



Norwegian University
of Life Sciences

Master's Thesis 2018 60 ECTS

Faculty of Chemistry, Biotechnology and Food Science (KBM)
Main Supervisor at NMBU Knut Rudi

Evaluation of Liquid Array Diagnostics for microbiota-based diagnosis of dysbiosis

Katrine Amlie

MSc Biotechnology, Microbiology

Preface

This master thesis was written at the Faculty of Chemistry, Biotechnology and Food Science at the Norwegian University of Life Sciences (NMBU) in Ås between January 2018 and November 2018, with Professor Knut Rudi and Professor Robert C. Wilson as supervisors. Most of the laboratory work was conducted at the Inland Norway University of Applied Sciences in Hamar (INN), while the remainder was done at NMBU. Transportation costs to Hamar were covered by NMBU.

I would like to express my gratitude towards my two wonderful supervisors Knut Rudi and Robert C. Wilson for their guidance throughout this process, for being so encouraging and positive, and for giving me the opportunity to work on such an exciting project. An extra thanks to Knut Rudi for helping me with the statistical analysis.

A very big thank you goes to Pranvera Hiseni for all her help, support and patience both in the lab and outside. It has been a true pleasure to work with you. Thanks also to the MiDiv group and the other master students for their assistance, and for making me feel like a part of the team even though most of my days were spent in Hamar. In addition, I would like to thank Janne Beate Utåker for all the excellent guidance I received during my studies.

I would like to thank Finn Terje Hegge at Genetic Analysis AS for trusting me to work with their probes, for providing me with the necessary bacterial templates and fecal samples, and for being so helpful.

I want to thank my former boss Tor Myhre for enabling and encouraging me to pick up my studies again, and my bosses Hans Petter Loose and Birgitte Øgrey for being so supportive and accommodating during this time. Also, my wonderful colleagues for being so kind and understanding. I am very grateful.

Finally, I would like to thank my dear family and friends for being so supportive throughout all of this, for believing in me when I couldn't believe in myself, and for reminding me that it's ok to take a break. I could not have done this without you!

Ås, November 2018

Katrine Amlie

Abbreviations

AUC	Area under curve
BLAST	Basic Local Alignment Search Tool
C	Cytosine
CD	Crohn's disease
ddCTP	Dideoxycytidine triphosphate
ddCTP-Q	Quencher-labelled dideoxycytidine triphosphate
ddNTP	Dideoxynucleotide triphosphate
ddNTP-Q	Quencher-labelled dideoxynucleotide triphosphate
ddTTP	Dideoxythymidine triphosphate
DI	Dysbiosis index
DNA	Deoxyribonucleic acid
FMT	Fecal microbiota transplantation
FODMAP	Fermentable oligosaccharides, disaccharides, monosaccharides and polyols
FRET	Fluorescence resonance energy transfer
G	Guanine
GA	Genetic Analysis
GA-map	Genetic Analysis microbiota array platform
gDNA	Genomic DNA
GI	Gastrointestinal
HRM	High Resolution Melting
HTS	High-throughput sequencing
IBD	Irritable bowel disease
IBS	Irritable bowel syndrome
LAD	Liquid Array Diagnostics

LP	Labelling probe
Maldi-TOF MS	Matrix-assisted laser desorption ionization time-of-flight mass spectrometry
MPA	Multiplex probe amplification
MS	Multiple sclerosis
NCBI	National Center for Biotechnology Information
NEC	Necrotizing enterocolitis
OLA	Oligonucleotide ligation assay
PCR	Polymerase chain reaction
PD	Parkinson's disease
QEXT	Quencher extension assay
qPCR	Quantitative polymerase chain reaction
ROC	Receiver operating characteristic
RP	Reporter probe
SCFA	Short-chain fatty acid
SDS	Sodium dodecyl sulfate
SNP	Single nucleotide polymorphism
SNuPE	Single nucleotide primer extension
T	Thymine
T _m	Melting temperature
UC	Ulcerative colitis

Sammendrag

Menneskets tarm inneholder et komplekst samfunn av mikroorganismer som har en direkte innvirkning på vår helse. Basert på observerte mikrobielle mønster kan en balansert sammensetning defineres, kjent som normobiose. Avvik fra en slik balansert tilstand kalles dysbiose og har blitt knyttet til flere sykdommer. En test som har blitt utviklet for påvisning av dysbiose er GA-map® Dysbiosis Test. Denne testen er svært spesifikk og effektiv, men også relativt arbeidskrevende og dyr, noe som har ført til ønsket om å supplementere fast-fase plattformen som brukes i dag med en flytende en. Hensikten med denne oppgaven var å undersøke muligheten for et slikt supplement, ved å teste 20 av innmerkingsprobene som blir brukt i dagens GA-map® Dysbiosis Test med en ny teknologi kalt Liquid Array Diagnostics (LAD). Denne teknologien er basert på en kombinasjon av Single Nucleotide Primer Extension og High Resolution Melting. Hele prosessen finner sted i løsning og inkluderer ikke et rensing- og separeringssteg, noe som gjør teknologien mer effektiv, samt mindre utsatt for forurensning, enn mange andre genotypingsmetoder på markedet i dag.

Det ble først laget komplementære deteksjonsprober til de 20 innmerkingsprobene, slik at hvert probepar hadde et spesifikt smeltepunkt. For å undersøke probenes kompatibilitet med LAD ble de testet på renkulturer av målbakterier. Deretter ble LAD utført på en mer kompleks prøve som inneholdt 11 forskjellige bakteriekulturer for å sikre at tydelige signaler fortsatt kunne oppnås. Til slutt ble LAD testet på avføringsprøver som tidligere hadde blitt diagnostisert av Genetic Analysis AS, samt at statistiske analyser ble utført for å undersøke LADs diagnostiserende evne.

Resultatene i denne oppgaven viste at LAD kan identifisere minst 15 mål samtidig, og det er forventet at videre optimalisering vil øke kapasiteten ytterligere. I tillegg ble det demonstrert at LAD kan skille mellom normobiose og dysbiose, og de fleste probene viste en sterk assosiasjon med en av de to tilstandene. Denne oppgaven har dermed vist at LAD kan være et lovende supplement til fast-fase plattformen som benyttes i dagens GA-map® Dysbiosis Test, selv om videre optimalisering er nødvendig. LAD krever ingen bakteriedyrking, er økonomisk, enkel og rask, og er kun avhengig av standardutstyr som er å finne i de fleste mikrobiologiske laboratorier. Teknologien har potensiale til å bli et svært verdifullt verktøy for diagnostisering, og kan i tillegg gi viktige bidrag til områder som personlig tilpasset behandling, ernæring, og presisjonsmedisin.

Abstract

The human gut contains a complex community of microorganisms that impact our health and general well-being. Based on observed microbial patterns, a balanced composition can be defined, known as normobiosis. Deviations from such a state is termed dysbiosis and have been linked to several medical conditions. A current assay designed for detection of dysbiosis is the GA-map® Dysbiosis Test. This test is quite specific and efficient, but also relatively labor-intensive and expensive, which has led to the desire to supplement the solid-phase platform that is currently being used with a liquid one. This thesis aimed to investigate the possibility of such a supplement by testing 20 of the labelling probes (LPs) used in the current GA-map® Dysbiosis Test with a novel technology called Liquid Array Diagnostics (LAD). This technology is based on a combination of Single Nucleotide Primer Extension and High Resolution Melting. The entire process takes place in solution, and does not include a purification and separation step, making the technology more efficient and less prone to contamination than many other genotyping methods on the market today.

Complementary reporter probes were designed for the 20 LPs, so that each probe duplex had a specific melting temperature. To investigate the compatibility of the probes with LAD they were tested on pure cultures of target bacteria. LAD was then performed on a more complex sample containing 11 different bacterial templates to ensure distinguishable probe signals could still be achieved. Finally, LAD was performed on fecal samples that had previously been diagnosed by Genetic Analysis AS, and statistical analyses were conducted to investigate the diagnostic ability of the technology.

The results showed that LAD can detect at least 15 targets simultaneously, and it is expected that further optimization will increase this number. In addition, it was demonstrated that LAD can distinguish between normobiosis and dysbiosis, and that most of the probes showed a strong association with one of the two states. This thesis has thus shown that LAD is a promising supplement to the solid-phase platform used in the GA-map® Dysbiosis Test today, although further optimization is needed before the technology can be implemented. LAD requires no bacterial growth, is economical, easy and rapid, and relies only on standard equipment found in most microbiology laboratories. The technology thus has the potential of becoming a highly valuable tool for diagnostics, and can provide important contributions to areas such as personalized treatments, nutrition and precision medicine.

Table of Contents

1. Introduction.....	1
1.1 The human body as an ecosystem	1
1.1.1 Development of the human gut microbiota	1
1.1.2 The adult gut microbiota	2
1.2 Normobiosis and dysbiosis	4
1.3 Investigating the composition of the gut microbiota	6
1.3.1 Hybridization-based methods for SNP genotyping.....	6
1.3.2 Single Nucleotide Primer Extension (SNUPE) technologies	8
1.3.3 Assay format	11
1.3.4 Liquid Array Diagnostics (LAD)	12
1.4 The GA-map® Dysbiosis Test	15
1.5 Aim of this thesis.....	17
2. Materials and methods	18
2.1 Design of reporter probes.....	19
2.2 Probe validation on pure cultures of target bacteria	20
2.2.1 Template preparation	20
2.2.2 Labelling of the LPs	21
2.2.3 LP-RP hybridization, signal detection and melting curve analysis	22
2.3 Validation of LAD on a mixed sample	23
2.4 Evaluation of LAD for diagnosis of dysbiosis.....	23
2.4.1 Performing LAD on fecal samples.....	23
2.4.2 Statistical analysis	24
3. Results	26
3.1 Probe compatibility with LAD.....	26
3.1.1 Probe validation on pure cultures of target bacteria	26
3.1.2 Investigation of probes that did not provide satisfactory signals.....	29
3.2 Validation of LAD on a mixed sample	30
3.3 Evaluation of LAD for diagnosis of dysbiosis.....	32
3.3.1 Probe signals in normobiotic versus dysbiotic samples	32
3.3.2 ROC curve analysis	34
3.3.3 Association of probes to dysbiosis or normobiosis	36
4. Discussion	37
4.1 The diagnostic performance of LAD	37
4.1.1 Probes associated with dysbiosis	38
4.1.2 Probes associated with normobiosis.....	42

4.1.3	Probes that did not provide satisfactory signals	49
4.2	Future work on LAD for use with the GA-map® Dysbiosis Test	51
4.2.1	Optimizing and improving the use of LAD as a dysbiosis test	51
4.2.2	Expanding the LAD dysbiosis test	54
4.3	Other potential uses for LAD	56
4.3.1	Tailor-made therapies	56
4.3.2	Screening of genes in industry and medicine	58
4.4	Conclusion	59
	Reference List	60
	Supplementary material	69

List of figures

Figure 1: The principle of the QEXT assay	10
Figure 2: Overview of the Liquid Array Diagnostics process	13
Figure 3: The process of the GA-map® Dysbiosis Test	16
Figure 4: Flow chart for the project.	18
Figure 5: Examples of melting curves from samples containing pure bacterial cultures.....	28
Figure 6: Examples of melting curves from a mixed sample containing 11 bacterial species	31
Figure 7: Comparison of average signals for each probe from normobiotic and dysbiotic samples	33
Figure 8: ROC plot showing the classification accuracy of LAD.....	35
Figure 9: Association between individual probes and normobiosis or dysbiosis.....	36
Figure S1: Melting curves with signals from all probes compatible with LAD on the HEX channel.....	69
Figure S2: Melting curves with signals from all probes compatible with LAD on the ROX channel.....	70
Figure S3: Melting curves with signals from all probes compatible with LAD on the CY5 and FAM channels.....	71

List of tables

Table 1: List of probes with bacterial targets, fluorophores and theoretical Tm	19
Table 2: Summary of results for all probes after testing on pure cultures and mixed sample	26
Table S1: Raw data of the probe peak values from the 15 normobiotic fecal samples.....	72
Table S2: Raw data of the probe peak values from the 17 dysbiotic fecal samples.....	72
Table S3: The p-values for each probe derived from a Welch Two Sample t-test.....	72

1. Introduction

1.1 The human body as an ecosystem

The human body is a highly complex system, with many elements affecting its health and general well-being. Mankind has always strived to understand the causes of diseases and conditions, and in recent years more attention has been directed at the importance of the gastrointestinal (GI) tract, and the organisms that reside there.

The type and abundance of microorganisms found in the various parts of the GI tract varies greatly based on the microbes' physiological needs and resistance to unfavorable elements. The mouth is nutrient-rich and contains a large number of microbes. This number rapidly declines when reaching the stomach, where the acidic gastric fluid makes for a very unfriendly environment for most microorganisms. The first part of the small intestine contains a limited microbiota due to stomach acid, bile and pancreatic secretions, while the last part contains a microbiota that is more similar to that of the large intestine (1).

The large intestine contains a substantial community of microbes with a great variety of bacterial genera (1). Many of the microorganisms found there are beneficial to us, as they perform such tasks as extracting energy and nutrients from food and drink, regulating the storage of fats, protecting against pathogens, performing detoxification, and participating in training, development and maintenance of our immune system (1, 2, 3, 4, 5, 6, 7).

1.1.1 Development of the human gut microbiota

Up until quite recently, it was believed that the development of the gut microbiota of an individual didn't start until birth. However, studies have found that colonization starts even before a baby is born (8, 9, 10, 11), as microbes have been found in the placenta (9, 10), umbilical cord (11), amniotic fluid (12, 13) and the meconium of healthy newborns (14, 15, 16). This may suggest that these first colonizers play a role in human development and preparation for a life outside of the womb (14). The pre-birth colonization seems to be affected by the length of gestation, as full-term infants are usually colonized by members from phyla Actinobacteria, Proteobacteria and Bacteroidetes, with low quantities of Firmicutes, while premature infants are often dominated by the Firmicutes and Tenericutes phyla, and less by Actinobacteria (8, 10, 17).

The first few days after birth entail major microbe colonization, and a difference in bacterial composition can be seen early on based on how the child was delivered. A vaginal birth will expose the infant to microbes found in and around the birth canal, and will mostly be from the mother. A child born by cesarean section, however, is not exposed to these vaginal microbes, and instead can be seen to be colonized by microorganisms associated with skin, many of which will stem from persons other than the mother, such as nurses, doctors, and family members (1, 3, 4, 17, 18). The early colonization is important for training and development of the immune system and, if disturbed, can lead to problems later in life in the form of, for example, allergies (1, 3, 19). In addition, the colonizers are important for digestion of dietary substances and xenobiotics, for triggering the intestine to mature, and for preventing pathogen colonization (1, 3, 18).

In the following days, weeks, months and years of an individual's life, the composition of the microbiota will change significantly based on many different factors such as genetics, hygiene, environment and nutrition. At around 2-5 years of age the microbiota reaches a more stable state associated with adults, and it is believed that the initial colonization of the infant gut affects the subsequent stages into the more complex adult microbiota (1, 3, 10, 18). When an adult-like microbial composition has been settled, evidence suggests that while short-term perturbations, caused by for example infections or use of antibiotics, may occur from time to time, the long-term composition usually remains stable and resistant to disturbance (1, 3, 4, 7, 8).

1.1.2 The adult gut microbiota

The gut microbiota of healthy adults consists mostly of facultative and obligate anaerobes (6, 8). The composition varies between individuals at lower taxonomic levels due to differences in for example genetics, age, gender, environment and diet, but at the higher taxonomic levels, certain compositional trends are usually shared (1, 3, 17, 20, 21, 22).

The large intestine is mainly dominated by members of the phyla Bacteroidetes and Firmicutes, followed by phyla Proteobacteria and Actinobacteria, with *Bifidobacterium* being the main genus (1, 5, 21, 22, 23). Other genera that are often detected are *Bacteroides*, *Eubacterium*, *Clostridium*, *Ruminococcus* and *Faecalibacterium* (1, 4, 5, 7, 17, 24).

Compared to the large intestine, there has been little investigation into the composition of the small intestine, mostly because this area of the body is harder to access. Studies that have been carried out on adults show that it is dominated by bacteria from the genera *Streptococcus*, *Clostridium* and *Veillonella* (4, 25, 26, 27). However, the composition fluctuates, and the bacterial profile can be quite different in the morning and the afternoon. The large intestine on the other hand has been shown to be quite stable over long periods of time, and the overall core composition of bacteria in the human gut seems to remain relatively constant within an individual over time (20, 27).

The tasks performed by the microbiota in the gut are numerous, and still more roles are being revealed. For example, *Bifidobacterium* species in the colon perform fermentation, which results in the products propionate and acetate. The former product affects the ability of the adaptive immune system to recognize foreign molecules by upregulating a type of white blood cells known as T cells, while the latter participates in reduction of inflammation, among other things (1, 6, 22, 24). Other bacteria commonly found in the colon have also been associated with anti-inflammatory effects (1, 6, 22, 24).

Another example of beneficial bacteria are members of the Bacteroidetes phylum, who fulfill several important functions such as producing exoenzymes that degrade certain complex carbohydrates like pectin and hemicellulose. These members, along with other microbes, are also needed in the development of a component of humans' pathogen defense called the gut-associated lymphoid tissue. In addition, they are linked to protection against food-related ailments, among other things (1).

It has been found that there is less variation in functional diversity between individuals than taxonomic diversity, indicating that different configurations of microorganisms in the gut lead to the same functions being fulfilled (3, 17, 28).

As the understanding of the patterns of microbial composition and function in the gut has increased, a recognition of what is considered a healthy gut and what is seen as a disturbed one have emerged, and with it, a realization of just how important a balanced microbiota is to our health.

1.2 Normobiosis and dysbiosis

Due to the observed shared trends and patterns among individuals, a healthy state of the gut, called normobiosis, can be defined for specific populations based on a characteristic composition and proportion of bacteria. A deviation from such a healthy state, in which the relationship between the microorganisms and the host has become dysfunctional, is termed dysbiosis, and can occur as a result of disturbances such as illness, change in diet, weight, aging, stress, or use of antibiotics and other medications (2, 3, 5, 6, 20, 29, 30). For some people a deviation from normobiosis has no noticeable negative associations. For others, however, dysbiosis has been linked to various medical conditions, indicating that the shift in the gut microbiota is somehow correlated with disease (3, 5, 6, 20, 28, 31, 32, 33, 34, 35). Examples of such medical conditions are irritable bowel syndrome (IBS), inflammatory bowel disease (IBD), neurological disorders, diabetes, malnutrition, obesity and cancer. In addition, dysbiosis has been suspected to affect behavior and mood, and disorders such as depression and autism (3, 5, 6, 20, 22, 28, 31, 32, 36).

Medical conditions associated with dysbiosis often demonstrate a reduced diversity in the gut microbiota of the individual, although how this is connected to a condition varies (22, 28, 33, 37). For example, in adults it has been found that if the abundance of certain Firmicutes species in the gut grows too large, it could be connected to both type 2 diabetes and obesity, although not all studies agree on this. In addition, a decrease in the Bacteroidetes phylum is often observed compared to in lean people (1, 3, 17, 22, 23). Certain metabolites linked to bacteria in the gut have also been found to play a role in these disorders (37).

IBS is a term used for conditions affecting mainly the large intestine with varying symptoms such as abdominal pain, constipation, diarrhea, cramps, bloating, flatulence and changes in bowel movement. The condition is quite common in the industrialized world and is caused by several different factors. Some cases seem to be linked to diet or the use of antibiotics, but anything that causes a disturbance in the microbial community in the gut can lead to symptoms of IBS (2, 3, 5, 6). While not all studies agree on the details, the general findings state that the bacterial composition is different in IBS patients than in healthy individuals. Examples are a reduction in genera *Lactobacilli* and *Bifidobacteria*, and an increase in the proportion of aerobes compared to anaerobes in persons suffering from IBS (5, 22, 24, 28).

IBD involves chronic inflammation of the digestive tract, and is a collective term for several diseases, the principal being Crohn's disease (CD) and ulcerative colitis (UC) (20, 22, 28, 38, 39). Although CD and UC both fall under the IBD definition, they can display significant differences in the microbial composition of the gut. Even within CD there are considerable differences between individuals, demonstrating the complexity of dysbiosis definitions (20, 22, 28, 38, 39). Some common trends found for patients suffering from CD is that abundance of phylum Firmicutes is lower, while certain members of phylum Proteobacteria, like *Escherichia coli*, are in increased abundance compared to healthy individuals. In addition, an increase in Bacteroidetes species and facultative anaerobes is often seen, although a decrease in the phylum Bacteroidetes has also been detected (22, 28).

Although dysbiosis is frequently associated with illness, it is often found to be an effect of a medical condition rather than the cause of one. However, in some cases a dysbiotic state is suspected of leading to serious diseases (3, 5, 6, 20, 22, 28, 31, 32, 33, 34, 36). An example of this is *Clostridium difficile* infections that sometimes follow an antibiotic treatment. This bacterium is an opportunistic pathogen that is commonly found in low amounts in the gut, and that is held in check by non-pathogenic bacteria. However, through the use of antibiotics and the subsequent disturbance of the microbial composition, *C. difficile* may grow more or less uninhibited and release toxins, leading to an infection in the individual (1, 3). Another example is necrotizing enterocolitis (NEC), which is a condition that is sometimes found in premature infants in which part of the intestine dies. In recent years studies have found that a dysbiotic state precedes the development of this condition, indicating that the disturbed microbiota could be the cause, rather than the effect of the disease (28, 33, 34, 35). Decreased diversity and a dominance of the phylum Proteobacteria can often be observed in patients suffering from NEC, and it is believed that certain bacterial communities, such as those dominated by a specific strain of *E. coli*, may in some cases trigger a hyperinflammatory response of the immune system, which in turn may lead to the development of the disease (33, 35). However, more studies are needed to confirm this and to further identify the microorganisms involved, as it is still difficult to determine whether dysbiosis in this case really is a cause or simply a reflection of the medical condition (3, 28, 34). There are still many unknown factors regarding dysbiosis, and compositional patterns and relationships in the gut are continuously being investigated (3, 5, 6, 20, 22, 31, 32, 36).

1.3 Investigating the composition of the gut microbiota

As the human gut contains mainly anaerobes, various genotyping methods, in which DNA sequences are examined, are frequently used for investigating the microbial composition. Such methods are largely based on using the marker gene 16S rRNA which is found in bacteria and archaea. This gene has been highly valuable for investigating microbial phylogenies and taxonomic assignments, as it is vital for the survival of the microbe and therefore tolerates very little mutation. In addition, it does not seem to participate in horizontal gene transfer. It consists of conserved regions, which allow for the comparison of more distantly related microbes, and hypervariable regions, which allow for lower level taxonomy (1). A limitation to the use of the 16S rRNA gene for genotyping, however, is that it has been unable to discriminate between closely related species and strains (1, 20, 40).

The widespread use of the 16S rRNA gene has led to the identification of variations in single nucleotides at specific sites in a genome that can be used to distinguish between organisms (1). Such single nucleotide polymorphisms (SNPs) are found in genes, intergenic regions, or other non-coding regions that are conserved, and can therefore provide information on evolution. They are found within all three domains of Bacteria, Archaea and Eukarya, and have become popular targets for detection technologies. In microbiology, the use of SNPs for genotyping has made it possible to distinguish between closely related species, and even strains, of bacteria (1).

1.3.1 Hybridization-based methods for SNP genotyping

Many SNP-based methods within microbiology rely on the use of probes, which are short sequences of nucleic acid that will hybridize to complementary strands. If a target DNA sequence is already known, these probes can be designed to fit specifically, and to produce a signal to inform if the target is present in the sample (1, 41, 42).

Different types of probes have been developed, and one example is the TaqMan probe. This type of probe contains a fluorophore (reporter) usually in the 5' end, and a quencher in the 3' end (41, 42). The proximity of the reporter to the quencher causes the fluorescence to get quenched due to a phenomenon known as fluorescence resonance energy transfer (FRET), in which the energy from the excited fluorophore is transferred to the quencher. During a quantitative polymerase chain reaction (qPCR), the probes bind to the target DNA

sequences, and when *Taq* polymerase starts elongating the primers, it cleaves the probe hybridized to the amplicon thanks to its exonuclease activity. This leaves the reporter and quencher physically separated, which results in a detectable increase in the intensity of the fluorescence. The qPCR machine can read this increase for each cycle, and creates an amplification plot that can be used later for analyses (41, 42, 43, 44).

