the refreshing chats with you that have nothing to do with science and PhD.

And of course my family, mom Eelika, dad Kunnar and big sister Angela. Without you I would not be at the place in my life where I am today. My dearest dad, despite you have probably been completely confused by what I have been doing during the past years you have always supported me through everything. I am forever grateful for your support, wisdom and wittiness. Also, thank you for always sending too many potatoes and cucumbers to Tartu for me to eat. I am truly blessed to have you as my father! My sweetest and too opinionated mom, I love you. Thank you for always believing in me and supporting me, it means the world to always have someone who believes that you can do it even if you do not believe it yourself. My dear sister Angela, thank you for making me an aunt to the two most precious little humans. I admire your strength as a mother and hope

to be half as good as you are at it. Words cannot express the gratitude and love I have for my family!

Finally, my favorite boys in the whole world – my love Endrik and my precious little Teemu. Sweet Teemu boy, thank you for picking us as your family and growing out of your tyrannosaurus teenager times to become the best dog there could ever be. You are the best! My dearest Endrik, thank you for your never dying support, encouraging words and ensuring that everything will go just fine in every aspect of life. I am truly grateful for you putting up with me even when I may have not been the most reasonable person. I love how fun and silly you can be and always manage to put a smile on my face. I am looking forward to spending the life with you! I love you more!

PUBLICATIONS

CURRICULUM VITAE

Name: Kreete Lüll
Date of birth: October 21, 1992
Contact: Estonian Genome Center, Institute of Genomics, University of Tartu, Riia 23B, 51010, Tartu, Estonia
E-mail: kreete.lull@ut.ee

Education:

2021–…	PhD in Gene Technology, Institute of Genomics, Faculty of Science and Technology, University of Tartu, Estonia (transferred due to a change in the structural unit)
2016–2021	PhD in Gene Technology, Institute of Molecular and Cell Biology, Faculty of Science and Technology, University of Tartu, Estonia
2014–2016	Master's degree in Gene Technology, Institute of Molecular and Cell Biology, Faculty of Science and Technology, University of Tartu, Estonia
2011–2014	Bachelor's degree in Biology, Institute of Molecular and Cell Biology, Faculty of Science and Technology, University of Tartu, Estonia

Professional employment:

2019–…	Estonian Genome Center, Institute of Genomics, University of Tartu, specialist
2014–2016	Medical Pharmacy Group OÜ, product developer

Administrative work:

2015–…	Member of Estonian Society of Human Genetics

Publications:

Aasmets, O; Krigul, KL; **Lüll, K**; Metspalu, A; Org, E (2022). Gut metagenome associations with extensive digital health data in a volunteer-based Estonian microbiome cohort. *Nat. Comm.* 13 (1), 869.

Lüll K, Saare M, Peters M, Kakhiani E, Zhdanova A, Salumets A, Boyarsky K, Org, E. (2022) Differences in microbial profile of endometrial fluid and tissue samples in women with IVF failure are driven by *Lactobacillus* abundance. *AOGS* 101 (2), 212–220.

Lüll K, Arffman RK, Sola-Leyva A, Molina NM, Aasmets O, Herzig K-H, Plaza-Díaz J, Franks S, Morin-Papunen L, Tapanainen JS, Salumets A, Altmäe S, Piltonen TT, Org E. (2021). The gut microbiome in polycystic ovary syndrome and its association with metabolic traits. *J. Clin. Endocrinol. Metab*. 106 (3), 858–871.

Aasmets, O; **Lüll, K**; Lang, JM; Pan, C; Kuusisto, J; Fischer, K; Laakso, M; Lusis, AJ.; Org, E (2021). Machine Learning Reveals Time-Varying Microbial Predictors with Complex Effects on Glucose Regulation. *mSystems*, 6 (1), ARTN e01191–20

Kurilshikov A, Medina-Gomez C, Bacigalupe R, Radjabzadeh D, Wang J, Demirkan A, Le Roy CI, Raygoza Garay JA, Finnicum CT, Liu X, Zhernakova DV, Bonder MJ, Hansen TH, Frost F, Rühlemann MC, Turpin W, Moon J-Y, Kim H-N, **Lüll K**, et al. (2021). Large-scale association analyses identify host factors influencing human gut microbiome composition. *Nat. Genet.* 53 (2), 156–165.

Krigul, KL; Aasmets, O; **Lüll, K**; Org, T; Org, E. (2021). Using fecal immunochemical tubes for the analysis of the gut microbiome has the potential to improve colorectal cancer screening. *Scientific Reports*, 11 (1), 19603.

Wang J, Kurilshikov A, Radjabzadeh D, Turpin W, Croitoru K, Bonder MJ, Jackson MA, Medina-Gomez C, Frost F, Homuth G, Rühlemann M, Hughes D, Kim HN; **MiBioGen Consortium Initiative**, Spector TD, Bell JT, Steves CJ, Timpson N, Franke A, Wijmenga C, Meyer K, Kacprowski T, Franke L, Paterson AD, Raes J, Kraaij R, Zhernakova A. (2018). Meta-analysis of human genome-microbiome association studies: the MiBioGen consortium initiative. *Microbiome*, 6 (1), 101.

Awards and stipends:

2021 Doctoral School of Biomedicine and Biotechnology Scholarship: poster presentation at the 8th International Human Microbiome Consortium virtual conference

2020 Doctoral School of Biomedicine and Biotechnology Scholarship: poster presentation at the Exploring Human Host-Microbiome Interactions in Health and Disease conference, Wellcome Genome Campus United Kingdom

2020 Doctoral School of Biomedicine and Biotechnology Scholarship: oral presentation at the Annual PCOS Meeting, Levi, Finland

2018 Dora Plus Scholarship for short-term study mobility: poster presentation at the Metagenomics Bioinformatics course, Hinxton, Great Britain

2016 Scholarship in smart specialization growth areas from Archimedes Foundation

ELULOOKIRJELDUS

Nimi: Kreete Lüll
Sünniaeg: 21. oktoober 1992
Aadress: Eesti Geenivaramu, genoomika instituut, Tartu ülikool, Riia 23B, 51010, Tartu, Eesti
E-post: kreete.lull@ut.ee

Hariduskäik:

2021–2022	Doktoriõpe, geenitehnoloogia erialal, genoomika instituut, loodus- ja täppisteaduste valdkond, Tartu ülikool, Eesti (viidud üle seoses struktuuriüksuse muutusega)
2016–2021	Doktoriõpe, geenitehnoloogia erialal, molekulaar- ja rakubioloogia instituut, loodus- ja täppisteaduste valdkond, Tartu ülikool, Eesti
2014–2016	Magistriõpe, geenitehnoloogia erialal, molekulaar- ja rakubioloogia instituut, loodus- ja täppisteaduste valdkond, Tartu ülikool, Eesti
2011–2014	Bakalaureuseõpe, bioloogia erialal, molekulaar- ja rakubioloogia instituut, loodus- ja täppisteaduste valdkond, Tartu ülikool, Eesti

Teenistuskäik:

2019–…	Eesti Geenivaramu, genoomika instituut, Tartu ülikool, spetsialist
2014–2016	Medical Pharmacy Group OÜ, tootearendaja

Teadusorganisatsiooniline ja- administratiivne tegevus:

2015–…	Eesti Inimesegeneetika ühingu liige

Teaduspublikatsioonid:

Aasmets, O; Krigul, KL; **Lüll, K**; Metspalu, A; Org, E (2022). Gut metagenome associations with extensive digital health data in a volunteer-based Estonian microbiome cohort. *Nat. Comm.* 13 (1), 869.

Lüll K, Saare M, Peters M, Kakhiani E, Zhdanova A, Salumets A, Boyarsky K, Org, E. (2022) Differences in microbial profile of endometrial fluid and tissue samples in women with IVF failure are driven by *Lactobacillus* abundance. *AOGS* 101 (2), 212–220.

Lüll K, Arffman RK, Sola-Leyva A, Molina NM, Aasmets O, Herzig K-H, Plaza-Díaz J, Franks S, Morin-Papunen L, Tapanainen JS, Salumets A, Altmäe S, Piltonen TT, Org E. (2021). The gut microbiome in polycystic ovary syndrome and its association with metabolic traits. *J. Clin. Endocrinol. Metab*. 106 (3), 858–871.

Aasmets, O; **Lüll, K**; Lang, JM; Pan, C; Kuusisto, J; Fischer, K; Laakso, M; Lusis, AJ.; Org, E (2021). Machine Learning Reveals Time-Varying Microbial Predictors with Complex Effects on Glucose Regulation. *mSystems*, 6 (1), ARTN e01191–20

Kurilshikov A, Medina-Gomez C, Bacigalupe R, Radjabzadeh D, Wang J, Demirkan A, Le Roy CI, Raygoza Garay JA, Finnicum CT, Liu X, Zhernakova DV, Bonder MJ, Hansen TH, Frost F, Rühlemann MC, Turpin W, Moon J-Y, Kim H-N, **Lüll K**, et al. (2021). Large-scale association analyses identify host factors influencing human gut microbiome composition. *Nat Genet* 53 (2), 156–165.

