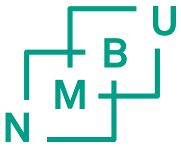

N M B U

Dedication and Acknowledgement

I would like to dedicate this humble work to my beloved mother, my younger sister, my dad's pure spirit, and my beloved wife, since they have been with me along in this project. I would like to acknowledge my advisor – Prof. Tor Lea, and the academic staff in this great establishment for their support and guidance, especially my academic advisor. I would like to thank all the other academic and non-academic staff in this fine establishment for their hospitality.

Farhad Mahmod

ÅS, mai 2014

Abstarct

This thesis discusses the culture-dependent and culture-independent methods for the analysis of the human intestinal microbiota, with a special focus on the New Generation Sequecing procedures. The discussion includes the comparission of the principles, advantages, and disadvantages of these techniques. These techniques include conventional methods and novel methods which depend on molecular sequencing, such as the first generation and next generation sequencing techniques with different platforms of next generation sequencing tecnologies. Library data collection methods were used such as books, textbooks, scientific journal articles, and online library databases such as ISI Web of Science, Bibsys, PubMed, Google Scholar, and Google were used for this literature review paper. In conclusion, choosing among the increasing number of methods of analyzing human intestinal microbiota depends upon the goal and the target that is needed to be achived; sinces advantages and disadvantages are associated with all of these methods.

Sammendrag

Denne oppgaven diskuterer kultur-avhengige og kultur-uavhengigmetoder for analyse av den menneskelige tarmfloraen(intestinal microbiota), med et spesielt fokus på den nye generasjonen Sequecing prosedyrer. Diskusjonen omfatter sammenligning av prinsippene , fordeler og ulemper ved disse teknikkene. Disse teknikkene omfatter tradisjonelle metoder og nye metoder som er avhengig av molekylær sekvensering , slik som de første generasjon og neste generasjons sekvensering teknikker med ulike plattformer av neste generasjons sekvense tecnologies . Datainnsamlingsmetoder ble brukt for eksempel bøker , lærebøker, vitenskapelige tidsskriftartikler og elektroniske databaser som ISI Web of Science , Bibsys , PubMed , Google Scholar og Google ble brukt for denne litteratur review papir . I konklusjonen , velg blant det økende antall metoder for å analysere menneskelige tarmfloraen avhenger av målet , og målet som er nødvendig for å bli oppnådde ,fordi fordeler og ulemper forbundet med alle disse metodene.

Table of Contents

Abstarct 2

Table of Contents 3

Summary 8

Abbreviations 11

1Introduction 12

1.1 Historical Background of Next Generation Sequencing (NGS) 14

1.2 Important contributors on microbiota methodology 17

1.2.1 O'Sullivan (2000) 17

1.2.2 Kleessen et al.(2000) 17

1.2.3 McCartney (2002) 17

1.2.4 Zoetendal et al.(2004) 18

1.2.5 Clarridg (2004) 18

1.2.6 Amor et al.(2007) 18

1.2.7 Woo et al.(2008) 19

1.2.8 Hamady and Knight (2009) 19

1.2.9 Rogers and Bruce (2010) 19

1.2.10 Nikkila, and de vos, (2010) 20

1.2.11 Paliy and Agans(2012) 20

1.2.12 Shokralla (2012) 20

1.2.13 Guinane and Cotter (2013) 21

1.2.14 Cani (2013) 21

1.2.15 Salipante(2013) 21

1.2.16 Becker(2013) 22

1.2.17 Raoult (2013) 22

1.3 Taxa of microbiota 23

1.4 Functions of microbiota 23

Vitamin synthesis 24

Protective function 24

Metabolic activity 24

Intestinal epithelial cell proliferation and differentiation 24

Immune stimulation: 24

Signaling from the periphery to the brain 24

1.5 The Role of Microbiota in Certain Diseases and Disorders 25
2 Aim of the study 25
3 Methods 27
4 Results of investigations and discussion……………………………………………………29
4.1Classical approach 29
4.1.1 Culture-dependent methods 29
4.1.2 Culture-independent methods 33
4.1.2.1 Direct microscopic analysis 33
4.1.2.2 Monitoring of specific enzymes or metabolites in faecal samples. 33
4.2Molecular (Novel) approach……………………………………………………..33
A historical background 33
Introduction to Molecular Approach ……………………………………..…34
4.2.1Culture-dependent methods.. 40
4.2.1.1Phenotypic fingerprinting analysis……………………………………40
4.2.1.1.1Polyacrylamide gel electrophoresis of soluble protein……………...41
4.2.1.1.2Fatty acid analysis………………………………………………...41
4.2.1.1.3Bacteriophage typing……………………………………………..41
4.2.1.1.4Serotyping ……………………………………………………….…41
4.2.1.2Genotypic fingerprinting analysis..…………………………………..…42
4.2.1.2.1Colony hybridization with nucleic acid probes ……………………42
4.2.1.2.2Pulsed field gel electrophoresis (PFGE) …………………………..…42
4.2.1.2.3Ribotyping…………………………………………………….... 42
4.2.2Culture- Independent molecular methods ………………..……..……………43
4.2.2.1Fluorescence in situ hybridization …………………………….……..44
4.2.2.2Flow cytometry…………………………………….……………….….. 45
4.2.2.3Quantitative dot blot………………………….……………………….. 45
4.2.2.4 Probes……………………………….……………………………..….46
4.2.2.5 PCR-based techniques………………………………………….….……47
4.2.2.6 Checkerboard hybridization………………………………………… 48
4.2.2.7 Microarrays………………………………………………………… 49
4.2.2.8 16S rRNA gene as microbial molecular marker …………………..…50
4.2.2.9 Random Amplified Polymorphic DNA (RAPD)……………………50
4.2.2.10RecA gene sequence analysis ………………………………………51

4.2.2.11 Multiplex-PCR……………………………………......................... 51
4.2.2.12 Arbitrary primed (AP) PCR……………………………………….52
4.2.2.13 Triplet arbitrary primed (TAP) PCR ……………………………...52
4.2.2.14 DGGE/TGGE……………………………………………………..… 52
4.2.2.15 Gene cloning and sequencing…………………………………...….. 53
4.2.2.16 Real-Time PCR…………………………………………………….. 53
4.2.2.17 T-RFLP………………………………………………………..…….54
4.3 Sequencing………………………………………………………..………….. 55
4.3.1Sanger sequencing……………………………………………………………60
4.3.2 Shotgun sequencing ………………………………………………………60
4.3.3 Next Generation Sequencing Platforms Capable of RNA Sequencing.… 60
4.3.3.1 Roche 454 genome sequencers (pyrosequencing technology)……….. .61
4.3.3.2 Illumina sequencers……………………………………………….... 63
4.3.3.3 Applied Biosystems SOLiD sequencer (Life Technologies) ……….…64
4.3.3.4 Life Technologies Ion Torrent……………………………………... 65
4.3.3.5 Single-molecule DNA-sequencing technologies…………………….... 66
4.3.3.5.1 Helicos Biosciences HeliScope…………………………………... 66
4.3.3.5.2Pacific Biosciences SMRT DNA sequencing…..……………….……… 66
4.4Advantages and Disadvantages of NGS Platforms ………………………………...67
4.4.1Illumina and SOLiD systems ………………………………………………68
4.4.2 PCR-based NGS systems ……………………………………….….69
4.4.3Improvements on NGS Technologies ………………………………………69
4.5Application of NGS for Analysing DNA IntestinalMicrobiota……………………..…70
4.6Tools for functional studies of intestinal microbiota………………………………… 71
4.6.1Stable isotope probing (SIP)……….. …………………………………….…72
4.6.2‘Omics’……………………………………………………………………….. 72
4.6.2.1Metagenomics………………………………………………...……….. 75
4.6.2.2Metatranscriptomics ……………………………………………..….75
4.6.2.3 Metaproteomics……………………………………………….……. 75
4.6.2.4Metabolomics ……………………………………………….…………76
4.7Insertion sequencing (InSeq) ……………………………………………………..76
4.8 Animal models ……………………………………………………………….….77
4.9 Data analyses and bioinformatics…………………………………………………….. 77

4.10Microbial culturomics …………………………………………………………………..78
5 Conclusion……………………………………………………………………………78
6 References…………………………………………………………………………... 81

Figures:

Figure 1: Abundance of bacterial phyla in each segment of the human gastrointestinal tract. ……………………………………………………………………..………………………16
Figure 2: Overview of common human intestinal microbiota, bacterial, archaea, viral, and eukaryota……………………………………………………………………………………...23
Figure 3: Diseases and physical disorders associated or correlated with human intestinal microbiota…………………………………………………………………………….....25
Figure4: Compared findings using different techniques in one stool sample………………...31
Figure 5: Sequencing by both methods:Sanger and NGS…………………………………….56
Figure 6: This diagram shows the different NGS platforms technologies……………………61
Figure 7: Principles of pyrosequencing………………………………………………………62
Figure 8: Shows Roche 454 workflow………………………………………………………..63
Figure 9: Shows Illumine / SOLiD sample preparation……………………………………...65
Figure 10. Shows single molecule Helicos BioScience immobilization principle…….……66
Figure 11. Shows Pacific Bioscience immobilization principle…………………..………67
Figure 12: Represent different platforms of NGS…...71
Figure 13: The overlap among the 'Omics' platfporems ……………………………..…...73

Tables:

Table1 Advantages and disadvantages of culture dependent techniques. …………………..32
Table 2 Advantages and disadvantages of culturing independent techniques. ……………..38
Table 3 Next Generation Sequencing platforms ……………………………….………..57
Table 4 Comparison of Sanger sequencing method with some next generation Sequencing technologies……..58
Table 5 Comparison of currently available next-generation sequencing technologies ...……59
Table 6 Culturing independent techniques 'omics'……………………………………………..74

Summary

After surveying literature from books, journals found and websites there are many methods of characterizing the microbiota resident in the human body, especially the gastrointestinal tract. These methods can be divided as revealed from references surveyed into:

Classical approach:

- Culture-dependent methods
- Culture-independent methods

Most of the microbiota present the in human intestine is anaerobic which necessitates special requirements for cultivation which sometimes is difficult to obtain. Most of the literature concludes that classical methods are inefficient in identifying all the microbiota revealed in examination. Furthermore, classical culture-dependent technologies are less sensitive, time consuming, not accurate and with high cost of performance.

Novel culture-dependent methods include the phenotypic fingerprinting analysis which in turn extends to other techniques like polyacrylamide gel electrophoresis of soluble proteins, fatty acid analysis, bacteriophage typing and serotyping. Anyhow, phenotypic fingerprints are found by some authors to be less sensitive than genotypic fingerprints, and a change in fingerprint does not necessarily mean a different organism. Furthermore, it was known that the most rapid methods of the above mentioned procedures is serotyping as colonies can be typed directly, without subculturing by colony hybridization with monoclonal antibodies specific for a particular genus, species or even strain.