Another example of probes are molecular beacons, which are probes that consist of an internal sequence complementary to a target, and regions on each side that are complementary to each other, allowing for a so-called hairpin loop structure to form. In its 5' end the probe carries a fluorophore, and in its 3' end a quencher. As with TaqMan, FRET occurs when the fluorophore and quencher are close to each other, which is what happens when the probe forms a hairpin structure. When the probe hybridizes to its target, it undergoes a conformational change from the hairpin loop, and the increased distance between the quencher and the fluorophore allows for fluorescence to be detected (41, 45).

A method that can use probes to look at SNPs is High Resolution Melting (HRM). This method is used to examine nucleic acid sequence variations by looking at their melting curves in a real-time PCR machine. When using probes for this, the probes can be designed to bind to a target sequence, and to dissociate from this target at a specific melting temperature (T_m) (46, 47, 48). This T_m can be adjusted by altering the length of the probe and the content of guanine (G) and cytosine (C) (due to the triple hydrogen bonds between the GC-base pair). To detect the dissociation, the probes can be used together with intercalating dyes that fluoresce when bound to double-stranded DNA, or the probes themselves can be labelled with molecules that fluoresce when the probe is hybridized to the target sequence. For both approaches, an increase in temperature leads to separation of the probe from the target, which in turn leads to a decrease in light intensity. This is registered by the qPCR machine and presented in melting curves as a drop in fluorescence near the designed T_m (46, 47, 48).

Oligonucleotide ligation assay, or OLA, is another example of a method that uses probes for SNP genotyping. It is based on using two oligonucleotides designed to hybridize adjacent to each other over the target SNP site on the DNA strand. DNA ligase will then covalently join the ends of the two probes. To detect the products, several methods can be used, such as labelling the probes with different fluorescent dyes. OLA can be combined with HRM to

increase multiplexing, by looking both at the dye color (first dimension) and the melting temperature (second dimension) in a technique known as 2D labelling. The level of multiplexing will then be dependent on the number of detection channels and the temperature resolution of the machine used (41, 49).

Another approach for detecting SNPs using probes is Multiplex Probe Amplification, or MPA (50). This method relies on probe pairs that are partially complementary to each other, and that are designed to be fully complementary to a specific target DNA sequence. Each probe pair carries a fluorophore and a quencher, and has a unique T_m . During a melting curve analysis, the two probes will denature as the temperature is increased, which based on the design will result in either an increase or a decrease in fluorescence, shown as a positive or negative peak near their unique T_m in the melting curve. If a target is present, however, one of the probes will bind to it rather than its complementary probe, because such a hybridization has a higher T_m than the two probes bound together. As a result, there will be no peak at that probe-duplex's T_m during the melting curve analysis, and this lack of peak is what reveals the presence of the target bacterium (50).

Hybridization-based methods for SNP genotyping are highly useful in many situations and have been found to have a better taxonomic resolution than many previous genotyping approaches (1, 20, 40, 41, 42, 46). However, as the human gut contains microbial communities and compositions that are quite complex, investigations of patterns that can provide important information for diseases and conditions require technologies that have a higher level of specificity than can be achieved with methods based predominantly on probe hybridization.

1.3.2 Single Nucleotide Primer Extension (SNUPE) technologies

Single nucleotide primer extension technologies, or SNUPE, are frequently used for detection of SNPs, and are often chosen because they are simple and rapid, and have a higher specificity and taxonomic resolution than techniques based mainly on hybridization. The principal behind SNUPE is that a specific primer gets extended in its 3' end with a single labelled terminating nucleotide, known as a ddNTP (dideoxynucleotide triphosphate), when it hybridizes upstream from the nucleotide position of interest (41, 51). The discriminating power of the technology comes both from the need for the primers to hybridize to the target DNA sequence, and from the fidelity of the polymerase ensuring it elongates the

primer only when the 3' end matches the template. This means that both internal and terminal mismatches can cause the extension to fail, as internal mismatches cause the primer-template hybridization to be less stable, while a mismatch at the 3' end of the primer affects the incorporation of the correct ddNTP. By designing primers with such mismatches in mind, non-specific extensions can be prevented (51).

Initially radioisotopes were used for labelling of the ddNTPs, followed by various other alternatives, but today it is more common to use fluorescent dyes (51). As the ddNTP is complementary to the nucleotide on the opposite strand, the label provides information on the SNP on the target sequence. By using different fluorophores on the four ddNTPs, several base variations can be detected in one reaction. Combined with primers that are designed with different lengths and for different targets, it is possible to analyze a relatively large number of SNPs in one reaction by reading the signals from the ddNTPs and separating the extended primers (51).

There are several methods available for separation of the primers following SNUPE. When primer extension assays were first developed, it was common to use gel electrophoresis, wherein the separation of macromolecules is based on size. Using an electric current, the negatively charged DNA fragments travel through the gel towards the positive charge, with smaller fragments travelling faster than larger ones. The final band pattern, visualized with UV light, reveals the length of the oligonucleotides and thus the identity of the fragments (1, 51). Later capillary electrophoresis was developed, which has a more accurate size separation than gel electrophoresis, resulting in higher resolution (51).

Another separation method used for SNUPE is the Matrix-assisted laser-desorption ionization-time-of flight-mass-spectrometry (MALDI-TOF-MS), which separates by mass rather than size by taking advantage of the fact that there is a slight mass difference in nucleotides, which is increased in chemically altered ddNTPs. This makes it possible to separate products that are labelled differently by measuring their molecular weight (41, 51). Alternatively, tags can be used for primer separation. If a tag is added to the 5' end of the primers, they can be sorted by hybridizing them to complementary "anti-tags" attached to a solid format such as an array or bead after the labelling has occurred. The signal can then be analyzed, and the anti-tags provide information on the target (51).

There are many assays that are based on the primer extension principle, and still more are being developed. Initially, primer extension technologies were mostly used for human genetics, but other disciplines have also embraced this approach. An example of an assay based on SNUPE within the field of microbiology is the quencher extension (QEXT) assay. This assay detects and quantifies SNPs by combining the use of ddNTPs with the use of reporters and quenchers in real-time in a closed, single-step tube (52, 53).

The QEXT principle is illustrated in figure 1 and is based on a probe containing a reporter dye which is complementary to the target DNA. This probe gets extended in its 3' end with a ddNTP containing a quencher when it hybridizes to the target sequence with the specific SNP. The proximity of the quencher to the reporter results in a change in fluorescence due to FRET. Some quencher molecules can be used either as fluorescence energy donors or acceptors depending on the reporter dye, and it is possible to detect the SNP either by an increase or decrease in fluorescence. By using a variety of fluorophores for different SNPs, the method has multiplexing possibilities (52, 53).

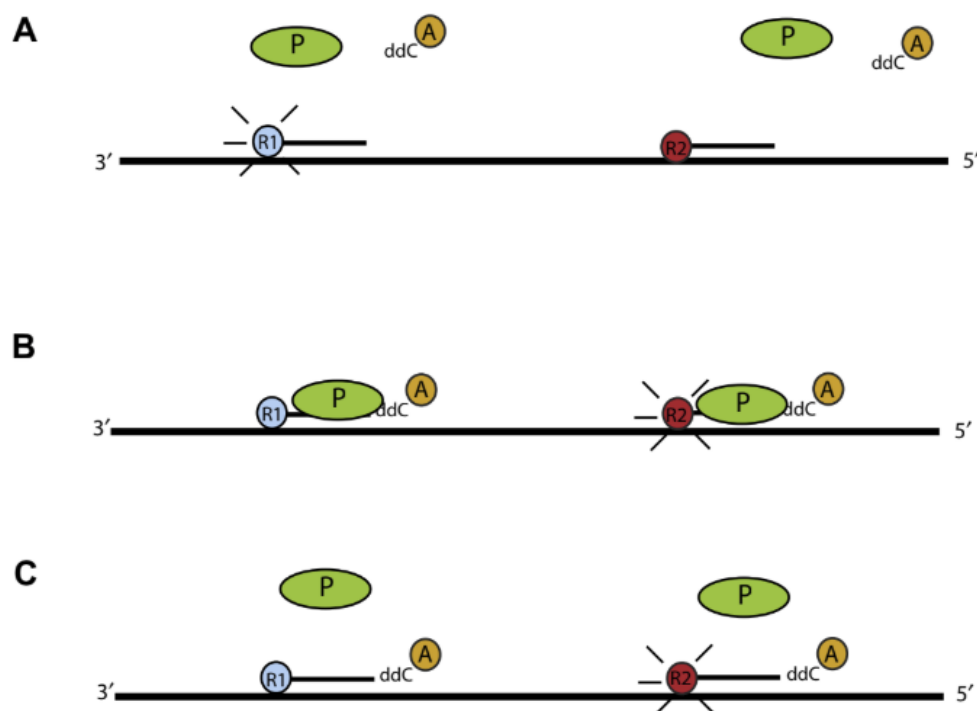


Figure 1: The principle of the QEXT assay. The black lines are DNA, the green oval (P) is the polymerase, the blue circle is reporter 1 (R1), the red circle is reporter 2 (R2), and the ddC with an orange circle (A) is the ddNTP with the energy donor/acceptor. Before an SNP is detected, R1 emits fluorescence while R2 does not, as shown in part A of the figure. The labelled ddCTP is incorporated into the probe by the polymerase if the probe is hybridized to the target sequence with the SNP of interest, as shown in part B. The QEXT signal is detected as reduced fluorescence for R1, while R2 shows increased fluorescence, both indicating that the target SNP is present, as shown in part B and C of the figure (53).

1.3.3 Assay format

For both hybridization-based methods and SNuPE technologies there are generally two types of reactions used for investigating SNPs when looking at assay format; in solution (homogenous), or with a solid-phase (e.g. chips, beads, or slides). Reactions that utilize solid-phase have greater possibilities for multiplexing than reactions in solutions, but homogenous assays have the advantage that no purification and separation is required. As a result, solution-based methods are more easily automated, as PCR and detection take place in the same sealed-tube reaction, removing the need for post-PCR handling and reducing the risk of contamination (41).

An example of a technology that detects SNPs using a solid-phase platform is the microarray technology. This technology allows for analysis of thousands of DNA sequences at the same time by first labelling amplified target fragments with fluorophores. These fragments then get hybridized to complementary capture probes that are attached on a solid surface, after which the fluorescence from the labels can be detected. Identity is revealed by the position of the capture probe on the surface, while relative quantitation is seen by the intensity of the signal (40, 41).

HRM and QEXT are two examples of methods that can use reactions in solution to detect SNPs. For these methods, the use of various fluorophores enables some multiplexing, but this is limited by the number of detection channels on the available instrument. Hence, although homogenous assays are superior to solid-phase platforms in several ways, the limited possibility for multiplexing has made them less useful in cases where this is important, such as when investigating complex microbial communities. New methods could therefore benefit greatly from being developed with this in mind.

1.3.4 Liquid Array Diagnostics (LAD)

A novel technology that aims to alleviate the multiplexing limitation of homogenous assay formats is Liquid Array Diagnostics (LAD). This technology enables rapid, high-volume screening of SNPs in various genetic markers. The increase in multiplexing is due to the use of a combination of SNUPE labelling and HRM analysis (54, 55).

An illustration of the LAD process can be seen in figure 2 on the next page. The principle behind the method is that when a target DNA sequence is present in a sample, a specific labelling probe (LP) will hybridize to it and get extended with a quencher-labelled ddNTP (ddNTP-Q) in its 3' end. This labelled LP will then denature from the target sequence and bind to a reverse complementary reporter probe (RP) that carries a fluorophore in its 5' end. The proximity of the quencher to the fluorophore results in a quenching of fluorescence (54, 55).

Each probe duplex is designed to have a specific T_m . By using HRM, a signal is detected near this T_m as the temperature is increased and the probe duplex denatures, separating the quencher from the fluorophore (54, 55). The differences in T_m between probe duplexes will result in signals that are distinguishable from one another. Thus, the temperature a signal is detected on reveals the identity of the target, and there is no need to separate the probes physically, for example by use of gel electrophoresis or mass spectroscopy (47, 54, 55).

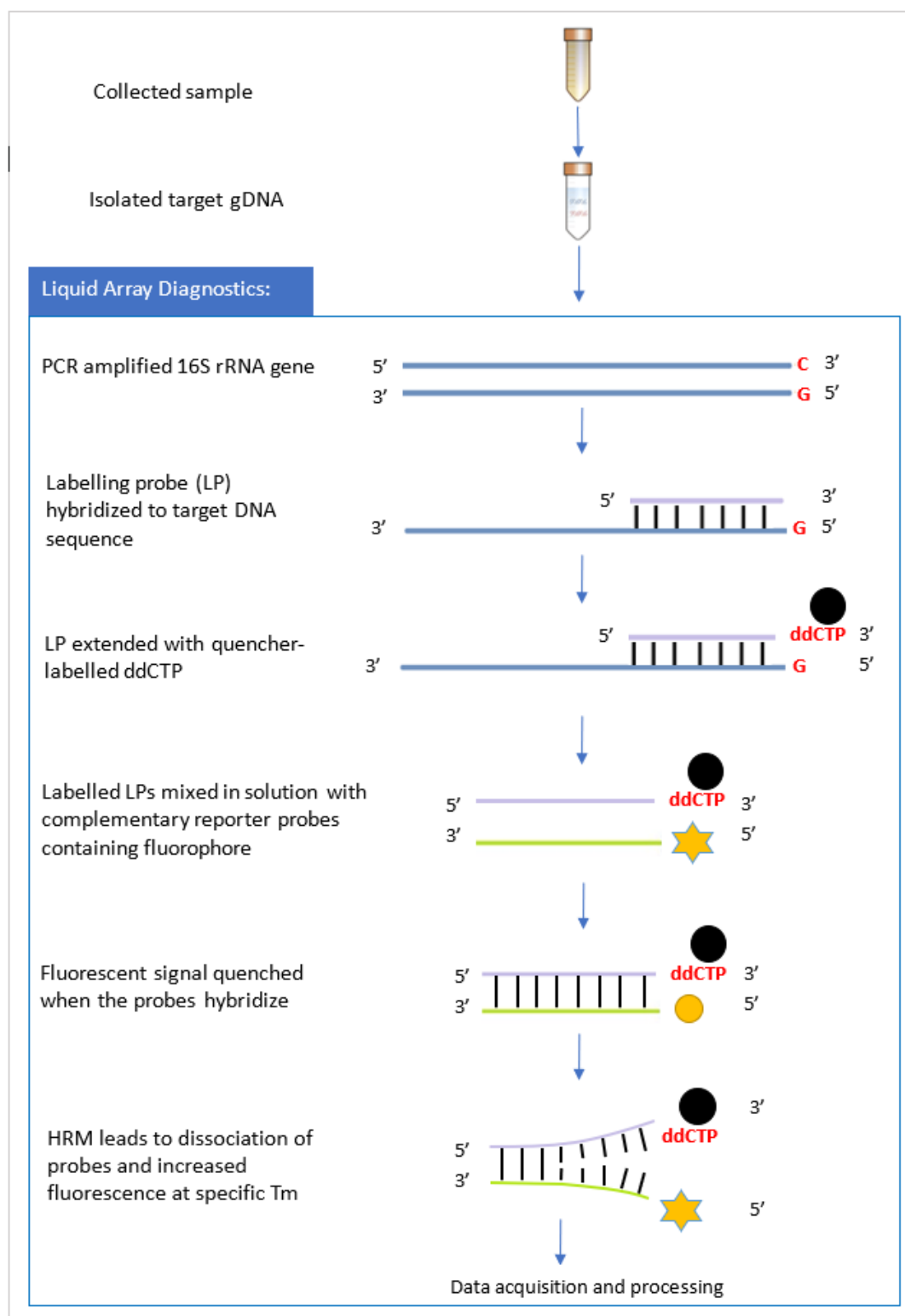


Figure 2: Overview of the Liquid Array Diagnostics process. First a sample is collected, genomic DNA (gDNA) is isolated, and 16S rRNA genes are amplified. Labelling probes (LPs) are added to the sample and get extended with a quencher-labelled ddNTP (in this case a ddCTP-Q) when hybridized to the target DNA sequence. The LPs are denatured from the target, and fluorescently labelled reporter probes (RPs) that are complementary to the LPs are added to the solution. If the RP hybridizes to an LP containing a ddCTP-Q, the fluorescence is quenched, indicating the presence of the target bacterium. Using High Resolution Melting (HRM), the probes dissociate at their specific melting temperature (T_m), and fluorescence increases as the fluorophore and quencher are separated. Data about fluorescence and quenching is then processed.

There are several different fluorescence molecules available on the market, and by labelling the RPs with different ones, they can be distinguished by using different detection channels on the qPCR machine. When this is combined with the variations in T_m of the probe duplexes, the number of targets that can be detected in one single reaction increases dramatically from previous solution-based technologies (54, 55).

Initial experiments with LAD have demonstrated it to be a very promising genotyping technology for investigations of complex microbial communities such as that of the gut, and a competitor to many of the current methods available (54, 55). Its use of SNUPE makes it highly specific, as the LP in LAD needs to bind to the target, and in addition needs to get the ddNTP-Q correctly incorporated. This sets LAD above technologies that only rely on hybridization, like OLA and MPA. These technologies can suffer from false positives due to mismatches between the probe and the SNP site being treated as a normal base-pairing, as it has been found that primers that have a G/T or T/G mismatch can work almost as efficiently as a normal primer with a G/C or C/G match (56, 57).

The combination of SNUPE with HRM enables LAD to have a one-tube assay which improves efficiency and reduces contamination risk compared to for example SNaPshot, which is a kit from Applied Biosystems (Waltham, Massachusetts, USA) that requires physical separation of the oligonucleotides to detect microbial identity. Additionally, HRM is cheaper than for example electrophoresis, and there is no need for hazardous chemicals like acrylamide, ethidium bromide or formamide, like in methods such as denaturing gradient gel electrophoresis (DGGE) (46, 47, 51).

As the probes in LAD remain intact, subsequent analysis is possible if necessary, for example by running the samples again. This makes it more advantageous than for example TaqMan-based assays, in which the probes are fragmented. A drawback of LAD is that at present it hasn't been optimized for quantitative use.

The different methods available today for investigating the gut microbiota all have their strengths and limitations, and together their use has led to an increased understanding of the main microbial patterns found in the gut. The link of these patterns to human health and well-being has resulted in a need for assays to be developed that can quickly identify a large number of microbes, and distinguish sufficiently between them to discriminate between normobiosis and dysbiosis.

1.4 The GA-map® Dysbiosis Test

As technologies based on SNuPE have a higher fidelity than those based mainly on hybridization, they can be of great value for detection of dysbiosis due to the complexity of the microbial communities found in the gut. An assay that uses the SNuPE concept successfully for this purpose is the Genetic Analysis microbiota array platform Dysbiosis Test, or GA-map® Dysbiosis Test, developed by Genetic Analysis AS, Oslo, Norway. This test is a well-established diagnostic tool, and currently the only clinically validated assay for determination of dysbiosis. The test uses a solid-phase platform and performs SNP detection by analyzing fecal samples using 54 highly specific SNuPE probes that cover more than 300 strains of bacteria within phyla Firmicutes, Proteobacteria, Bacteroidetes, Actinobacteria, Tenericutes and Verrucomicrobia (19, 20, 58). The probes, known as labelling probes (LPs), have been designed to be complementary to bacteria that can be expected to be found in the adult gut. Of the 54 probes, 26 are species specific, 19 detect at genus level, and the remaining 9 detect bacteria at higher taxonomic levels (19, 20, 58). The probes target conserved and hypervariable regions V3 to V9 of the 16S rRNA gene, which gives the test higher resolution than for example high-throughput sequencing (HTS), which usually covers V3 and/or V4 of the 16S rRNA gene (20). The effect of this difference was demonstrated by Casén et al. (20) when their experiment managed to detect the genus *Mycoplasma* in a majority of the 188 samples investigated using the GA-map® Dysbiosis Test, but only in one when using MiSeq Illumina sequencing on the same samples. As a result, certain bacteria that may have great importance for dysbiosis can go undetected when using methods such as HTS.

An illustration of the GA-map® Dysbiosis Test process is presented in figure 3 on the following page. The principle behind the test is that the LPs hybridize to a target sequence present in a sample collected from the patient, and get extended with a labelled ddNTP. These labelled LPs will then separate from the template and bind to reverse complementary oligonucleotides that are attached to uniquely barcoded magnetic beads. Signals from the labelled probes are detected, and the barcode on the bead will inform which target the signal refers to (19, 20). Using an algorithm, the bacterial composition of the fecal sample is then compared to a normobiotic microbiota profile model which has previously been made from a reference population based on individuals that are non-symptomatic.

Deviations from the profile are detected, and a dysbiosis score is calculated. Genetic Analysis divides the degree of dysbiosis into an index of 5. If a fecal sample is given a dysbiosis index (DI) of 1 or 2, the microbiota is considered normobiotic. If it is given a score of 3-5, there is dysbiosis, with DI5 being the most deviant from normobiosis (20).

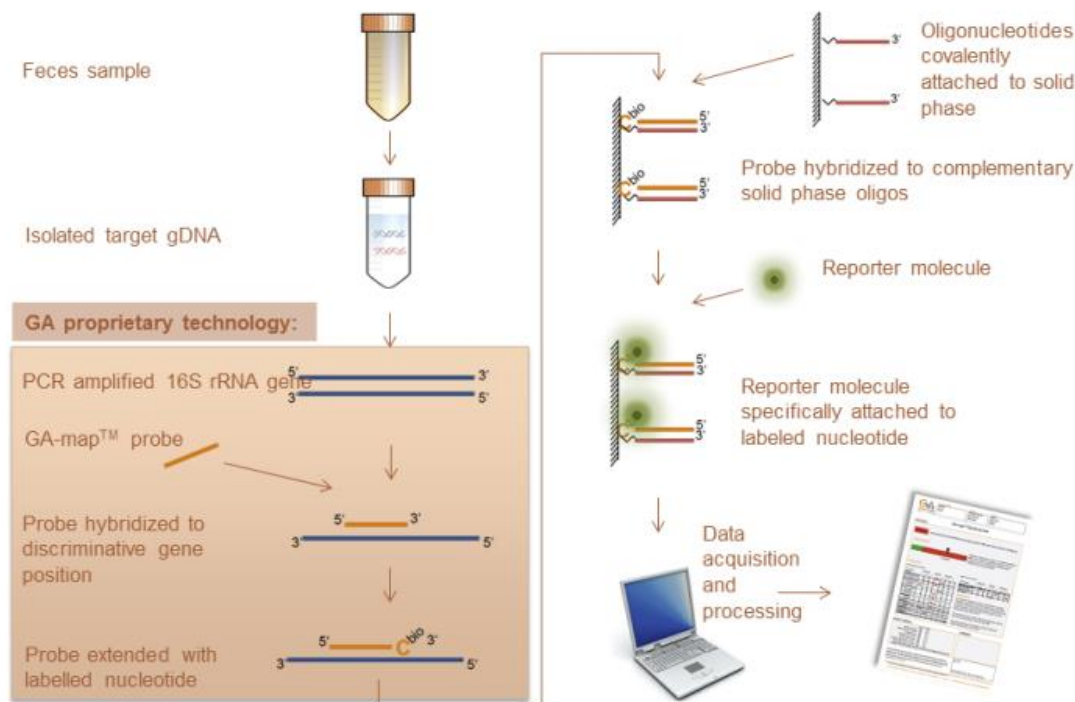


Figure 3: The process of the GA-map® Dysbiosis Test. Fecal samples are collected, genomic DNA is isolated, and 16S rRNA genes are amplified using PCR. Labelling probes (LPs) hybridize to complementary bacterial DNA targets present in the sample and get extended with a labelled ddCTP. Magnetic beads containing oligonucleotides that are complementary to the LPs are added, to which the LPs hybridize. Signals are detected from the labelled LPs for bacteria that are present, and the magnetic beads inform which bacterium a signal refers to. The data collected from the test is processed, and dysbiosis index (DI) level is determined, where a DI of 1 or 2 indicates normobiosis, and a DI of 3-5 indicates dysbiosis.

(<http://www.genetic-analysis.com/wp-content/uploads/2014/10/GA-map-principle-1024x576.png>)

The GA-map® Dysbiosis Test has proven to be an efficient and accurate way to determine degree of dysbiosis in an individual, and is very useful for deciding the best course of action for treatment, and for tracking said treatment (19, 20). However, it is a relatively labor-intensive and expensive process, as the probes need to bind to the beads, washing is needed to remove any unmarked probes, and the barcodes of the beads need to be identified. Consequently, there is a desire to improve the test by supplementing the currently used solid-phase platform with a homogenous one, as such an implementation is believed to reduce cost and increase efficiency of the test.