Krigul, KL; Aasmets, O; **Lüll, K**; Org, T; Org, E. (2021). Using fecal immunochemical tubes for the analysis of the gut microbiome has the potential to improve colorectal cancer screening. *Scientific Reports*, 11 (1), 19603.

Wang J, Kurilshikov A, Radjabzadeh D, Turpin W, Croitoru K, Bonder MJ, Jackson MA, Medina-Gomez C, Frost F, Homuth G, Rühlemann M, Hughes D, Kim HN; **MiBioGen Consortium Initiative**, Spector TD, Bell JT, Steves CJ, Timpson N, Franke A, Wijmenga C, Meyer K, Kacprowski T, Franke L, Paterson AD, Raes J, Kraaij R, Zhernakova A. (2018). Meta-analysis of human genome-microbiome association studies: the MiBioGen consortium initiative. *Microbiome*, 6 (1), 101.

Stipendiumid:

2021 Biomeditsiini ja biotehnoloogia doktorikooli stipendium: posterettekanne kaheksandal *International Human Microbiome Consortium* virtuaalkonverentsil

2020 Biomeditsiini ja biotehnoloogia doktorikooli stipendium: posterettekanne *Exploring Human Host-Microbiome Interactions in Health and Disease* konverentsil, Wellcome Genome Campus, Suurbritannia

2020 Biomeditsiini ja biotehnoloogia doktorikooli stipendium: suuline ettekanne *Annual PCOS Meeting* konverentsil, Levi, Soome

2018 Dora Plus doktorantide lühiajalise õpirände stipendium: posterettekanne *Metagenomics Bioinformatics* kursusel, Hinxton, Suurbritannia

2016 Sihtasutus Archimedes nutika spetsialiseerumise kasvuvaldkondade doktorandistipendium

DISSERTATIONES BIOLOGICAE UNIVERSITATIS TARTUENSIS

1. **Toivo Maimets**. Studies of human oncoprotein p53. Tartu, 1991, 96 p.
2. **Enn K. Seppet**. Thyroid state control over energy metabolism, ion transport and contractile functions in rat heart. Tartu, 1991, 135 p.
3. **Kristjan Zobel**. Epifüütsete makrosamblike väärtus õhu saastuse indikaatoritena Hamar-Dobani boreaalsetes mägimetsades. Tartu, 1992, 131 lk.
4. **Andres Mäe**. Conjugal mobilization of catabolic plasmids by transposable elements in helper plasmids. Tartu, 1992, 91 p.
5. **Maia Kivisaar**. Studies on phenol degradation genes of *Pseudomonas* sp. strain EST 1001. Tartu, 1992, 61 p.
6. **Allan Nurk**. Nucleotide sequences of phenol degradative genes from *Pseudomonas sp.* strain EST 1001 and their transcriptional activation in *Pseudomonas putida.* Tartu, 1992, 72 p.
7. **Ülo Tamm**. The genus *Populus* L. in Estonia: variation of the species biology and introduction. Tartu, 1993, 91 p.
8. **Jaanus Remme**. Studies on the peptidyltransferase centre of the *E.coli* ribosome. Tartu, 1993, 68 p.
9. **Ülo Langel**. Galanin and galanin antagonists. Tartu, 1993, 97 p.
10. **Arvo Käärd**. The development of an automatic online dynamic fluorescense-based pH-dependent fiber optic penicillin flowthrought biosensor for the control of the benzylpenicillin hydrolysis. Tartu, 1993, 117 p.
11. **Lilian Järvekülg**. Antigenic analysis and development of sensitive immunoassay for potato viruses. Tartu, 1993, 147 p.
12. **Jaak Palumets**. Analysis of phytomass partition in Norway spruce. Tartu, 1993, 47 p.
13. **Arne Sellin**. Variation in hydraulic architecture of *Picea abies* (L.) Karst. trees grown under different enviromental conditions. Tartu, 1994, 119 p.
13. **Mati Reeben**. Regulation of light neurofilament gene expression. Tartu, 1994, 108 p.
14. **Urmas Tartes**. Respiration rhytms in insects. Tartu, 1995, 109 p.
15. **Ülo Puurand**. The complete nucleotide sequence and infections *in vitro* transcripts from cloned cDNA of a potato A potyvirus. Tartu, 1995, 96 p.
16. **Peeter Hõrak**. Pathways of selection in avian reproduction: a functional framework and its application in the population study of the great tit (*Parus major*). Tartu, 1995, 118 p.
17. **Erkki Truve**. Studies on specific and broad spectrum virus resistance in transgenic plants. Tartu, 1996, 158 p.
18. **Illar Pata**. Cloning and characterization of human and mouse ribosomal protein S6-encoding genes. Tartu, 1996, 60 p.
19. **Ülo Niinemets**. Importance of structural features of leaves and canopy in determining species shade-tolerance in temperature deciduous woody taxa. Tartu, 1996, 150 p.

20. **Ants Kurg**. Bovine leukemia virus: molecular studies on the packaging region and DNA diagnostics in cattle. Tartu, 1996, 104 p.
21. **Ene Ustav**. E2 as the modulator of the BPV1 DNA replication. Tartu, 1996, 100 p.
22. **Aksel Soosaar**. Role of helix-loop-helix and nuclear hormone receptor transcription factors in neurogenesis. Tartu, 1996, 109 p.
23. **Maido Remm**. Human papillomavirus type 18: replication, transformation and gene expression. Tartu, 1997, 117 p.
24. **Tiiu Kull**. Population dynamics in *Cypripedium calceolus* L. Tartu, 1997, 124 p.
25. **Kalle Olli**. Evolutionary life-strategies of autotrophic planktonic microorganisms in the Baltic Sea. Tartu, 1997, 180 p.
26. **Meelis Pärtel**. Species diversity and community dynamics in calcareous grassland communities in Western Estonia. Tartu, 1997, 124 p.
27. **Malle Leht**. The Genus *Potentilla* L. in Estonia, Latvia and Lithuania: distribution, morphology and taxonomy. Tartu, 1997, 186 p.
28. **Tanel Tenson**. Ribosomes, peptides and antibiotic resistance. Tartu, 1997, 80 p.
29. **Arvo Tuvikene**. Assessment of inland water pollution using biomarker responses in fish *in vivo* and *in vitro*. Tartu, 1997, 160 p.
30. **Urmas Saarma**. Tuning ribosomal elongation cycle by mutagenesis of 23S rRNA. Tartu, 1997, 134 p.
31. **Henn Ojaveer**. Composition and dynamics of fish stocks in the gulf of Riga ecosystem. Tartu, 1997, 138 p.
32. **Lembi Lõugas**. Post-glacial development of vertebrate fauna in Estonian water bodies. Tartu, 1997, 138 p.
33. **Margus Pooga**. Cell penetrating peptide, transportan, and its predecessors, galanin-based chimeric peptides. Tartu, 1998, 110 p.
34. **Andres Saag**. Evolutionary relationships in some cetrarioid genera (Lichenized Ascomycota). Tartu, 1998, 196 p.
35. **Aivar Liiv**. Ribosomal large subunit assembly *in vivo*. Tartu, 1998, 158 p.
36. **Tatjana Oja**. Isoenzyme diversity and phylogenetic affinities among the eurasian annual bromes (*Bromus* L., Poaceae). Tartu, 1998, 92 p.
37. **Mari Moora**. The influence of arbuscular mycorrhizal (AM) symbiosis on the competition and coexistence of calcareous grassland plant species. Tartu, 1998, 78 p.
38. **Olavi Kurina**. Fungus gnats in Estonia (*Diptera: Bolitophilidae, Keroplatidae, Macroceridae, Ditomyiidae, Diadocidiidae, Mycetophilidae*). Tartu, 1998, 200 p.
39. **Andrus Tasa**. Biological leaching of shales: black shale and oil shale. Tartu, 1998, 98 p.
40. **Arnold Kristjuhan**. Studies on transcriptional activator properties of tumor suppressor protein p53. Tartu, 1998, 86 p.
41. **Sulev Ingerpuu**. Characterization of some human myeloid cell surface and nuclear differentiation antigens. Tartu, 1998, 163 p.