Other culture-dependent methods include also genotype fingerprinting analysis, which in turn comprises: colony hybridization with nucleic acid probes, pulsed field gel electrophoresis, and ribotyping. Noteworthy, pulsed field gel electrophoresis (PFGE) is based on an electrical pulse system which causes migration of very large DNA fragments through an agarose gel.

Regarding culture-independent molecular methods they are known as a standard phylogenetic classical technology which were applied, and they are more reliable for identification than culture-dependent because identification dependents on nucleic acids rather than depending on a given culture media, and the DNA can be detected from living or dead cells not depending on the need for growing bacteria. Anyhow the benefit of such technology

is the recognition and detection of single bacterial species in complex macrobiotics, and to analyze evolutionary relatedness of bacteria.

It has been pointed out by many authors that molecular methods can include fluorescence *in situ* hybridization, flow cytometry, quantitative dot blot hybridization and PCR-based techniques. From the literature it is revealed that the use of a genetic probe and fluorescence microscopy makes it possible to count the total number of bacteria present in human intestinal gut, and hybridization with fluorescent probes targeting 16S rRNA and inspection of hybridized bacteria with fluorescence microscopy is a good tool for examining multi species bacterial samples. An additional advantage of 16S rRNA hybridization is that it allows analyzing the bacterial cells that are intact morphologically. The usage of flow cytometry enables fast analysis of bacteria and is more accurate and reliable compared to microscopy. The Quantitative Dot blot technique was introduced to investigate bacterial diversity within samples from various environmental conditions also indicated the status of bacterial metabolic activity.

The Polymerase Chain Reaction procedure is a rapid as well a wide range procedure to detect bacterial species and become a consistent technique to detect microorganisms. The advantage is that the PCR-based finger printing techniques can use few cells, which means there is no need for culturing. Anyhow, as some researchers found that faeces, which contain bilirubin and bile salts, can inhibit PCR analysis for such phenomena bacteria in faeces usually requires total DNA or RNA purification.

Microarray is a perfect tool, which helps in analysis of both RNA and DNA of thousand of genes or same gene from thousand of organisms. A typical microarray experiment involves the hybridization of an mRNA (or DNA) molecule to the DNA template from which is originated. In addition this technology besides used in molecular biology can be used also in medicine.

As regarding recA gene sequences analysis is useful for determining intrageneric phylogenetic relationship. Anyhow the disadvantage of this technique is that prior sequence knowledge is required. Denaturing gradient gel electrophoresis (DGGE) is suitable method to analyze complex microbial ecosystem diversity. In this method the DNA fragments can be separated even if they are of the same length. While the other closely related technique, i.e. temperature gradient gel electrophoresis (TGGE), DNA fragments are separated in temperature gradient. Furthermore, now it is known that temperature gradient gel electrophoresis allows the analysis of predominant bacteria that are difficult to culture.

Real-time PCR or quantitative real-time Polymerase Chain Reaction (qPCR) is a culture-independent molecular based method. This technique, as concluded by many authors, is used for the detection and the quantification of a strain without using further post PCR analysis. The main useful characters of this technique is its speed and its possibility to detect minor populations of bacteria within a large population, and from other finding a conclusion was drawn to that both qPCR or rT-q PCR are low cost and suitable for daily routine analysis.

RNA sequencing is a high throughput sequencing method to obtain sequence cDNA to get information about RNA content of sample. RNA sequencing technique offers two advantages: first it offers more than the detection of transcripts corresponding to the existing genomic sequences. The second advantage is that it has a large dynamic range of expression levels allowing the transcripts to be detected.

Several recent techniques were developed to study microbiota through its molecular function. These techniques includes fingerprinting of 16S r RNA gene amplicons, DNA sequencing of 16 S rRNA gene clones, FISH, flow cytometry, DNA microarrays and high throughput sequencing with 16 S rRNA genes as the target. The use of targeting 16 S rRNA genes has provided insights into the function of microbiota and their influence human health.

“Omics” is a group of techniques that are useful for a wide range of microbiota communities. This technique as found by some authors recently include metagenomics, metatranscriptomics, metaproteomics and metabolomics. These techniques can be used to analyze proteins, DNA, mRNA and different metabolites of gut microbiota and generally analyzing the functions of ecosystem of gut microbiota.

Metagenomics was used to analyze the composition and function of gut microbiota by sequencing information from the combined genomes of the microbiota.The resulting advantage of this technique is its high throughput and capacity to recognize new functional genes, while the disadvantage of this method is that it cannot distinguish DNA from dead cell and DNA from live cells.

Insertion Sequencing (INseq) is a mixed technique from genome (wide transposon metagenesis) and parallel sequencing on a larger scale. INseq is, used for functional genome, which is a wide analysis of microbiota. Transposons with recognizable DNA bar code were used to introduce mutations into thousands of bacteria.

Data analysis and bioinformatics, this method is valid for the study of determining the evolutionary relations between microorganisms resident the gut. As found by some researches that the in the alignment based sequence alignment against database such as ARB, and Ribosomal database project II (RDP II). Furthermore most popular approach for making

alignments is the CLUSTAL online software and database such as NCBI. The advantage of this method is its accuracy to obtain a map of phylogenetic relationship, while the disadvantage is not useful for analysis of large set of data.

Abbreviations

16S	16 Svedberg
AFLP	Amplified fragment length polymorphism
BBA	Brucella blood agar
BHI	Brain heart infusion
CF	Cystic fibrosis
DGGE	Denaturing gradient gel electrophoresis
DNA	Deoxyribonucleic acid
DsDNA	Double stranded DNA
ELISA	Enzyme-Linked immunosorbent assay
FCM	Flow cytometry
FISH	Fluorescent in situ hybridization
GALT	Gut associated lymphoid tissue
GF	Germfree
GI	Gastrointestinal
GIT	Gastrointestinal tract
HTS	High throughput sequencing
INseq	Insertion sequencing
NGS	Next generation sequencing
NMR	Nuclear magnetic resonance
PCR	Polymerase chain reaction
PFGE	Pulsed field gel electrophoresis
PGM	Personal genome machine
PYG	Peptone-yeast extra-glucose
rDNA	Ribosomal deoxyribonucleic acid
RDPII	Ribosome database project II
rRNA	Ribosomal ribo nucleic acid
RAPD	Random Amplified Polymorphic DNA

SCFA Short-chain fatty acid

SMS Single molecular sequencing

SNP Single nucleotide polymorphism

TGGE Temperature Gradient Gel Electrophoresis

T-RFLP Terminal restriction fragment length polymorphism

1. Introduction

The human body harbors a highly complex mixture of microorganisms. Most investigators in this field estimate the number of this complex to be about 10^{14} which is 10 times the total number of human cells in an individual. The most common sites in which microorganisms reside are skin, mouth, nose, ears, vagina, and gastrointestinal tract. However it should be mentioned in this regard that each site or location in human body is harbored by a specialized group of microbes depending on the physiochemical characteristics of the body site (Dethlefsen et al., 2007). Moreover, the majority of the microbiota of humans has been found to reside in the GI tract, where microbial abundance is known to be in the colon, accounting for about 10^{11} cells per ml. Bike (2009) found that this complex ecosystem consists of a variety of microorganisms (e.g. bacteria, archaea, yeasts and other eukaryotes).

Qin and his coworker (2010) concluded in their study on the microbiota of the gastrointestinal tract that it is made of approximately 500-1000 species. The anaerobic bacteria constituting the majority of these organisms and they outnumber the aerobic and facultative anaerobic bacteria by 100 -1000 folds. It has been reported by Qin et al. (2010) that the most common organisms encountered belong to two phyla, namely Firmicutes and Bacteriodetes. However, other species present are members of the phyla Proteobacteria, Verrumicrobia, Actinobacteria, Fusobacteria and Cyanobacteria.

The investigations were expanded by Sekirov and his associates (2010). The study showed that two gradients of microbial distribution can be found in the gastrointestinal tract. The first gradient includes the increase in the microbial density both from the proximal to the distal gut (e.g. the stomach content is 10^2 cells/g, the duodenum 10^3 cells/g, the jejunum 10^4 cells/g, the ileum 10^7 cells/g and the colon up to 10^{12} cells/g) (Figure1). The second gradient revealed increases in the bacterial diversity in the same axes and manner as microbial density.The colonizing microbiota provides resistance to pathogenic bacteria and also

supports the development of the immune system of the host. This includes the development of gut-associated lymphoid tissue (GALT), mucosal immunity and the oral tolerance. The microbiota population in the gut, in one way or another influences the nutritional, physiological and immunological status of the host. However, much is unknown about their composition as well as their interaction with each other and with the host cell physiology, and much remains to fully understand and uncover about these microorganisms. This calls for thorough and comprehensive methods and procedures that could allow us to understand this ecosystem thouroughly. Advanced methods and techniques are essential to understand fully, the composition, activities, and their relationship to each other and with the host.

The main finding and investigation of microbiota have been achieved by application of three techniques, namely the culture -dependent methods, culture- independent methods and the development of germ-free animal models. There are many culture-dependent methods and procedures employed to isolate and identify microbiota with the aim of studying the abundance, composition and interactions with each other and the host. These conventional culture-dependent isolation and identification methods are among the most important procedures to study the microbiota. They contribute greatly to the understanding of the microbiota. However, they are time and labor consuming and are relatively costly. Above all, there are some viable but uncultivable microorganisms in the microbiota population, which could not be identified by these conventional methods. Zoetendal et al., (2004) described the reasons for such limitations of the technique as follows: “Some microbiota growth requirements are not well identified and could not be isolated in the media and the selectivity of some media favour only the growth of some specific microbiota by depressing the others”.

The above stated disadvantages of the culture-dependent methods called for the development of culture-independent methods. These culture-independent methods have undergone extensive developments and have been applied widely in studying and investigating microbiota for better understanding. One of the widely applied molecular techniques for the study of microbiota diversity and composition is sequencing of 16S rRNA genes. The sequencing of this gene is not only a helpful molecular technique to study the diversity and composition, but it also helps the classification of the microbiota according to its genetic relationships (Clarridge & Jill, 2004). To date, several molecular culture-independent techniques are developed and applied to investigate the microbiota abundance and composition. The premolecular microbiota research over the last 40 years was based on the classical bacterial culture methods and the use of a relatively simple microscopic examination.

DNA sequencing, a method developed by Gilbert and Frederick in (1977), caused a rapid change in the development of vaccines, medical treatments and diagnostic methods (Sanger et al., 1977). Recent advances in studying the gut microbiota have made revolutionary technologies characterized by high efficiency and rapid identification of microbiota. These advanced techniques and methods include quantitative polymerase chain reaction (Q-PCR) analysis, PCR-based DNA profiling techniques, DNA microarray, flow cytometry, insertion sequencing, and particularly next-generation DNA sequencing.

In addition, the above mentioned techniques provide some reliable data leading to further extensive and comprehensive studying of the intestine microbiota and its functions in the host. Moreover, the genetic approach may further improve our understanding of the gastrointestinal tract microbiota (Gong and Yang, 2012). The goal of this review is to provide a brief summary of the early work based on the classic microbiological techniques that provide a context for the molecular work published over the last 10 years, which will then be summarized in the following section of the review.