1.5 Aim of this thesis

Due to the growing need for efficient assays that are capable of determining dysbiosis, the aim of this thesis was to investigate the possibility of improving the GA-map® Dysbiosis Test by supplementing the solid-phase platform currently employed, with a homogenous platform. The new and emerging Liquid Array Diagnostics technology has been shown to fulfill the demands that are required to detect point mutations sufficiently for detection of dysbiosis, and has previously been proven to be rapid, simple, and economical. Owing to its combination of SNuPE and HRM, the purification and separation steps are no longer needed, and the limited multiplexing capabilities of previous homogenous technologies have been mitigated. This technology has therefore been chosen as a potential supplement, as a successful implementation of LAD as a diagnostic tool is believed to greatly improve the GA-map® Dysbiosis Test.

To accomplish the main objective of investigating if LAD can be incorporated into the GA-map® Dysbiosis Test, several sub-goals were set:

- Design reporter probes that were complementary to 20 of the labelling probes designed by Genetic Analysis for their current GA-map® Dysbiosis Test
- Determine if the probe duplexes could be used with LAD to detect target bacteria
- Examine if LAD could detect signals from 20 probes simultaneously
- Investigate if LAD could distinguish between normobiosis and dysbiosis
- Determine which of the probes tested best separate normobiosis from dysbiosis

This thesis focused on bacteria in relation to dysbiosis, and while the gut contains viruses, and members of the domains Archaea and Eukarya as well (5, 6, 28, 59, 60), this falls outside the scope of this thesis.

2. Materials and methods

This thesis is part of a larger project led by Genetic Analysis AS (GA), which aims to investigate the use of LAD as a supplement to the GA-map® Dysbiosis Test. It was theorized that LAD could detect 20 targets simultaneously, and the 54 LPs used in the GA-map® Dysbiosis Test were therefore divided by GA into three subsets. The contribution of this thesis was to test LAD with one of the subsets, which contained 20 probes selected by GA. The probes in this subset were the least associated with normobiosis or dysbiosis of the 54 probes. The remaining two subsets will be investigated by GA. A flow chart for the project is presented in figure 4, where the blue boxes demonstrate the process of this thesis.

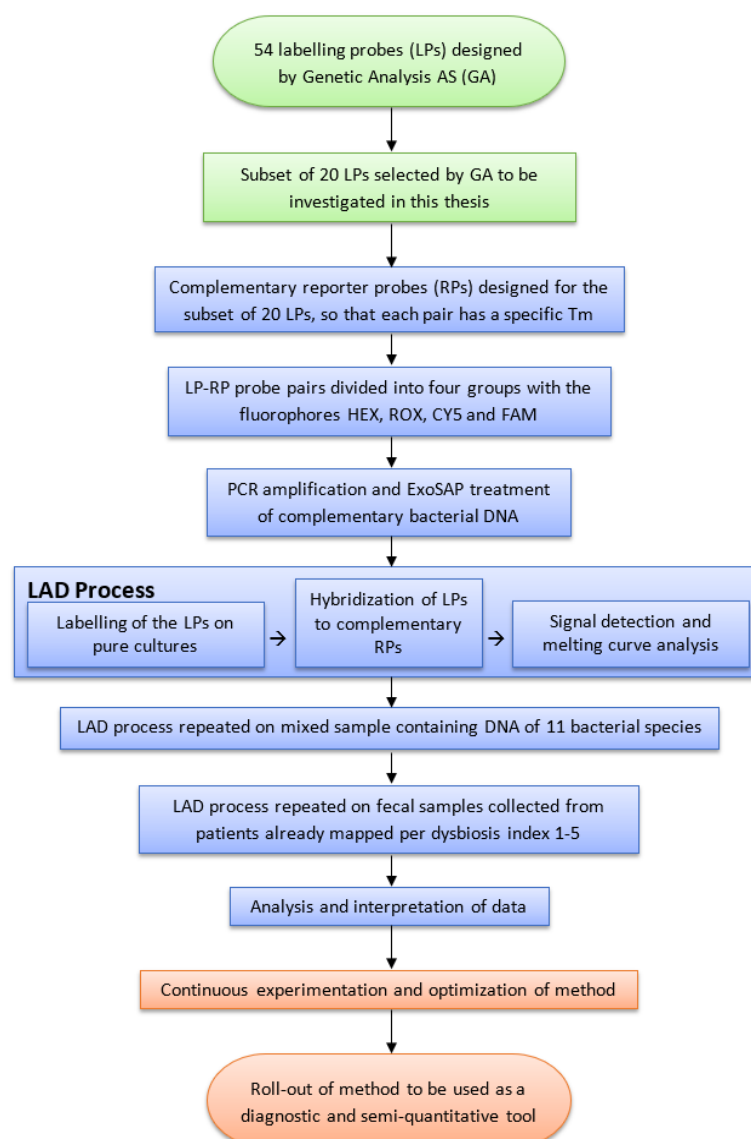


Figure 4: Flow chart for the project. The green boxes show the process handled by GA prior to the beginning of this thesis. The blue boxes show the process of this thesis, where a subset of 20 LPs designed by GA was investigated. Complementary RPs were designed, and the LP-RP duplexes were tested on pure cultures of target bacteria, on a mixed sample, and finally on fecal samples. The red boxes show how the project may continue after the completion of this thesis.

2.1 Design of reporter probes

The LPs from the GA-map® Dysbiosis Test had been designed by GA to get labelled with a ddCTP-Q in their 3' end when hybridized to a target DNA sequence. For this thesis, the first step was to design RPs that were complementary to the 20 LPs being investigated.

When designing the RPs, the goal was to enable detection of 20 targets simultaneously by using four detection channels on the qPCR machine, with five distinguishable melting temperatures per channel. The 20 LPs were therefore split into four groups of five probes each. The fluorophores used for each group were HEX, ROX, CY5 and FAM. Within each group the RPs were designed so that the LP-RP duplexes had a theoretical T_m of roughly 30 °C, 40 °C, 50 °C, 60 °C and 70 °C. One of the LPs was a universal (UNI) probe, and the complementary RP for this probe was designed by GA. A list of the probe pairs with the detection target, fluorophore and theoretical T_m is presented in table 1.

Table 1: List of the probes investigated in this thesis, along with the targets that they were designed to detect¹, the associated fluorophore, and the theoretical T_m for the probe duplexes in Celsius.

Probe	Designed for detection of ¹	Fluorophore	Theoretical T_m
Bac230	<i>Bacteroides fragilis</i>	HEX	28,8
Fir224	Firmicutes (various)	HEX	38,8
Eub206	<i>Eubacterium hallii</i>	HEX	46,8
Ali239	<i>Alistipes</i>	HEX	59,3
UNI	Universal 16S rRNA	HEX	64,8
Bac155	<i>Bacteroides Stercoris</i>	ROX	29
Dor124	<i>Dorea</i> spp.	ROX	40,3
Str138	<i>Streptococcus</i> spp.	ROX	49
Cat187	<i>Catenibacterium mitsuokai</i>	ROX	58,8
Eub119	<i>Eubacterium rectale</i>	ROX	69,7
Pse30	<i>Pseudomonas</i> spp.	CY5	29,5
Lac71	<i>Lactobacillus</i> spp.	CY5	40,2
Lac47	<i>Lactobacillus ruminis</i> & <i>Pediococcus acidilactici</i>	CY5	50
Aci57	<i>Acinetobacter junii</i>	CY5	59,5
Bac91	<i>Bacteroides</i> spp.	CY5	70
Par89	<i>Parabacteroides johnsonii</i>	FAM	28,8
Bac109	<i>Bacteroides pectinophilus</i>	FAM	39,9
Act39	Actinomycetales	FAM	49,2
Str78	<i>Streptococcus</i> spp.	FAM	59,4
Bac88	<i>Bacteroides zoogloformans</i>	FAM	70

¹ Certain content of this project is considered to be trade secrets and will therefore not be revealed in this thesis. This includes the sequences of the probes used, and any bacterial information related to the probes, except what has already been made publicly available by Genetic Analysis AS.

For designing the RPs, Oligoanalyzer 3.1 web-based bioinformatics tool (Integrated DNA Technologies, Coralville, Iowa, USA) was used. The parameters were set to: Oligo concentration 0.1 M, Na⁺ concentration 0 mM, Mg⁺⁺ concentration 1 mM and dNTP concentration 0.0008 mM.

The RPs were designed to have the fluorophore attached to their 5' end. To adjust the probe duplex T_ms, the RPs were therefore shortened in the 3' end to ensure the 5' end with the fluorophore would hybridize close to the quencher at the 3' end of the LP.

All of the LPs designed by GA had a T_m significantly lower than 70 °C. As each group was to have one probe duplex at roughly this temperature, four of the LPs needed to be altered to increase the T_m. This was done by adding nucleotides to the 5' end, leaving the 3' end intact for labelling with ddCTP-Q.

All designed probes were checked for potential hairpin structures and self-dimerization, and all duplexes were checked for hetero-dimer formations. Blunt ends were avoided in all duplexes.

The probes were ordered in dry form from Biomers.net, Ulm/Donau, Germany. They were prepared by adding pH7 of Ambion® molecular biology grade 0.1 M EDTA (Thermo Fisher Scientific, Waltham, Massachusetts, USA) and Ambion® molecular biology grade 10 mM Tris (Thermo Fisher Scientific).

2.2 Probe validation on pure cultures of target bacteria

2.2.1 Template preparation

To investigate if the probes were compatible with LAD, they were tested on pure cultures of target bacteria. Twenty species were therefore selected based on their complementarity to the probes, and their 16S rRNA genes were amplified using PCR. Each amplification reaction consisted of 2.5 mM MgCl₂ (Solis Biodyne, Tartu, Estonia), 0.05 U/μl HOT FIREPol DNA Polymerase (Solis Biodyne), 1X B1 Buffer (Solis Biodyne), 0.2 mM dNTPs (Solis Biodyne), 0.2 μM Mangala F1 forward primer (biomers.net), 0.2 μM 16S UR (biomers.net) and 2 μl bacterial template (provided by Genetic Analysis AS), in a total reaction volume of 50 μl.

The PCR program began with an activation step of 15 minutes at 95 °C. Following this was 30 cycles that started with 30 seconds at 95 °C for denaturation, 30 seconds at 55 °C for annealing, and 1 minute and 20 seconds at 72 °C for elongation. A final elongation step took place for 7 minutes at 72 °C. The process was performed using an Applied Biosystems 2720 Thermal Cycler (Thermo Fisher Scientific).

The DNA concentrations of the amplified templates were checked using the Quant-iT™ assay kit with the Qubit Fluorometer (Invitrogen Detection Technologies, Carlsbad, California, USA), and gel electrophoresis at 80 volts for 30 minutes with 5 µl of template mixed with 1 µl Gel Loading Dye, Purple (6X) (New England BioLabs, Ipswich, Massachusetts, USA) on a 1% agarose gel containing 1X TAE buffer.

To remove excess primers and nucleotides, the PCR products were treated with 0.12 U/µl Exonuclease I (Thermo Fisher Scientific), and 0.32 U/µl FastAP Thermosensitive Alkaline Phosphatase (Thermo Fisher Scientific). The reaction was performed using an Applied Biosystems Veriti™ 96 Well Thermal Cycler (Thermo Fisher Scientific), starting with two hours of incubation at 37 °C, followed by 15 minutes at 85 °C.

2.2.2 Labelling of the LPs

To examine if the LPs would hybridize to target DNA sequences and get labelled with a ddCTP-Q when used with LAD, 20 separate reactions were prepared which contained 2 µl of pure amplified DNA template, 0.1 µM of each LP (Table 1), 1x buffer C (Solis Biodyne), 1 mM MgCl₂ (Solis Biodyne), 0.25 U/µl HOT TERMIPol DNA Polymerase (Solis Biodyne), 0.8 µM ddCTP^{DYQ660} (2',3'-dideoxycytidine-5'-triphosphate labelled with quencher DYQ660, Jena Bioscience, Jena, Germany), and 20 µM ddTTP (2',3'-dideoxythymidine-5'-triphosphate, Jena Bioscience). Each reaction had a total volume of 20 µl. In addition, two negative controls were prepared, in which 2 µl of water was added instead of DNA template. The UNI probe was used as a positive control.

The labelling reaction was performed with Applied Biosystems Veriti™ 96 Well Thermal Cycler (Thermo Fisher Scientific). The program started at 95 °C for 12 minutes, followed by 40 cycles of 96 °C for 20 seconds and 60 °C for 40 seconds.

2.2.3 LP-RP hybridization, signal detection and melting curve analysis

To inactivate the polymerase and detect fluorescent signal, 10 µl of each labelling reaction was mixed with all RPs with a final concentration of 0.005 µM each, 0.1% SDS (Thermo Fisher Scientific), and 0.3X buffer C (Solis Biodyne). Total reaction volume in each sample was 15 µl. Signal detection and melting curve analysis was performed on Applied Biosystems 7500/7500 Fast Real-Time PCR Systems (Thermo Fisher Scientific). The process started with 15 seconds at 95 °C, followed by one minute at 20 °C, another 15 seconds at 95 °C and finally 15 seconds at 60 °C.

The qPCR software prepared one melting curve per sample from each channel, which means that since four channels were used, each sample had four melting curves. These melting curves were presented as the negative first derivative of fluorescence with respect to temperature ($-dF/dT$), in which the signals from the probes were shown as distinct negative peaks near the designed theoretical T_m s on an otherwise stable baseline. Melting curves for negative controls should show no peaks, as there should be no major changes in fluorescence.

To determine if a probe was compatible with LAD, all the melting curves from the pure culture tests were visually inspected for all samples to see if the probes provided a satisfactory signal in expected samples, and no signal in samples containing non-complementary, or no, templates. The criteria for a satisfactory signal was that it needed to be observed near the designed theoretical T_m on the correct channel, it needed to be clearly distinguishable from the baseline, noise and other signals, and it should not be seen on other channels at the same temperature as that would imply noise. In addition, the tests were repeated with all probes at least three times, and the signal needed to fulfill the above criteria each time.

DNA templates that had been used to test probes that did not provide satisfactory signals were sent to Eurofins Scientific (Luxembourg) for GATC Biotech's LightRun Sanger sequencing to investigate if there were any unexpected mutations in the 16S rRNA gene sequences.

2.3 Validation of LAD on a mixed sample

To investigate if the probes would provide distinguishable signals with LAD in a mixed sample, 11 of the bacterial templates that had led to good results during probe validation on pure cultures were selected and pooled together.

The labelling reaction was performed as described under section 2.2.2 on page 21, but with 1 µl of each template (11 µl in total). All reagent concentrations remained the same as under section 2.2.2, and the total reaction volume was 20 µl. LP-RP hybridization, signal detection and melting curve analysis of the mixed sample was performed as described in section 2.2.3 on the previous page.

Based on the bacterial templates selected, the probes that were expected to give a signal in the mixed sample were Fir224, Eub206, Ali239, UNI, Str138, Cat187, Eub119, Pse30, Lac47, Aci57, Par89, and Bac109. For more information on the probes, see table 1 on page 19.

2.4 Evaluation of LAD for diagnosis of dysbiosis

2.4.1 Performing LAD on fecal samples

To test if LAD could be used with the current probeset to correctly differentiate between normobiosis and dysbiosis in real patients, the process was performed on fecal samples (provided by Genetic Analysis AS) that had previously been indexed using the GA-map® Dysbiosis Test. GA's dysbiosis index ranges from 1 to 5, where DI1 and DI2 are considered normobiotic, and DI3 to DI5 are considered dysbiotic. Thirty-two fecal samples were investigated, of which 15 had been classified by GA as normobiotic (eight as DI1 and seven as DI2), and 17 samples as dysbiotic (seven as DI3, three as DI4, and seven as DI5). The samples had been collected from adults in Norway, and the ones that had been classified as dysbiotic were collected from individuals suffering from IBS and IBD.

Labelling was performed as described under section 2.2.2 on page 21, but with 5 µl of fecal sample in a total reaction volume of 20 µl. Probe hybridization and signal detection was performed as described in section 2.2.3 on page 22, but due to the complexity of the fecal samples compared to the pure cultures and mixed sample, visual determination of probe performance using the melting curves was not possible. Instead the raw data was extracted, and the peak values for each probe were used for further investigation (see table S1 and S2 in the supplementary material for the extracted peak values).

2.4.2 Statistical analysis

The average detected signal from a probe in the normobiotic samples was compared against the average detected signal from the same probe in the dysbiotic samples to investigate the difference between the two states. This was done for all probes except the UNI probe, as that was used as a positive control. In addition, a Welch Two Sample t-test was performed with a 95% confidence interval using R commander (open source software), version R 3.5.1. This was done to investigate the statistical significance of the differences in detected signals from each probe with respect to normobiosis and dysbiosis, where a p-value of less than 0.05 was considered significant.

Before further analyses were conducted, the raw data was normalized by dividing the extracted peak values for each probe by the signal from the UNI probe.

The following analyses were performed by Professor Knut Rudi using the multi-paradigm numerical computing environment MatLab (MathWorks, Natick, USA).

The accuracy of LAD as a diagnostic tool for dysbiosis was investigated by using a receiver operating characteristic (ROC) curve analysis. This analysis shows the true positive rate (the proportion of dysbiotic samples that are correctly classified as such) against the true negative (the proportion of normobiotic samples that are correctly classified as normobiotic), shown as sensitivity and specificity respectively. The analysis is a representation of how well LAD, with the probes used in this thesis, can distinguish between healthy and sick individuals and correctly classify a sample.

When creating the ROC curve, the best-case scenario was first estimated by the MatLab software based on the entire dataset. For this the software had been informed which samples were normobiotic and which were dysbiotic, from which it prepared an ideal classifier. Following this, a cross-validated model was prepared, in which the software had no knowledge of the diagnoses of the samples. This model was made to predict the true performance of the LAD technology when using these probes. To reach this model, the software randomly partitioned the data into complementary subsets, and then analyzed one such subset to make a preliminary model. The remaining subset was then used to validate the model by seeing how well they fit. This was repeated many times by the software, each time with a new random partitioning of samples. Finally, the software made a final cross-validated model based on an average of all the preliminary cross-validated models.

For both the ideal classifier and the final cross-validated model the optimal points with regards to sensitivity and specificity for the dataset were calculated. These points were determined by the software to give the best trade-off between correctly classifying a sample as dysbiotic, against the cost of wrongly diagnosing a healthy individual as sick.

Following the creation of the ROC curve, the probes were investigated to see which of them were more strongly associated with either normobiosis or dysbiosis to get an indication of which probes, if any, may best separate the two states. To achieve this, the MatLab software was informed of which samples were dysbiotic and which were normobiotic. First the software identified the regression vectors for each probe. The signals detected for the probes during the LAD process were then multiplied with the regression vectors, and the sum informed the association of the probe, as a positive value indicated dysbiosis, while a negative value would indicate normobiosis.

3. Results

3.1 Probe compatibility with LAD

An overview of the 20 probe pairs investigated in this thesis, with a summary of their performance during testing on pure cultures and the mixed sample, is presented in table 2. The background for the results shown in the table will be presented on the following pages.

Table 2: Summary of the findings for each of the 20 probe pairs during testing on pure cultures and the mixed sample

	Pure Culture		Mixed sample				
Probe	Signal on intended target(s)?	Comment ¹	Signal ² expected?	Signal detected?	Comment ¹	Compatible with LAD? ³	
HEX	Bac230	No	Weak signal for unintended target	No	Yes	Weak signal for unintended target	No
	Fir224	Yes	Medium signal	Yes	Yes	Medium signal	Yes
	Eub206	Yes	Strong signal	Yes	Yes	Strong signal	Yes
	Ali239	Yes	Strong signal	Yes	Yes	Strong signal	Yes
	UNI	Yes	Strong signal	Yes	Yes	Strong signal	Yes
ROX	Bac155	No	No signal	No	No	N/A	No
	Dor124	No	No signal	No	No	N/A	No
	Str138	Yes	Strong signal	Yes	Yes	Strong signal	Yes
	Cat187	Yes	Strong signal	Yes	Yes	Strong signal	Yes
	Eub119	Yes	Strong signal	Yes	Yes	Strong signal	Yes
CY5	Pse30	Yes	Strong signal	Yes	Yes	Strong signal	Yes
	Lac71	Yes	Medium signal	No	No	N/A	Yes
	Lac47	Yes	Strong signal	Yes	Yes	Strong signal	Yes
	Act57	Yes	Strong signal	Yes	Yes	Strong signal	Yes
	Bac91	Yes	Weak signal	No	No	N/A	Yes
FAM	Par89	Yes	Strong signal	Yes	Yes	Strong signal	Yes
	Bac109	Yes	Strong signal	Yes	Yes	Strong signal	Yes
	Act39	Yes	Medium signal	No	No	N/A	Yes
	Str78	No	Target bacteria not available	No	No	N/A	Inconclusive
	Bac88	No	Target bacteria not available	No	No	N/A	Inconclusive

¹ Signals were defined as weak, medium or strong based on a comparison with other signals on the same channel

² The mixed sample did not contain bacterial templates complementary to all probes, and some probes were therefore not expected to provide a signal in this sample.

³ Probes were evaluated against LAD based on the experimental setup used in this thesis. The probes found to be compatible showed true positive signals throughout testing, while the probes that were deemed non-compatible with LAD either provided no signal during testing, or false positive signals.

3.1.1 Probe validation on pure cultures of target bacteria

For a probe pair to be determined to be compatible with LAD it needed to fulfill the criteria for a satisfactory signal as defined under section 2.2.3 on page 22. For this to happen, the LP first needed to bind to its complementary DNA target and get extended with a ddCTP-Q. This labelled LP would then need to hybridize to its complementary fluorophore-labelled RP, which would result in quenching of fluorescence. During the melting curve analysis, the gradually rising temperature would lead to dissociation of the LP-RP duplex, resulting in increased fluorescence. This would be presented as a negative peak near the theorized T_m in

the dissociation curve from the channel corresponding to the fluorophore attached to the RP (HEX, ROX, CY5 or FAM). Dissociation curves from all four channels were provided by the qPCR software for all 22 samples (of which 20 contained template and two were negative controls) giving a total of 88 melting curves per run. The test was run on each pure culture sample at least three times, and due to this large number of melting curves, only a representative selection is presented in this thesis. However, the results for all the probes from the tests on the pure cultures are listed in table 2 on the previous page.

The negative controls confirmed that none of the probes led to a signal in the absence of bacterial template. The UNI probe, which had been included in all tests as a positive control, confirmed that bacteria were present in all samples except the negative controls, as all the samples containing DNA template had a negative peak at roughly 65 °C on the melting curve from the HEX channel (demonstrated in figure 5 on the next page and in figure S1 in the supplementary material).

Figure 5 presents two representative examples of melting curves achieved from testing on pure cultures. In figure 5a a melting curve from the HEX channel for a sample containing *Eubacterium hallii* is shown, where two signals can be seen. The first signal at roughly 40 °C was from probe Eub206, which was designed to detect *E. hallii*. The second peak at 65 °C was from the UNI probe. No peaks were seen on the melting curves from the ROX, CY5 and FAM channels for this sample (not included in this thesis), as none of the probes carrying those fluorophores were complementary to the 16S rRNA gene of *E. hallii*.

Figure 5b shows a melting curve collected from the ROX channel from a sample containing *Eubacterium rectale*, and demonstrates a signal from the complementary probe Eub119 at 70 °C. This sample also showed a peak on the melting curve from the HEX channel at 65 °C due to the UNI probe, but no peaks on the melting curves collected from the CY5 and FAM channels, as no probes on those channels were complementary to the *E. rectale* template (melting curves not shown).

For information on the probes and their corresponding targets, fluorophore and T_m, see table 1 on page 19.