42. **Veljo Kisand**. Responses of planktonic bacteria to the abiotic and biotic factors in the shallow lake Võrtsjärv. Tartu, 1998, 118 p.
43. **Kadri Põldmaa**. Studies in the systematics of hypomyces and allied genera (Hypocreales, Ascomycota). Tartu, 1998, 178 p.
44. **Markus Vetemaa**. Reproduction parameters of fish as indicators in environmental monitoring. Tartu, 1998, 117 p.
45. **Heli Talvik**. Prepatent periods and species composition of different *Oesophagostomum* spp. populations in Estonia and Denmark. Tartu, 1998, 104 p.
46. **Katrin Heinsoo**. Cuticular and stomatal antechamber conductance to water vapour diffusion in *Picea abies* (L.) karst. Tartu, 1999, 133 p.
47. **Tarmo Annilo**. Studies on mammalian ribosomal protein S7. Tartu, 1998, 77 p.
48. **Indrek Ots**. Health state indicies of reproducing great tits (*Parus major*): sources of variation and connections with life-history traits. Tartu, 1999, 117 p.
49. **Juan Jose Cantero**. Plant community diversity and habitat relationships in central Argentina grasslands. Tartu, 1999, 161 p.
50. **Rein Kalamees**. Seed bank, seed rain and community regeneration in Estonian calcareous grasslands. Tartu, 1999, 107 p.
51. **Sulev Kõks**. Cholecystokinin (CCK) – induced anxiety in rats: influence of environmental stimuli and involvement of endopioid mechanisms and serotonin. Tartu, 1999, 123 p.
52. **Ebe Sild**. Impact of increasing concentrations of O_3 and CO_2 on wheat, clover and pasture. Tartu, 1999, 123 p.
53. **Ljudmilla Timofejeva**. Electron microscopical analysis of the synaptonemal complex formation in cereals. Tartu, 1999, 99 p.
54. **Andres Valkna**. Interactions of galanin receptor with ligands and G-proteins: studies with synthetic peptides. Tartu, 1999, 103 p.
55. **Taavi Virro**. Life cycles of planktonic rotifers in lake Peipsi. Tartu, 1999, 101 p.
56. **Ana Rebane**. Mammalian ribosomal protein S3a genes and intron-encoded small nucleolar RNAs U73 and U82. Tartu, 1999, 85 p.
57. **Tiina Tamm**. Cocksfoot mottle virus: the genome organisation and translational strategies. Tartu, 2000, 101 p.
58. **Reet Kurg**. Structure-function relationship of the bovine papilloma virus E2 protein. Tartu, 2000, 89 p.
59. **Toomas Kivisild**. The origins of Southern and Western Eurasian populations: an mtDNA study. Tartu, 2000, 121 p.
60. **Niilo Kaldalu**. Studies of the TOL plasmid transcription factor XylS. Tartu, 2000, 88 p.
61. **Dina Lepik**. Modulation of viral DNA replication by tumor suppressor protein p53. Tartu, 2000, 106 p.
62. **Kai Vellak**. Influence of different factors on the diversity of the bryophyte vegetation in forest and wooded meadow communities. Tartu, 2000, 122 p.

63. **Jonne Kotta**. Impact of eutrophication and biological invasionas on the structure and functions of benthic macrofauna. Tartu, 2000, 160 p.
64. **Georg Martin**. Phytobenthic communities of the Gulf of Riga and the inner sea the West-Estonian archipelago. Tartu, 2000, 139 p.
65. **Silvia Sepp**. Morphological and genetical variation of *Alchemilla L.* in Estonia. Tartu, 2000. 124 p.
66. **Jaan Liira**. On the determinants of structure and diversity in herbaceous plant communities. Tartu, 2000, 96 p.
67. **Priit Zingel**. The role of planktonic ciliates in lake ecosystems. Tartu, 2001, 111 p.
68. **Tiit Teder**. Direct and indirect effects in Host-parasitoid interactions: ecological and evolutionary consequences. Tartu, 2001, 122 p.
69. **Hannes Kollist**. Leaf apoplastic ascorbate as ozone scavenger and its transport across the plasma membrane. Tartu, 2001, 80 p.
70. **Reet Marits**. Role of two-component regulator system PehR-PehS and extracellular protease PrtW in virulence of *Erwinia Carotovora* subsp. *Carotovora*. Tartu, 2001, 112 p.
71. **Vallo Tilgar**. Effect of calcium supplementation on reproductive performance of the pied flycatcher *Ficedula hypoleuca* and the great tit *Parus major,* breeding in Nothern temperate forests. Tartu, 2002, 126 p.
72. **Rita Hõrak**. Regulation of transposition of transposon Tn*4652* in *Pseudomonas putida*. Tartu, 2002, 108 p.
73. **Liina Eek-Piirsoo**. The effect of fertilization, mowing and additional illumination on the structure of a species-rich grassland community. Tartu, 2002, 74 p.
74. **Krõõt Aasamaa**. Shoot hydraulic conductance and stomatal conductance of six temperate deciduous tree species. Tartu, 2002, 110 p.
75. **Nele Ingerpuu**. Bryophyte diversity and vascular plants. Tartu, 2002, 112 p.
76. **Neeme Tõnisson**. Mutation detection by primer extension on oligonucleotide microarrays. Tartu, 2002, 124 p.
77. **Margus Pensa**. Variation in needle retention of Scots pine in relation to leaf morphology, nitrogen conservation and tree age. Tartu, 2003, 110 p.
78. **Asko Lõhmus**. Habitat preferences and quality for birds of prey: from principles to applications. Tartu, 2003, 168 p.
79. **Viljar Jaks**. p53 – a switch in cellular circuit. Tartu, 2003, 160 p.
80. **Jaana Männik**. Characterization and genetic studies of four ATP-binding cassette (ABC) transporters. Tartu, 2003, 140 p.
81. **Marek Sammul**. Competition and coexistence of clonal plants in relation to productivity. Tartu, 2003, 159 p
82. **Ivar Ilves**. Virus-cell interactions in the replication cycle of bovine papillomavirus type 1. Tartu, 2003, 89 p.
83. **Andres Männik**. Design and characterization of a novel vector system based on the stable replicator of bovine papillomavirus type 1. Tartu, 2003, 109 p.

84. **Ivika Ostonen**. Fine root structure, dynamics and proportion in net primary production of Norway spruce forest ecosystem in relation to site conditions. Tartu, 2003, 158 p.
85. **Gudrun Veldre**. Somatic status of 12–15-year-old Tartu schoolchildren. Tartu, 2003, 199 p.
86. **Ülo Väli**. The greater spotted eagle *Aquila clanga* and the lesser spotted eagle *A. pomarina*: taxonomy, phylogeography and ecology. Tartu, 2004, 159 p.
87. **Aare Abroi**. The determinants for the native activities of the bovine papillomavirus type 1 E2 protein are separable. Tartu, 2004, 135 p.
88. **Tiina Kahre**. Cystic fibrosis in Estonia. Tartu, 2004, 116 p.
89. **Helen Orav-Kotta**. Habitat choice and feeding activity of benthic suspension feeders and mesograzers in the northern Baltic Sea. Tartu, 2004, 117 p.
90. **Maarja Öpik**. Diversity of arbuscular mycorrhizal fungi in the roots of perennial plants and their effect on plant performance. Tartu, 2004, 175 p.
91. **Kadri Tali**. Species structure of *Neotinea ustulata*. Tartu, 2004, 109 p.
92. **Kristiina Tambets**. Towards the understanding of post-glacial spread of human mitochondrial DNA haplogroups in Europe and beyond: a phylogeographic approach. Tartu, 2004, 163 p.
93. **Arvi Jõers**. Regulation of p53-dependent transcription. Tartu, 2004, 103 p.
94. **Lilian Kadaja**. Studies on modulation of the activity of tumor suppressor protein p53. Tartu, 2004, 103 p.
95. **Jaak Truu**. Oil shale industry wastewater: impact on river microbial community and possibilities for bioremediation. Tartu, 2004, 128 p.
96. **Maire Peters**. Natural horizontal transfer of the *pheBA* operon. Tartu, 2004, 105 p.
97. **Ülo Maiväli**. Studies on the structure-function relationship of the bacterial ribosome. Tartu, 2004, 130 p.
98. **Merit Otsus**. Plant community regeneration and species diversity in dry calcareous grasslands. Tartu, 2004, 103 p.
99. **Mikk Heidemaa**. Systematic studies on sawflies of the genera *Dolerus, Empria,* and *Caliroa* (Hymenoptera: Tenthredinidae). Tartu, 2004, 167 p.
100. **Ilmar Tõnno**. The impact of nitrogen and phosphorus concentration and N/P ratio on cyanobacterial dominance and N_2 fixation in some Estonian lakes. Tartu, 2004, 111 p.
101. **Lauri Saks**. Immune function, parasites, and carotenoid-based ornaments in greenfinches. Tartu, 2004, 144 p.
102. **Siiri Rootsi**. Human Y-chromosomal variation in European populations. Tartu, 2004, 142 p.
103. **Eve Vedler**. Structure of the 2,4-dichloro-phenoxyacetic acid-degradative plasmid pEST4011. Tartu, 2005. 106 p.
104. **Andres Tover**. Regulation of transcription of the phenol degradation *pheBA* operon in *Pseudomonas putida*. Tartu, 2005, 126 p.
105. **Helen Udras**. Hexose kinases and glucose transport in the yeast *Hansenula polymorpha*. Tartu, 2005, 100 p.