1.1 Historical Background of Next Generation Sequencing (NGS)

The analysis of the genome is a comprehensive, sensitive, and efficient tool for evaluating evolution, function, ecology, and biodiversity, spatially of the intestinal microbiota. It is less time consuming because it doesn't need laboratory cultivation and/or isolation of individual specimens. Although, some techniques like Sanger DNA sequencing technology is inefficient with complex samples because it sequence specimens individually, but it led to advances in this field to be more comprehensive of the specimen analysis. The more advanced DNA barcoding produces DNA library in order to identify an unknown specimen that could work as standardized species-specific genomic regions (DNA barcodes). Other DNA sequencing techniques are efficient with complex environmental samples and more suitable for larger scale studies than Sanger DNA sequencing technique.

Conventional DNA sequencing helped to produce and develop large DNA barcode reference libraries so that the next generation sequencing (NGS) can identify an environmental sample and read DNA from multiple templates in parallel, which is beyond the capacity of the conventional DNA sequencing methods .This process of simultaneous multiple reading of DNA reduces time and cost. NGS platforms were introduced in 2005 and can recover DNA sequence data directly from environmental samples, these data have been applied in studies like the comparison of microbiota between healthy individuals and patients (Hajibabaei et al., 2011).

Sanger et al., (1977) first introduced the conventional DNA sequencing approach. It can recover up to 1 kb of sequence data from a single specimen at a time and the most advanced version can perform sequencing of up to 1 kb for 96 individual specimens at a time. Based on different chemistries and detection techniques, NGS techniques are capable to generate a massive amount of sequencing reads in parallel. There are several approaches of NGS for example; genome sequencing that can generate reads from fragmented libraries of a specific genome. Some NGS can generate reads from a pool of cDNA library fragments generated through reverse transcription of RNA molecules. Some can generate reads from a pool of PCR-amplified molecules. NGS platforms don't use the vector-based procedure which is used to amplify and isolate DNA templates, therefore some of the cloning bias issues can be avoided that could affect sequencing evenness. Despite the strengths, NGS platforms can have their own associated limitations. The first challenge is the length and accuracy of the sequencing output, the second is the cost and the labor expended of the total output of the sequencing experiment, the third is the pre sequencing amplification step, finally, sources of PCR bias.

There are two categories of NGS technologies: the first group is PCR based technologies and single molecule sequencing (SMS) technologies presequencing amplification (Zhang et al., 2011). This paper will briefly describe the available NGS platforms in each category and their strengths and weaknesses. Figure 1 shows the gastrointestinal part of human body and of which the intestinal part is cover by this paper.

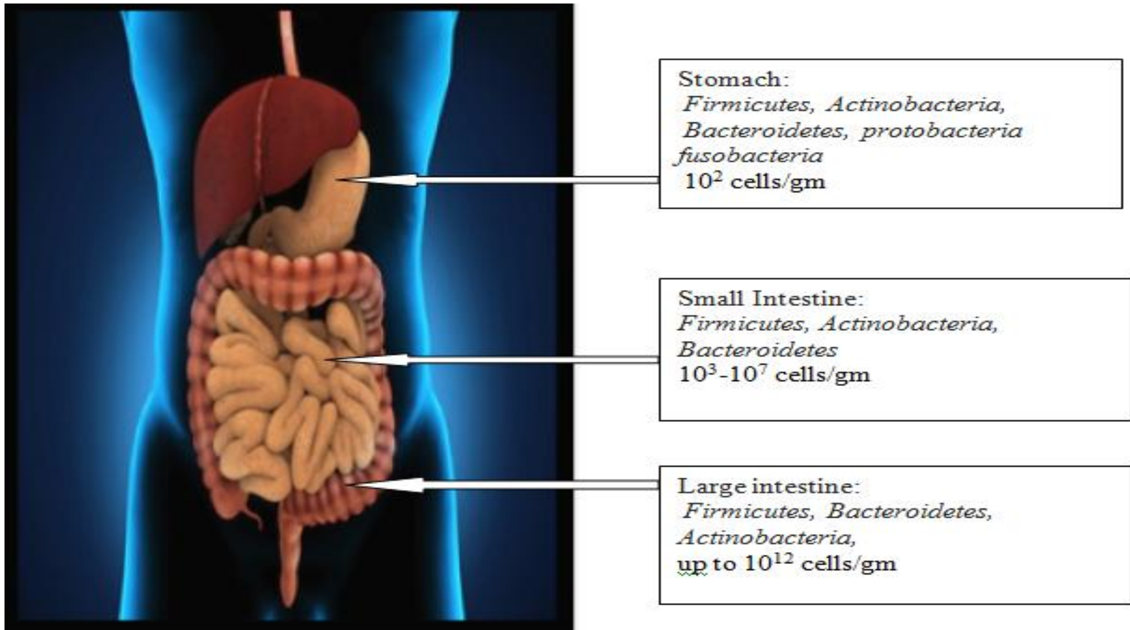


Figure 1: Abundance of bacterial phyla in each segment of the human gastrointestinal tract. (Sekirov et al., 2010).

1.2 Important contributors on microbiota methodology

1.2.1 OSullivan, (2000)

Osullivan, (2000) in a survey of methods of analysis reviewed some traditional methodologies in order to analyze the normal intestinal flora. In addition he emphasized the approaches of the development of modern molecular studies including the diversity and phylogeny of this flora. Furthermore, he worked on the rapid molecular processes for the detection of certain strains of the microorganisms encountered in the gastrointestinal tract, in an attempt to study their metabolic activities.

1.2.2 Kleessen, (2000)

The studies carried out by Kleessen et al., (2000) focusing on the culture-based knowledge on biodiversity and the development and stability of human gastrointestinal microflora have concentrated on certain species of *Bacteroides, Eubacterium, Clostridium, Bifidobacterium* anaerobic cocci. They also reported the factors influencing the intestinal flora and the effect of age, diet and environment on the microbiota. Their final conclusion was that culture methods have to be applied concurrently with recent techniques built on molecular ecology based on nucleic acids so that an acceptable evaluation of the normal intestinal microbial flora can be obtained.

1.2.3 McCartney, (2002)

McCartney,(2002) reviewed the application of molecular biological methods for studying probiotics and the gut flora, such methods covered 16S rRNA gene sequencing, genetic probing strategies, genetic fingerprinting PCR-typing and other molecular techniques. It should be mentioned that these methods were applied on some microorganisms, mainly *Lactobacillus and Bidobacterium* species. The author stated that characterization of such intestinal flora could be performed genetically by DNA fingerprinting or by 16S rRNA gene sequencing. However after applying this technique, differentiation of species are rather limited. Therefore the combination of polyphasic strategy based on genetic techniques is essential to approach accurate results concerning the gastrointestinal microbial flora.

1.2.4 Zoetendal, (2004)

Zoetendal et al., (2004) reviewed several entitles including: transition from cultivation to molecular analysis quantification of SSU (Small Sub Unit) rDNA and SSU rRNA, fluorescence *in situ* hybridization (FISH), diversity microarrays, and non-SSU rRNA-based profiling. The authors concluded that the use of SSU rRNA-based analysis has given novel insights into the contents and structure of the gastrointestinal microbiota; this has directly retrieved the number of SSU rDNA sequences from the gastrointestinal tract of a variety of animals. Furthermore, characterization of gut microbiota provides limited information in regard to the interaction betwen bacteria - bacteria and host.

The conclusions drawn are that the measurement of functional genes could be one of the important approaches to determine the *in situ* activity of bacteria in an ecosystem. Also, it should be indicated that despite the value of complete genome sequences and the application of DNA microarrays to study transcriptional responses of microorganisms, these approaches are still in early stage of development and are expensive. Finally, they emphasize that application of isotopes has been found to be another method to obtain information on the functional aspects of certain microorganisms inhabiting the gastrointestinal tract.

1.2.5 Clarrigde, (2004)

In a detailed review article published by Clarrigde (2004) on the impact of 16SrRNA gene sequence analysis on identification capacity of bacteria several studies were indicated. A mechanism of bacterial 16SrRNA gene sequence analysis is fully described with its impact and potential contribution so that an understanding of microbiological aspects and infectious diseases could be made. Although the technique is sophisticated but it played a limited role in the identification of microorganisms mainly due to its high cost and requires technical skill. Despite these the 16SrRNA gene sequencing provides accurate data in regard to the genotypic identification in the clinical microbiology laboratory.

1.2.6 Amor, (2007)

Modern molecular tools at the advanced level are pointed out by Kaouther Ben Amor et al., (2007) for identifying lactic acid their activity. The authors claimed that this can be divided on the basis of nucleic acids and other macromolecules. The most frequent tools used are the PCR and hybridization with DNA, RNA or peptide nucleic acids, including 16SrRNA sequences.

Furthermore the validation of the 16SrRNA targeted oligonucleotides probes for the identification of some bacteria (e.g. *Lactobacillus, Leuconostoc, Pediococcus and Enterococcus spp*.) could be provided by the application of this method. In conclusion it can be regarded as one of the advanced tool for studying the intestinal microbiota.

1.2.7 Woo, (2008)

In 2008 Woo and his colleagues described fully the advantages and disadvantages by the use of the 16S rDNA gene sequencing for bacterial identification and discovery of novel bacteria. In their review of the literature concerning this problem they found that several bacterial genera and species have been classified and renamed, they also established phylogenetic relationship which could be determined. In addition they were able to discover and classify novel bacterial species besides the identification or detection of bacteria which cannot be cultivated in the microbiology laboratory by the culture-dependent methods. They cited some examples of such bacteria as: *Streptococcus sinensis, Laribacter hongkongensis, Clostridium hathewayi and Borrelia spielmanii.*

1.2.8 Hamady and Knight, (2009)

Hamady and Knight, (2009) in their review of the microbial community profiling for human microbiome projects, mentioned some details on the tools techniques and challenges. The studies included some aspects on the requirement of the human genome to characterize the microbiome,i.e. the collection of genes in the microbiota.The main emphasis was on the application of two principal methods for carrying out this characterization that are culture-independent. These are small subunit ribosomal RNA (rRNA) studies or the 18SrRNA gene sequences. The former may be used for archea and bacteria for eukaryotes. They also extended their review to involve metagenomic studies by which small subunit rRNA-based problems are sometimes regarded to be “metgenomic” since they analyze or determine the abundance of each organism. In conclusion the authors stated that through the use of metagenomic and rRNA-based techniques, much progress has been made in terms of characterization of human microbiome and their role in health and disease.

1.2.9 Rogers and Bruce, (2010)

The review published by Rogers and Bruce, (2010) described the essential consideration for clinical application in connection with the next-generation sequencing in the analysis of human microbiota. Some years back the analysis of the complex microbiota was

performed by the selective isolation (i.e. culture-dependent). However, some microorganisms such as those inhabiting the gastrointestinal tract are difficult to isolate or to identify by culture-dependent techniques. Thus, culture-independent methods have to be performed especially surveys of 16S rRNA gene diversity which have indicated that the majority (>75%) of the phenotypes in the human GI tract are not in full correspondence with the original cultured species. The review has been extended to the development of next-generation sequencing (NGS) and their strategies which are used to identify and or characterize various human microbiota.