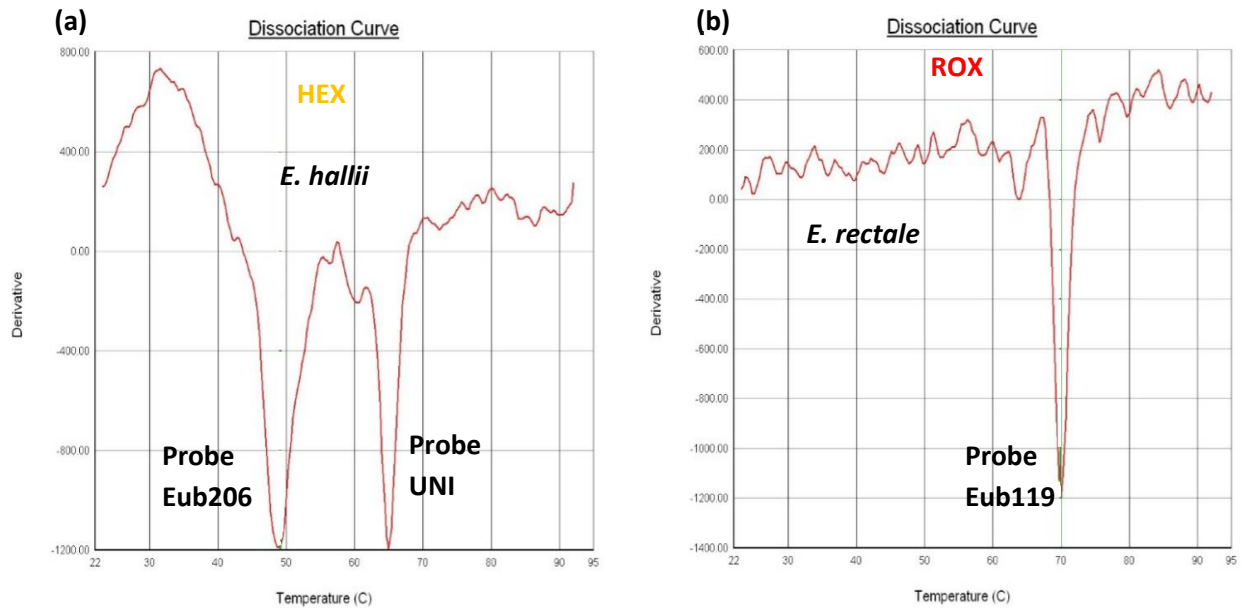


Figure 5: Two representative examples of melting curves from samples containing pure bacterial cultures. The temperature is presented on the horizontal axis, starting at 22 °C and ending at 95 °C. The negative derivative of the fluorescence and temperature measurements ($-dF/dT$) is on the vertical axis, ranging from -1200 to 800 on the HEX channel (a), and from -1400 to 600 on the ROX channel (b). The melting curve from the HEX channel is collected from a sample containing *Eubacterium hallii*, which is targeted by probe Eub206 (seen as a negative peak just below 50 °C). The signal for the UNI probe can be seen at roughly 65 °C. The melting curve from the ROX channel is from a sample containing *Eubacterium rectale*, which is targeted by probe Eub119, shown as a peak at 70 °C.

After the tests on pure cultures had been completed, 15 of the 20 LPs investigated in this thesis were determined to be compatible with LAD under the current experimental setup (table 2). Representative melting curves demonstrating true positive signals for all 15 probes can be seen in figure S1, S2 and S3 in the supplementary material. The five probes that did not provide satisfactory signals when tested with LAD were Bac230, Bac155, Dor124, Str78 and Bac88 (see table 1 for more information on these probes).

3.1.2 Investigation of probes that did not provide satisfactory signals

Probe Bac230 (designed to detect *Bacteroides fragilis*), never provided a signal in response to its target bacterium. However, it gave a weak false positive signal in response to *Lactobacillus ruminis*, for which it was not intended, each time the test was performed (demonstrated in figure S1 in the supplementary material). Due to this false positive signal, the sequence of the Bac230 LP was checked against reference sequences of *L. ruminis*' 16S rRNA gene using the Basic Alignment Search Tool (BLAST, National Center for Biotechnology Information (NCBI), Bethesda, Maryland, USA). This confirmed that the sequences were not complementary. As the probe never gave a signal for its target bacterium, the sample containing *B. fragilis* was sent for sequencing. This confirmed that the template that had been used for testing was a perfect match to the Bac230 LP, and thus provided no information on the lack of true positive signal. The *L. ruminis* sample was not sent for sequencing because probe Lac47, which was designed to detect *L. ruminis*, provided a satisfactory signal in all tests performed. It is therefore unknown whether the sample contained any contamination that Bac230 could be reacting to.

The bacterial templates used to test the probes Bac155 and Dor124 were sent for sequencing, as no signal was detected for these two probes in any of the tests. The results from this confirmed that the templates used were perfectly complementary to the LPs, and therefore provided no information on the lack of detected signals.

Self-dimerization and hairpin structures were found to be highly unlikely for the three probes Bac230, Bac155 and Dor124, as all possible formations had a near zero or positive delta G. The probes themselves were not sequenced, and it is therefore unknown if they contained any unexpected errors.

Several bacterial species were used to test the probes Bac88 and Str78, but neither of them provided a detectable signal during testing. However, this was expected, because none of the bacterial species used were a perfect match to the LPs, and were only chosen because fully complementary species were not available during the time of testing.

3.2 Validation of LAD on a mixed sample

Following the probe validation on pure cultures, LAD was performed on a mixed sample containing 11 bacterial templates whose 16S rRNA gene sequences were complementary to 12 of the 15 probes (one of which was the UNI probe) that had been determined to be compatible with LAD. This was to ensure that the technology could provide distinguishable signals in more complex bacterial communities as well. The results from this test confirmed that LAD showed the expected satisfactory signals (based on the criteria listed under section 2.2.3 on page 22) for the probes that were complementary to the bacteria included in the mix, and no signals for the probes whose target templates were not included. The exception to this was probe Bac230, which gave a false positive signal in response to *L. ruminis*.

As the mixed sample contained several bacterial templates, the melting curves from each channel presented a baseline with several peaks. These peaks were found near the theoretical Tms of all the probes designed for that channel that were complementary to the species included in the mix. This is demonstrated in figure 6 on the following page, where representative examples of melting curves from all four channels collected from a test performed on the mixed sample are presented.

On the HEX channel shown in figure 6a, signals can be seen for probes Fir224, Eub206, Ali239 and UNI. These signals were all expected based on the species included in the mixed sample. In addition, a false positive signal from Bac230 in response to *L. ruminis* can be seen at 30 °C. The ROX channel is presented in 6b, where probes Str138, Cat187 and Eub119 demonstrate peaks near their designed theoretical Tms, all of which were expected.

On the CY5 channel, shown in figure 6c, probes Pse30, Lac47 and Aci57 can all be seen to provide a signal, as expected based on the templates used in the mix. The baseline on this channel can be seen to decline with increasing temperature, and this was observed in all tests performed (also demonstrated in figure S3a in the supplementary material).

Figure 6d shows the FAM channel, where probes Par89 and Bac109 both show deep signals in response to the mixed sample, as was expected.

For an overview of the fluorophore, theoretical Tm and bacterial target for each probe pair, see table 1 on page 19. For a list of the probes expected to provide a signal in the mixed sample, see table 2 on page 26.

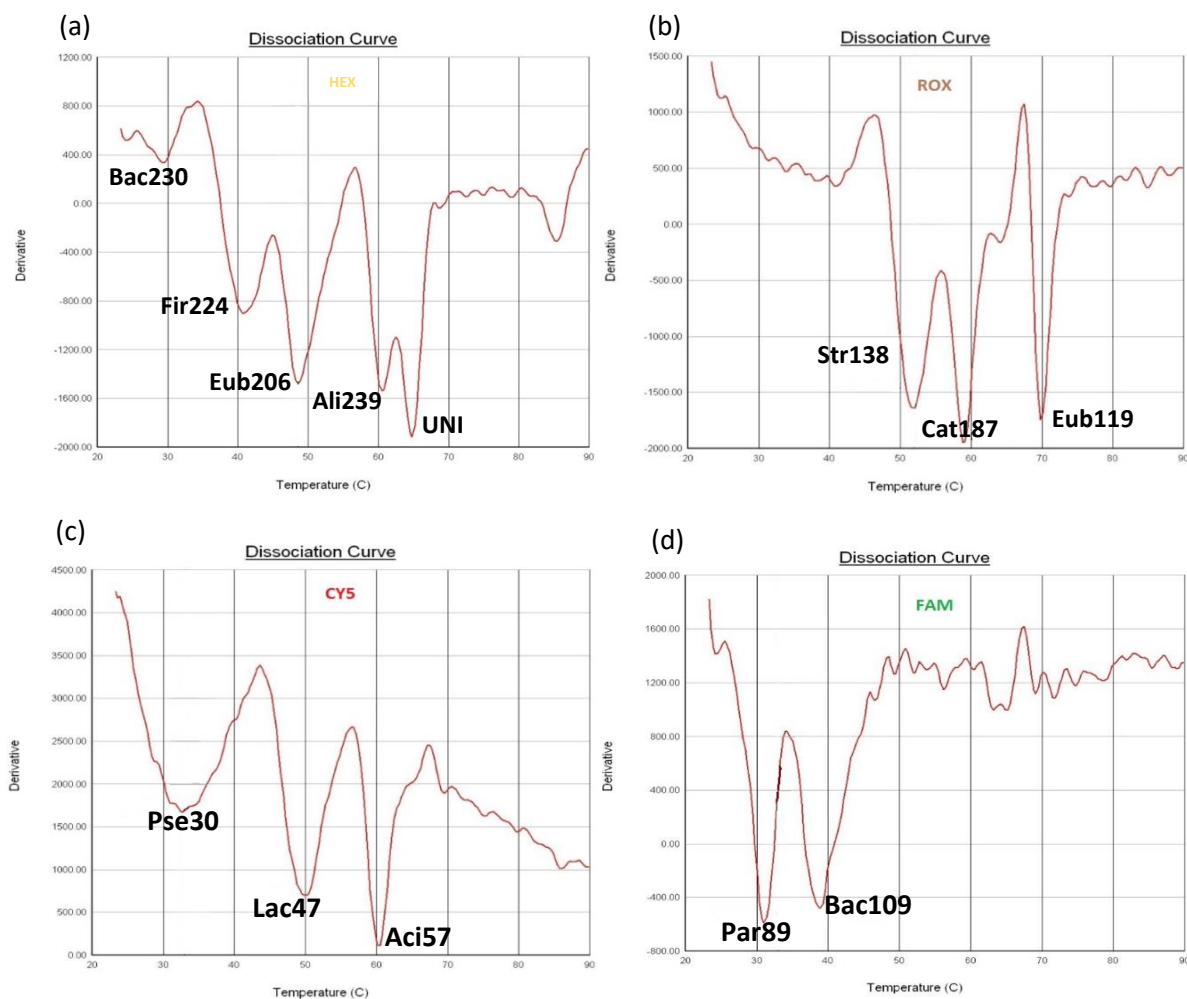


Figure 6: Representative melting curves from a test performed on a mixed sample containing pure DNA templates from 11 bacterial species. The temperature is presented on the horizontal axis and goes from 20 °C to 90 °C for all four melting curves. The negative first derivative of the fluorescence and temperature measurements ($-dF/dT$) is on the vertical axis, and goes from -2000 to 1200 on the HEX channel (a), from -2000 to 1500 on the ROX channel (b), from 0 to 4500 on the CY5 channel (c), and from -800 to 2000 on the FAM channel (d). The signal seen for probe Bac230 on the HEX channel was a false positive signal. In addition, a contaminating signal can be seen on this channel at around 85 °C. This is outside the temperature range used for this thesis. All other observed peaks in the figure were true positive signals.

3.3 Evaluation of LAD for diagnosis of dysbiosis

As an aim of this thesis was to investigate if LAD can be used to correctly define a sample as normobiotic or dysbiotic, the entire LAD process was performed on fecal samples collected from 32 individuals. These samples had previously been mapped by GA according to their dysbiosis index, and of the 32 samples, 15 were classified by GA as normobiotic and 17 as dysbiotic. After LAD had been completed, the UNI probe confirmed that all fecal samples contained bacterial DNA.

3.3.1 Probe signals in normobiotic versus dysbiotic samples

The extracted peak values collected from the test on the fecal samples were first used to examine the differences in the average signal detected for each probe in the normobiotic samples compared to the dysbiotic samples. In addition, a Welch Two Sample t-test was performed to investigate if the differences observed were statistically significant, indicated by a p-value < 0.05. The signal comparisons are visualized in figure 7 on the following page, with the statistical significance indicated by asterisks. In the figure, a strong signal (given as low magnitude) indicates a higher bacterial abundance as more LPs have been labelled.

The figure shows that the average detected signals for probes Fir224, Eub206, Ali239, Str138, Cat187, Eub119, Lac71, Lac47, Aci57 and Bac91 were stronger in the normobiotic samples than in the dysbiotic ones. The dysbiotic samples show on average a stronger signal for probes Pse30, Par89, Bac109 and Act39. The results from the Welch Two Sample t-test shows that the increased signal strengths detected for probes Ali239 and Cat187 in the normobiotic samples were statistically significant (p-value = 0.02664 and p-value = 0.02263 respectively). For the remaining probes, observed differences in signal strength were not considered statistically significant with respect to normobiosis and dysbiosis ($p > 0.05$). The p-values for all probes can be seen in table S3 in the supplementary material.

The columns for the probes carrying a CY5 fluorophore (figure 7c) show a similar trend as the melting curves from the CY5 channel in figures 6c and S3a, with a decline in signal depth from probe Pse30 (designed for a theoretical T_m of 29.5 °C), to Bac91 (designed for a theoretical T_m of 70 °C).

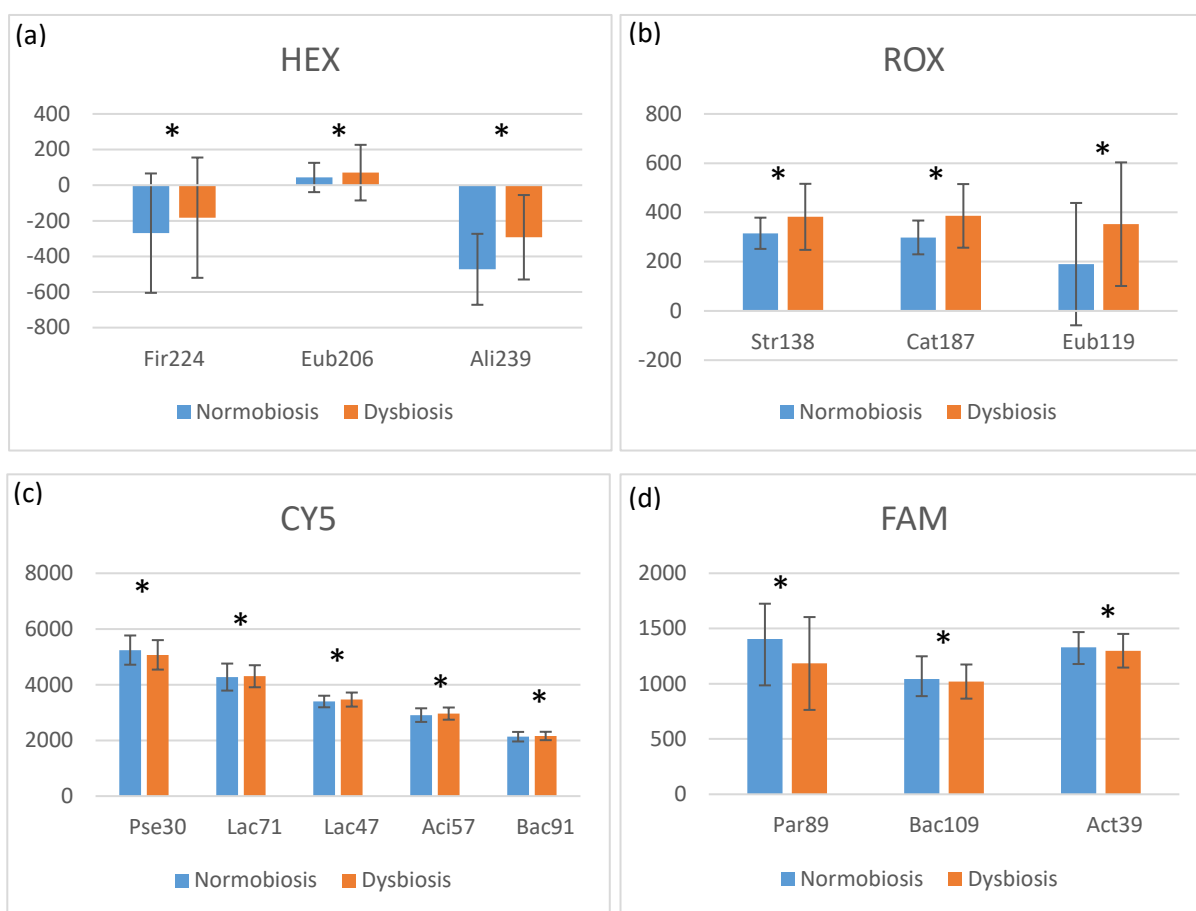


Figure 7: Presentation of the average detected signals with standard deviation for each probe from the normobiotic and dysbiotic samples. The signals from the normobiotic samples are shown by the blue columns, while the signals detected from the dysbiotic samples are shown by the orange columns. The probe signals are the negative first derivative of the fluorescence and temperature measurements ($-dF/dT$) and are shown by the y-axis. As the signal is detected as low magnitude, a lower number on the y-axis indicates a stronger signal. The probes Fir224, Eub206, Ali239, Str138, Cat187, Eub119, Lac71, Lac47, Aci57, and Bac91 show a stronger signal in the normobiotic samples than in the dysbiotic samples, while for probes Pse30, Par89, Bac109 and Act39, the dysbiotic samples display stronger signals than the normobiotic samples. The statistical significance of the differences in average probe signals between normobiotic and dysbiotic samples is shown by asterisks, where one asterisk represents a non-significant p-value ($p > 0.05$), while two asterisks represent a significant p-value ($p < 0.05$).

3.3.2 ROC curve analysis

The ROC curves investigating the accuracy of LAD as a diagnostic tool for dysbiosis is presented in figure 8 on the following page. In this plot, the blue curve represents the theoretical best-case scenario based on this dataset, while the green curve shows the cross-validated model, which indicates the realistic diagnostic performance of the LAD technology with the current probes.

The red rings on the curves show the optimal results with regards to sensitivity and specificity as determined by the MatLab software, and indicate the proportion of each group being tested that get classified correctly. For the blue curve sensitivity of the red ring is just below 0.9 (or 90%), as seen from the y-axis. Specificity is read from the x-axis as 1-specificity, and in the figure this can be found at 0.1 for the red ring, which translates into a specificity of 90%. The best-case scenario for the entire dataset from this thesis can therefore be seen to have both sensitivity and specificity at around 90%. This means that roughly 90% of the dysbiotic samples (seen by the sensitivity) will be classified correctly as dysbiotic, while about 10% of the dysbiotic samples get mistaken for healthy. Similarly, 90% of normobiotic samples (seen by the specificity) will be correctly classified as normobiotic, while around 10% will get wrongly diagnosed as dysbiotic according to the ideal classifier.

For the cross-validated model represented by the green curve, the red ring indicates that the realistic optimal result has a sensitivity a little under 0.8, indicating that around 80% of samples that are dysbiotic will be classified correctly. This means that a little more than 20% of dysbiotic individuals will be wrongly classified as healthy using the LAD technology with the current probeset. Of normobiotic individuals, the red circle indicates that about 75% will be correctly classified as healthy. Around 25% of the normobiotic samples will be classified wrongly as dysbiotic, as demonstrated by the specificity of the red ring which is found at roughly 0.25 on the x-axis.

The area under the curve (AUC) shows the accuracy of classification for all the samples, and indicates how well LAD can distinguish between dysbiosis and normobiosis when using the current probeset. The blue curve has an AUC of 0.9373 (94%), showing that an ideal result from this probeset would mean that 94% of all samples were classified correctly. The cross-validated green line has a value of 0.8157 (82%), indicating that the realistic result using the LAD method with the current probeset is that 82% of all samples investigated will get classified correctly as either dysbiotic or normobiotic, while 18% will be incorrectly classified.

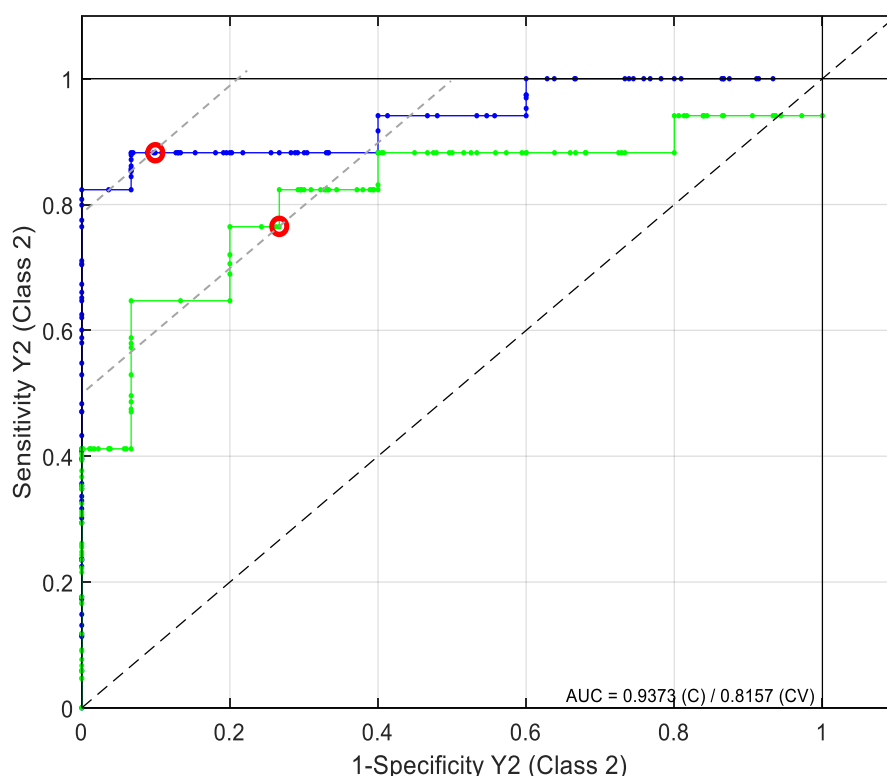


Figure 8: ROC plot showing the classification accuracy of LAD when using the probe subset tested in this thesis.

The blue curve represents the estimated best-case scenario, while the green curve shows the cross-validated model. Sensitivity can be read from the vertical axis, while specificity is read as 1-specificity on the horizontal axis. The red ring on the blue line informs what the optimal result for the dataset is, where both sensitivity and specificity are at around 0.9. The red circle on the green line shows the optimal realistic result, where a sensitivity was achieved of just below 0.8, while specificity can be found at roughly 0.75. These optimal points give the best trade-offs between correctly classifying a sample as dysbiotic, against the cost of wrongly diagnosing a normobiotic sample as dysbiotic. The location of the ring is based on a straight line whose gradient is dependent on previously determined accepted thresholds for sensitivity and specificity. When the cost of misclassifying normobiotic and dysbiotic samples is the same, and it is assumed that there is roughly the same number of healthy and sick samples, this line has a slope of 1, making it angled at 45 degrees. In the figure, this is presented as the two grey dashed lines going through the red circles. When using such a line, the optimal result can be found at the point where the ROC curve and the dashed line cross closest to the top left corner of the figure.

3.3.3 Association of probes to dysbiosis or normobiosis

The association of individual probes to either dysbiosis or normobiosis is presented in figure 9, and indicates which probes in the subset best separate the two states. In this figure, a positive value on the vertical axis means the probe is associated with dysbiosis, while a negative value means it is associated with normobiosis.

The figure shows that for dysbiosis, probe Eub206 is the most strongly associated of the subset, closely followed by Par89, Bac109 and Act39. The final probe seen associated with dysbiosis is Fir224 with a regression vector a little under 0.2.

Of the probes associated with normobiosis, Eub119 has the strongest association, closely followed by Lac47. Ali239 has a regression vector value of about -0.3, followed by Aci57 and Cat 187 near -0.2. Bac91 can be found at -0.15, while Str138 is at just below -0.1. The probes Lac71 and Pse30 are the probes with the least association with normobiosis at just below zero.