106. **Ave Suija**. Lichens and lichenicolous fungi in Estonia: diversity, distribution patterns, taxonomy. Tartu, 2005, 162 p.
107. **Piret Lõhmus**. Forest lichens and their substrata in Estonia. Tartu, 2005, 162 p.
108. **Inga Lips**. Abiotic factors controlling the cyanobacterial bloom occurrence in the Gulf of Finland. Tartu, 2005, 156 p.
109. **Krista Kaasik**. Circadian clock genes in mammalian clockwork, metabolism and behaviour. Tartu, 2005, 121 p.
110. **Juhan Javoiš**. The effects of experience on host acceptance in ovipositing moths. Tartu, 2005, 112 p.
111. **Tiina Sedman**. Characterization of the yeast *Saccharomyces cerevisiae* mitochondrial DNA helicase Hmi1. Tartu, 2005, 103 p.
112. **Ruth Aguraiuja**. Hawaiian endemic fern lineage *Diellia* (Aspleniaceae): distribution, population structure and ecology. Tartu, 2005, 112 p.
113. **Riho Teras**. Regulation of transcription from the fusion promoters generated by transposition of Tn*4652* into the upstream region of *pheBA* operon in *Pseudomonas putida*. Tartu, 2005, 106 p.
114. **Mait Metspalu**. Through the course of prehistory in India: tracing the mtDNA trail. Tartu, 2005, 138 p.
115. **Elin Lõhmussaar**. The comparative patterns of linkage disequilibrium in European populations and its implication for genetic association studies. Tartu, 2006, 124 p.
116. **Priit Kupper**. Hydraulic and environmental limitations to leaf water relations in trees with respect to canopy position. Tartu, 2006, 126 p.
117. **Heili Ilves**. Stress-induced transposition of Tn*4652* in *Pseudomonas Putida.* Tartu, 2006, 120 p.
118. **Silja Kuusk**. Biochemical properties of Hmi1p, a DNA helicase from *Saccharomyces cerevisiae* mitochondria. Tartu, 2006, 126 p.
119. **Kersti Püssa**. Forest edges on medium resolution landsat thematic mapper satellite images. Tartu, 2006, 90 p.
120. **Lea Tummeleht**. Physiological condition and immune function in great tits (*Parus major* l.): Sources of variation and trade-offs in relation to growth. Tartu, 2006, 94 p.
121. **Toomas Esperk**. Larval instar as a key element of insect growth schedules. Tartu, 2006, 186 p.
122. **Harri Valdmann**. Lynx (*Lynx lynx*) and wolf (*Canis lupus*) in the Baltic region: Diets, helminth parasites and genetic variation. Tartu, 2006. 102 p.
123. **Priit Jõers**. Studies of the mitochondrial helicase Hmi1p in *Candida albicans* and *Saccharomyces cerevisia*. Tartu, 2006. 113 p.
124. **Kersti Lilleväli**. Gata3 and Gata2 in inner ear development. Tartu, 2007, 123 p.
125. **Kai Rünk**. Comparative ecology of three fern species: *Dryopteris carthusiana* (Vill.) H.P. Fuchs, *D. expansa* (C. Presl) Fraser-Jenkins & Jermy and *D. dilatata* (Hoffm.) A. Gray (Dryopteridaceae). Tartu, 2007, 143 p.

126. **Aveliina Helm**. Formation and persistence of dry grassland diversity: role of human history and landscape structure. Tartu, 2007, 89 p.
127. **Leho Tedersoo**. Ectomycorrhizal fungi: diversity and community structure in Estonia, Seychelles and Australia. Tartu, 2007, 233 p.
128. **Marko Mägi**. The habitat-related variation of reproductive performance of great tits in a deciduous-coniferous forest mosaic: looking for causes and consequences. Tartu, 2007, 135 p.
129. **Valeria Lulla**. Replication strategies and applications of Semliki Forest virus. Tartu, 2007, 109 p.
130. **Ülle Reier**. Estonian threatened vascular plant species: causes of rarity and conservation. Tartu, 2007, 79 p.
131. **Inga Jüriado**. Diversity of lichen species in Estonia: influence of regional and local factors. Tartu, 2007, 171 p.
132. **Tatjana Krama**. Mobbing behaviour in birds: costs and reciprocity based cooperation. Tartu, 2007, 112 p.
133. **Signe Saumaa**. The role of DNA mismatch repair and oxidative DNA damage defense systems in avoidance of stationary phase mutations in *Pseudomonas putida*. Tartu, 2007, 172 p.
134. **Reedik Mägi**. The linkage disequilibrium and the selection of genetic markers for association studies in european populations. Tartu, 2007, 96 p.
135. **Priit Kilgas**. Blood parameters as indicators of physiological condition and skeletal development in great tits (*Parus major*): natural variation and application in the reproductive ecology of birds. Tartu, 2007, 129 p.
136. **Anu Albert**. The role of water salinity in structuring eastern Baltic coastal fish communities. Tartu, 2007, 95 p.
137. **Kärt Padari**. Protein transduction mechanisms of transportans. Tartu, 2008, 128 p.
138. **Siiri-Lii Sandre**. Selective forces on larval colouration in a moth. Tartu, 2008, 125 p.
139. **Ülle Jõgar**. Conservation and restoration of semi-natural floodplain meadows and their rare plant species. Tartu, 2008, 99 p.
140. **Lauri Laanisto**. Macroecological approach in vegetation science: generality of ecological relationships at the global scale. Tartu, 2008, 133 p.
141. **Reidar Andreson**. Methods and software for predicting PCR failure rate in large genomes. Tartu, 2008, 105 p.
142. **Birgot Paavel**. Bio-optical properties of turbid lakes. Tartu, 2008, 175 p.
143. **Kaire Torn**. Distribution and ecology of charophytes in the Baltic Sea. Tartu, 2008, 98 p.
144. **Vladimir Vimberg**. Peptide mediated macrolide resistance. Tartu, 2008, 190 p.
145. **Daima Örd**. Studies on the stress-inducible pseudokinase TRB3, a novel inhibitor of transcription factor ATF4. Tartu, 2008, 108 p.
146. **Lauri Saag**. Taxonomic and ecologic problems in the genus *Lepraria* (*Stereocaulaceae*, lichenised *Ascomycota*). Tartu, 2008, 175 p.

147. **Ulvi Karu**. Antioxidant protection, carotenoids and coccidians in greenfinches – assessment of the costs of immune activation and mechanisms of parasite resistance in a passerine with carotenoid-based ornaments. Tartu, 2008, 124 p.
148. **Jaanus Remm**. Tree-cavities in forests: density, characteristics and occupancy by animals. Tartu, 2008, 128 p.
149. **Epp Moks**. Tapeworm parasites *Echinococcus multilocularis* and *E. granulosus* in Estonia: phylogenetic relationships and occurrence in wild carnivores and ungulates. Tartu, 2008, 82 p.
150. **Eve Eensalu**. Acclimation of stomatal structure and function in tree canopy: effect of light and CO_2 concentration. Tartu, 2008, 108 p.
151. **Janne Pullat**. Design, functionlization and application of an *in situ* synthesized oligonucleotide microarray. Tartu, 2008, 108 p.
152. **Marta Putrinš**. Responses of *Pseudomonas putida* to phenol-induced metabolic and stress signals. Tartu, 2008, 142 p.
153. **Marina Semtšenko**. Plant root behaviour: responses to neighbours and physical obstructions. Tartu, 2008, 106 p.
154. **Marge Starast**. Influence of cultivation techniques on productivity and fruit quality of some *Vaccinium* and *Rubus* taxa. Tartu, 2008, 154 p.
155. **Age Tats**. Sequence motifs influencing the efficiency of translation. Tartu, 2009, 104 p.
156. **Radi Tegova**. The role of specialized DNA polymerases in mutagenesis in *Pseudomonas putida.* Tartu, 2009, 124 p.
157. **Tsipe Aavik**. Plant species richness, composition and functional trait pattern in agricultural landscapes – the role of land use intensity and landscape structure. Tartu, 2009, 112 p.
158. **Kaja Kiiver**. Semliki forest virus based vectors and cell lines for studying the replication and interactions of alphaviruses and hepaciviruses. Tartu, 2009, 104 p.
159. **Meelis Kadaja**. Papillomavirus Replication Machinery Induces Genomic Instability in its Host Cell. Tartu, 2009, 126 p.
160. **Pille Hallast**. Human and chimpanzee Luteinizing hormone/Chorionic Gonadotropin beta (*LHB/CGB*) gene clusters: diversity and divergence of young duplicated genes. Tartu, 2009, 168 p.
161. **Ain Vellak**. Spatial and temporal aspects of plant species conservation. Tartu, 2009, 86 p.
162. **Triinu Remmel**. Body size evolution in insects with different colouration strategies: the role of predation risk. Tartu, 2009, 168 p.
163. **Jaana Salujõe**. Zooplankton as the indicator of ecological quality and fish predation in lake ecosystems. Tartu, 2009, 129 p.
164. **Ele Vahtmäe**. Mapping benthic habitat with remote sensing in optically complex coastal environments. Tartu, 2009, 109 p.
165. **Liisa Metsamaa**. Model-based assessment to improve the use of remote sensing in recognition and quantitative mapping of cyanobacteria. Tartu, 2009, 114 p.