1.2.10 Nikkila and de Vos, (2010)

Nikkila and de Vos, (2010) carried out a study on the advanced approaches for the characterization of the human intestinal microbiota by computational meta-analysis. Their investigations were based on phylogenetic microarray analysis addressing over a million data points. They reported convinced evidence on the feasibility of the advanced computation meta-analysis of the datasets obtained from the gastrointestinal microbiota.

1.2.11 Paliy and Agans, (2012)

In a minireview published by Paliy and Agans, (2012), some details are pointed out on the application of phylogenetic microarrays, several ones have been used successfully to identify and characterize the composition and function of various microbial communities including genome arrays and phylogenetic microarrays. The main advantages of phylogenetic microarrays when compared with other methodologies were also discussed in this review. On the other hand the design, use and analysis of microarrays require extensive testing, skill and validation for having reliable knowledge of microbial communities under study.

1.2.12 Shokralla, (2012)

Shokralla and his co-workers (2012) reviewed advantages and limitations of current next-generation sequences technologies in relation to their application for environmental DNA analysis they pointed out that it is becoming an important tool needed in laboratory techniques for the isolation and or characterization of various specimens. It should be mentioned in this regard that the traditional DNA-sequencing method is rather limited or inadequate for dealing or analyzing a large number of samples, which contain DNA obtained from hundreds or thousands of individuals.

1.2 .13 Guinane and Cotter, (2013)

By studying the role of gut microbiota in health and chronic gastrointestinal diseases, Guinane and Cotter, (2013) described the tools for investigating such microbiome. They stated that the traditional culture-based methods were used to investigate the gut microbiota, but such techniques emphasized mainly on what they called “easy-to-culture” microorganisms and become less popular due to their limitations in culturing just about 10-50% of the gut microbes. It is, therefore, evident that culture-independent approaches are more advanced and yield more rapid and accurate data about the gut microbiota. Examples of these approaches are: DNA sequencing methods 16S rRNA gene. The high through put sequencing (HTS) has widely been used to study the complexcity of the gut microbioya due to its speed, accurate or precise results provided.

The article has also made some descriptions and information about the relationships between the gut microbiota and diseases. Firstly, it starts with the evolution of the intestinal microbiota starting from infancy and all the consequences which take place afterwards. Two main diseases namely irritable bower syndrome and inflammatory bowel diseases are discussed briefly with some microorganism associated with such diseases.

1.2.14 Cani, (2013)

The review article of Cani, (2013) entitled “gut microbiota and obesity: lessons from the microbiome” reflects its complexicity and its contribution in the development of several diseases (e.g. obesity, type2 diabetes, steatosis, cardiovascular diseases and inflammatory bowel diseases). The culture-dependent and culture-independent methods have revealed global changes in the gastrointestinal microbial flora. However, the culture-independent methodology involved 16S rRNA gene analysis and DNA sequencing of the microbiota. The latter methods have facilitated the identification of the moved microorganisms involved in addition to the finding which have indicated the association of the metabolic activities with obesity and type 2diabetes.

1.2.15 Salipante, (2013)

Salipante and his colleagues (2013) performed a rapid 16S rRNA next-generation sequencing of polymicrobial clinical specimens for diagnosis of polymicrobial clinical specimens for diagnosis of complex bacterial infections. Their classification into individual bacterial species results in a challenge for culture-dependent, also at the molecular level. Therefore, they were concerned with metagenomic techniques to approach rapid identification

of the complex bacterial composition of clinical specimens collected from patients, without practicing culturing methods. The investigators were able to obtain sequences that can be used to perform reliable taxonomic assignment upon combining a semiconductor deep sequencing protocol that produces reads spanning 16S ribosomal RNA gene variable region 1 and 2. The above technique was applied on a collection of sputum specimens collected from cystic fibrosis (CF) patients revealing well-identified pathogens. Their final conclusion suggests that metagenomic profiling may prove to be valuable for diagnostic purposes.

1.2.16 Becker, (2013)

The research project carried out by Becker et al., (2013) revealed the role of intestinal microflora in regulating cell differentiation factors both *in vitro* and *in vivo.* They described them as being the most favorable environment for harboring a large number and microorganisms and their complexcity. Furthermore, the study was expanded to the role of the intact mucosal barrier and secretion of mucin, e.g.Muc1 and Muc2 as structural proteins covering the gastrointestinal tract. In addition, the production of broad-spectrum antimicrobial peptides, including defusing was also included in this study.

For investigating the possible role of several microorganisms of gut microbial flora in the regulation of epithelial differentiation, the contributors emphasized on the regulatory effects of these organisms on the expression factors Hes 1, Hath1 and KLF4. Moreover, the effects of certain bacteria on mucins Muc1, and Muc2 as well as defending HBD2 were also studied. Their conclusions can be summarized by the following: finding out that the intestinal microbiota influence the intestinal epithelial differentiation factors Hes1, Hath1 and KLF4, as well as Muc1 and HBD2 *in vitro* and *in vivo.*

1.2.17 Raoult, (2013)

Is a biologist from Aix-Marseilles University and developed a revolutionary approach called culturomics in early 2010s. This mixed method combines culture and rapid identification method such as spectrometry. This technique allowed researchers to identify 31 more gut microbiota, of bacteria, virus, and Archea. It allows for the isolation and identification uncluturable microorganisms and that cannot be identified and isolated using molecular methods.

1.3 Taxa of microbiota

The diverse microbial community can be characterized in all and each higher animal. They can be associated with a community of bacteria, archea, viruses, fungi and protozoa. The range and the number of the mocrobiota in the human GIT is ten times more than the number of the body cells. Predominating within the human gastrointestinal tract microbiota is composed of: Firmicutes, Proteobacteria, Bacteroidetes, Actinobacteria, Fusobacteria, Verromicrobia, and Cyanobacteria (Maukonen, 2012). Figure 2 shows the taxonomical divisions of microorganism's precent in the intestine.

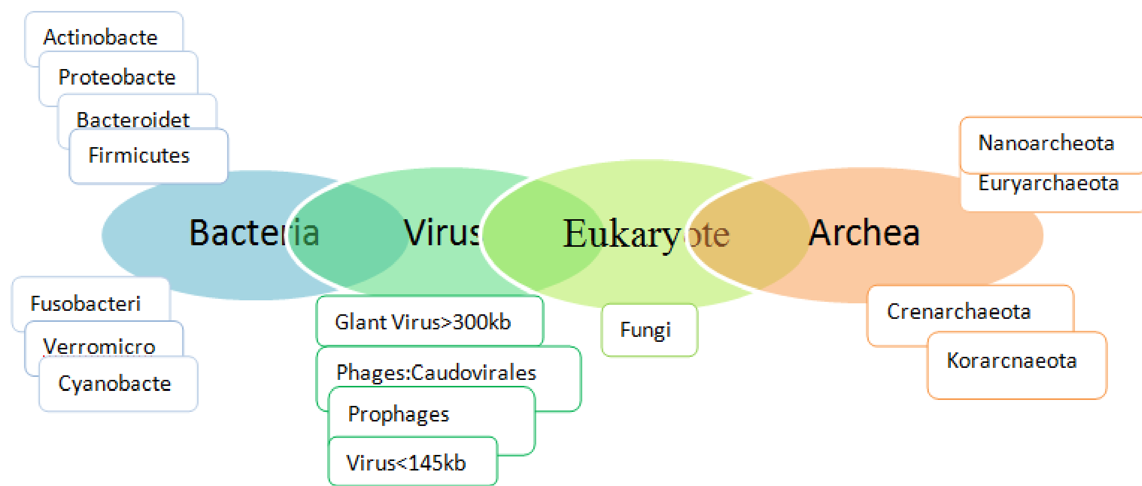


Figure 2: Overview of common human intestinal microbiota, bacteria, archaea, viruses, and eukaryota (Lagier et al., 2012; Searchhigh., 2012).

1.4 Functions of microbiota

The gut microbiota may play important roles in nutrient digestion and synthesis energy metabolism, vitamin synthesis, epithelial development and immune responses. These functions are summarized in the following:

1-Vitamin synthesis

The significance of the gut microbiota relating to vitamin synthesis has been reported in relation to vitamin K and some B vitamins, being produced by several intestinal microorganisms, e.g. *Bacteroides* and *Eubacterium* (Bik, 2009).

2-Protective function

Building resistance to infection by increasing the activity of the immune response, therefore creating a natural barrier against colonization by exogenous pathogenic bacteria. It may be based on the production of bacteriocins and organic acids which lower the pH and therefore inhibiting pathogenic microorganisms to flourish (Rakoff-Nahoum et al., 2004).

3-Metabolic activity

The intestinal microorganisms have a beneficial effect on the metabolic activity of the organism. They are important in proper functioning of the whole organism by carrying out fermentation of undigested debris in the large intestine. The metabolic activity leads to the acquisition of energy and absorbable substrates for the host organism and to provide energy required for growth of bacteria (Bik, 2009).

4-Intestinal epithelial cell proliferation and differentiation

Further studies have indicated that some microorganisms like *E.coli*, *Bifidobacterium, and Lactobacillus* species increase the survival of the intestinal epithelial cells by certain mechanisms and pathways when invaded by pathogenic organisms. The normal intestinal microbial flora may also contribute to or act as a barrier for keeping the integrity of the intestinal epithelial cells and for other functions (Bik, 2009).

5- Immune stimulation**:**

The commensally gut microbiota has been reported to play an important role in the development humeral and cellular mucosal immune systems initiating from neonatal life and maintained physiologically steady throughout life (Rakoff-Nahoum et al., 2004).

6- Signaling from the periphery to the brain:

Recent advances in medicine have concluded that the gut microbiota influence the enteric nervous system, therefore, it may contribute into signaling process to the brain. The

cytokine production and other immunological reactions can affect the peripheral and central nervous system especially in its modulation. (Collins & Bercik, P, 2009).

1.5 The Role of Microbiota in Certain Diseases and Disorders

Recently gut microbiota has been associated or correlated with a number of diseases and physical disorders rather than causal (Maukonen, 2012) as shown in Figure (3) the implication of analysis of intestinal microbiota due to OTS impact on the different health concerns.