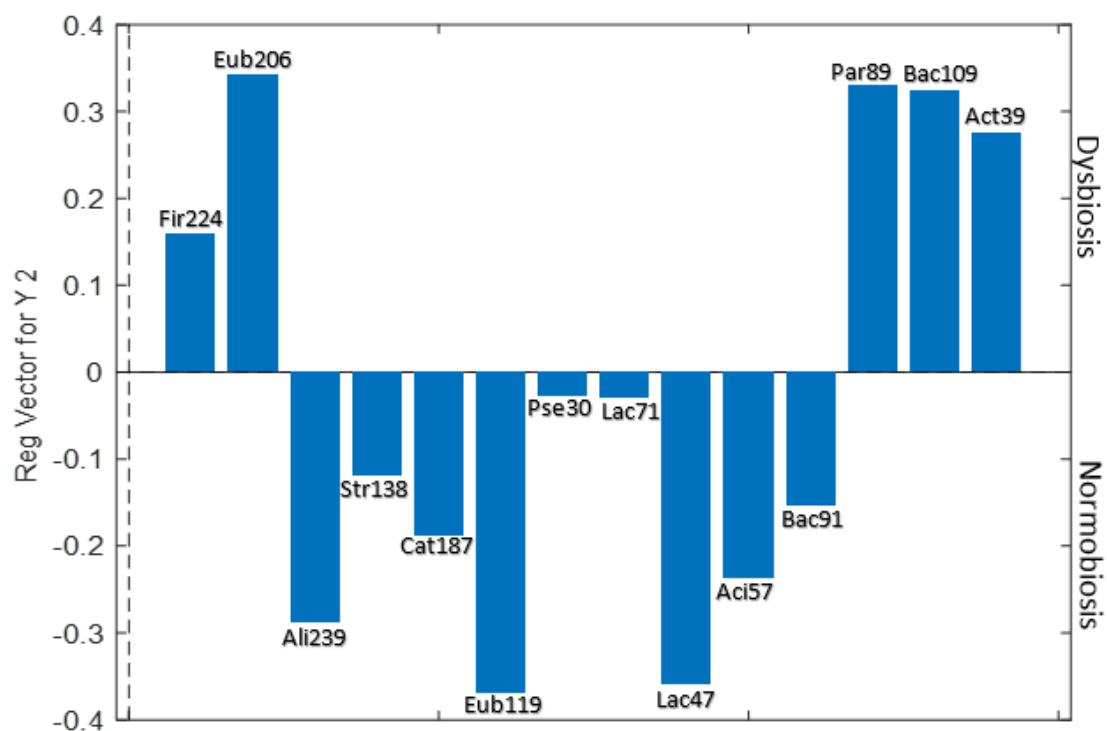


Figure 9: Association of individual probes to either dysbiosis or normobiosis. A positive value on the vertical axis, presented as the regression vector, means the probe is associated with dysbiosis. A negative value means it is associated with normobiosis. The probe Eub206 is the most strongly associated probe for dysbiosis in this probe subset, while Eub119 is the most associated to normobiosis. Fir224 is the least associated for dysbiosis of the dysbiosis-associated probes, while Pse30 is the least associated to normobiosis, followed closely by Lac71.

4. Discussion

4.1 The diagnostic performance of LAD

The aim of this thesis was to investigate whether the Liquid Array Diagnostics technology can detect target bacteria and differentiate between dysbiosis and normobiosis sufficiently to be a valuable supplement to the GA-map® Dysbiosis Test. As it was assumed that LAD could detect 20 targets simultaneously, a subset of 20 labelling probes designed by Genetic Analysis AS was used for the study. These probes were first tested with LAD on pure cultures of target bacteria to ensure they could provide true positive signals when used with this technology. In addition, LAD was performed on a mixed sample to confirm that probe signals would also be distinguishable in more complex bacterial communities. At the end of the thesis 15 of the 20 probes were deemed to be compatible with LAD under the current experimental setup.

After probe compatibility had been established, LAD was performed on previously classified fecal samples to investigate how well the technology could distinguish between dysbiosis and normobiosis. Using a ROC curve analysis (figure 8) it was estimated that LAD had an overall accuracy of 82% for diagnosing patients correctly. The analysis also indicated that a slightly larger proportion of dysbiotic samples would be classified correctly compared to the normobiotic ones.

As five probes did not provide distinguishable signals in this thesis, and the UNI probe was used as a positive control, the evaluation of LAD's diagnostic ability was based on the remaining 14 probes that gave true positive signals to their target bacteria. Due to the complexity of factors influencing dysbiosis, it is assumed that using 20 compatible probes would increase the accuracy of the test. In addition, the probes investigated are the least associated with dysbiosis or normobiosis of GA's complete probeset, and it is expected that probes that have a stronger association would achieve an even higher accuracy using LAD.

Of the 14 probes used to evaluate LAD's diagnostic performance, five were associated with dysbiosis in this thesis, while nine were associated with normobiosis (figure 9). To examine the validity of these associations, investigations were made into previous studies on the topic, the biology of the bacteria targeted by the probes, and the relationships between the bacteria and the host.

4.1.1 Probes associated with dysbiosis

The probes that were identified as associated with dysbiosis in this study were Fir224, Eub206, Par89, Bac109 and Act39 (figure 9).

Probe Fir224 was designed to detect species within the Firmicutes phylum. This is one of the most dominant phyla in the gut, and studies have found that dysbiosis often involves a shift within this phylum. For example, as was mentioned in the introduction of this thesis type 2 diabetes and obesity are both associated with the abundance of Firmicutes species growing too large compared to healthy microbiota profiles (1, 17, 22, 23). On the other hand, a decrease in the abundance of certain Firmicutes species is frequently found in individuals suffering from IBD. An example of this is in CD patients, where a reduction is often seen in *Faecalibacterium prausnitzii*, one of several bacteria within the phylum that has anti-inflammatory effects (20, 22, 23, 28, 38, 39, 61, 62, 63, 64). Such decreases in Firmicutes species is, together with a reduction in functional diversity, a strong indicator of IBD (22, 28, 38, 61, 62). This agrees with the data presented in figure 7 in this thesis which suggests that on average the dysbiotic samples, which were collected from individuals suffering from IBS and IBD, had a lower abundance of Firmicutes species than the normobiotic samples. Based on the above, it is as expected that probe Fir224 was associated with dysbiosis in this thesis. This finding also agrees with a previous study on the topic which found that for the IBS and IBD cohorts investigated the bacteria that contributed the most to dysbiosis were within the Firmicutes phylum (20).

The probe Eub206 targets *Eubacterium hallii*, which is a species within the Firmicutes phylum that is commonly found in human feces. This bacterium is important to the intestinal metabolic balance due to its ability to utilize glucose, acetate and lactate to produce the short-chain fatty acid (SCFA) butyrate. Its use of lactate in this production is particularly advantageous, because an accumulation of this intermediate has been linked to several diseases in humans (3, 6, 22, 24, 65, 66, 67, 68).

Butyrate, along with some other SCFAs, is believed to have anti-inflammatory properties (6, 22, 24). In addition, butyrate is a source of energy for epithelial colonic cells, it participates in the proliferation and apoptosis of said cells, and it is seen as an important contributor to the prevention of colon cancer. Epithelial cells protect the host against invading microorganisms and their toxins, and a disturbance in the production of butyrate can therefore lead to a

poorer immune response. Butyrate also increases the growth of species within the *Lactobacillus* genus, some of which are important for a healthy gut (5, 22, 24, 67, 69, 70, 71). Studies have found that butyrate-producing bacteria such as *E. hallii* are often depleted in IBS and IBD patients, particularly in CD patients, compared to healthy subjects. A low abundance of *E. hallii* has also been linked to obesity and type 2 diabetes, and there is some indication that a decrease in butyrate producers is associated with the pathogenesis of multiple sclerosis (MS), a pro-inflammatory disease that affects the central nervous system (6, 22, 64, 65, 66, 67, 72, 73, 74).

Along with butyrate, *E. hallii* produces H_2 , which normally gets efficiently consumed by methanogenic and sulfate-reducing microorganisms in the gut (3, 6, 66, 75). H_2 has been found to neutralize certain toxic hydroxyl radicals, downregulate expression of proinflammatory factors, and maintain cerebrovascular reactivity. In addition, hydrogen sulfide, which is produced from H_2 by the sulfate-reducers, helps modulate smooth muscle and neuronal functions, protects cells from harmful agents, and can function as an anti-inflammatory (3, 5, 76). A decrease in the production of H_2 can therefore have a negative effect on human health. For example, a recent study theorized that such a reduction could be linked to Parkinson's disease (PD), a condition that is often accompanied by dysbiosis (77).

Based on the above information, the recognition in this thesis of Eub206 as a probe associated with dysbiosis could suggest that in the dysbiotic samples investigated, the abundance of the bacterium was lowered compared to the normobiotic samples. This also agrees with the findings presented in figure 7, which indicate an average lower amount of *E. hallii* in the dysbiotic samples compared to the non-dysbiotic ones. However, as can be seen from the raw data presented in table S2 in the supplementary material, some of the dysbiotic samples display relatively strong signals for Eub206, and the strongest detected signal for this bacterium comes from a DI3 sample. Such an elevated amount of *E. hallii* could also be linked to dysbiosis, as it could lead to an abundance of H_2 , which in turn could lead to an overproduction of hydrogen sulfide and methane. While hydrogen sulfide has been found to be beneficial to human health, it can also lead to stomach pain and have toxic effects if the concentration grows too high. Methane has been linked to both constipation and constipation-predominant IBS, likely because it causes a delay in the movement of food in the small intestine. In diseases that include symptomatic diarrhea, like CD or UC,

the production of methane has generally been found to be low (3, 6, 28, 66, 75, 76). Hydrogen has been suspected to increase the small intestinal motility, but as it is easily consumed, its effect on this is usually minor (3, 6, 28, 66, 75, 76). However, there is a possibility that a significant increase in H₂ production beyond what the microbiota could efficiently consume could be associated with diarrheal diseases. In addition, hydrogen removal is needed for fermentation to remain efficient. It is also worth mentioning that intestinal gases, such as H₂, are frequently associated with abdominal pain, bloating, and changes in bowel movement, all of which are symptoms of IBS, further underlining *E. hallii*'s link to dysbiosis (5, 6, 75).

A disturbance is frequently seen in the abundance of species within the Bacteroidetes phylum in individuals suffering from dysbiosis-related disorders compared to healthy subjects (20, 22, 28, 61, 62, 74). This agrees with the finding in this thesis that probes Par89 and Bac109 were associated with dysbiosis, as they target *Parabacteroides johnsonii* and *Bacteroides pectinophilus* respectively, which are both species within this phylum.

It has been hypothesized that MS is linked with dysbiosis, and investigations of fecal samples collected from patients suffering from this disease have revealed a decrease within the *Parabacteroides* genus compared to healthy subjects (72, 73, 74). A lower abundance of this genus has also been found in patients suffering from chronic hepatitis B, along with a decrease in the *Bacteroides* genus (78).

Many members of the *Bacteroides* genus ferment pectin and other plant cell wall polysaccharides in our food, including *B. pectinophilus* which is frequently isolated from human feces. Degraded pectin is absorbed and used for energy by the host or bacteria in the colon, and both pectin and its fermentation products help develop and maintain a healthy microbial balance in the gut (62, 79, 80). It has been found that pectin can positively affect the success of a fecal microbiota transplantation (FMT), which is sometimes used as a treatment for patients suffering from UC, while also delaying the deterioration of diversity in the gut (5, 62, 80). In addition, pectin participates in modifying blood glucose and insulin response, and has been found to be useful in the treatment of diabetes and other carbohydrate-related disorders (62, 80).

One of the main products from pectin degradation, acetate, is used as an energy source both for bacteria and for the colon epithelial cells. In addition, it acts as a substrate in cholesterol- and lipid biosynthesis, has anti-inflammatory effects, and is used by some bacteria to inhibit enteropathogens (1, 24, 67). Acetate is also used in the synthesis of butyrate, and a decrease in pectin-degradation can therefore lead to a decrease in the production of butyrate, which, along with a general decrease in microbial diversity, is a key trait in dysbiosis (6, 80).

Based on the above information, it could be assumed that since Par89 and Bac109 were associated with dysbiosis in this thesis, the abundances of the bacteria detected by these two probes were lower in the dysbiotic samples than in the normobiotic ones. This is contradictory to the findings presented in figure 7, which indicate that both of these probes on average detected a stronger signal (and thus theoretically a higher abundance) in the dysbiotic samples. However, for Bac109 this difference was minor. For Par89 it could be theorized that the stronger signal detected in the dysbiotic samples was because those samples came from IBS and IBD patients, while the cases of reduced abundance described above was in relation to MS and chronic hepatitis B.

The last probe that was associated with dysbiosis in this thesis was Act39, which targets the order Actinomycetales within the Actinobacteria phylum. This association agrees with a study performed by Casén et al. (20), which found that members of the Actinobacteria phylum were some of the predominant bacteria linked to dysbiosis within their IBS cohort, displaying a stronger signal in the dysbiotic samples than in the non-dysbiotic ones. The results presented in figure 7 in this thesis indicate the same, although the difference in average signal between the dysbiotic and normobiotic samples is small. This could be because the dysbiotic cohort in this thesis consisted of both IBS and IBD samples, rather than purely IBS.

A change in Actinomycetales abundance has been recorded in several dysbiosis-related cases. For example, taking proton pump inhibitors to reduce the production of stomach acid can lead to a decrease in bacterial diversity and to moderate dysbiosis, and in such cases an increase in the Actinomycetales order is often seen. An increase has also been found in obese individuals compared to lean ones (23, 81, 82, 83).

In a study by Valeur et al. (58) it was found that IBS patients that responded well to a low FODMAP diet (Fermentable oligosaccharides, disaccharides, monosaccharides and polyols. For more information on the diet, see section 4.3.1 on page 56) had lower levels of Actinomycetales species than the patients that did not respond well. However, whether this is an indication of its link to dysbiosis, or simply a trait that affects response to this specific diet, is not clear.

Based on the above information, it could seem that an increased presence of the order Actinomycetales in the human gut is negative. However, the order contains many members, several of which are frequently found on mucosal surfaces, on skin, and in the human gut. Some of its members produce antiproliferative and immunosuppressant, and it has been theorized that this enables them to prevent allergies, autoimmunity and IBD because they suppress both excessive proliferation of certain microorganisms and local immune responses (1, 84). Hence, while certain trends can be seen with regards to specific bacterial orders and dysbiosis, the above serves as a demonstration of how methods that can distinguish down to species level is necessary to truly discover the underlying patterns of a disturbed gut.

4.1.2 Probes associated with normobiosis

The probes shown to be associated with normobiosis in this thesis were Aci57, Pse30, Cat187, Ali239, Bac91, Eub119, Lac71, Lac47, and Str138 (figure 9).

Probe Aci57 targets the bacterium *Acinetobacter junii*, which is an opportunistic pathogen that sometimes causes infections in individuals with a compromised immune system (1, 73, 85). According to Pepper et al. (85), up to 25% of adults carry this bacterium in their respiratory tract and on their skin, and during outbreaks in hospitals, the source is often human skin, sinks or taps (73, 85).

A research group recently found in a study that the abundance of the *Acinetobacter* genus was much higher in MS patients than in healthy individuals, indicating a negative association to the genus. However, the nature of this connection remains to be verified, and it is unclear whether *A. junii* was involved in this, as the study was unable to distinguish at species level (73). A positive association to the genus has also been observed, as a study on patient response to a low FODMAP diet found that individuals with higher levels of *Acinetobacter* species responded well to the diet (58).

Figure 7 in this thesis indicates that the normobiotic samples on average contained a slightly higher abundance of *A. junii* than the dysbiotic ones, which could have led the probe to be associated with normobiosis. However, this difference in probe signals was relatively small. *Acinetobacter* species are generally rare in the healthy human gut, and it may be suggested that probe Aci57 is not consistently associated with either normobiosis or dysbiosis (73, 85). Instead it could be a probe to use in certain situations, for example when there is suspicion of MS development, for investigation of potential contamination during sampling (e.g. from human skin), to check potential response to a low FODMAP diet, or to check for the opportunistic pathogen.

The same might apply to probe Pse30, which showed only a weak association to normobiosis in this thesis. This probe targets *Pseudomonas* spp., a genus that is very prevalent in nature and can be found in for example snow, rain, lakes, plants, soil and animals (1, 86). Humans frequently get exposed to members of the genus from for example swimming pools, hot tubs, and taps (5, 72, 85). As for the *Acinetobacter* genus, a higher abundance of the *Pseudomonas* genus has been found in patients that have MS, and the probe might therefore be used to investigate MS development. In addition, the genus contains several opportunistic pathogens (5, 72, 85). Based on this, it may be that in the samples tested, the bacteria targeted by Pse30 had little influence on normobiosis, leading to only a weak association. In addition, because the genus is so common due to its widespread habitations, it could have been detected almost equally in both the normobiotic and dysbiotic samples. This theory is supported by figure 7, in which the average signals detected for probe Pse30 are of similar strength for dysbiosis and normobiosis.

Probe Cat187 targets *Catenibacterium mitsuokai*, a bacterial species belonging to the class Erysipelotrichia which is found frequently in the gut (87). The species performs fermentation, producing lactic, acetic, butyric and isobutyric acids (88).

Acetic and butyric acids are among the most abundant SCFAs in the gut. As has been mentioned earlier in this thesis, acetate, derived from acetic acid, helps reduce inflammation, inhibits pathogens, and can be used as an energy source. In addition, it is involved in the production of butyrate and lipids. Butyrate was discussed during the investigation of *E. hallii*, and in brief functions as an anti-inflammatory, is very important for

epithelial cells, and promotes the growth of beneficial bacteria. Based on this, the association of probe Cat187 to normobiosis seems correct.

The gut microbiota is influenced by diet, and it has been found that in mice fed with a western diet, which usually consists of large amounts of animal protein and saturated fats, there is an elevation in the abundance of members of the Erysipelotrichia class, including *C. mitsuokai*. Other bacteria that have been seen to increase in individuals consuming such a diet are within the genera *Alistipes* and *Bacteroides*, covered by probes Ali239 and Bac91 respectively (1, 3, 22, 23, 80). This could perhaps partly explain why these three probes were seen in this thesis to be associated with normobiosis, even though a western diet is not necessarily associated with good health. The fecal samples used in this project were from Norway, and it may therefore be assumed that the normobiotic individuals most likely followed a western diet. Individuals suffering from dysbiosis-related illnesses often make changes to their diets, and may exclude certain food items, which will affect the bacterial composition (1, 17, 20, 22). This is further supported by figure 7 in this thesis, which suggests that there was on average a higher abundance of *C. mitsuokai*, *Alistipes* spp., and *Bacteroides* spp. in the normobiotic samples compared to the dysbiotic ones. This link to diet may also partly explain why the differences in average signals detected for probes Cat187 and Ali239 were found to be statistically significant in this thesis, as diet may be one of the most consistent differences between the normobiotic and dysbiotic samples.

In addition to being affected by a western diet, certain species within the *Alistipes* genus have been found to be shared among individuals in general, suggesting its importance in the human gut, and supporting the link of probe Ali239 to normobiosis (1, 24, 39). Its positive effect has been indicated in studies on cancer patients' response to immunotherapy, where it has been found that in the individuals responding well to the treatment, there was a higher abundance of *Alistipes* species, along with certain other bacteria (22, 89). In addition, it has been found that obese individuals who succeed in weight loss have a microbiota baseline which is elevated in members from this genus, although it is worth noting that an increase in the genus has generally been observed in obese individuals compared to lean ones (90, 91). Species within the *Alistipes* genus have also been found to be more abundant in healthy individuals compared to people suffering from non-alcoholic liver diseases, chronic hepatitis B and HIV (78, 91, 92, 93).

On the other hand, some studies have observed adverse effects in relation to elevated amounts of the *Alistipes* genus. For example, a high abundance of certain of its species has been found to be correlated with an increase in abdominal pain in children with IBS, and to major depressive disorder (36, 90, 94). Therefore, one could also argue that the probe could be associated with dysbiosis. However, all samples used in this project were collected from adults, and seeing as the genus is shared among individuals, especially those that follow a western diet, its link to normobiosis seems correct for the samples investigated.

Additionally, both in the cases of obesity and major depressive disorder, it was found that a high abundance of the genus could be connected to a better response to treatment, thus presenting the question of whether the elevated presence of the bacteria is in fact negative or positive (36, 90, 91). Indeed, it is a good example of the complexity of the gut microbiota and the roles of the microbes.

The *Bacteroides* genus, which is one of the most common genera found in the large intestine, contains several species that have been found to be shared among individuals (1, 39). According to Willey et al. (1) about 30% of the bacteria isolated from human feces belong to this genus, and the association of probe Bac91 (which targets *Bacteroides* spp.) to normobiosis is therefore as expected.

The probe was first tested on one species from the *Bacteroides* genus, for which it gave no signal. By checking the sequence of the LP against the 16S rRNA gene sequence of the bacterium used, it was revealed that the probe had TCC at the 3' end, while the bacterial sequence had TTC, meaning the reverse complementary strand had AAG. In theory, the probe could get extended with a ddCTP-Q, as the complementary bacterial strand had a G in its 5' end. The lack of signal could therefore be due to the mismatch in the penultimate position. The second species within the *Bacteroides* spp. that was tested with Bac91 did not have this penultimate mismatch and provided a satisfactory signal. This is therefore an example of the specificity of LAD. It could be interesting to investigate further how penultimate mismatches are linked to the specificity of LAD, and potentially to test other polymerases to see if the issue persists.

If a goal is to detect more species within the genus with probe Bac91, changes to the protocol could be investigated. For example, lowering the temperature of the annealing step

during labelling may enable the probes to bind to more templates despite a penultimate mismatch. However, this can also lead to more incidents of nonspecific binding.

Eubacterium rectale is one of the most dominant species in the colon, and is an important bacterium for butyrate production and for synthesis of the biologically active form of vitamin B6 called Pyridoxal 5'-phosphate, or PLP. This active form of the vitamin is involved in the metabolism and synthesis of lipids, proteins, and carbohydrates, in addition to being important for the immune system (24, 61).

People suffering from IBD and IBS are frequently found to have reduced levels of butyrate-producers, including *E. rectale*, compared to healthy subjects, and several studies have found that healthy individuals have up to a 10-fold higher abundance of this bacterium (6, 22, 24, 61, 63). This agrees with figure 7 in this thesis, which indicates that on average the normobiotic samples had a higher abundance of *E. rectale* compared to the dysbiotic ones. The strong association of probe Eub119, targeting *E. rectale*, to normobiosis is therefore as expected.

Probes Lac71 and Lac47 both target members of the *Lactobacillus* genus that are normal constituents of the human microbiota (1, 24, 68). Probe Lac71 covers *Lactobacillus* spp., and the genus contains species that reduce the symptoms of IBD, alleviate certain obesity-related complications linked to metabolism, protect against pathogens, have anti-inflammatory effects, and are suspected of having anti-cancer activities (22, 68, 83). In addition, several species are involved in the production of SCFAs, and of vitamins such as vitamin K, B12 and B9 (also known as folic acid) (22, 61, 68, 83, 95). Vitamin B9 is important for DNA repair, cell division and cell growth, and a deficiency in this vitamin has been seen in patients suffering from CD (61, 95).

Probe Lac47 targets *Lactobacillus ruminis*, and the genome of this bacterium has revealed that while it is unable to produce folic acid itself, it has the genes needed for chorismite biosynthesis, which is a precursor to folic acid. A hypothesis is that *L. ruminis* contributes to the production of vitamin B9 by secreting chorismite to the environment, allowing for other bacteria in the gut to use that to synthesize folic acid (68). Lac47 also targets *Pediococcus acidilactici*, which is often involved in food fermentation products like sausages (1). Various members of the *Lactobacillus* genus are also frequently used in food, such as in yoghurts, fermented vegetable foods, sourdough bread and more. In addition, due to their beneficial

role in the human gut, several *Lactobacillus* species are commonly used in probiotics (1, 5, 22, 68).

A reduction in abundance of *Lactobacillus* species is often observed in cases of IBD and IBS compared to in healthy individuals, and some species have been associated with obesity. In addition, certain species have been found to express receptors for vitamin B12, which can lead to a decreased bioavailability of this vitamin for the host (5, 6, 22, 68). These negative associations are not immediately reflected in the results of this experiment, as the two probes used in this thesis that are directed at members of this genus, Lac71 and Lac47, have both been demonstrated as more strongly associated with normobiosis. However, this could be because the probes do not target, or only target a few of the *Lactobacillus* species that are linked to an imbalanced gut. The genus is part of the Firmicutes phylum, and the species associated with dysbiosis could possibly fall under the Fir224 probe. In addition, GA has designed another probe directed at *Lactobacillus* spp., and there is a possibility that this probe would have been associated with dysbiosis using LAD had it been tested in this thesis. Hence, as the bacteria targeted by Lac71 and Lac47 are normal members of the gut microbiota, are used as probiotics, and are frequently found in food products, an association to normobiosis is not surprising. In figure 9 in this thesis, probe Lac71 barely shows on the normobiotic side, suggesting that the species targeted by this probe are somewhat equally distributed among the samples. This is also reflected in figure 7, which indicates that there was only a slight increase in abundance in the normobiotic samples compared to the dysbiotic ones. This could perhaps be because the probe targets some species within the genus that are associated with normobiosis, and others that are associated with dysbiosis. In addition, it could be theorized that the dysbiotic samples investigated in this thesis came from individuals who had recently consumed probiotics or food products containing *Lactobacillus* species, perhaps as a treatment for their dysbiosis-related symptoms, which could veil a larger difference in the abundance between healthy and sick individuals.