166. **Pille Säälik**. The role of endocytosis in the protein transduction by cell-penetrating peptides. Tartu, 2009, 155 p.
167. **Lauri Peil**. Ribosome assembly factors in *Escherichia coli*. Tartu, 2009, 147 p.
168. **Lea Hallik**. Generality and specificity in light harvesting, carbon gain capacity and shade tolerance among plant functional groups. Tartu, 2009, 99 p.
169. **Mariliis Tark**. Mutagenic potential of DNA damage repair and tolerance mechanisms under starvation stress. Tartu, 2009, 191 p.
170. **Riinu Rannap**. Impacts of habitat loss and restoration on amphibian populations. Tartu, 2009, 117 p.
171. **Maarja Adojaan**. Molecular variation of HIV-1 and the use of this knowledge in vaccine development. Tartu, 2009, 95 p.
172. **Signe Altmäe**. Genomics and transcriptomics of human induced ovarian folliculogenesis. Tartu, 2010, 179 p.
173. **Triin Suvi**. Mycorrhizal fungi of native and introduced trees in the Seychelles Islands. Tartu, 2010, 107 p.
174. **Velda Lauringson**. Role of suspension feeding in a brackish-water coastal sea. Tartu, 2010, 123 p.
175. **Eero Talts**. Photosynthetic cyclic electron transport – measurement and variably proton-coupled mechanism. Tartu, 2010, 121 p.
176. **Mari Nelis**. Genetic structure of the Estonian population and genetic distance from other populations of European descent. Tartu, 2010, 97 p.
177. **Kaarel Krjutškov**. Arrayed Primer Extension-2 as a multiplex PCR-based method for nucleic acid variation analysis: method and applications. Tartu, 2010, 129 p.
178. **Egle Köster**. Morphological and genetical variation within species complexes: *Anthyllis vulneraria* s. l. and *Alchemilla vulgaris* (coll.). Tartu, 2010, 101 p.
179. **Erki Õunap**. Systematic studies on the subfamily Sterrhinae (Lepidoptera: Geometridae). Tartu, 2010, 111 p.
180. **Merike Jõesaar**. Diversity of key catabolic genes at degradation of phenol and *p*-cresol in pseudomonads. Tartu, 2010, 125 p.
181. **Kristjan Herkül**. Effects of physical disturbance and habitat-modifying species on sediment properties and benthic communities in the northern Baltic Sea. Tartu, 2010, 123 p.
182. **Arto Pulk**. Studies on bacterial ribosomes by chemical modification approaches. Tartu, 2010, 161 p.
183. **Maria Põllupüü**. Ecological relations of cladocerans in a brackish-water ecosystem. Tartu, 2010, 126 p.
184. **Toomas Silla**. Study of the segregation mechanism of the Bovine Papillomavirus Type 1. Tartu, 2010, 188 p.
185. **Gyaneshwer Chaubey**. The demographic history of India: A perspective based on genetic evidence. Tartu, 2010, 184 p.
186. **Katrin Kepp**. Genes involved in cardiovascular traits: detection of genetic variation in Estonian and Czech populations. Tartu, 2010, 164 p.

187. **Virve Sõber**. The role of biotic interactions in plant reproductive performance. Tartu, 2010, 92 p.
188. **Kersti Kangro**. The response of phytoplankton community to the changes in nutrient loading. Tartu, 2010, 144 p.
189. **Joachim M. Gerhold**. Replication and Recombination of mitochondrial DNA in Yeast. Tartu, 2010, 120 p.
190. **Helen Tammert**. Ecological role of physiological and phylogenetic diversity in aquatic bacterial communities. Tartu, 2010, 140 p.
191. **Elle Rajandu**. Factors determining plant and lichen species diversity and composition in Estonian *Calamagrostis* and *Hepatica* site type forests. Tartu, 2010, 123 p.
192. **Paula Ann Kivistik**. ColR-ColS signalling system and transposition of Tn*4652* in the adaptation of *Pseudomonas putida.* Tartu, 2010, 118 p.
193. **Siim Sõber**. Blood pressure genetics: from candidate genes to genome-wide association studies. Tartu, 2011, 120 p.
194. **Kalle Kipper**. Studies on the role of helix 69 of 23S rRNA in the factor-dependent stages of translation initiation, elongation, and termination. Tartu, 2011, 178 p.
195. **Triinu Siibak**. Effect of antibiotics on ribosome assembly is indirect. Tartu, 2011, 134 p.
196. **Tambet Tõnissoo**. Identification and molecular analysis of the role of guanine nucleotide exchange factor RIC-8 in mouse development and neural function. Tartu, 2011, 110 p.
197. **Helin Räägel**. Multiple faces of cell-penetrating peptides – their intracellular trafficking, stability and endosomal escape during protein transduction. Tartu, 2011, 161 p.
198. **Andres Jaanus**. Phytoplankton in Estonian coastal waters – variability, trends and response to environmental pressures. Tartu, 2011, 157 p.
199. **Tiit Nikopensius**. Genetic predisposition to nonsyndromic orofacial clefts. Tartu, 2011, 152 p.
200. **Signe Värv**. Studies on the mechanisms of RNA polymerase II-dependent transcription elongation. Tartu, 2011, 108 p.
201. **Kristjan Välk**. Gene expression profiling and genome-wide association studies of non-small cell lung cancer. Tartu, 2011, 98 p.
202. **Arno Põllumäe**. Spatio-temporal patterns of native and invasive zooplankton species under changing climate and eutrophication conditions. Tartu, 2011, 153 p.
203. **Egle Tammeleht**. Brown bear (*Ursus arctos*) population structure, demographic processes and variations in diet in northern Eurasia. Tartu, 2011, 143 p.
205. **Teele Jairus**. Species composition and host preference among ectomycorrhizal fungi in Australian and African ecosystems. Tartu, 2011, 106 p.
206. **Kessy Abarenkov**. PlutoF – cloud database and computing services supporting biological research. Tartu, 2011, 125 p.

207. **Marina Grigorova**. Fine-scale genetic variation of follicle-stimulating hormone beta-subunit coding gene (*FSHB*) and its association with reproductive health. Tartu, 2011, 184 p.
208. **Anu Tiitsaar**. The effects of predation risk and habitat history on butterfly communities. Tartu, 2011, 97 p.
209. **Elin Sild**. Oxidative defences in immunoecological context: validation and application of assays for nitric oxide production and oxidative burst in a wild passerine. Tartu, 2011, 105 p.
210. **Irja Saar**. The taxonomy and phylogeny of the genera *Cystoderma* and *Cystodermella* (Agaricales, Fungi). Tartu, 2012, 167 p.
211. **Pauli Saag**. Natural variation in plumage bacterial assemblages in two wild breeding passerines. Tartu, 2012, 113 p.
212. **Aleksei Lulla**. Alphaviral nonstructural protease and its polyprotein substrate: arrangements for the perfect marriage. Tartu, 2012, 143 p.
213. **Mari Järve**. Different genetic perspectives on human history in Europe and the Caucasus: the stories told by uniparental and autosomal markers. Tartu, 2012, 119 p.
214. **Ott Scheler**. The application of tmRNA as a marker molecule in bacterial diagnostics using microarray and biosensor technology. Tartu, 2012, 93 p.
215. **Anna Balikova**. Studies on the functions of tumor-associated mucin-like leukosialin (CD43) in human cancer cells. Tartu, 2012, 129 p.
216. **Triinu Kõressaar**. Improvement of PCR primer design for detection of prokaryotic species. Tartu, 2012, 83 p.
217. **Tuul Sepp**. Hematological health state indices of greenfinches: sources of individual variation and responses to immune system manipulation. Tartu, 2012, 117 p.
218. **Rya Ero**. Modifier view of the bacterial ribosome. Tartu, 2012, 146 p.
219. **Mohammad Bahram**. Biogeography of ectomycorrhizal fungi across different spatial scales. Tartu, 2012, 165 p.
220. **Annely Lorents**. Overcoming the plasma membrane barrier: uptake of amphipathic cell-penetrating peptides induces influx of calcium ions and downstream responses. Tartu, 2012, 113 p.
221. **Katrin Männik**. Exploring the genomics of cognitive impairment: whole-genome SNP genotyping experience in Estonian patients and general population. Tartu, 2012, 171 p.
222. **Marko Prous**. Taxonomy and phylogeny of the sawfly genus *Empria* (Hymenoptera, Tenthredinidae). Tartu, 2012, 192 p.
223. **Triinu Visnapuu**. Levansucrases encoded in the genome of *Pseudomonas syringae* pv. tomato DC3000: heterologous expression, biochemical characterization, mutational analysis and spectrum of polymerization products. Tartu, 2012, 160 p.
224. **Nele Tamberg**. Studies on Semliki Forest virus replication and pathogenesis. Tartu, 2012, 109 p.