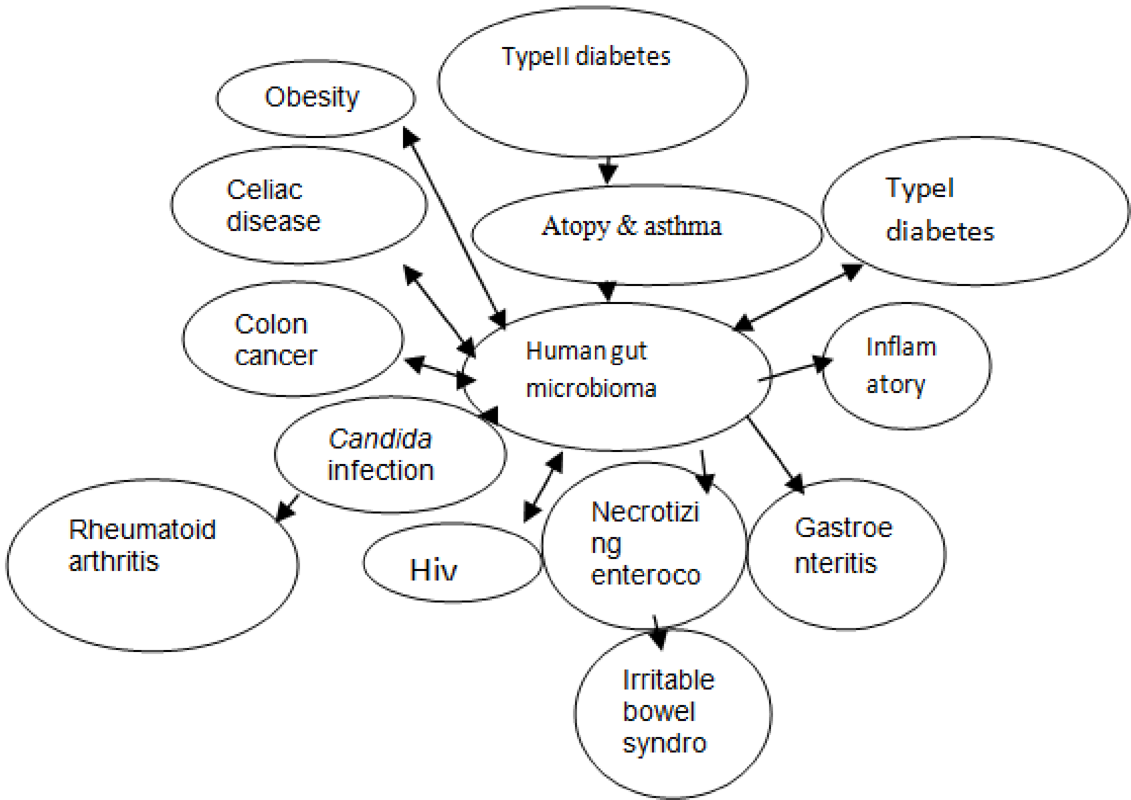


Figure 3: Diseases and physical disorders associated or correlated with human intestinal microbiota (Maukonen, 2012).

2 Aim of the study

The knowledge about mocrobiota is increasing because the techniques for studying of the composition, isolation, and identification of microbiota are an ongoing development process. Therefore, it is crucial to review, compare and discuss the different techniques. The estimation and identification of microbiota started with classical methods and now scientists develop other recent techniques getting benefit from recently innovated technology. In

general the conventional methods include culturing and biochemical investigation some biochemical tests include monitoring of specific enzymes and metabolites samples. The classical methods applied by scientist from 17th century and went through different improvements till the present time as reported by Ursell et al., (2012). Therefore, the aim of the study is to compare and contrast the different diagnostic tools for the investigation of microbiota. This paper will include classical and contemporary methods and their advantages and disadvantages.

In general, the diagnostic tools for the investigation of microbiota can be divided into two main methods:

1. Culture-dependent methods
2. Culture-independent method

Novel methods are the most recent methods that are promising due to their accuracy, quick, cost efficiency and efficiency. These methods involve sequencing of RNA and DNA and gene functions. Furthermore, making comparative studies between various classical and novel methods concerning accuracy, sensitivity, time consumption and economical aspect duration. Therefore this will lead to discover the advantages and disadvantages of each of the above-mentioned methods.

3. Methods

This review paper has been prepared based on literature available and obtained from scientific (both review and original) journals that are peer reviewed and highly cited, and newly released articles. Reference list of the articles were also used for more comprehensive and extensive search. The following medical, biological, and scientific library databases were used such as ISI Web of Science, Bibsys, PubMed, Google Scholar, and Google in a non systematic fashion. The collected data have been arranged in a chronological order. Several keyword combinations were used in the search as in following: "gastrointestinal microbiota", "Next Generation DNA Sequencing (NGS)", "Omics", "16S rDNA", "Culture-dependent Methods", "16S rRNA gene library", "molecular techniques", "Culture-Independent Methods", "Microbial ecology", and "Pyrosequencing".

After extensive review of all possibly obtained data from journals information's concerning the methods of investigating human intestinal microbiota can be categorized as in the following:

3.1 Classical approaches

3.1.1. Culture -dependent methods

3.1.2. Culture -independent methods

3.1.2.1. Direct microscopic analysis

3.1.2.2. Monitoring of specific enzymes or metabolites in faecal samples.

3.2 Molecular (novel) approaches

3.2.1. Culture-dependent methods

3.2.1.1. Phenotypic fingerprinting analysis

Polyacrylamide gel electrophoresis of soluble proteins

Fatty acid analysis

Bacteriophage typing

Serotyping

3.2.1.2. Genotypic fingerprinting analysis

Colony hybridization with nucleic acid probes

Pulsed field gel electrophoresis (PFGE)

Ribotyping

3.2.2. Culture- independent molecular methods

3.2.2.1 Fluorescence in situ hybridization

3.2.2.2. Flow cytometry

3.2.2.3. Quantitative dot blot

3.2.2.4 PCR-based techniques

Checkerboard hybridization

Microarrays

16S rRNA

RAPD

RecA gene

Multiplex-PCR

Arbitrary primed (AP) PCR

Triplet arbitrary primed (TAP) PCR

DGGE/TGGE

Gene cloning and sequencing

Real-Time PCR

T-RFLP

3.3. Tools for functional studies of gut microbiota

3.3.1. Stable isotope probing (SIP)

3.3.2. ‘Omics’

3.3.2.1. Metagenomics

3.3.2.2. Metatranscriptomics

3.3.2.3. Metaproteomics

3.3.2.4. Metabolomics

3.4. Insertion sequencing

3.5. Animal models

3.6. Data analyses and bioinformatics

4 Results of investigations and discussion

Traditionally, diagnostic bacteriology has been depending on phenotypic characters of the bacteria, for instance cell morphology and function of biochemical reactions. This characterization requires series of biochemical tests described in numerous manuals. For the identification we can use several tests, such as the API system (Rautio, 2002).

The classical approaches to study microbiota are classified into:

Culture-dependent methods

Culture-independent methods

4.1 Classical approach

4.1.1 Culture- dependent Methods

The culture- dependent techniques are used to isolate cultivatable bacteria from faecal or intestinal samples. Normally, culturing techniques are based on plating fresh or conserved faecal material or intestinal content on selective medium (such as Eosin Methylene Blue agar (EMB agar), MacConkey agar, Mannitol salt agar, Phenylethyl Alcohol) and non-selective medium (such as Brain heart infusion (BHI), Brucella blood agar (BBA), GAM (Gifu Anaerobic Medium) agar, Peptone-yeast extra-glucose (PYG), Plate count agar), medium incubated under a variety of conditions (Rautio, 2002). So the bacteria can be classified according to the selectivity of these media.

Genus and species identification is very essential after performing isolation of bacterial colonies. Both morphological and biochemical tests are required to know the characteristics of each genus or species. Some researchers such as O'Sullivan, (2000) state that these tools are ineffective in finding the relationship between species from different individuals. The second disadvantage of these techniques is that 40–90% of microorganisms cannot be cultivated under laboratory conditions (Zoetendal, Collier, Koike, Mackie, & Gaskins, 2004). Therefore, their growth requirements are unknown, the conditions of *in vitro* culturing and the stress caused by its procedure, and anaerobic are the only conditions necessary for most microbiota species (Zoetendal et al., 2004).

Methods for culture-dependent and other traditional methods are often laborious, time consuming and not reliable for distinguishing microbiota species or strains. Many bacteria are morphologically and biochemically similar and often the phenotypic traits are unstable being linked to extrachromosomal mobile genetic elements (plasmids) which can be lost after

various cultivations on synthetic media. In addition, various genetic recombinant events such as point mutations, chromosomal rearrangements, duplication, infection by bacteriophages, and horizontal gene transfer might have a role in changing phenotypes among bacteria (Cano-Gomez et al., 2010). Furthermore, bacteria identified using these methods can represent only a small part of the natural microbial communities (Pond et al., 2006). Since they were established, culture-dependent methods were useful to obtain knowledge about intestinal microbiota, however, its limitations are highlighted when it come to ecological studies and a comprehensive overview of intestinal microbiota (Gong, & Yang, 2012).

Nocker, Burr, and Camper (2009) noted that traditional methods cannot replicate the natural and complex intestinal conditions, including the biochemical interactions between host cells and microbiota and between the bacteria themselves. However, with all the above stated disadvantages, culture- dependent methods still considered a useful tool for in depth study of the physiology of specific isolated microorganism (Gong, & Yang, 2012). A brief discription, advantages and disadvantages are listed in table 1.

100%

80%

60%

40%

20%

0%

Gram staining (×100 oil immersion) ***Electron microscopy (×7100)*** **Pyrosequencing**

53% bacteria Gram-+ve 60% bacteria Gram-positive 80% bacteria Gram-+ve

47% bacteria Gram--ve 40% bacteria Gram-negative 9% bacteria Gram--ve

11% not available

Phylum	%
Firmicutes	71.12
Actinobacteria	9.21
Other	10.53
Bacteroidetes	6.33
Proteobacteria	2.78
Cyanobacteria	0.03
Verrucomicrobia	00.1
Total	100

Figre 4: Comparison of findings using different techniques in one stool sample. (Red represent gram –ve and blue represents gram +ve bacteria) (Lagier, J. C. et al., 2012).

Table 1: Advantages and disadvantages of culture--dependent techniques (Fraher et al., 2012).

Technique	Description	Advantages	Disadvantages
Culture	microorganisms are isolated from a sample using selective media	• cost efficient, • semi-quantitative, • commonly used, • used for biochemical and physiological studies • can be used combined with other techniques as a priori	• laborious, • less than 30% of gut microbiota have been cultured till now • immediate processing of sample required, • only cultivable microorganisms can be isolated, • results are affected by the selection of growth media, • most bacteria cannot be recovered, • post isolation, a number of techniques required for identification, • good expertise and sophisticated equipment needed to isolate microorganisms.

4.1.2 Culture-independent method

These techniques include:

- Direct microscopic analysis;
- Monitoring of specific enzymes and/or metabolites in faecal samples.

4.1.2.1 Direct microscopic analysis

Direct microscopic analysis has been a useful approach to estimate the number of Bacteria in faeces. This method is a valuable aid to assess how useful a culture methodology may be for investigation of the intestinal microbiota. But, the microscopic technique is not reliable and may significantly misrepresent the accurate numbers. This technique includes heat fixation and staining (Rautio. 2002).

4.1.2.2 Monitoring of specific enzymes and/ or metabolites in faecal samples

This technique can detect the presence of microbiota indirectly in a faecal sample or detect even metabolic activity of certain groups of microorganisms rapidly and using massive number of samples. O'Sullivan (2000) concluded that measurements of principal fatty acid such as propionate, butyrate, and acetate. Measuring these principle fatty acids can be directly correlated with the metabolism of a specific bacterium.