The probe Str138, covering *Streptococcus* spp., was the last probe that showed a relation to normobiosis in this thesis. This association was expected, as members of this genus dominate the small intestine, in addition to being some of the few types of bacteria that can survive in the harsh environment of the stomach (1, 24, 26, 96). Thanks to their lactic acid producing ability, *Streptococcus* species are used in food and as probiotics, and are

important for promoting and maintaining a healthy gut (1, 5). Figure 7 further supports the association, as the normobiotic samples on average appear to have had a higher abundance of *Streptococcus* spp. than the dysbiotic samples.

Valeur et al. (58) found in their study that patients that responded well to a low FODMAP diet had higher levels of *Streptococcus* species than those that did not, although it is not yet confirmed if this response was a reflection of the positive effect of *Streptococcus*. In previous studies, a negative health association to the genus has been found, as higher levels of *Streptococcus* spp. have been linked to an increase in pro-inflammatory cytokine IL-6 (58). Some species within this genus are pathogens that can cause disease, but as that is not necessarily linked to dysbiosis, it does not contradict the finding that the probe is associated with normobiosis. On the contrary, it could contribute to the normobiosis association, as an estimated 25% of people exposed to the pathogenic species become carriers of the bacterium, resulting in it becoming part of their microbiota profile, which could be the case for some of the fecal samples in this thesis (1, 96).

As has already been mentioned, dysbiosis is related to complex microbial patterns and relationships in the gut rather than one specific microbe, and the above information on the bacteria related to the investigated probes further demonstrates this. This is why a test like the GA-map® Dysbiosis Test is needed, where it is the collected information from all the probes that determines the health of an individual's gut. This is further demonstrated by the Welch Two Sample t-test performed in this thesis, which informed that apart from probes Ali239 and Cat187, there were no differences in the average signals detected for each probe in the normobiotic versus the dysbiotic samples that were statistically significant. This indicates that one probe alone can't reveal dysbiosis, but rather the pattern shown by all the probes in a probe-set.

A relatively small sample set was investigated in this thesis, and the distribution between normobiotic and dysbiotic samples was slightly uneven. Different links may therefore be seen in a larger sample pool. However, although some of the probe associations in this thesis could be debatable, most of the findings were as expected based on previous studies, the physiology of the target bacteria, and the relationship between the bacteria and the host.

4.1.3 Probes that did not provide satisfactory signals

As there is a desire to incorporate the LAD technology as a supplement to the GA-map® Dysbiosis Test, it would be very practical if all the probes were compatible with both platforms. In addition, the five probes that did not give satisfactory signals with LAD in this thesis all target bacteria that can be expected to be found in the colon, and it could therefore be very useful for determining the health of a gut to ensure that these probes provide true positive signals with the technology.

Probe Bac230 was designed to detect *B. fragilis*, which is a common resident in the human gut. The bacterium has been found to function protectively in patients suffering from inflammation of the colon by modulating immune responses. In addition, patients that respond well to a low FODMAP diet have been found to have a higher level of *B. fragilis* than those that do not respond well (1, 32, 58).

A next step to get the Bac230 probe pair to provide a satisfactory signal with LAD could be to order a fresh batch of both the LP and the RP, as there could be errors in their sequences preventing the LP from binding to the bacterial template and get extended, or the RP from binding to the LP and get quenched. Errors in the RP sequence could also potentially enable it to bind to the Lac47 LP, resulting in the observed peak in the samples containing the pure culture of *L. ruminis*. If a fresh batch of the probes still does not improve the results, new probes could be designed. However, the LP has already been validated by GA for their current GA-map® Dysbiosis Test and has proven to satisfactorily detect *B. fragilis* without detecting *L. ruminis*, indicating that the results in this thesis could be related to the current LAD set-up. Hence, it underlines the need for further investigation and optimization before the technology can be used as a supplement to the current solid-phase platform.

Dorea spp., which is targeted by probe Dor124, is one of the dominant genera in the colon. The genus has been linked to obesity, and to IBS in children, and some members are occasionally found in higher abundances in UC patients compared to healthy subjects (24, 37, 94).

As the bacterial template used in this thesis was found to be a perfect match to the Dor124 LP, no immediate explanation for the lack of signal was apparent. However, one theory is related to the design of the RP, as this probe had four thymine nucleobases added to its 5' end that were not complementary to the LP. These were added to cover up a G which would

otherwise be at the 5' end of the RP, because guanine has a quenching effect and should not get too close to the fluorophore (97). The additional Ts could perhaps have made the distance from the fluorophore on the RP to the quencher on the LP too great for quenching to occur.

Bacteroides stercoris, targeted by probe Bac155, has been found to be strongly associated to dysbiosis in some cases. For example, in MS patients the proportion of the bacteria has been found to be decreased compared to in healthy subjects (20, 74).

When designing the Bac155 RP, three mismatched Ts were included in the 5' end to adjust the melting temperature. This had also been done for several other RPs who all provided satisfactory signals to their target bacteria. However, for Bac155, the mismatched nucleotides could have made the probe duplex too unstable to provide a detectable signal, as it was found that due to the short length of the probes, only 9 bases matched between the LP and the RP.

For both Dor124 and Bac155, a next step for future research could therefore be to design new RPs that match the LPs better.

As probes Bac88 and Str78 were never tested on the target bacteria in this thesis, it is impossible to determine at this point whether or not they are compatible with LAD. The next step for these two probes should therefore be to test them on the intended targets. The second bacterial template used for Str78 had only one mismatch to the probe, located at a central position, and could therefore be a demonstration of the specificity of LAD. However, this conclusion cannot be drawn until a true positive signal for the intended target has been achieved.

4.2 Future work on LAD for use with the GA-map® Dysbiosis Test

In this thesis it has been demonstrated that LAD, by use of liquid reactions, is able to differentiate between dysbiosis and normobiosis with high accuracy, and that it has great multiplexing potential. Based on this it is believed that LAD can be a good addition to the GA-map® Dysbiosis Test, offering increased efficiency and reduced cost. However, there is still work to be done before the technology can be used as a valid supplement, and optimization is needed before further improvements and expansions should be addressed.

4.2.1 Optimizing and improving the use of LAD as a dysbiosis test

As discussed under section 4.1.3 on page 49 and 50, one way to optimize the LAD dysbiosis test could be to try to get the five probes that did not provide satisfactory signals in this thesis to work with LAD. In addition, any of the remaining LPs of the 54 probes in GA's complete probeset that have not yet been tested with LAD by GA should also be investigated.

While the results from this thesis demonstrated that LAD can differentiate between the two states dysbiosis and normobiosis, the GA-map® Dysbiosis Test is able to separate samples into subgroups of these two states. This is done according to their dysbiosis index, and a next step to optimize LAD for use as a supplement to the GA-map® Dysbiosis Test could be to investigate how well the technology can distinguish between the five degrees of dysbiosis defined by GA.

It was shown in this thesis that detecting five probes per channel is possible with LAD, as all five probes designed for the CY5 and HEX channels provided distinguishable signals (although the signal from the Bac230 probe was a false positive). It is therefore still assumed that LAD can detect all 20 probes simultaneously when using four detection channels, as long as all the probes are compatible with the technology. Furthermore, it is believed that the multiplexing level can be increased beyond 20 targets. One way to do this could be to add more probes to each channel. This can for example be attempted by reducing the difference in theoretical T_m between each probe pair. In this thesis, the difference was roughly 10 °C between each probe pair in a channel, except between Ali239 and UNI which had about 5 °C difference between them. As these two probe pairs gave distinguishable signals, it can be theorized that such a small difference in T_m is possible for all the probes in

a channel, thus allowing for more probes to be used per channel. If this is combined with the use of a qPCR machine that has more than four detection channels, it can be expected that LAD will be able to detect as many as 30-40 targets simultaneously in the near future. However, this would need to be tested with all the probes in a probeset to ensure that satisfactory signals, as described under section 2.2.2 on page 22, are achieved for all the probes.

Further optimizations of LAD can be attempted by experimenting with the various parts of the protocol, such as temperatures, cycle numbers, or the reagents used. For example, some false positive probe signals were experienced during this thesis, which could be due to an incomplete polymerase inactivation by sodium dodecyl sulfate (SDS). The false peaks were easy to distinguish from the real ones, and were diminished if the plate was run through the qPCR machine twice. This was therefore not investigated further, as optimization was not within the scope of this thesis. However, as running the plate twice is both time-consuming and less efficient, and the signals can potentially be deceiving, alternatives or additions to SDS should be investigated. An idea is to use heparin for the deactivation. While SDS is an anionic detergent that inactivates the polymerase by denaturing it and covering it with a negative charge (1), heparin is believed to occupy the DNA binding site of the enzyme so that the polymerase is unable to bind to the template (98). Another alternative could be to use a combination, for example by using SDS with EDTA, a chelating agent that binds up the magnesium ions needed by the polymerase to function (1). EDTA on its own has been found to be ineffective for this use with LAD (Pranvera Hiseni, pers. comm.), but together with SDS could potentially get rid of some of the false positive peaks.

The melting curve from the CY5 channel always showed a continuously declining baseline in this thesis (figure 6c and S3a), and it could be worth looking more into this fluorophore to try and improve it, or to look at alternatives. The dropping baseline could theoretically lead to a peak not being acknowledged, as the baseline may end up being at a lower point than a relatively weak peak, or the peak could be interpreted as noise. In this thesis, the peaks on the CY5 channel were always easily distinguishable, but the peak for probe Bac91, found near 70 °C, seemed weak as the baseline had dropped dramatically by the time it reached this temperature. This means the peak's depth may appear smaller than it really was.

A declining baseline could also make quantification difficult. This is not yet something LAD is optimized for, but as dysbiosis is often related to a change in abundance of bacteria, rather than simple presence or absence of species, it is desirable to use the technology quantitatively in the future. For this to be possible, the relationship between the fluorescent signal intensity and the abundance of template needs to be established and validated (51).

The probes used in this thesis detect bacteria based on their 16S rRNA genes. The use of this gene is a good choice as it is found in all bacteria, and there are extensive databases that contain 16S rRNA gene sequences. Its combination of conserved and hypervariable regions makes it easy to amplify all bacterial DNA in a sample using universal primers, and it makes probe design relatively easy. In addition, it allows for detection at various taxonomic levels. However, the 16S rRNA gene is partly why LAD can only be used semi-quantitatively, as bacteria have different numbers of copies of the gene, some having more than 15 copies per genome, making them overrepresented. Copy numbers mostly appear to be specific to taxon, allowing for this to be taken into consideration during testing. However, this is not consistent, and variations have even been found among strains of the same species (99, 100, 101).

Another issue with the 16S rRNA gene which can affect the registered abundance is that the genes from different species don't get amplified with the same efficiency during PCR, which can lead to the genes of some microbes being amplified more than others. By designing primers that target highly conserved regions, this issue can be overcome, but it is then important to ensure that the part of the gene that is complementary to the probes used in LAD is intact and sufficient for the necessary classification accuracy (99, 101).

As has been mentioned earlier in this thesis, it can be difficult to differentiate between certain bacteria when using the 16S rRNA gene, and some species even have identical 16S rRNA genes (101). If such species are found to be associated with dysbiosis, being able to distinguish between them would be very useful.

Based on the above information, other genes could be suggested to be used instead of the 16S rRNA gene. For example, some studies claim that the single-copy protein coding genes *cpn60* and *rpoB*, coding for type I chaperonin protein and the β subunit of DNA polymerases respectively, have a phylogenetic resolution as good as, or even better than, the 16S rRNA gene. However, more work still needs to be conducted on such replacement genes,

and the limited number of sequences found in databases of those genes compared to that of the 16S rRNA gene makes it less preferable at the moment (99, 100, 102, 103). Nevertheless, it could be interesting to examine such genes in the future as an alternative, or addition, to the 16S rRNA gene to improve the LAD dysbiosis test even further, both in relation to resolution of the test, and the quantification ability of the technology.

4.2.2 Expanding the LAD dysbiosis test

The probes of the GA-map® Dysbiosis Test have been designed based on microbial compositions that are associated with IBS and IBD in a Scandinavian setting. To increase the diagnostic scope of the test, investigations could be made into additional dysbiosis-related disorders, for which specific probe sets could be designed. For example, as H₂ may be involved in PD pathogenicity, it could be interesting to investigate further the producers and consumers involved in this, along with other contributors to the disease. Or for those suffering from MS, more research could be done on *Parabacteroides* and *Acinetobacter* abundance, as well as on any other patterns that may be shared among the affected individuals. By designing appropriate probesets for various disorders, LAD can become a highly useful tool for the diagnosis and treatment of a large number of medical conditions. In addition, the technology can prove useful for predictions of diseases that may develop in the future. For example, in the introduction of this thesis it was mentioned that NEC is believed to be preceded by bacterial compositions in the gut that trigger bodily responses which may lead to development of the disease. In addition, both PD and MS have been seen to display distinct microbial patterns prior to the onset of disease, which may function as warning signs (34, 35, 58, 72, 73, 74, 77). By detecting such patterns early by using an appropriate probeset with LAD, treatments can commence before the disease has a chance to develop. The same is true for any illness where it is believed that a disturbed gut microbiota may precede the actual disease, whether the microbes are a cause or simply a reflection of the development of the condition.

As the probesets in the GA-map® Dysbiosis Test are based on gut bacteria that are common in adults in Scandinavia, bacterial compositions could also be mapped out for other age groups and populations, further developing the scope of the LAD dysbiosis test.

Another potential expansion for the LAD dysbiosis test is to develop a catalogue of functional genes, as studies have found that while the composition of the gut microbiota in individuals differ, the functions remain more or less the same in everyone (3, 6, 28, 38, 61, 62, 64). This means that different species are performing similar functions in our guts, and as a disturbance in function can be a trait of dysbiosis, it may be useful to investigate the genes involved in this, in addition to species composition, to truly investigate dysbiosis and its related diseases. For example, the metabolism of carbohydrates is an important fermentative process performed by microorganisms, providing the host with energy and nutrients. By investigating functional groups, it has been found that certain types of IBS are characterized by an imbalance in the microbial community that performs this important task, and that some IBS patients suffer from changes in carbohydrate metabolism. In addition, as mentioned earlier shifts in H₂ metabolism and butyrate-production are often observed (3, 6, 28, 64). One study showed that NEC is often linked to functionally related strains of *E. coli*. Such a link could be missed if the investigation only looks at *E. coli* as an entire group (35). The actual mapping of the functions could be done by other methods than LAD, such as HTS. However, once patterns are known, appropriate probes can be designed to detect these functional genes using LAD.

For this thesis, the focus was to investigate the bacteria associated with dysbiosis. However, Eukaryotes and Archaea are also found in the human gut, and studies have found that they are also involved in dysbiosis-related diseases. For example, both fungi and methane-production from archaea have been linked to IBD (5, 6, 28, 59, 60). The composition and interactions of all three domains could therefore be investigated for a more complete picture of the elements involved in the health of an individual's gastrointestinal tract. Appropriate probes, for example ones targeting the 16S rRNA genes found in archaea, or the internal transcribed spacer (ITS, a common marker used in mycology) region in fungi, can then be designed and included in the LAD dysbiosis test catalogue (6, 59, 60, 104).

By expanding LAD's diagnostic scope, the dysbiosis test could prove to be an invaluable tool in the world of medicine.

4.3 Other potential uses for LAD

4.3.1 Tailor-made therapies

LAD's ability to accurately and rapidly detect selected DNA sequences in the human gut makes it suitable for use also with other aims than diagnosing dysbiosis. For example, as has already been indicated in this thesis there are strong indications that the gut microbiota is connected to a person's response to drugs and treatments. Some treatments are ineffective for certain individuals, and can even be harmful if their microbes metabolize the drug into metabolites that are detrimental. A quick and easy test like LAD could be used to investigate the composition of the microorganisms in the gut prior to the administration of a drug, enabling the creation of tailor-made therapies formulated to eliminate undesirable reactions to treatment (1, 3).

For dysbiosis-related ailments there are several treatments available, one of which is FMT (1, 3, 5). The efficiency of FMT is still debated, but while there are some contradictory findings, positive effects have been noted in various studies, especially in individuals suffering from UC. The microbiota of the donor has been seen to affect the result of an FMT, with the most successful donors often having an increased proportion of species that can ferment pectin, and of species from the family *Lachnospiraceae*, of which *E. hallii* is a member (1, 5, 62, 64, 80, 105). With FMT it is normal to use transplants from healthy family members, as it is assumed that the gut microbiota is similar due to being exposed to the same environments and diets (1, 5). However, by using the LAD technology, the microbial profile of both the recipient and potential donors can quickly be mapped, increasing the chance for success.

Patients suffering from IBS often experience abdominal pain, and as mentioned on the previous page, a disturbance in the microbial community involved in metabolizing carbohydrates is suggested as one factor involved in this (5, 6, 106). Lactose, fructose, polyols and other fermentable carbohydrates can be difficult for the body to absorb, and have been seen as a culprit for such symptoms. To treat this, it has been common to restrict intake of one or a few types of these carbohydrates. Lately the low FODMAP diet has gained increasing support, in which intake of several such fermentable carbohydrates is greatly reduced. Such a diet has been found to relieve abdominal pain quickly, probably at least partly due to the reduced fermentation by microbes in the gut which leads to less gas and

osmotically active metabolites (22, 58, 106, 107). However, this is a very challenging diet to follow, and often includes resource-demanding nutritional counseling. The diet does not work for all IBS subjects, and potential negative effects of long-term use are less known. Following extreme diets like this will change the composition in the gut by selecting for certain microorganisms, and detrimental effects can be experienced (58, 106, 107). Putting a patient through such a diet if they will not have a good response to it should therefore be avoided. As was mentioned during the discussion of the probe associations in this study certain bacteria have been linked to the success of following the low FODMAP diet, and patients that respond well have been found to have a different bacterial profile in their gut than the non-responders. Using LAD to determine if the regime is likely to have any positive effect on an individual could therefore be very useful (22, 58, 106, 107, 108).

Lately, links have been drawn between gut microbiota and the efficacy of cancer treatments such as immunotherapy. This kind of treatment uses antibodies to stimulate the immune system to attack cells that are cancerous, and some patients respond better to this than others (3, 22, 89, 109). As has been mentioned earlier in this thesis, one study recently found that people who responded well to the treatment had a higher abundance of species within the *Alistipes* genus. In addition, a higher abundance of the bacteria *Akkermansia muciniphila* was detected. When *A. muciniphila* was transferred to mice with tumors using FMT, a positive correlation to the immunotherapy treatment was observed (89). In addition, both patients and mice that had been given antibiotics had worse cancer outcomes than those that had not used antibiotics recently, which may be linked to the reduction in microbial diversity that often follows an antibiotic treatment (2, 3, 28, 29, 30, 89).

Another study on gut microbiota and cancer immunotherapy treatment found a link between *Bacteroides* species, particularly *B. fragilis*, and the efficacy of a treatment involving the use of antibodies targeting and blocking the protein CTLA-4, which functions to downregulate immune response (109). They found that tumors in mice that had been treated with antibiotics, or were raised germ-free, did not respond to blockade of CTLA-4. However, this was alleviated by gavage with *B. fragilis*, immunization using *B. fragilis* polysaccharides, or by transferring *B. fragilis* specific T cells (109).

The probe designed for *B. fragilis* used in this thesis, Bac230, did not work as expected with LAD, but the above example demonstrates how important the probe could be for the

success of cancer treatments. Although the results from these two studies are far from conclusive, and more research is needed to determine the cause and effect of all factors involved, it is an example of the exciting potential of LAD. By creating a probeset to target *A. muciniphila*, *B. fragilis* and other bacteria found to be involved in the efficacy of cancer treatment, LAD can be used to quickly determine the best path forward, for example by performing an FMT prior to immunotherapy.

4.3.2 Screening of genes in industry and medicine

While this thesis has focused on bacteria in the gut and diagnosis of dysbiosis, LAD is by no means limited to this. For example, the technology might prove useful during outbreaks of food poisoning, or for screening of known pathogens in raw or processed food prior to becoming available to the consumer. In addition, transfusion of blood and blood products to patients that have reduced immune systems can be dangerous, partly due to bacterial contamination. Many of the methods used for detecting contamination today require incubation time to allow for bacterial growth, but could perhaps be improved by using LAD (110, 111).

As has already been mentioned, LAD can also be used on eukaryotes. This was demonstrated by Hiseni (54) when she succeeded in using the technology to genotype polymorphisms in the *MSTN* gene in beef cattle to reveal unfavorable mutations that lead to double muscling. In her thesis, she proved that LAD can be used to detect allele variants as well as indels. This suggests that LAD can potentially be used to detect genetic diseases in humans and animals. In addition, genetically modified organisms (GMOs) could perhaps be investigated with LAD to check if labelling is correct by targeting known genes that are frequently used, such as insect resistant genes. As new genetic elements are continuously emerging, a high-throughput method such as LAD to target a large set of genes could be highly useful for qualitative assurance of GMOs (47).

These are just a few suggestions for how the LAD technology may prove useful, and they would all need to be tested and verified, which falls outside the scope of this thesis. However, it is clear that the specificity, speed and low cost of LAD makes it a highly versatile tool, with great potential in many areas.

4.4 Conclusion

In this thesis, it has been successfully proven that LAD can detect 15 target DNA sequences simultaneously, and it is expected that at least 30 should be possible in the near future. In addition, it has been demonstrated that LAD can distinguish between normobiosis and dysbiosis. Although further optimization is still needed before LAD can be used in a clinical setting, the results from this thesis indicate that the technology can be a highly valuable supplement to the current solid-phase platform used in the GA-map® Dysbiosis Test.

The use of LAD requires no culturing of bacteria, and no designated expensive equipment, as the process can be performed on qPCRs found in most laboratories. As no separation and purification step is needed, LAD is more efficient and has less risk of contamination than many other genotyping methods on the market today. The technology is inexpensive, easy, specific and rapid compared to other methods, and can provide important contributions to personalized treatments, nutrition and precision medicine. By adjusting the probes to fit age groups, populations, target of investigation, and more, LAD can become a highly versatile and valuable high-throughput technology.

Reference List

1. Willey JM, Sherwood LM, Woolverton CJ. 2014. Prescott's Microbiology, ninth edition. McGraw-Hill Education, New York.
2. Pérez-Cobas AE, Gosalbes MJ, Friedrichs A, Knecht H, Artacho A, Eismann K, Otto W, Rojo D, Bargiela R, von Bergen M, Neulinger SC, Däumer C, Heinsen FA, Latorre A, Barbas C, Seifert J, dos Santos VM, Ott SJ, Ferrer M, Moya A. 2012. Gut microbiota disturbance during antibiotic therapy: a multi-omic approach. *Gut*. 62:1591-1601.
3. Lozupone CA, Stombaugh JI, Gordon JI, Jansson JK, Knight R. 2012. Diversity, stability, and resilience of the human gut microbiota. *Nature*. 489:220-230.
4. Palmer C, Bik EM, DiGiulio DB, Relman DA, Brown PO. 2007. Development of the Human Infant Intestinal Microbiota. *PLoS Biology*. 5:1556-1573.
5. Distrutti E, Monaldi L, Ricci P, Fiorucci S. 2016. Gut microbiota role in irritable bowel syndrome: New therapeutic strategies. *World Journal of Gastroenterology*. 22:2219-2241.
6. Chassard C, Dapoigny M, Scott KP, Crouzet L, Del'homme C, Marquet P, Martin JC, Pickering G, Ardid D, Eschalier A, Dubray C, Flint HJ, Bernalier-Donadille A. 2012. Functional dysbiosis within the gut microbiota of patients with constipated-irritable bowel syndrome. *Aliment Pharmacol Ther*. 35:828-838.
7. Rodríguez JM, Murphy K, Stanton C, Ross RP, Kober OI, Juge N, Avershina E, Rudi K, Narbad A, Jenmalm MC, Marchesi JR, Collado MC. 2015. The composition of the gut microbiota throughout life, with an emphasis on early life. *Microbial Ecology in Health & Disease*. 26:26050.
8. Avershina E, Lundgård K, Sekelja M, Dotterud C, Storrø O, Øien T, Johnsen R, Rudi K. 2016. Transition from infant- to adult-like gut microbiota. *Society for Applied Microbiology and John Wiley & Sons Ltd, Environmental Microbiology*. 18:2226-2236.
9. Satokari R, Grönroos T, Laitinen K, Salminen S, Isolauri E. 2008. *Bifidobacterium* and *Lactobacillus* DNA in the human placenta. *Society for Applied Microbiology, Letters in Applied Microbiology*. 48:8-12.
10. Aagaard K, Ma J, Antony KM, Ganu R, Petrosino J, Versalovic J. 2014. The Placenta Harbors a Unique Microbiome. *Sci Transl Med*. 6:237ra265.
11. Jiménez E, Fernández L, Marín ML, Martín R, Odriozola JM, Nueno-Palop C, Narbad A, Olivares M, Xaus J, Rodríguez JM. 2005. Isolation of Commensal Bacteria from Umbilical Cord Blood of Healthy Neonates Born by Cesarean Section. *Current Microbiology*. 51:270-274.
12. DiGiulio DB, Romero R, Amogan HP, Kusanovic JP, Bik EM, Gotsch F, Kim CJ, Erez O, Edwin S, Relman DA. 2008. Microbial Prevalence, Diversity and Abundance in Amniotic Fluid During Preterm Labor: A Molecular and Culture-Based Investigation. *PLoS One*. 3:e3056.
13. DiGiulio DB. 2012. Diversity of microbes in amniotic fluid. *Seminars in Fetal & Neonatal Medicine*. 17:2-11.