225. **Tõnu Esko**. Novel applications of SNP array data in the analysis of the genetic structure of Europeans and in genetic association studies. Tartu, 2012, 149 p.
226. **Timo Arula**. Ecology of early life-history stages of herring *Clupea harengus membras* in the northeastern Baltic Sea. Tartu, 2012, 143 p.
227. **Inga Hiiesalu**. Belowground plant diversity and coexistence patterns in grassland ecosystems. Tartu, 2012, 130 p.
228. **Kadri Koorem**. The influence of abiotic and biotic factors on small-scale plant community patterns and regeneration in boreonemoral forest. Tartu, 2012, 114 p.
229. **Liis Andresen**. Regulation of virulence in plant-pathogenic pectobacteria. Tartu, 2012, 122 p.
230. **Kaupo Kohv**. The direct and indirect effects of management on boreal forest structure and field layer vegetation. Tartu, 2012, 124 p.
231. **Mart Jüssi**. Living on an edge: landlocked seals in changing climate. Tartu, 2012, 114 p.
232. **Riina Klais**. Phytoplankton trends in the Baltic Sea. Tartu, 2012, 136 p.
233. **Rauno Veeroja**. Effects of winter weather, population density and timing of reproduction on life-history traits and population dynamics of moose (*Alces alces*) in Estonia. Tartu, 2012, 92 p.
234. **Marju Keis**. Brown bear (*Ursus arctos*) phylogeography in northern Eurasia. Tartu, 2013, 142 p.
235. **Sergei Põlme**. Biogeography and ecology of *alnus*- associated ectomycorrhizal fungi – from regional to global scale. Tartu, 2013, 90 p.
236. **Liis Uusküla**. Placental gene expression in normal and complicated pregnancy. Tartu, 2013, 173 p.
237. **Marko Lõoke**. Studies on DNA replication initiation in *Saccharomyces cerevisiae*. Tartu, 2013, 112 p.
238. **Anne Aan**. Light- and nitrogen-use and biomass allocation along productivity gradients in multilayer plant communities. Tartu, 2013, 127 p.
239. **Heidi Tamm**. Comprehending phylogenetic diversity – case studies in three groups of ascomycetes. Tartu, 2013, 136 p.
240. **Liina Kangur**. High-Pressure Spectroscopy Study of Chromophore-Binding Hydrogen Bonds in Light-Harvesting Complexes of Photosynthetic Bacteria. Tartu, 2013, 150 p.
241. **Margus Leppik**. Substrate specificity of the multisite specific pseudouridine synthase RluD. Tartu, 2013, 111 p.
242. **Lauris Kaplinski**. The application of oligonucleotide hybridization model for PCR and microarray optimization. Tartu, 2013, 103 p.
243. **Merli Pärnoja**. Patterns of macrophyte distribution and productivity in coastal ecosystems: effect of abiotic and biotic forcing. Tartu, 2013, 155 p.
244. **Tõnu Margus**. Distribution and phylogeny of the bacterial translational GTPases and the Mqsr/YgiT regulatory system. Tartu, 2013, 126 p.
245. **Pille Mänd**. Light use capacity and carbon and nitrogen budget of plants: remote assessment and physiological determinants. Tartu, 2013, 128 p.

246. **Mario Plaas**. Animal model of Wolfram Syndrome in mice: behavioural, biochemical and psychopharmacological characterization. Tartu, 2013, 144 p.
247. **Georgi Hudjašov**. Maps of mitochondrial DNA, Y-chromosome and tyrosinase variation in Eurasian and Oceanian populations. Tartu, 2013, 115 p.
248. **Mari Lepik**. Plasticity to light in herbaceous plants and its importance for community structure and diversity. Tartu, 2013, 102 p.
249. **Ede Leppik**. Diversity of lichens in semi-natural habitats of Estonia. Tartu, 2013, 151 p.
250. **Ülle Saks**. Arbuscular mycorrhizal fungal diversity patterns in boreo-nemoral forest ecosystems. Tartu, 2013, 151 p.
251. **Eneli Oitmaa**. Development of arrayed primer extension microarray assays for molecular diagnostic applications. Tartu, 2013, 147 p.
252. **Jekaterina Jutkina**. The horizontal gene pool for aromatics degradation: bacterial catabolic plasmids of the Baltic Sea aquatic system. Tartu, 2013, 121 p.
253. **Helen Vellau**. Reaction norms for size and age at maturity in insects: rules and exceptions. Tartu, 2014, 132 p.
254. **Randel Kreitsberg**. Using biomarkers in assessment of environmental contamination in fish – new perspectives. Tartu, 2014, 107 p.
255. **Krista Takkis**. Changes in plant species richness and population performance in response to habitat loss and fragmentation.Tartu, 2014, 141 p.
256. **Liina Nagirnaja**. Global and fine-scale genetic determinants of recurrent pregnancy loss. Tartu, 2014, 211 p.
257. **Triin Triisberg**. Factors influencing the re-vegetation of abandoned extracted peatlands in Estonia. Tartu, 2014, 133 p.
258. **Villu Soon**. A phylogenetic revision of the *Chrysis ignita* species group (Hymenoptera: Chrysididae) with emphasis on the northern European fauna. Tartu, 2014, 211 p.
259. **Andrei Nikonov**. RNA-Dependent RNA Polymerase Activity as a Basis for the Detection of Positive-Strand RNA Viruses by Vertebrate Host Cells. Tartu, 2014, 207 p.
260. **Eele Õunapuu-Pikas**. Spatio-temporal variability of leaf hydraulic conductance in woody plants: ecophysiological consequences. Tartu, 2014, 135 p.
261. **Marju Männiste**. Physiological ecology of greenfinches: information content of feathers in relation to immune function and behavior. Tartu, 2014, 121 p.
262. **Katre Kets**. Effects of elevated concentrations of CO_2 and O_3 on leaf photosynthetic parameters in *Populus tremuloides*: diurnal, seasonal and inter-annual patterns. Tartu, 2014, 115 p.
263. **Külli Lokko**. Seasonal and spatial variability of zoopsammon communities in relation to environmental parameters. Tartu, 2014, 129 p.
264. **Olga Žilina**. Chromosomal microarray analysis as diagnostic tool: Estonian experience. Tartu, 2014, 152 p.

265. **Kertu Lõhmus**. Colonisation ecology of forest-dwelling vascular plants and the conservation value of rural manor parks. Tartu, 2014, 111 p.
266. **Anu Aun**. Mitochondria as integral modulators of cellular signaling. Tartu, 2014, 167 p.
267. **Chandana Basu Mallick**. Genetics of adaptive traits and gender-specific demographic processes in South Asian populations. Tartu, 2014, 160 p.
268. **Riin Tamme**. The relationship between small-scale environmental heterogeneity and plant species diversity. Tartu, 2014, 130 p.
269. **Liina Remm**. Impacts of forest drainage on biodiversity and habitat quality: implications for sustainable management and conservation. Tartu, 2015, 126 p.
270. **Tiina Talve**. Genetic diversity and taxonomy within the genus *Rhinanthus*. Tartu, 2015, 106 p.
271. **Mehis Rohtla**. Otolith sclerochronological studies on migrations, spawning habitat preferences and age of freshwater fishes inhabiting the Baltic Sea. Tartu, 2015, 137 p.
272. **Alexey Reshchikov**. The world fauna of the genus *Lathrolestes* (Hymenoptera, Ichneumonidae). Tartu, 2015, 247 p.
273. **Martin Pook**. Studies on artificial and extracellular matrix protein-rich surfaces as regulators of cell growth and differentiation. Tartu, 2015, 142 p.
274. **Mai Kukumägi**. Factors affecting soil respiration and its components in silver birch and Norway spruce stands. Tartu, 2015, 155 p.
275. **Helen Karu**. Development of ecosystems under human activity in the North-East Estonian industrial region: forests on post-mining sites and bogs. Tartu, 2015, 152 p.
276. **Hedi Peterson**. Exploiting high-throughput data for establishing relationships between genes. Tartu, 2015, 186 p.
277. **Priit Adler**. Analysis and visualisation of large scale microarray data, Tartu, 2015, 126 p.
278. **Aigar Niglas**. Effects of environmental factors on gas exchange in deciduous trees: focus on photosynthetic water-use efficiency. Tartu, 2015, 152 p.
279. **Silja Laht**. Classification and identification of conopeptides using profile hidden Markov models and position-specific scoring matrices. Tartu, 2015, 100 p.
280. **Martin Kesler**. Biological characteristics and restoration of Atlantic salmon *Salmo salar* populations in the Rivers of Northern Estonia. Tartu, 2015, 97 p.
281. **Pratyush Kumar Das**. Biochemical perspective on alphaviral nonstructural protein 2: a tale from multiple domains to enzymatic profiling. Tartu, 2015, 205 p
282. **Priit Palta**. Computational methods for DNA copy number detection. Tartu, 2015, 130 p.
283. **Julia Sidorenko**. Combating DNA damage and maintenance of genome integrity in pseudomonads. Tartu, 2015, 174 p.
284. **Anastasiia Kovtun-Kante**. Charophytes of Estonian inland and coastal waters: distribution and environmental preferences. Tartu, 2015, 97 p.