4.2 Molecular (novel) approaches

A Historical Background

The introduction of the molecular methods has expanded our knowledge in discovering and developing reliable information concerning identification of isolates, also calculating the evolutionary relationships between strains of these isolates. It has been possible to locate accurate species of unknown isolates by applying sequence analysis of 16S rRNA. This method was first developed by Woese and his coworkers (1987) for identifying and classifying organisms, also establishing their evolutionary relationships. This was followed by the use of the databases of rRNA sequences which have been covered by several gene banks, such as GenBank and the ribosomal database project (Maidak et al., 2001) this allows future studies to be carried out on the phylogenetic position of unknown isolates. From the technical point of view, such studies can be performed by the use of polymerase chain reaction (PCR) to elicit the 16S rRNA gene directly from colonies using primers.

Amann and his collaborates (1995) showed that the entire PCR amplicon can then be directly sequenced and compared to the rRNA database. Further work made by Leblond-Bourget and his coworkers (1996) to discover the region between the 16S and 23S rRNA

genes (called the internal transcribed spacer). The sequence analysis of this molecule was proved far more sensitive and accurate than the rRNA analysis. The demonstration of the above stated method is the identification of intestinal *Lactobacillus* species.

This work was later performed by other investigators (Tannock *et al*., 1999) who confirmed the validity of this method. The determination of intrageneric phylogenetic relationships was achieved as a result of the emergence of a short segment of the recA gene as a potential candidate for sensitive molecules (Eisen, 1995; Karlin *et al*.1995). It was claimed that the recA gene encodes the recA protein, which plays an important role in recombination, DNA repair (Roca and Cox, 1997).This study was extended and applied to member of the genus *Bifidobacterium (*Kullen *et al.,* 1997) with an approach that the resulting molecule from the other strain types and intestinal *Bifidobacterium* isolates and the phylogenic relationship obtained by recA sequence analysis are in accordance with rRNA gene analysis.

Several detailed studies and review articles have been published (Collins and Gibson, 1999; O'Sullivany 2000; Vaughan *et al*.2000) dealing with the increased application of molecular biological methods for studying the ecology of the gastrointestinal tract microbiota. As mentioned earlier, these methods involve the use of PCR for the amplification of the 16S rRNA genes (16S DNA) in microbial DNA extracted directly from the samples. These amplified sequences are cloned and, therefore, should contain copies of the genes from all the species present in the sample. Sequencing of 16S DNA clones permits the identification of certain uncultivable species by the ordinary conventional methods.

Introduction to molecular approaches

Although culturing of intestinal microbiota has contributed in the early stages of the field in the studying, analysis, quantification and identification of microorganisms, it has several limitations. Due to the observational nature of these methods it is possible that microorganisms with similar morphology can be misdiagnosed especially in complex ecosystems like the human intestinal tract. The majority of the intestinal microorganisms are uncultivable therefore it is not possible to be detected by this technology. Therefore the majority of unknown microorganisms in the intestinal microbiota remained unknown.

The human intestinal ecosystems are complex, and it had characteristics in terms of changes in levels of oxygen, pH and solutions in additions to continuous intake of food and other host-producing solutions. The interactions within the intestinal ecosystem were also challenging to replicate in the laboratory therefore it is not possible to study the ecosystem

using the cultivating methodology. In addition to the laborious nature of the process, it can be time and labor intensive and costly due to the need for the use of sophisticated machinery.

Therefore researchers introduced molecular methods relying on the genome which is the study of biomarkers, including metabolites, proteins, RNA, DNA, and cells. These molecular methods depend mainly on the SSU rRNA gene to study the phylogenetic background on micororganizms in complex ecosystem. The SSU rRNA gene has many characteristics to become the principal study unit: due to its high levels of functional constancy, it is present in every organism, it changes which the change in the phylogenetic relationship, it allows for direct sequencing, and diverse ecosystems can be studied through SSU rRNA gene sequence databases. The molecular techniques have allowed for the rapid identification, quantification, and profiling of the human intestinal microbiota.

The combination of the SSU rRNA gene sequencing and the ever-expanding clone libraries are the main principle for these molecular techniques that gives these techniques considerable diagnostic capabilities, but it also can produce library biases and incomplete sequence coverages. These technologies include the following techniques: fluorescent *in situ* hybridisation (FISH) that can be combined with microscopy, flow cytometry or quantitative real time PCR (qRTPCR) for a powerful diagnostic process. It also includes fingerprinting denaturing gradient gel electrophoresis (DGGE) and terminal restriction fragment length polymorphism (T-RFLP) and phylogenetic microarrays.

Findings using the molecular technologies have contributed vastly to the understanding of the human intestinal microbiota and its ecosystems during the last decade. Findings showed that the human intestinal ecosystem is more complex than it was anticipated before. Findings also indicated that the Firmicutes phylum is the most dominant microorganism in the intestinal ecosystems. Also findings indicated that the microbiota composition is mostly affected by host’s genotype. Despite these major contributions and advantages of these molecular methodologies, there are disadvantages associated with them and described/discussed in detail in the following sections in this paper. The next sections also discuss the characteristics of these technologies and their useful contributions.

Each of the various molecular technologies is specific for targeting a specific purpose. For example, the advantage of the phylogenetic microarray analysis is that it is carried out on the phylotype level and it considers high resolution of profiling of complex intestinal ecosystems. This gives the advantage to phylogenetic microarrays over the qPCR or FISH, that will allow for the performance of the analysis of high levels of phylotype comprehensively and with high throughput. *Clostridium leptum* or *Clostridium coccoides*

taxonomical groups abundant in the complex intestinal ecosystems. These two taxonomical groups share 85% of their SSU rRNA gene sequence similarity that could lead to generating functionally different bacterial genera and families when cultivated.

Whereas on the lower phylogenetic levels, studies showed effective application of these diagnostic technologies and the better study of the intestinal microbiota ecological diversity in the analysis of the SSU rRNA gene clone libraries on a wide scale. The disadvantages of the phylogenetic microarray is that it is laborious and expensive and can affect its application for wider studies of the various factors affecting intestinal microbiota for example health status like diet, age, geographic, and genetic origin. However, combined with classical culture- dependent and other molecular techniques can provide a powerful tool for this field.

The difference in the functional changes of various microorganisms in the intestinal tract can be a fundamental factor in searching for more novel approaches for studying the complex and various ecologies. For example strains of same microorganism like *E. coli* can exhibit entirely different functionality in different ecological situations. Molecular is a powerful priori for sequencing (O'Sullivan, 2000).

TGGE and DGGE are developed mainly for the analysis of the intestinal microbiota. The most predominant bacterial flora could be detected in faecal sample by TGGE technique (Zoetendal et al., 1998). However, the DGGE technique could be used for the qualitative analysis of intestinal microbiota. Moreover, monitoring of succession of bacterial population in neonates may also be investigated by the latter technique (Favier et al., 2002).

The techniques of TGGE and DGGE have been explained in that 16S rDNA is amplified by PCR from DNA of microbial cells and then the various molecular forms of 16SrDNA can be separated from each other by these two techniques. A temperature or chemical gradient is established in polyacrylamide gels parallel to the electric field. Migration of the DNA samples take place through the gradient from low to high temperature or chemical gradient. This is followed by partial denaturation of the double-stranded DNA when the migration of the fragment is drastically retarded and sequences of the same length but of different thermal or chemical stability, can be separated. The resulting 16S rDNA bands eluted from the gel may be utilized for further investigations, e.g. amplification by PCR and then sequencing, leading to relatively reliable information for the identification and characterization of the microorganism to be tested.

Amongst these tools is the PCR, which is known to amplify rDNA molecules from mixed populations with different degrees of efficiency (Suzuki and Giovannoni, 1996). It has also been reported that PCR-based analysis of fecal samples is difficult due to the presence of inhibitors of polymerase reactions (Satake et al., 1997). It has been found that only the predominant bacteria in the complex community could be detected by TGGE (Zoetendal et al., 1998).

Other molecular techniques discussed in the flowing sections in this part of the paper are the Culture-dependent molecular methods that are divided into Phenotypic fingerprinting analysis and Genotypic fingerprinting analysis. These techniques are mainly DNA based. However the importance of these techniques is growing due to the ever-expanding established databases. The advantage of these techniques includes the ability of conducting analysis of high throughput of unknown microorganisms and rapidly.

The disadvantage of these techniques is that they are not sensitive in separating between different strains and closely related microorganisms and identification of phylogenetic relationships. This low sensitivity can affect the effectiveness of this tool for evaluation the phylogenetic background and relationships of unknown microorganisms. But it is a very useful tool for diagnosing and monitoring known intestinal microbiota and tracking the prevalence of intestinal microbita within a community. It can also provide information on the range of different types of strains and their relativity. Although they can be performed rapidly, they are less sensitive than the sequencing technologies and usually used prior to sequencing for the isolations and dividing of the microbiota.

Phenotypic fingerprints are less sensitive and changes in the fingerprint may indicate change in expression of the phenotypic traits rather than different organism. Types of phenotypic fingerprints are polyacrylamide gel electrophoresis of soluble proteins, fatty acid analysis, bacteriophage typing and serotyping. The later is the most time saving and useful type due to the capacity of direct typing and does not need subculturing. Monoclonal antibody that is specific for specific microorganism is used for colony hybridization. Corthier et al. ,(1996) applied these methods for analyzing two Bacteroides species in different human intestines. The molecular techniques started with the hybridization nucleic acid probe targeting specific DNA sequence later more sophisticated techniques have been developed and this field is ever-expanding with the advancement of technology. A brief description, advantages and disadvantages of the molecular methods are listed in table 2.

Table 2: Advantages and disadvantages of culturing- independent techniques (Fraher et al., 2012).