14. Jiménez E, Marín ML, Martín R, Odriozola JM, Olivares M, Xaus J, Fernández L, Rodríguez JM. 2008. Is meconium from healthy newborns actually sterile? *Research in Microbiology*. 159:187-193.
15. Ardisson AN, de la Cruz DM, Davis-Richardson AG, Rechcigl KT, Li N, Drew JC, Murgas-Torrazza R, Sharma R, Hudak ML, Triplett EW, Neu J. 2014. Meconium Microbiome Analysis Identifies Bacteria Correlated with Premature Birth. *PLoS One*. 9:e90784.
16. Moles L, Gómez M, Heilig H, Bustos G, Fuentes S, de Vos W, Fernández L, Rodríguez JM, Jiménez E. 2013. Bacterial diversity in meconium of preterm neonates and evolution of their fecal microbiota during the first month of life. *PLoS One*. 8:e66986.
17. Rosenbaum M, Knight R, Leibel RL. 2015. The gut microbiota in human energy homeostasis and obesity. *Trends in Endocrinology & Metabolism*. 26:493-501.
18. Dominguez-Bello MG, Costello EK, Contreras M, Magris M, Hidalgo G, Fierer N, Knight R. 2010. Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns. *Proceedings of the National Academy of Sciences of the United States of America*. 107:11971-11975.
19. Vebø HC, Sekelja M, Nestestog R, Storrø O, Johnsen R, Øien T, Rudi K. 2011. Temporal Development of the Infant Gut Microbiota in Immunoglobulin E-Sensitized and Nonsensitized Children Determined by the GA-Map Infant Array. *Clinical and Vaccine Immunology*. 18:1326-1335.
20. Casén C, Vebø HC, Sekelja M, Hegge FT, Karlsson MK, Cierniejewska E, Dzankovic S, Frøyland C, Nestestog R, Engstrand L, Munkholm P, Nielsen OH, Rogler G, Simrén M, Öhman L, Vatn MH, Rudi K. 2015. Deviations in human gut microbiota: a novel diagnostic test for determining dysbiosis in patients with IBS or IBD. *Alimentary Pharmacology and Therapeutics*. 42:71-83.
21. Eckburg PB, Bik EM, Bernstein CN, Purdom E, Dethlefsen L, Sargent M, Gill SR, Nelson KE, Relman DA. 2005. Diversity of the human intestinal microbial flora. *Science*. 308:1635– 1638.
22. Singh RK, Chang HW, Yan D, Lee KM, Ucmak D, Wong K, Abrouk M, Farahnik B, Nakamura M, Zhu TH, Bhutani T, Liao W. 2017. Influence of diet on the gut microbiome and implications for human health. *Journal of Translational Medicine*. 15:73.
23. Clarke SF, Murphy EF, Nilaweera K, Ross RP, Shanahan F, O'Toole PW, Cotter PD. 2012. The gut microbiota and its relationship to diet and obesity. *Gut microbes*. 3:186-202.
24. Rivièrè A, Selak M, Lantin D, Leroy F, De Vuyst L. 2016. *Bifidobacteria* and Butyrate-Producing Colon Bacteria: Importance and Strategies for Their Stimulation in the Human Gut. *Frontiers in Microbiology*. 7:979.
25. Leimena MM, Ramiro-Garcia J, Davids M, van den Bogert B, Smidt H, Smid EJ, Boekhorst J, Zoetendal EG, Schaap PJ, Kleerebezem M. 2013. A comprehensive

- metatranscriptome analysis pipeline and its validation using human small intestine microbiota datasets. *BMC Genomics*. 14:530.
26. Zoetendal EG, Raes J, van den Bogert B, Arumugam M, Booijink CCGM, Troost FJ, Bork P, Wels M, de Vos WM, Kleerebezem M. 2012. The human small intestinal microbiota is driven by rapid uptake and conversion of simple carbohydrates. *The ISME Journal*. 6:1415-1426.
 27. Booijink CCGM, El-Aidy S, Rajilić-Stojanović M, Heilig HGHJ, Troost FJ, Smidt H, Kleerebezem M, de Vos WM, Zoetendal EG. 2010. High temporal and inter-individual variation detected in the human ileal microbiota. *Environmental Microbiology*. 12:3213-3227.
 28. Walker AW, Sanderson JD, Churcher C, Parkes GC, Hudspith BN, Rayment N, Brostoff J, Parkhill J, Dougan G, Petrovska L. 2011. High-throughput clone library analysis of the mucosa-associated microbiota reveals dysbiosis and differences between inflamed and noninflamed regions of the intestine in inflammatory bowel disease. *BMC Microbiology*. 11:7
 29. Cochetière MFDL, Durand T, Lepage P, Bourreille A, Galmiche JP, Doré J. 2005. Resilience of the dominant human fecal microbiota upon short-course antibiotic challenge. *Journal of clinical microbiology*. 43:5588-5592.
 30. Panda S, El Khader I, Casellas F, Vivancos JL, Cors MG, Santiago A, Cuenca S, Guarner F, Manichanh C. 2014. Short-term effect of antibiotics on human gut microbiota. *PLoS One*. 9:e95476.
 31. Walters SS, Quiros A, Rolston M, Grishina I, Li J, Fenton A, DeSantis TZ, Thai A, Andersen GL, Papathakis P, Nieves R, Prindiville T, Dandekar S. 2014. Analysis of Gut Microbiome and Diet Modification in Patients with Crohn's Disease. *SOJ Microbiology & Infectious Diseases*. 2:1-13.
 32. Xuan C, Shamonki JM, Chung A, DiNome ML, Chung M, Sieling PA, Lee DJ. 2014. Microbial Dysbiosis Is Associated with Human Breast Cancer. *PLoS One*. 9:e83744.
 33. Wang Y, Hoenig JD, Malin KJ, Qamar S, Petrof EO, Sun J, Antonopoulos DA, Chang EB, Claud EC. 2009. 16S rRNA gene-based analysis of fecal microbiota from preterm infants with and without necrotizing enterocolitis. *The ISME Journal*. 3:944-954.
 34. Warner BB, Deych E, Zhou Y, Hall-Moore C, Weinstock GM, Sodergren E, Shaikh N, Hoffmann JA, Linneman LA, Hamvas A, Khanna G, Rouggy-Nickless LC, Ndao IM, Shands BA, Escobedo M, Sullivan JE, Radmacher PG, Shannon WD, Tarr PI. 2016. Gut bacteria dysbiosis and necrotizing enterocolitis in very low birthweight infants: a prospective case-control study. *Lancet*. 387:1928-1936.
 35. Ward DV, Scholz M, Zolfo M, Taft DH, Schibler KR, Tett A, Segata N, Morrow AL. 2016. Metagenomic Sequencing with Strain-Level Resolution Implicates Uropathogenic *E.coli* in Necrotizing Enterocolitis and Mortality in Preterm Infants. *Cell Reports*. 14:2912-2924.

36. Jiang H, Ling Z, Zhang Y, Mao H, Ma Z, Yin Y, Wang W, Tang W, Tan Z, Shi J, Li L, Ruan B. 2015. Altered fecal microbiota composition in patients with major depressive disorder. *Brain, Behavior, and Immunity*. 48:186-194.
37. Ottosson F, Brunkwall L, Ericson U, Nilsson PM, Almgren P, Fernandez C, Melander O, Orho-Melander M. 2018. Connection between BMI related plasma metabolite profile and gut microbiota. *The Journal of Clinical Endocrinology & Metabolism*. 103:1491-1501.
38. Carding S, Verbeke K, Vipond DT, Corfe BM, Owen LJ. 2015. Dysbiosis of the gut microbiota in disease. *Microbial Ecology in Health and Disease*. 26:26191.
39. Latella G, Fiocchi C, Caprili R. 2010. News from the “5th international meeting on inflammatory bowel diseases” CAPRI 2010. *Journal of Crohn’s and Colitis*. 4:690-702.
40. Namsolleck P, Thiel R, Lawson P, Holmstrøm K, Rajilic M, Vaughan EE, Rigottier-Gois L, Collins MD, de Vos WM, Blaut M. 2004. Molecular Methods for the Analysis of Gut Microbiota. *Microbial Ecology in Health and Disease*. 16:71-85.
41. Sobrino B, Brión M, Carracedo A. 2005. SNPs in forensic genetics: a review of SNP typing methodologies. *Forensic science international*. 154:181-194.
42. Rains D. 2013. Ask TaqMan Episode 13 – How Taqman Works. Life Technologies Corporation. ThermoFisher Scientific website.
43. Nagy A, Černíková L, Vitásková E, Křivda V, Dán Á, Dirbáková Z, Jiřincová H, Procházka B, Sedlák K, Havlíčková M. 2016. MeltMan: Optimization, Evaluation and Universal Application of a qPCR System Integrating the TaqMan qPCR and Melting Analysis into a Single Assay. *PLoS One*. 11:e0151204.
44. Nagy A, Vitásková E, Černíková L, Křivda V, Jiřincová H, Sedlák K, Horníčková J, Havlíčková M. 2017. Evaluation of TaqMan qPCR system integrating two identically labelled hydrolysis probes in single assay. *Scientific Reports*. 7:41392.
45. Tyagi S, Kramer FR. 1996. Molecular Beacons: Probes that Fluoresce upon Hybridization. *Nature Biotechnology*. 14:303-308.
46. Tong SYC, Giffard PM. 2012. Microbial Application of High -Resolution Melting Analysis. *Journal of Clinical Microbiology*. 50:3418-3421.
47. Druml B, Cichna-Markl M. 2014. High resolution melting (HRM) analysis of DNA – Its role and potential in food analysis. *Food chemistry*. 158:245-254.
48. Capper RL, Jin YK, Lundgren PB, Peplow LM, Matz MV, van Oppen MJH. 2015. Quantitative high resolution melting: two methods to determine SNP allele frequencies from pooled samples. *PMC Genetics*. 16:62.
49. Liao Y, Wang X, Sha C, Xia Z, Huang Q, Li Q. 2012. Combination of fluorescence color and melting temperature as a two-dimensional label for homogenous multiplex PCR detection. *Nucleic Acid Research*. 41:e76
50. Fu G, Miles A, Alpey L. 2012. Multiplex Detection and SNP Genotyping in a Single Fluorescence Channel. *PLoS One*. 7:e30340.
51. Nikolausz M, Chatzinotas A, Táncsics A, Imfeld G, Kästner M. 2009. The single-nucleotide primer extension (SNUPE) method for the multiplex detection of various

- DNA sequences: from detection of point mutations to microbial ecology. *Biochemical Society Transactions*. 37:454-459.
52. Rudi K, Holck AL. 2003. Real-time closed tube single nucleotide polymorphism (SNP) quantification in pooled samples by quencher extension (QEXT). *Nucleic Acids Research*. 31:e117.
 53. Rudi K, Zimonja M, Hannevik SE, Drømtorp SM. 2006. Multiplex real-time single nucleotide polymorphism detection and quantification by quencher extension. *BioTechniques*. 40:323-329.
 54. Hiseni P. 2016. Liquid Array Diagnostics (LAD)-based multiplexed genotyping tests for double-muscling mutations in beef cattle. Hedmark University of Applied Sciences. Master Thesis.
 55. Tibyehabwa RR. 2015. Evaluate the diagnostic potential of the Liquid Array Diagnostics (LAD) technology for single nucleotide polymorphism (SNP) detection. Hedmark University College. Master thesis.
 56. Beck IA, Deng W, Payant R, Hall R, Bumgarner RE, Mullins JI, Frenkel LM. 2014. Validation of an Oligonucleotide Ligation Assay for Quantification of Human Immunodeficiency Virus Type 1 Drug-Resistant Mutants by Use of Massively Parallel Sequencing. *Journal of Clinical Microbiology*. 52:2320-2327.
 57. Simsek M, Adnan H. 2000. Effect of single mismatches at 3'-end of primers on polymerase chain reaction. *Journal for Scientific Research, Medical Sciences*. 2:11-14.
 58. Valeur J, Småstuen MC, Knudsen T, Lied GA, Røseth AG. 2018. Exploring Gut Microbiota Composition as an Indicator of Clinical Response to Dietary FODMAP Restriction in Patients with Irritable Bowel Syndrome. *Digestive Diseases and Sciences*. 63:429-436.
 59. Chehoud C, Albenberg LG, Judge C, Hoffmann C, Grunberg S, Bittinger K, Baldassano RN, Lewis JD, Bushman FD, Wu GD. 2015. A Fungal Signature in the Gut Microbiota of Pediatric Patients with Inflammatory Bowel Disease. *Inflammatory Bowel Diseases*. 21:1948-1956.
 60. Hoffmann C, Dollive S, Grunberg S, Chen J, Li H, Wu GD, Lewis JD, Bushman FD. 2013. Archaea and Fungi of the Human Gut Microbiome: Correlations with Diet and Bacterial Residents. *PLoS One* 8:e66019.
 61. Statovci D, Aguilera M, MacSharry J, Melgar S. 2017. The Impact of Western Diet and Nutrients on the Microbiota and Immune Response at Mucosal Interfaces. *Frontiers in Immunology*. 8:838.
 62. Wei Y, Gong J, Zhu W, Tian H, Ding C, Gu L, Li N, Li J. 2016. Pectin enhances the effect of fecal microbiota transplantation in ulcerative colitis by delaying the loss of diversity of gut flora. *BMC microbiology*. 16:255
 63. Magnusson MK, Strid H, Sapnara M, Lasson A, Bajor A, Ung KA, Öhman L. 2016. Anti-TNF Therapy Response in Patients with Ulcerative Colitis Is Associated with Colonic Antimicrobial Peptide Expression and Microbiota Composition. *Journal of Crohn's and Colitis*. 10:943-952.

64. Geirnaert A, Calatayud M, Grootaert C, Laukens D, Devriese S, Smagghe G, De Vos M, Boon N, Van de Wiele T. 2017. Butyrate-producing bacteria supplemented in vitro to Crohn's disease patient microbiota increased butyrate production and enhanced intestinal epithelial barrier integrity. *Scientific Reports*. 7:11450.
65. Udayappan S, Manneras-Holm L, Chaplin-Scott A, Belzer C, Herrema H, Dallinga-Thie GM, Duncan SH, Stroes ESG, Groen AK, Flint HJ, Backhed F, de Vos WM, Nieuwdorp M. 2016. Oral treatment with *Eubacterium hallii* improves insulin sensitivity in db/db mice. *Nature partner journals*. 2:16009.
66. Engels C, Ruscheweyh HJ, Beerenwinkel N, Lacroix C, Schwab C. 2016. The Common Gut Microbe *Eubacterium hallii* also Contributes to Intestinal Propionate Formation. *Frontiers in Microbiology*. 7:713.
67. Ríos-Covián D, Ruas-Madiedo P, Margolles A, Gueimonde M, de los Reyes-Gavilán CG, Salazar N. 2016. Intestinal Short Chain Fatty Acids and their Link with Diet and Human Health. *Frontiers in Microbiology*. 7:185.
68. Le Roy CI, Štšepetova J, Sepp E, Songisepp E, Claus SP, Mikelsaar M. 2015. New insights into the impact of *Lactobacillus* population on host-bacteria metabolic interplay. *Oncotarget*. 6:30545-30556.
69. Corrêa-Oliveira R, Fachi JL, Vieira A, Sato FT, Vinolo MAR. 2016. Regulation of immune cell function by short-chain fatty acids. *Clinical & Translational Immunology*. 5:e73.
70. Velázquez M, Davies C, Marett R, Slavin JL, Feirtag JM. 2000. Effect of Oligosaccharides and Fibre Substitutes on Short-chain Fatty Acid Production by Human Faecal Microflora. *Anaerobe*. 6:87-92.
71. Roy CC, Kien CL, Bouthillier L, Levy E. 2006. Short-chain fatty acids: ready for prime time? *Nutrition in Clinical Practice*. 21:351-366.
72. Chen J, Chia N, Kalari KR, Yao JZ, Novotna M, Soldan MMP, Luckey DH, Marietta EV, Jeraldo PR, Chen X, Weinshenker BG, Rodriguez M, Kantarci OH, Nelson H, Murray JA, Mangalam AK. 2016. Multiple sclerosis patients have a distinct gut microbiota compared to healthy controls. *Scientific reports*. 6:28484.
73. Cekanaviciute E, Yoo BB, Runia TF, Debelius JW, Singh S, Nelson CA, Kanner R, Bencosme Y, Lee YK, Hauser SL, Crabtree-Hartman E, Sand IK, Gracias M, Zhu Y, Casaccia P, Cree BAC, Knight R, Mazmanian SK, Baranzini SE. 2017. Gut bacteria from multiple sclerosis patients modulate human T cells and exacerbate symptoms in mouse models. *PNAS*. 114:10713-10718.
74. Miyake S, Kim S, Suda W, Oshima K, Nakamura M, Matsuoka T, Chihara N, Tomita A, Sato W, Kim SW, Morita H, Hattori M, Yamamura T. 2015. Dysbiosis in the Gut Microbiota of Patients with Multiple Sclerosis, with a Striking Depletion of Species Belonging to Clostridia XIVA and IV Clusters. *PLoS One*. 10:e0137429.
75. Jahng J, Jung IS, Choi EJ, Conklin JL, Park H. 2012. The effects of methane and hydrogen gases produced by enteric bacteria on ileal motility and colonic transit time. *Neurogastroenterol Motility*. 24:185-e92.

76. Linden DR. 2014. Hydrogen Sulfide Signaling in the Gastrointestinal Tract. *Antioxidants & Redox Signaling*. 20:818-830.
77. Ostojic SM. 2018. Inadequate Production of H₂ by Gut Microbiota and Parkinson Disease. *Trends in endocrinology and metabolism*. 29:286-288.
78. Wang J, Wang Y, Zhang X, Liu J, Zhang Q, Zhao Y, Peng J, Feng Q, Dai J, Sun S, Zhao Y, Zhao L, Zhang Y, Hu Y, Zhang M. 2017. Gut Microbial Dysbiosis Is Associated with Altered Hepatic Functions and Serum Metabolites in Chronic Hepatitis B Patients. *Frontiers in Microbiology*. 8:2222.
79. Jensen NS, Canale-Parola E. 1986. *Bacteroides pectinophilus* sp. nov. and *Bacteroides galacturonicus* sp. nov.: Two Pectinolytic Bacteria from the Human Intestinal Tract. *Applied and environmental microbiology*. 52:880-887.
80. Bang SJ, Kim G, Lim MY, Song EJ, Jung DH, Kum JS, Nam YD, Park CS, Seo DH. 2018. The influence of in vitro pectin fermentation on the human fecal microbiome. *AMB Express*. 8:98.
81. Jackson MA, Goodrich JK, Maxan ME, Freedberg DE, Abrams JA, Poole AC, Sutter JL, Welter D, Ley RE, Bell JT, Spector TD, Steves CJ. 2015. Proton pump inhibitors alter the composition of the gut microbiota. *Gut*. 65:749-756.
82. Imhann F, Bonder MJ, Vila AV, Fu J, Mujagic Z, Vork L, Tigchelaar EF, Jankipersadsing SA, Cenit MC, Harmsen HJM, Dijkstra G, Franke L, Xavier RJ, Jonkers D, Wijmenga C, Weersma RK, Zhernakova A. 2015. Proton pump inhibitors affect the gut microbiome. *Gut*. 65:740-748.
83. Bastard QL, Al-Ghalith GA, Grégoire M, Chapelet G, Javaudin F, Dailly E, Batard E, Knights D, Montassier E. 2017. Systematic review: human gut dysbiosis induced by non-antibiotic prescription medicine. *Alimentary Pharmacology and Therapeutics*. 47:332-345.
84. Bolourian A, Mojtahedi Z. 2018. *Streptomyces*, shared microbiome member of soil and gut, as "old friends" against colon cancer. *FEMS Microbial Ecology*. doi: 10.1093/femsec/fiy120.
85. Pepper IL, Gerba CP, Gentry TJ. 2015. *Environmental Microbiology*. Academic Press, London.
86. Özen AI, Ussery DW. 2012. Defining the *Pseudomonas* Genus: Where Do We Draw the Line with *Azotobacter*? *Microbial Ecology*. 63:239-248.
87. Hakansson A, Molin G. 2011. Gut Microbiota and Inflammation. *Nutrients*. 3:637-682.
88. Kageyama A, Benno Y. 2000. *Catenibacterium mitsuokai* gen. nov., sp. nov., a Gram-positive anaerobic bacterium isolated from human faeces. *International Journal of Systematic and Evolutionary Microbiology*. 50:1595-1599.
89. Routy B, Le Chatelier E, Derosa L, Duong CPM, Alou MT, Daillère R, Fluckiger A, Messaoudene M, Rauber C, Roberti MP, Fidelle M, Flament C, Poirier-Colame V, Opolon P, Klein C, Iribarren K, Mondragón L, Jacquelot N, Qu B, Ferrere G, Clémenson C, Mezquita L, Masip JR, Naltet C, Brosseau S, Kaderbhai C, Richard C, Rizvi H, Levenez F, Galleron N, Quinquis B, Pons N, Ryffel B, Minard-Colin V, Gonin P, Soria JC,

- Deutsch E, Lorient Y, Ghiringhelli F, Zalcman G, Goldwasser F, Escudier B, Hellman MD, Eggermont A, Raoult D, Albiges L, Kroemer G, Zitvogel L. 2018. Gut microbiome influences efficacy of PD-1-based immunotherapy against epithelial tumors. *Science*. 359:91-97.
90. Clarke SF, Murphy EF, O'Sullivan O, Ross RP, O'Toole PW, Shanahan F, Cotter PD. 2013. Targeting the Microbiota to Address Diet-Induced Obesity: A Time Dependent Challenge. *PLoS One*. 8:e65790.
 91. Louis S, Tappu RM, Damms-Machado A, Huson DH, Bischoff SC. 2016. Characterization of the Gut Microbial Community of Obese Patients Following a Weight-Loss Intervention Using Whole Metagenome Shotgun Sequencing. *PLoS One*. 11:e0149564.
 92. Dinh DM, Volpe GE, Duffalo C, Bhalchandra S, Tai AK, Kane AV, Wanke CA, Ward HD. 2015. Intestinal Microbiota, Microbial Translocation, and Systemic Inflammation in Chronic HIV Infection. *The Journal of Infectious Diseases*. 211:19-27.
 93. Jiang W, Wu N, Wang X, Chi Y, Zhang Y, Qiu X, Hu Y, Li J, Liu Y. 2015. Dysbiosis gut microbiota associated with inflammation and impaired mucosal immune function in intestine of humans with non-alcoholic fatty liver disease. *Scientific Reports*. 5:8096.
 94. Saulnier DM, Riehle K, Mistretta TA, Diaz MA, Mandal D, Raza S, Weidler EM, Qin X, Coarfa C, Milosavljevic A, Petrosino JF, Highlander S, Gibbs R, Lynch SV, Shulman RJ, Versalovic J. 2011. Gastrointestinal Microbiome Signatures of Pediatric Patients With Irritable Bowel Syndrome. *Gastroenterology*. 141:1782-1791.
 95. LeBlanc JG, Milani C, de Giori GS, Sesma F, van Sinderen D, Ventura M. 2013. Bacteria as vitamin suppliers to their host: a gut microbiota perspective. *Current Opinion in Biotechnology*. 24:160-168.
 96. Rørvik LM. 2015. Andre Gram-positive bakterier, p. 208-209. *In: Granum PE (ed.), Matforgiftning: smitte gjennom mat og vann*. Cappelen Damm AS, Norge.
 97. Torimura M, Kurata S, Yamada K, Yokomaku T, Kamagata Y, Kanagawa T, Kurane R. 2001. Fluorescence-Quenching Phenomenon by Photoinduced Electron Transfer between a Fluorescent Dye and a Nucleotide Base. *Analytical Sciences*. 17:155-160.
 98. Pfeffer SR, Stahl SJ, Chamberlin MJ. 1977. Binding of *Escherichia coli* RNA Polymerase to T7 DNA. Displacement of Holoenzyme from Promoter Complexes by Heparin. *The Journal of Biological Chemistry*. 252:5403-5407.
 99. Větrovský T, Baldrian P. 2013. The Variability of the 16S rRNA Gene in Bacterial Genomes and Its Consequences for Bacterial Community Analyses. *PLoS ONE*. 8:e57923.
 100. Kembel SW, Wu M, Eisen JA, Green JL. 2012. Incorporating 16S Gene Copy Number Information Improves Estimates of Microbial Diversity and Abundance. *PLoS Computational Biology*. 8:e1002743.
 101. Carlos N, Tang YW, Pei Z. 2012. Pearls and pitfalls of genomics-based microbiome analysis. *Emerging Microbes & Infections*. 1:e45.