285. **Ly Lindman**. The ecology of protected butterfly species in Estonia. Tartu, 2015, 171 p.
286. **Jaanis Lodjak**. Association of Insulin-like Growth Factor I and Corticosterone with Nestling Growth and Fledging Success in Wild Passerines. Tartu, 2016, 113 p.
287. **Ann Kraut**. Conservation of Wood-Inhabiting Biodiversity – Semi-Natural Forests as an Opportunity. Tartu, 2016, 141 p.
288. **Tiit Örd**. Functions and regulation of the mammalian pseudokinase TRIB3. Tartu, 2016, 182. p.
289. **Kairi Käiro**. Biological Quality According to Macroinvertebrates in Streams of Estonia (Baltic Ecoregion of Europe): Effects of Human-induced Hydromorphological Changes. Tartu, 2016, 126 p.
290. **Leidi Laurimaa**. *Echinococcus multilocularis* and other zoonotic parasites in Estonian canids. Tartu, 2016, 144 p.
291. **Helerin Margus**. Characterization of cell-penetrating peptide/nucleic acid nanocomplexes and their cell-entry mechanisms. Tartu, 2016, 173 p.
292. **Kadri Runnel**. Fungal targets and tools for forest conservation. Tartu, 2016, 157 p.
293. **Urmo Võsa**. MicroRNAs in disease and health: aberrant regulation in lung cancer and association with genomic variation. Tartu, 2016, 163 p.
294. **Kristina Mäemets-Allas**. Studies on cell growth promoting AKT signaling pathway – a promising anti-cancer drug target. Tartu, 2016, 146 p.
295. **Janeli Viil**. Studies on cellular and molecular mechanisms that drive normal and regenerative processes in the liver and pathological processes in Dupuytren's contracture. Tartu, 2016, 175 p.
296. **Ene Kook**. Genetic diversity and evolution of *Pulmonaria angustifolia* L. and *Myosotis laxa sensu lato* (Boraginaceae). Tartu, 2016, 106 p.
297. **Kadri Peil**. RNA polymerase II-dependent transcription elongation in *Saccharomyces cerevisiae*. Tartu, 2016, 113 p.
298. **Katrin Ruisu**. The role of RIC8A in mouse development and its function in cell-matrix adhesion and actin cytoskeletal organisation. Tartu, 2016, 129 p.
299. **Janely Pae**. Translocation of cell-penetrating peptides across biological membranes and interactions with plasma membrane constituents. Tartu, 2016, 126 p.
300. **Argo Ronk**. Plant diversity patterns across Europe: observed and dark diversity. Tartu, 2016, 153 p.
301. **Kristiina Mark**. Diversification and species delimitation of lichenized fungi in selected groups of the family Parmeliaceae (Ascomycota). Tartu, 2016, 181 p.
302. **Jaak-Albert Metsoja**. Vegetation dynamics in floodplain meadows: influence of mowing and sediment application. Tartu, 2016, 140 p.
303. **Hedvig Tamman**. The GraTA toxin-antitoxin system of *Pseudomonas putida*: regulation and role in stress tolerance. Tartu, 2016, 154 p.

304. **Kadri Pärtel**. Application of ultrastructural and molecular data in the taxonomy of helotialean fungi. Tartu, 2016, 183 p.
305. **Maris Hindrikson**. Grey wolf (*Canis lupus*) populations in Estonia and Europe: genetic diversity, population structure and -processes, and hybridization between wolves and dogs. Tartu, 2016, 121 p.
306. **Polina Degtjarenko**. Impacts of alkaline dust pollution on biodiversity of plants and lichens: from communities to genetic diversity. Tartu, 2016, 126 p.
307. **Liina Pajusalu**. The effect of CO_2 enrichment on net photosynthesis of macrophytes in a brackish water environment. Tartu, 2016, 126 p.
308. **Stoyan Tankov**. Random walks in the stringent response. Tartu, 2016, 94 p.
309. **Liis Leitsalu**. Communicating genomic research results to population-based biobank participants. Tartu, 2016, 158 p.
310. **Richard Meitern**. Redox physiology of wild birds: validation and application of techniques for detecting oxidative stress. Tartu, 2016, 134 p.
311. **Kaie Lokk**. Comparative genome-wide DNA methylation studies of healthy human tissues and non-small cell lung cancer tissue. Tartu, 2016, 127 p.
312. **Mihhail Kurašin**. Processivity of cellulases and chitinases. Tartu, 2017, 132 p.
313. **Carmen Tali**. Scavenger receptors as a target for nucleic acid delivery with peptide vectors. Tartu, 2017, 155 p.
314. **Katarina Oganjan**. Distribution, feeding and habitat of benthic suspension feeders in a shallow coastal sea. Tartu, 2017, 132 p.
315. **Taavi Paal**. Immigration limitation of forest plants into wooded landscape corridors. Tartu, 2017, 145 p.
316. **Kadri Õunap**. The Williams-Beuren syndrome chromosome region protein WBSCR22 is a ribosome biogenesis factor. Tartu, 2017, 135 p.
317. **Riin Tamm**. In-depth analysis of factors affecting variability in thiopurine methyltransferase activity. Tartu, 2017, 170 p.
318. **Keiu Kask**. The role of RIC8A in the development and regulation of mouse nervous system. Tartu, 2017, 184 p.
319. **Tiia Möller**. Mapping and modelling of the spatial distribution of benthic macrovegetation in the NE Baltic Sea with a special focus on the eelgrass *Zostera marina* Linnaeus, 1753. Tartu, 2017, 162 p.
320. **Silva Kasela**. Genetic regulation of gene expression: detection of tissue- and cell type-specific effects. Tartu, 2017, 150 p.
321. **Karmen Süld**. Food habits, parasites and space use of the raccoon dog *Nyctereutes procyonoides*: the role of an alien species as a predator and vector of zoonotic diseases in Estonia. Tartu, 2017, p.
322. **Ragne Oja**. Consequences of supplementary feeding of wild boar – concern for ground-nesting birds and endoparasite infection. Tartu, 2017, 141 p.
323. **Riin Kont**. The acquisition of cellulose chain by a processive cellobiohydrolase. Tartu, 2017, 117 p.
324. **Liis Kasari**. Plant diversity of semi-natural grasslands: drivers, current status and conservation challenges. Tartu, 2017, 141 p.

325. **Sirgi Saar**. Belowground interactions: the roles of plant genetic relatedness, root exudation and soil legacies. Tartu, 2017, 113 p.
326. **Sten Anslan**. Molecular identification of Collembola and their fungal associates. Tartu, 2017, 125 p.
327. **Imre Taal**. Causes of variation in littoral fish communities of the Eastern Baltic Sea: from community structure to individual life histories. Tartu, 2017, 118 p.
328. **Jürgen Jalak**. Dissecting the Mechanism of Enzymatic Degradation of Cellulose Using Low Molecular Weight Model Substrates. Tartu, 2017, 137 p.
329. **Kairi Kiik**. Reproduction and behaviour of the endangered European mink (*Mustela lutreola*) in captivity. Tartu, 2018, 112 p.
330. **Ivan Kuprijanov**. Habitat use and trophic interactions of native and invasive predatory macroinvertebrates in the northern Baltic Sea. Tartu, 2018, 117 p.
331. **Hendrik Meister**. Evolutionary ecology of insect growth: from geographic patterns to biochemical trade-offs. Tartu, 2018, 147 p.
332. **Ilja Gaidutšik**. Irc3 is a mitochondrial branch migration enzyme in *Saccharomyces cerevisiae*. Tartu, 2018, 161 p.
333. **Lena Neuenkamp**. The dynamics of plant and arbuscular mycorrhizal fungal communities in grasslands under changing land use. Tartu, 2018, 241 p.
334. **Laura Kasak**. Genome structural variation modulating the placenta and pregnancy maintenance. Tartu, 2018, 181 p.
335. **Kersti Riibak**. Importance of dispersal limitation in determining dark diversity of plants across spatial scales. Tartu, 2018, 133 p.
336. **Liina Saar**. Dynamics of grassland plant diversity in changing landscapes. Tartu, 2018, 206 p.
337. **Hanna Ainelo**. Fis regulates *Pseudomonas putida* biofilm formation by controlling the expression of *lapA*. Tartu, 2018, 143 p.
338. **Natalia Pervjakova**. Genomic imprinting in complex traits. Tartu, 2018, 176 p.
339. **Andrio Lahesaare**. The role of global regulator Fis in regulating the expression of *lapF* and the hydrophobicity of soil bacterium *Pseudomonas putida*. Tartu, 2018, 124 p.
340. **Märt Roosaare**. *K*-mer based methods for the identification of bacteria and plasmids. Tartu, 2018, 117 p.
341. **Maria Abakumova**. The relationship between competitive behaviour and the frequency and identity of neighbours in temperate grassland plants. Tartu, 2018, 104 p.
342. **Margus Vilbas**. Biotic interactions affecting habitat use of myrmecophilous butterflies in Northern Europe. Tartu, 2018, 142 p.
343. **Liina Kinkar**. Global patterns of genetic diversity and phylogeography of *Echinococcus granulosus* sensu stricto – a tapeworm species of significant public health concern. Tartu, 2018, 147 p.