Technique	Description	Advantages	Disadvantages
qPCR	16S rRNA is amplified and quantified. Fluorescent used to bind to double-stranded DNA.	• Phylogenetic identification, • quantitative, • rapidly conducted • highly sensitive	• PCR bias, • In applicable for unknown species • Individual targeting • Unknown microorganisms cannot be detected
DGGE/TGGE	Denaturant/temperature gel used for the separation of 16S rRNA amplicons.	• rapidly conducted, • semi-quantitative, • samples can be reused for further tests	• No phylogenetic identification, • PCR bias
T-RFLP	Primers are labeled using Fluorescent and amplified then restriction enzymes are used to digest the 16S rRNA amplicon. Gel electrophoresis can be used for separating fragments.	• rapidly conducted, • semi-quantitative, • cost effective	• No phylogenetic identification, • PCR bias, • low resolution
FISH	Oligonucleotide probes are labeled with fluorescent then hybridize to target complementary 16S rRNA sequences. Then the enumeration of the fluorescence is performed using flow cytometry	• Phylogenetic identification, • semi-quantitative, • no PCR bias • highly sensitive • in situ identification	• probe sequences dependent • unable to identify unknown species
DNA microarrays	Oligonucleotide probes are labeled with fluorescent then hybridize with complementary nucleotide sequences. Then laser is used for the detection of fluorescence.	• Phylogenetic identification, • semi-quantitative, • rapidly conducted	• Cross hybridization, • PCR bias, • species present in low levels can be difficult to detect

Technique	Description	Advantages	Disadvantages
Cloned 16S rRNA gene sequencing	Cloning of full-length 16S rRNA amplicon, then Sanger sequencing and capillary electrophoresis are performed.	• Phylogenetic identification, • quantitative	• PCR bias, • laborious, • expensive, • cloning bias
Direct sequencing of 16S rRNA amplicons	enormous parallel sequencing of partial 16S rRNA amplicons is performed e.g., 454 Pyrosequencing (amplicon fixed on beads, then amplified by emulsion PCR, chemoluminescent signal is generated by additing of luciferase)	• Phylogenetic identification, • quantitative, • fast, • identification of unknown bacteria	• PCR bias, • expensive, laborious
Flow cytometry	Suspension then bacterial fixing then hybridized with antibodies labeled with flourescein, probes labeled with flourescein and targeted with rRNA oligonucleotide, or contrasting stain with a general DNA-binding dye then flow cytometry identification.	• Quantitative and qualitative • Morphological, density, and metabolic analysis • Time efficient • Massive parallel analysis of cells • No need for DNA extraction and amplification • Highly accurate	• Complex data analysis is needed • Cell size bias
Quantitative dot blot	16S rRNA of a targeted microorganism devided by number of 16S rRNA of all microbiota present in a sample. rRNA is isolated then ploting the known numbers on two parallel nylon membranes then hybridized with 32P or fluorescently labeled probes. Then quantification of the signal intensity of each spot.	• Representation of metabolic activity • Groups comparison • Specify of probes can be determined	• Application bias • Reading bias

It is possible to divide the molecular methods collection of analysis in two major divisions namely:

Culture-dependent molecular methods.

Culture-independent molecular methods.

4.2.1 Culture- dependent molecular methods.

Molecular methods give the opportunity to analyze a great number of isolates and provide good portion of information about the genera. It is still a valuable first step to divide the isolates into broader groups, prior to sequencing (O'Sullivan, 2000). Several molecular genetics-based methods for the detection and identification of microbial flora in different samples have been developed, which have greater accuracy and are faster than classical methods (Seidave;2012).

Culture-dependent molecular methods can primary be divided in two major divisions:

Phenotypic fingerprinting analysis.

Genotypic fingerprinting analysis.

4.2.1.1 Phenotypic fingerprinting analysis:

Fingerprinting techniques have been developed intended for the analysis of both phenotypic and genotypic traits. While phenotypic fingerprints can be obtained, these are less sensitive, and changes in the fingerprint may not essentially mean a different organism. Phenotypic fingerprints are polyacrylamide gel electrophoresis of soluble proteins, fatty acid analysis, bacteriophage typing and serotyping. The most express of these procedures is serotyping, as colonies can be directly typed, without sub-culturing, by colony hybridization with a monoclonal antibody specific for a particular genus, species or strain. This strategy has been applied for the study of microbiota (Corthier et al., 1996).

Examples of phenotypic fingerprints are.

Polyacrylamide gel electrophoresis of soluble proteins.

Fatty acid analysis.

Bacteriophage typing.

Serotyping.

4.2.1.1.1 Polyacrylamide gel electrophoresis of soluble proteins:

Moore et al., (1994) think that this method is used to compare cellular proteins from bacterial isolate of gingival crevice floras to screen isolated bacteria. Furthermore, the method includes reducing the analytical time and immolating accuracy. The scientific principle depends on the movement of charged DNA and RNA protein molecules in an electronic field and in a gel medium. The gel medium allows for shorter protein molecules which are negatively charged to migrate toward a positive pole faster than longer protein molecules.

4.2.1.1.2 Fatty acid analysis:

Types and relative quantities of long chain fatty acids (9-20 carbon atoms) that are present in bacterial membranes can be used to identify and characterize microorganisms. Each bacterial species has a characteristic profile of fatty acids (Rautio, 2002). In the Gram positive bacteria, branched chain acids are common, whereas in Gram negative bacteria short chain hydroxy acids illustrate lipopolysaccharides. Fused silica capillary columns are now used for the recovery of many isomers of the hydroxy acids and resolution. Gas chromatography is commonly used to identify many ranges or microorganisms. Techniques such as The Sherlock Microbial Identification System are found to be cost effective and produce reproducible results.

4.2.1.1.3 Bacteriophage typing:

Bacteriophages can be used for classification of bacteria (phage typing). Strains with a particular serotype can be distinguished and differentiated with different types of phages. Different lyses activity exist between different species of bacteria as each bacteriophage has its own specificity as such we can differentiate the bacterial species (O'Sullivan, 2000).

4.2.1.1.4 Serotyping

The advantages of this tool are that its results are acquired rapidly. It is useful for the direct identification of colonies without sub-culturing. This can be performed by colony hybridization with a monoclonal antibody specific for a particular genus, species or strain (O'Sullivan, 2000). Furthermore, these methods are used for identification of microbiota depend on an enzyme-linked immunosorbent assay (ELISA) based on the use of polyclonal antibodies against live cells (Phianphak et al., 2005)

4.2.1.2 Genotypic fingerprint analysis.

The development of multiple genotypic fingerprinting methodologies has been a major advantage for deciphering the complex human intestinal ecosystem (O'Sullivan, 2000).

The following methodologies are described in this paragraph:

Colony hybridization with nucleic acid probes.

Pulsed field gel electrophoresis (PFGE).

Ribotyping.

4.2.1.2.1 Colony hybridization with nucleic acid probes

It is defined by Brown (2006) as a "technique that uses labeled nucleic acids molecule as a probe to identify complimentary or homologues molecules to which it base-pairs" (p. 664). Taxonomic groups of species have their complementary probes. A single-stranded nucleic acid can purposely hybridize with its complementary sequence; therefore, they can target specific sequences in a genome. It can be used alone or after a PCR step. Hybrids of single-stranded DNA/RNA combined with a probe can be identified depending on the type of label, with radioactive signal, fluorescence, or color reaction. The number of the target microorganism is identified by the intensity of the hybridization signal.

4.2.1.2.2 Pulse field gel electrophoresis (PFGE)

PFGE is a diagnostic method that is used in epidemiological studies and public health surveillance with standardized protocols for microbiota includes *E. coli, Listeria,* and *Campylobacter*. Its typing technique is highly discriminative and it is based on the variability of movement of large DNA restriction fragments. The process is performed in an electrical field of alternating polarity in an agarose gel medium. DNA fragments with similar strains are separated by comparing fingerprints. The highly discriminatory element of this technique comes from that this method detects DNA changes over time and its results are highly reproducible. The disadvantages of this method are that it is laborious and time consuming because it requires culturing to obtain enough cells (Applied Math, n.d.).

4.2.1.2.3 Ribotyping

It is one of the RFLP, rRNA gene containing restriction fragments of a specific genome. After culturing to obtain enough cells, and then isolating of DNA, by using a restriction enzyme with a frequently occurring recognition sequence (6 bp), it is then restricted into fragments, sizes between 1 kb to 20 kb. The restricted fragments separated by

agarose gel electrophoresis then 16S (the most common), 23S, or 5S rRNA genes are targeted with the hybridized probe. Alternately, hybridization can be performed using in gel hybridisation technique, on a nylon or nitrocellulose membrane. After the detection of the probes, characteristic is represented in fingerprint, restriction bands carrying copies of the rRNA genes. Patterns of sizes of the bands represent species and genus of microorganisms.

The advantage of this method is that for typing all bacteria a single rRNA probe can be used; it is highly reproducible and has proven it effectiveness to analyze intestinal microbiota (McCartney et al., 1996). These methods disadvantages are that it is less discriminative than PFGE it's laborious and requires bacterial culturing. Bacteria usually contain more than eight copies of rRNA genes which help for the obtaining of RFLP; one factor that can limit the effectiveness of ribotyping for fingerprinting is that some bacteria can contain fewer than eight copies of rRNA genes.

4.2.2. Culture- independent molecular methods.

Culture- independent molecular methods are known as a standard phylogenetic classification tool. Which are rapid and reliable for the identification than culturing because the identification is dependent on the nucleic acids rather than on the genomic expression under a given cultural condition, and the DNA can be created from living or dead cells .The applicability of molecular methods ranges from recognition or detection of single bacterial species to characterization of complex macrobiotics, and molecular techniques have been applied in analyzing evolutionary relatedness of several types of bacteria (McCartney, 2002). Molecular methods are based mainly on the detection of ribosomal RNA (rRNA) or ribosomal DNA (rDNA; DNA encoding the rRNA). The probes for target DNA sequences used for designing and detecting bacteria on diverse phylogenetic levels from major genera or the group level to the species or even strain-specific level depending on the type of the study ranging from gut ecology studies to tracking specific probiotics or pathogens (Charteris et al., 1997, Franks et al., 1998).

Highly conserved regions of ribosome can be used for designing universal probes and different variable regions for specific/targeted probes. Several thousands of 16S rDNA sequences counting many uncultured bacteria are freely available in genomic databanks. The target DNA can be detected using various PCR-based methods or dot blot hybridisation with specific artificial oligonucleotide probes or by fixed bacterial cells by fluorescent *in situ* hybridisation combined with flow cytometry or microscopic analysis (Wilson & Blitchington, 1996; Lin et al., 1997). Quantitative PCR allows for the quantification of all DNA fragments

detected by PCR using specific controls of known quantity giving an estimate of the number of target microorganisms in the sample (Sanz et al., 2004). Using multiplex PCR several target regions can be multiplied in single reaction with all necessary primers (García et al., 1998).

There two limitations listed by O'Sulivan (2000), the first are related to the underestimation of microorganisms with rRNA fewer than eight in the estimation of the bacterial representation of a natural habitat. The second limitation is related to the universal primers used in the process of the magnification of the rRNA which might not be equally efficient with all rRNA products.

Molecular methods include:

Fluorescence in situ hybridization (FISH)

Flow cytometry

Quantitative Dot Blot Hybridization

PCR- based technique (Checkerboard hybridization, Microarrays, 16S RNA, recA gene, multiplex-PCR, AP-PCR, TAP-PCR, DGGE/TGGE, gene cloning and sequencing, Real-Time PCR, T-RFLP, RAPD).

4.2.2.1 Fluorescence *in situ* hybridization (FISH)

FISH is one of the florescence marking techniques by which DNA and RNA is targeted in site by molecular probes with fluorescent labels. Using the signals emitted by the marker (florescence) the location of the DNA or RNA molecule on nitrocellulose, nylon membrane, or in gel. Fluorescent markers are common alternatives to radioactive markers for environmental and health concerns. 16S rRNA sequences are usually what the probes are designed for. Individual bacteria contain 10^{3}-10^{5} ribosomes for resulting in cell fluorescence. The total number of microorganisms in a natural medium can be counted with of a genetic probe and fluorescence microscope. It can be used a culture-independent method with, 16S rRNA-targeted oligonucleotide probes for complex mixed populations (Franks et al., 1998), which is one of the advantages of this approach for complex samples with multiple species (Amann et al., 1995).