102. Vos M, Quince C, Pijl AS, de Hollander M, Kowalchuk GA. 2012. A Comparison of *rpoB* and 16S rRNA as Markers in Pyrosequencing Studies of Bacterial Diversity. *PLoS One*. 7:e30600.
103. Kryachko Y, Semler D, Vogrinetz J, Lemke M, Irvine R, Davidson J, Links MG, McCarthy EL, Haug B, Hemmingsen SM. 2017. Analyses of 16S rRNA and *cpn60* gene sequences provide complementary information about potentially useful and harmful oil field microbiota. *International Biodeterioration & Biodegradation*. 123:320-327.
104. Lindahl BD, Nilsson RH, Tedersoo L, Abarenkov K, Carlsen T, Kjølner R, Kõljalg U, Pennanen T, Rosendahl S, Stenlid J, Kauserud H. 2013. Fungal community analysis by high-throughput sequencing of amplified markers – a user's guide. *The New Phytologist*. 199:288-299.
105. Laszlo M, Ciobanu L, Andreica V, Pascu O. 2016. Fecal transplantation indications in ulcerative colitis. Preliminary study. *Clujul Medical*. 89:224-228.
106. Halmos EP, Power VA, Shepherd SJ, Gibson PR, Muir JG. 2014. A diet low in FODMAPs reduces symptoms of irritable bowel syndrome. *Gastroenterology*. 146:67-75.
107. Chumpitazi BP, Cope JL, Hollister EB, Tsai CM, McMeans AR, Luna RA, Versalovic J, Shulman RJ. 2015. Randomised critical trial: Gut microbiome biomarkers re associated with Clinical Response to a Low FODMAP Diet in children with irritable bowel syndrome. *Aliment Pharmacol Ther*. 42:418-427.
108. Daïen CI, Pinget GV, Tan JK, Macia L. 2017. Detrimental impact of microbiota-accessible carbohydrate-deprived diet on gut and immune homeostasis: An overview. *Frontiers in Immunology*. 8:548.
109. Vétizou M, Pitt JM, Daillère R, Lepage P, Waldschmitt N, Flament C, Rusakiewicz S, Routy B, Roberti MP, Duong CPM, Poirier-Colame V, Roux A, Becharef S, Formenti S, Golden E, Cording S, Eberl G, Schlitzer A, Ginhoux F, Mani S, Yamazaki T, Jacquilot N, Enot DP, Bérard M, Nigou J, Opolon P, Eggermont A, Woerther PL, Chachaty E, Chaput N, Robert C, Mateus C, Kroemer G, Raoult D, Boneca IG, Carbonnel F, Chamaillard M, Zitvogel L. 2015. Anticancer immunotherapy by CTLA-4 blockade relies on the gut microbiota. *Science*. 350:1079-1084.
110. Ribault S, Harper K, Grave L, Lafontaine C, Nannini P, Raimondo A, Faure B. 2004. Rapid Screening Method for Detection of Bacteria in Platelet Concentrates. *Journal of Clinical Microbiology*. 42:1903-1908.
111. McDonald C, Allen J, Brailsford S, Roy A, Ball J, Moule R, Vasconcelos M, Morrison R, Pitt T. 2017. Bacterial screening of platelet components by National Health Service Blood and Transplant, an effective risk reduction measure. *Transfusion*. 57:1122-1131.

Supplementary material

Demonstrations of true positive signals detected for all probes carrying the HEX fluorophore that were found to be compatible with LAD in this thesis can be seen in figure S1. This figure shows a combination of representative melting curves collected from 20 pure culture samples and two negative controls.

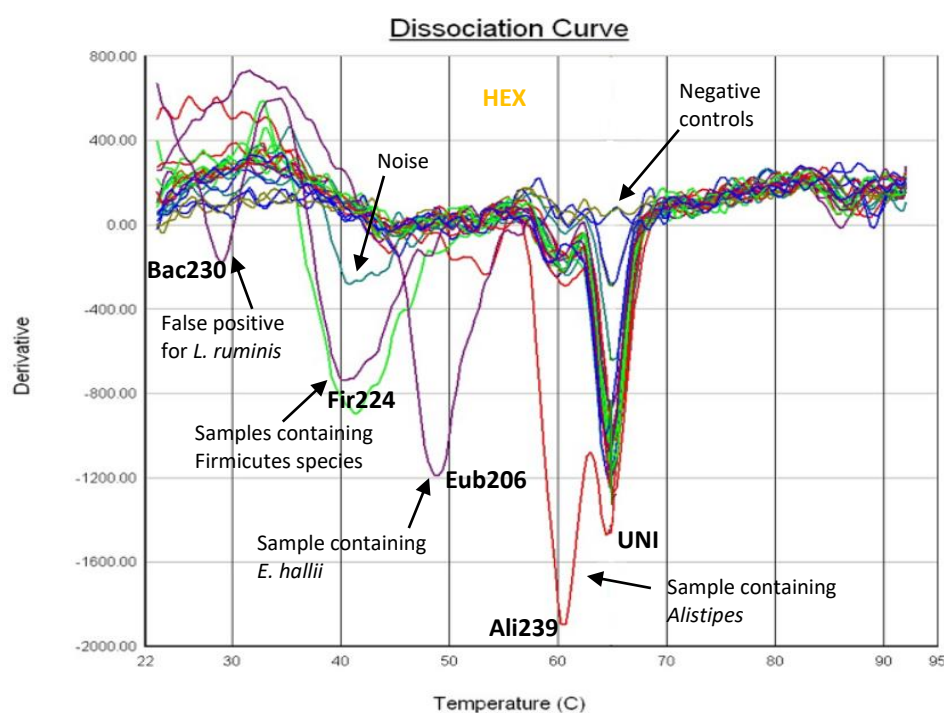


Figure S1: Combination of representative melting curves demonstrating signals from all the probes found to be compatible with LAD from the HEX channel. The temperature is presented on the x-axis, and goes from 22 °C to 95 °C. The negative first derivative of the fluorescence and temperature measurements ($-dF/dT$) is shown on the y-axis, and goes from -2000 to 800. Negative peaks indicate signal detection. Each colored line represents one pure culture sample, and as 22 samples were investigated (in which two were negative controls), 22 colored lines are visible. The software had a limited number of colors available, and some samples are therefore presented with very similar, or the same, colors, making them hard to distinguish. However, the temperature of a peak reveals the identity of the probe. A small negative peak on a violet line is visible just below 30 °C, corresponding to probe Bac230. This signal was a false positive in response to *L. ruminis*, observed in all tests performed. The signal for probe Fir224, designed for the Firmicutes phylum, can be seen at around 40 °C, and shows detection in two samples (violet and light green lines). In addition, a weak signal can be seen in a third sample (dark green line) which did not contain a complementary bacterium. This signal was determined to be noise, as it was not seen in the sample when the test was repeated. The peak for probe Eub206, designed to detect *E. hallii* at a theoretical T_m of 46.8 °C, is visible just below 50 °C in the figure (purple line). This is the same melting curve as the one seen in figure 5a. Probe Ali239 has a peak at almost exactly 60 °C (red line) and is detected in a sample containing an *Alistipes* species. This red line first shows a peak at 60 °C for Ali239, before a second peak is seen at roughly 65 °C for the UNI probe.

Figure S2 presents a combination of representative melting curves from the ROX channel collected from 22 samples (of which 20 were pure cultures and two were negative controls), and shows the three probes designed for that channel that were found to be compatible with LAD in this thesis.

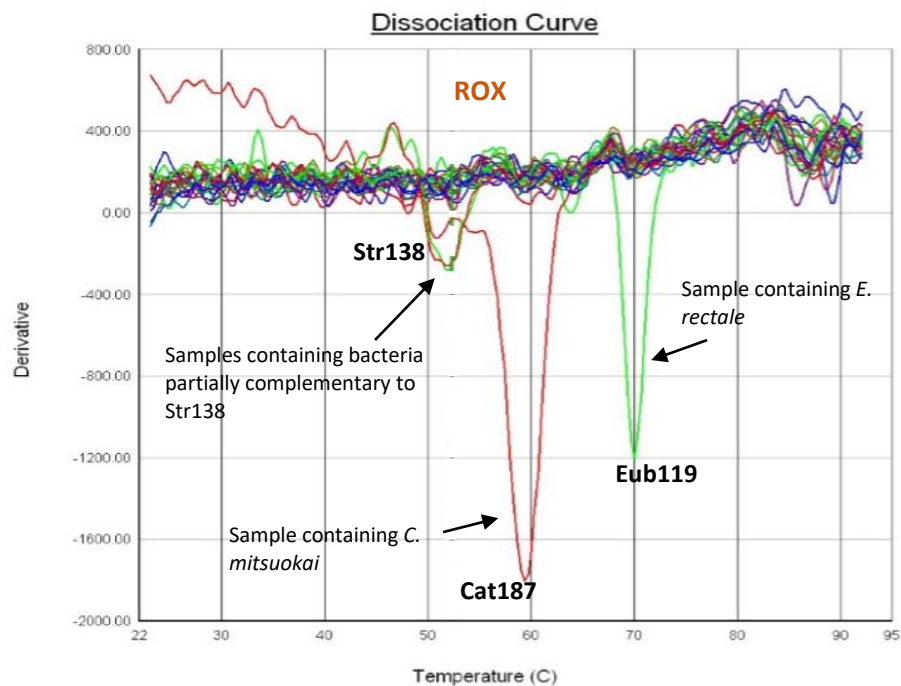


Figure S2: Combination of representative melting curves showing the detected signals of all the probes that were compatible with LAD on the ROX channel. The temperature is presented on the x-axis and goes from 22 °C to 95 °C. The negative derivative of the fluorescence and temperature measurements ($-dF/dT$) is on the y-axis and goes from -2000 to 800. Signal detection is indicated by negative peaks. Each sample is represented by a colored line. The software had a limited number of colors available, and some samples are therefore presented with very similar, or the same, colors, making them hard to distinguish. However, the temperature of a peak reveals the identity of the probe. Probe Str138 can be seen to give a weak signal for two samples (red and light green lines) just after 50 °C, both of which were expected based on the bacterial species used. The bacteria were not perfectly complementary to the probe, and later a species that was a better fit was obtained, which was used in the mixed sample. This provided a deeper signal for Str138, which can be seen in figure 6b in this thesis. Cat187 (crimson line), designed for a theoretical T_m of 58.8 °C, demonstrates a deep signal just below 60 °C in response to a sample containing *C. mitsuokai*. Probe Eub119 (green line) provides a deep peak at exactly 70 °C, and is the same melting curve as the one seen in figure 5b, collected from a sample containing *E. rectale*. The probes Bac155 at theoretical T_m 29 °C, and Dor124 at theoretical T_m 40.3 °C, both designed for the ROX channel, did not provide a distinguishable signal at any point during testing.

Figure S3 presents a demonstration of detected signals for all the probes found to be compatible with LAD on the CY5 (a) and FAM (b) channels. The representative melting curves shown in the figure are collected from a mixed sample, four pure cultures, and a negative control.

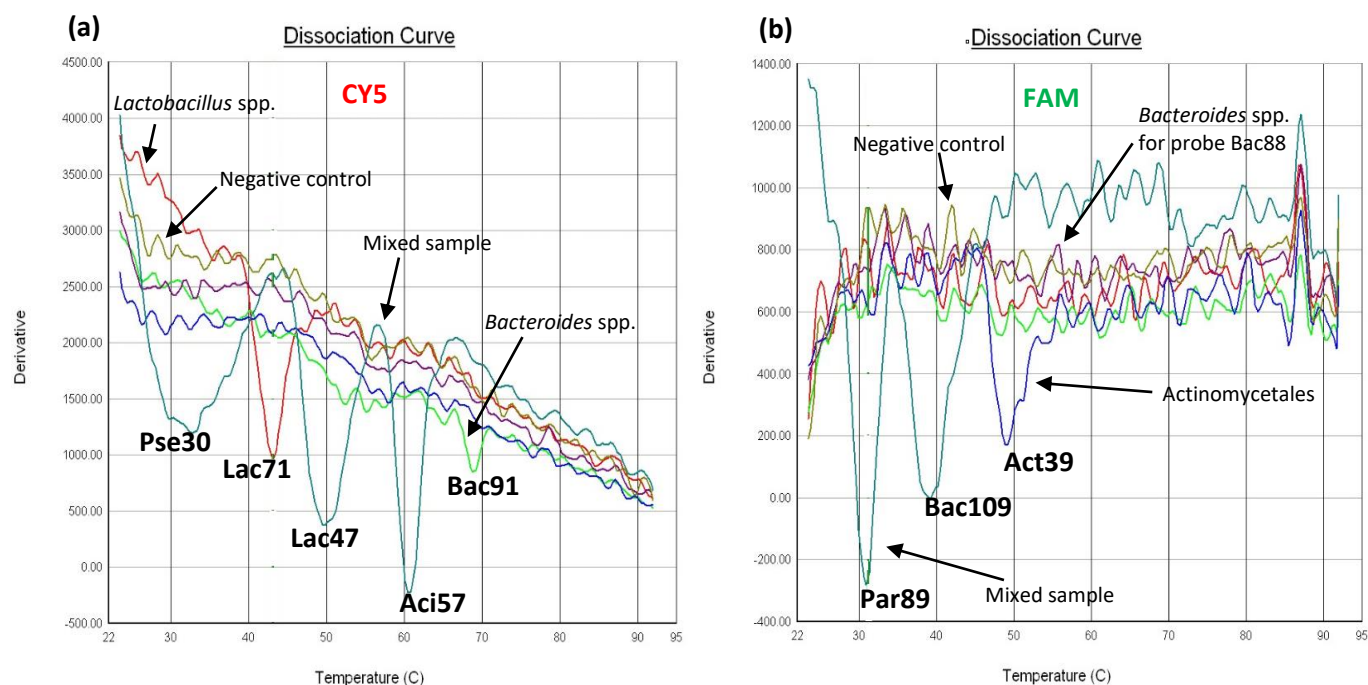


Figure S3: Representative examples of melting curves showing the signals of all the probes for the CY5 (a) and FAM (b) channels that were found to be compatible with LAD under the current experimental setup. Both panels consist of a combination of melting curves collected from a mixed sample (turquoise line), four pure cultures (red, purple, blue and green lines), and one negative control (yellow line). The temperature is presented on the x-axis and goes from 22 °C to 95 °C for both panels. The negative derivative of the fluorescence and temperature measurements ($-dF/dT$) is shown on the y-axis, and goes from -500 to 4500 on the CY5 channel, and from -400 to 1400 on the FAM channel. All five probes designed for the CY5 channel were found to be compatible with LAD under the current experimental setup, as can be seen in (a). On this channel the mixed sample contained species that lead to a signal with probes Pse30 (30 °C), Lac47 (50 °C) and Aci57 (60 °C), as seen by the turquoise line. In addition, two of the four pure cultures contained species targeted by CY5-labelled probes, where one contained bacteria complementary to probe Lac71 (43 °C, red line), and the other to Bac91 (69 °C, light green line). The purple and blue lines represent samples that were not expected to provide a signal on the CY5 channel. On the FAM channel, presented in panel (b), Par89 (31 °C) and Bac109 (39 °C) both show deep signals on the turquoise line from the mixed sample. In addition, probe Act39 shows a peak on the blue line at 49.3 °C, which represents a sample containing a pure culture of an Actinomycetales species. The purple line represents a sample containing a *Bacteroides* species aimed at probe Bac88 (70 °C) on the FAM channel. This line shows no peak, as the species used was not perfectly complementary to the probe. The species was chosen because the target bacteria was not available during testing.

Table S1 and S2 presents the raw data of the probe peak values detected in the fecal samples investigated in this thesis. Table S1 shows the values for the normobiotic samples (DI1 and DI2), while table S2 shows the values for the dysbiotic samples (DI3-DI5).

Table S1: Raw data of the probe peak values from the 15 normobiotic fecal samples. The samples have previously been classified by GA. Only the probes that were determined to be working with LAD were included in the table.

Sample	Dysbiosis index	Fir224	Eub206	Ali239	UNI	Str138	Cat187	Eub119	Pse30	Lac71	Lac47	Act57	Bac91	Par89	Bac109	Act39
1	DI1	-56,33	-35,98	-670,95	-1418,08	212,78	288,39	59,88	4368,11	4119,26	3480,35	2669,34	1885,81	1030,32	1030,30	1363,12
2	DI1	22,48	5,79	-428,79	-1504,56	307,85	321,97	386,68	4958,67	4412,23	3361,38	2899,34	2036,91	1146,29	1252,79	1453,27
3	DI1	87,98	102,39	-273,71	-1664,84	272,46	253,91	436,65	6011,76	4902,58	3679,03	3343,14	2323,23	1969,01	1355,07	1347,63
4	DI1	-113,32	-16,89	-531,01	-1553,46	260,52	310,46	305,06	4782,32	3796,48	3303,20	2737,08	2050,08	1664,17	864,86	1233,63
5	DI1	-241,52	114,30	-534,90	-1620,00	344,46	299,93	190,27	5089,61	4461,14	3499,66	3038,20	2177,06	1326,86	1273,74	1469,10
6	DI1	-325,64	90,09	-730,34	-1740,24	390,50	247,04	245,36	5164,33	4109,22	3328,78	2707,23	2141,71	1507,24	970,38	1319,49
7	DI1	-473,58	-15,76	-844,92	-1509,17	294,90	296,44	37,44	5129,02	3957,00	3396,40	2795,56	2062,51	1499,97	977,93	1254,82
8	DI1	-158,01	8,82	-721,19	-1741,60	268,70	174,23	70,10	5302,97	3811,98	3207,78	2679,47	2000,30	848,73	800,68	1311,22
9	DI2	-87,28	-82,85	-515,38	-1556,90	395,78	306,62	412,78	4930,90	4603,42	3511,15	3068,70	2184,61	1001,82	1060,85	1437,45
10	DI2	290,10	221,31	-219,76	-1524,19	347,95	321,61	532,46	6209,26	5204,52	3671,84	3323,94	2585,43	1457,61	1385,08	1579,07
11	DI2	-890,34	77,86	-203,64	-1636,05	324,19	406,85	-17,01	5042,41	3949,82	3263,37	3003,11	2170,88	1304,08	1002,51	1326,63
12	DI2	-486,55	83,16	-375,87	-1361,85	254,31	248,94	10,29	4788,55	3271,16	2827,09	2463,62	1963,06	1545,81	642,10	997,19
13	DI2	-352,87	-22,83	-405,35	-1726,56	323,08	363,33	212,97	5897,37	4560,67	3540,37	3088,58	2271,41	1582,77	1073,13	1323,26
14	DI2	-935,92	138,23	-317,57	-1453,50	453,27	430,90	-443,85	5065,38	4522,44	3464,17	2900,17	2138,67	1292,39	1046,58	1373,70
15	DI2	-322,59	-21,78	-314,81	-1566,30	278,78	206,12	411,05	5889,48	4424,68	3410,59	2905,60	2009,77	1910,08	905,24	1178,49

Table S2: Raw data of the probe peak values from the 17 dysbiotic fecal samples. The samples have previously been classified by GA. Only the probes that were determined to be working with LAD were included in the table.

Sample	Dysbiosis index	Fir224	Eub206	Ali239	UNI	Str138	Cat187	Eub119	Pse30	Lac71	Lac47	Act57	Bac91	Par89	Bac109	Act39
16	DI3	-1006,22	81,74	-234,38	-1728,29	381,63	282,94	-357,56	5756,81	4832,58	3744,76	3128,13	2221,43	1394,84	1208,57	1382,99
17	DI3	-27,32	154,11	-547,44	-1547,20	416,61	365,45	526,69	5464,44	4631,66	3836,63	3350,81	2476,97	1301,24	966,47	1442,95
18	DI3	-421,33	-139,44	-600,27	-1669,57	167,09	132,08	273,26	5493,60	4541,21	3845,19	3129,12	2151,20	1461,01	865,44	1102,39
19	DI3	-83,32	129,58	-608,32	-1560,31	490,08	552,80	273,48	5101,59	4941,00	3838,81	3377,77	2334,11	92,84	1289,38	1497,89
20	DI3	-483,44	-315,00	-637,57	-1432,00	263,96	371,14	-35,81	4496,52	4117,84	3415,09	2867,70	2209,34	726,84	845,80	1198,83
21	DI3	-16,81	-61,03	-520,29	-1507,17	307,65	226,94	463,45	5462,34	4073,67	3193,75	2757,59	2031,47	1689,00	962,82	1260,11
22	DI3	-63,51	3,71	-235,97	-1628,39	254,74	444,22	477,44	5204,41	4294,65	3452,88	3073,53	2214,98	1333,78	963,93	1201,85
23	DI4	-296,89	-108,43	-516,21	-1566,08	451,10	279,49	234,61	4583,10	3939,83	3335,23	2849,66	1995,88	711,54	967,87	1346,09
24	DI4	116,97	55,68	-115,64	-1300,48	405,74	372,47	482,38	3883,01	4148,20	3384,60	2822,42	2161,30	1044,28	1067,44	1347,75
25	DI4	237,60	161,37	-466,45	-1380,73	525,26	528,75	523,90	5457,29	4913,94	3764,69	3200,57	2417,96	1303,40	1152,73	1617,80
26	DI5	-806,93	107,99	-80,48	-1381,25	161,71	345,83	420,40	4595,35	4080,77	3262,92	2726,37	1954,71	1046,91	1099,82	1185,07
27	DI5	-124,02	77,32	-60,68	-1307,58	392,21	328,35	359,25	5091,59	4089,12	3379,00	2852,14	2014,23	1340,51	935,69	1125,75
28	DI5	-63,71	195,85	-49,53	-1138,13	385,14	294,48	398,15	4339,31	3692,99	3143,03	2630,22	1991,38	1009,95	822,52	1195,87
29	DI5	180,02	299,39	-74,37	-1258,43	518,51	566,84	597,30	5060,23	4677,16	3438,68	2919,13	2096,44	1190,88	1252,03	1431,83
30	DI5	102,41	237,33	-87,05	-1093,82	425,33	469,11	602,86	5085,05	4125,59	3256,90	2933,87	2105,57	1404,81	1065,51	1238,84
31	DI5	-282,52	235,47	-22,40	-1333,23	675,89	609,77	160,59	5330,70	4386,26	3530,68	3000,12	2244,08	1096,07	1118,22	1424,37
32	DI5	-63,59	80,31	-117,41	-1337,59	272,04	390,26	588,19	5808,06	3678,54	3105,12	2739,20	2086,05	1973,76	755,20	1081,24

Table S3 lists the p-values for all the probes from the Welch Two Sample t-test performed in this thesis.

Table S3: p-values derived from the Welch Two Sample t-test. This test was performed to investigate the statistical significance of the differences in average signals detected for each probe in the normobiotic versus the dysbiotic samples. A p-value higher than 0.05 indicated no statistical significance, while a p-value below 0.05 meant the result was considered statistically significant.

Probe name	Fir224	Eub206	Ali239	Str138	Cat187	Eub119	Pse30	Lac71	Lac47	Act57	Bac91	Par89	Bac109	Act39
p-value	0.4709	0.5354	0.02664	0.07975	0.02263	0.07654	0.3678	0.8504	0.3966	0.517	0.6557	0.1004	0.7292	0.5319



Norges miljø- og biovitenskapelige universitet
Noregs miljø- og biovitenskapelige universitet
Norwegian University of Life Sciences

Postboks 5003
NO-1432 Ås
Norway