344. **Teivi Laurimäe**. Taxonomy and genetic diversity of zoonotic tapeworms in the species complex of *Echinococcus granulosus* sensu lato. Tartu, 2018, 143 p.
345. **Tatjana Jatsenko**. Role of translesion DNA polymerases in mutagenesis and DNA damage tolerance in Pseudomonads. Tartu, 2018, 216 p.
346. **Katrin Viigand**. Utilization of α-glucosidic sugars by *Ogataea* (*Hansenula*) *polymorpha*. Tartu, 2018, 148 p.
347. **Andres Ainelo**. Physiological effects of the *Pseudomonas putida* toxin grat. Tartu, 2018, 146 p.
348. **Killu Timm**. Effects of two genes (DRD4 and SERT) on great tit (*Parus major*) behaviour and reproductive traits. Tartu, 2018, 117 p.
349. **Petr Kohout**. Ecology of ericoid mycorrhizal fungi. Tartu, 2018, 184 p.
350. **Gristin Rohula-Okunev**. Effects of endogenous and environmental factors on night-time water flux in deciduous woody tree species. Tartu, 2018, 184 p.
351. **Jane Oja**. Temporal and spatial patterns of orchid mycorrhizal fungi in forest and grassland ecosystems. Tartu, 2018, 102 p.
352. **Janek Urvik**. Multidimensionality of aging in a long-lived seabird. Tartu, 2018, 135 p.
353. **Lisanna Schmidt**. Phenotypic and genetic differentiation in the hybridizing species pair *Carex flava* and *C. viridula* in geographically different regions. Tartu, 2018, 133 p.
354. **Monika Karmin**. Perspectives from human Y chromosome – phylogeny, population dynamics and founder events. Tartu, 2018, 168 p.
355. **Maris Alver**. Value of genomics for atherosclerotic cardiovascular disease risk prediction. Tartu, 2019, 148 p.
356. **Lehti Saag**. The prehistory of Estonia from a genetic perspective: new insights from ancient DNA. Tartu, 2019, 171 p.
357. **Mari-Liis Viljur**. Local and landscape effects on butterfly assemblages in managed forests. Tartu, 2019, 115 p.
358. **Ivan Kisly**. The pleiotropic functions of ribosomal proteins eL19 and eL24 in the budding yeast ribosome. Tartu, 2019, 170 p.
359. **Mikk Puustusmaa**. On the origin of papillomavirus proteins. Tartu, 2019, 152 p.
360. **Anneliis Peterson**. Benthic biodiversity in the north-eastern Baltic Sea: mapping methods, spatial patterns, and relations to environmental gradients. Tartu, 2019, 159 p.
361. **Erwan Pennarun**. Meandering along the mtDNA phylogeny; causerie and digression about what it can tell us about human migrations. Tartu, 2019, 162 p.
362. **Karin Ernits**. Levansucrase Lsc3 and endo-levanase BT1760: characterization and application for the synthesis of novel prebiotics. Tartu, 2019, 217 p.
363. **Sille Holm**. Comparative ecology of geometrid moths: in search of contrasts between a temperate and a tropical forest. Tartu, 2019, 135 p.

364. **Anne-Mai Ilumäe**. Genetic history of the Uralic-speaking peoples as seen through the paternal haplogroup N and autosomal variation of northern Eurasians. Tartu, 2019, 172 p.
365. **Anu Lepik**. Plant competitive behaviour: relationships with functional traits and soil processes. Tartu, 2019, 152 p.
366. **Kunter Tätte**. Towards an integrated view of escape decisions in birds under variable levels of predation risk. Tartu, 2020, 172 p.
367. **Kaarin Parts**. The impact of climate change on fine roots and root-associated microbial communities in birch and spruce forests. Tartu, 2020, 143 p.
368. **Viktorija Kukuškina**. Understanding the mechanisms of endometrial receptivity through integration of 'omics' data layers. Tartu, 2020, 169 p.
369. **Martti Vasar**. Developing a bioinformatics pipeline gDAT to analyse arbuscular mycorrhizal fungal communities using sequence data from different marker regions. Tartu, 2020, 193 p.
370. **Ott Kangur**. Nocturnal water relations and predawn water potential disequilibrium in temperate deciduous tree species. Tartu, 2020, 126 p.
371. **Helen Post**. Overview of the phylogeny and phylogeography of the Y-chromosomal haplogroup N in northern Eurasia and case studies of two linguistically exceptional populations of Europe – Hungarians and Kalmyks. Tartu, 2020, 143 p.
372. **Kristi Krebs**. Exploring the genetics of adverse events in pharmacotherapy using Biobanks and Electronic Health Records. Tartu, 2020, 151 p.
373. **Kärt Ukkivi**. Mutagenic effect of transcription and transcription-coupled repair factors in *Pseudomonas putida.* Tartu, 2020, 154 p.
374. **Elin Soomets**. Focal species in wetland restoration. Tartu, 2020, 137 p.
375. **Kadi Tilk**. Signals and responses of ColRS two-component system in *Pseudomonas putida.* Tartu, 2020, 133 p.
376. **Indrek Teino**. Studies on aryl hydrocarbon receptor in the mouse granulosa cell model. Tartu, 2020, 139 p.
377. **Maarja Vaikre**. The impact of forest drainage on macroinvertebrates and amphibians in small waterbodies and opportunities for cost-effective mitigation. Tartu, 2020, 132 p.
378. **Siim-Kaarel Sepp**. Soil eukaryotic community responses to land use and host identity. Tartu, 2020, 222 p.
379. **Eveli Otsing**. Tree species effects on fungal richness and community structure. Tartu, 2020, 152 p.
380. **Mari Pent**. Bacterial communities associated with fungal fruitbodies. Tartu, 2020, 144 p.
381. **Einar Kärgenberg**. Movement patterns of lithophilous migratory fish in free-flowing and fragmented rivers. Tartu, 2020, 167 p.
382. **Antti Matvere**. The studies on aryl hydrocarbon receptor in murine granulosa cells and human embryonic stem cells. Tartu, 2021, 163 p.
383. **Jhonny Capichoni Massante**. Phylogenetic structure of plant communities along environmental gradients: a macroecological and evolutionary approach. Tartu, 2021, 144 p.

384. **Ajai Kumar Pathak.** Delineating genetic ancestries of people of the Indus Valley, Parsis, Indian Jews and Tharu tribe. Tartu, 2021, 197 p.
385. **Tanel Vahter.** Arbuscular mycorrhizal fungal biodiversity for sustainable agroecosystems. Tartu, 2021, 191 p.
386. **Burak Yelmen.** Characterization of ancient Eurasian influences within modern human genomes. Tartu, 2021, 134 p.
387. **Linda Ongaro.** A genomic portrait of American populations. Tartu, 2021, 182 p.
388. **Kairi Raime.** The identification of plant DNA in metagenomic samples. Tartu, 2021, 108 p.
389. **Heli Einberg.** Non-linear and non-stationary relationships in the pelagic ecosystem of the Gulf of Riga (Baltic Sea). Tartu, 2021, 119 p.
390. **Mickaël Mathieu Pihain.** The evolutionary effect of phylogenetic neighbourhoods of trees on their resistance to herbivores and climatic stress. Tartu, 2022, 145 p.
391. **Annika Joy Meitern.** Impact of potassium ion content of xylem sap and of light conditions on the hydraulic properties of trees. Tartu, 2022, 132 p.
392. **Elise Joonas.** Evaluation of metal contaminant hazard on microalgae with environmentally relevant testing strategies. Tartu, 2022, 118 p.