One of the advantages of this method is that it allows for the analysis of complete bacterial cells compared to PCR-based methods, because samples on glass slides can be stained with DAPI or hybridized with probes. Based on the chosen probe, this technique is used for the detection of bacteria on different phylogenetic levels (Franks et al., 1998).

Oligonucleotide probes are designed for most species affiliated to the intestinal medium. These probes are used in the hybridization of 16S rRNA, which are also targeted by phylogenetic approach and the detection of specific species or groups (Amann et al., 1995). Ninety percent of intestinal microbiota can be detected few probes. The advantage of also include that it allows the study of deferential abundance of microorganisms, it is fast and semi quantitative; but unknown species and strain cannot be detected with this method (Fraher et al., 2012).

4.2.2.2 Flow cytometry (FCM)

Since it was developed in the 1960s, FCM has been a powerful tool to analyze intestinal microbiota especially when combined with a powerful tool such as FISH. It was first developed for the counting and studying mammalian cells (Festin et al., 1987), but now it has gone through powerful technical developments and has been used to study for example the intestinal microbita (Collado et al., 2008). The principle is that the cells are streamed through a capillary, during which a laser beam is used to detect the cells. A laminar flow effect is produced by a differential pressure system which prevents the fluid sample to mix with the sheath fluid by helping the sample fluid to flow in the central core.

The advantages of this technique are that it allows for the analysis of microbiota quantitively and qualitatively. It allows analyzing different parameters of the cells such as cell size, metabolic state, and density. It is time efficient, it allows for processing more than 10,000 cells/s. cells can be used for other studies. It does not require DNA extraction and does not require amplification and it has high accuracy. One disadvantage of this technique that it relies on liquid samples and it requires complex data analysis (Wang et al., 2009). Since the size of bacteria cells is smaller than the mammalians this may affect the accuracy of the results.

4.2.2.3 Quantitative dot blot

This technique was first introduced to analyze bacterial biodiversity (Amman et al., 1995). The number of 16S rRNA of a targeted microorganism in relation to the total number of 16S rRNA of all microbiota present in a sample. The rRNA is isolated then the known numbers will be plotted on two parallel nylon membranes then hybridized with 32P or fluorescently labeled probes. This helps to quantify the signal intensity of each spot.

The advantages of this method are that they can represent the status of bacterial metabolic activity. It is also commonly used in the determination of the specificity of newly

introduced oligonucleotide probes and the best conditions of temperature and formamide concentration of the hybridization conditions. It is also been used to compare groups of bacteria in samples. This method was compared with FISH and FCM combined and findings indicated that in the detection of some bacteria there were no statistically significant differences of the results of the number of bacteria whereas there were differences in the results of the detection of some other bacteria. This can be due to the difference of the principles of these methods (Namsolleck et al., 2004). Examples for bacteria that results were similar between the methods are Bifidobacterium spp. (3.9%) and Enterobacteria (2.8%); whereas results were different between the methods for the Bacteroides (41.7% dot blot, 9.1% FCM) and Atopobium (0.3% dot blot, 2.8%) (Namsolleck et al., 2004).

The disadvantages of the methods are the application and reading biases. Application bias means that the process includes applying the sample directly to the membraine without using gel electrophorsis which can lead to obtaining quantitative data and only abandance of microorganism is detected no the molecular wight which leads to reading bias .

4.2.2.4 Probes

There are several types of probes which used for targeting a specific gene or genetic segment. The use of oligonucleotide probes based on the 16S rDNA sequences (*in situ* hybridization) were widely used for the direct enumeration of species in samples. Highly specific probes have been designed since the microbiota of the intestine are complex and numerous. The specificity of the probe and its design can be adjusted to fit any species or subspecies (Raskin et al., 1994). Accordingly numerous genus and species-specific PCR primers and probes have been developed for intestinal microbiota.

The detection of bound probes may be achieved by radioactive, enzymatic, fluorescent or chemiluminecent means depending on the compound labeling the probe. If fluorescent dyes used for labeling, then the procedure is known as “fluorescent *in situ* hybridization (FISH).” Tannock (1999) reported that the lowest level of detection (microscopically) using FISH is 10^6 bacterial cells per gram. However, when combination of this method and automated microscopy is used, the expected value should be about 10^7 bacterial cells per gram (Jansen et al., 1999).Another technique, namely dot blot hybridization, was introduced by other workers (Vaughan et al., 2000) for measuring the specific 16SrRNA in a mixture relative to the total amount of rRNA. They briefly explained their technique as follows: the total DNA and RNA are first isolated from the sample, bound to a filter using a dot or slot manifold device and hybridized with labeled oligonucleotide probes. The amount of label bound to their filter is

the measure of specific rRNA target present and the relative amount of rRNA is calculated by dividing the amount of the labeled universal hybridized probe.

Concerning the development of DNA probes, it is to be mentioned here that it has provided workers in this field with great amounts of concepts for detection and identification of microorganisms. However, the hybridization assays have been limited by the number of probes to be tested simultaneously with large number of samples. Tannock (1999) reported that there are at least 30-40 predominant species of bacteria in the human fecal material. In fact it does not seem to be feasible to use one oligonucleotide probe for each species, but the work may be facilitated in the future by applying DNA chip technology (Tannock, 1999). Therefore, suggested that it is more convenient to have probes for the major groups of organisms present in the GI tract. The identification and characterization of such groups ,was carried out by hybridization of fecal .rRNA using three oligonucleotide probes targeted against Bifidobacteria, Enterobacteria and Bacteroides- porphyromonas -Prevotella (Hopkins et al ., 2001). The design consists of a panel of four 16S r RNA –targeted probes specific for four major or predominant groups in the fecal samples. Similar work was previously Carried out (Sghir et al., 2000) who used six probes. They were able to show 70% of the total 16S r RNA detected by the bacterial domain probe.

The design of the probes is based on currently available 16S r RNA sequences (Maidak et al., 2001). Moreover, Amann and Ludwig (2000) claimed that "the specificity is not guaranteed for unknown intestinal microbes". It is, therefore, suggested that further investigations are needed to either confirm or contradict this statement.

4.2.2.5 Polymerase Chain Reaction (PCR) based techniques

Polymerase Chain Reaction (PCR) was developed in 1993 by Nobel Prize winner Kary Mullis. It is used for the amplification of a defined DNA sequence to over a billion times from a single copy. Thermostable DNA polymerase is used for the amplification purpose especially Taq DNA polymerase, deoxynucleotides (dNTPs) and two primers, with complementary sequences to either ends of the targeted DNA. The process includes applying 30-40 cycles of the PCR. The reaction tube is heated to 94° C at the beginning of each PCR cycle to separate the double stranded DNA. Then temperatures are dropped to 55° C or lower allowing the attaching to their target sequences. The temperatures are increased to 72° C enabling the extension of polymers from both primes. This will lead to the duplication copies of regions of DNA between both primers, which become the template for the following PCR cycle. The amplified signal is exponential where one copy can be amplified to twice the time

of the amplification process for example a 35 PCR cycle can generate around 3.4×10^{10} copies.

One advantage of PCR-based technique is that it does not need culturing as it can be performed on small numbers of cells (O'Sullivan, 2000). These methods allow for *in situ* examination of bacteria (Wang et al., 2009). It allows the rapid and specific detection of a wide range of bacterial species and it has become a key procedure for detecting microorganisms even the slow growing ones. It can detect unviable microorganisms and the ones that cannot be cultured or difficult to grow. One disadvantage is that bilirubin and bile salts present in faeces samples can restrain PCR analysis therefore total DNA or RNA purification is required for the process. The occurrence of gram positive and gram negative in one sample can lead to the deferential lysis of the samples due to the difference in their cell walls that can affect the final result of the analysis. A major disadvantage of this method is that primers need to be designed to target all phyla (Fraher et al., 2012).

The PCR-based techniques divided in to:

Checkerboard hybridization

Microarrays

16S rRNA

RAPD

recA gene sequence analysis

Multiplex-PCR

AP-PCR

TAP-PCR

DGGE/TGGE

Gene cloning and sequencing

Real-Time PCR

T-RFLP

4.2.2.6 Checkerboard hybridization

This method is considered highly sensitive which uses universal rRNA primers to magnify the rRNA parts from a sample that allows for probing the amplicon with oligonucleotide probes that are species specific. Its advantages are that it has the ability to detect specific microbes quickly that helps to investigate the distribution of certain

microorganism within a large population of microbiota. This approach can be adapted to analyze multiple samples including multiple probes simultaneously using checkerboard hybridization. The effectiveness of the technique depends on using specific probes (O'Sullivan, 2000).

4.2.2.7 Microarrays

This method is powerful and with high throughput, allowing for the analysis of thousands of genes in a test. It was developed to monitor whole-genome gene expression. It has been also used for systematic and quantitative analysis of microbiota (Paliy et al., 2012; Palmer et al., 2007; Rajilic-Stojanovic et al., 2007). Agilent or Affymetrix platforms are used to support the printed and synthesized oligonucleotide probes, which is based on the entire synthesized genomic DNA, or 16S rRNA genes directly and *in situ* on the microarray solid surface. A reference and labeled sample with fluorescence (Cy3 and Cy5) are fragmented, pooled and then hybridized in a microarray. The presence of the target bacteria can be assessed and quantified by scanning the intensity of the fluorescence. In order to evaluate human intestinal microorganisms, phylogenetic microarrays with high density based on 16S rRNA and Small Subunit (SSU) rRNA genes are used (Paliy et al., 2009; Rajilic-Stojanovic et al., 2007).

In addition to high throughput the advantages of this method can be summarized in the following: it is cost-effective, it allows for direct identification of phylogenetic labels, it is used for more comprehensive study of microbita composition in a sample. It has the effectiveness of detecting low levels of abundance of microorganisms and has been used to compare bacterial communities. The disadvantages can be summarized as the following: some sequences can be hybridized more easily than others, known as hybridization biases (Sekirov et al., 2010). Cross hybridization is also a concern and it does not recognize novel or unknown strains (species) (Wu et al., 2001)

Depending on sample used and the obtained information, this technique can be applied in the following purposes:

1. Microarray expression analysis: this is for the detection of expression of large numbers of genes instead of single gene.
2. Microarray for mutation analysis: the detection of mutations, genomic deletions, and amplifications. It is used to detect Single Nucleotide Polymorphism" (SNP), which is the difference of a single base between two sequences

3. Microarray for protein modification analysis: it is the Detection of protein and protein modifications.
4. Microarrays for microorganisms and antibiotic resistance analysis: Detection and subtyping of microorganisms. The identification and detection of microorganisms is one of the applications of microarray (Rudi & Isaksen, 2012).

4.2.2.8 16S rRNA gene as microbial molecular marker

16S ribosomal RNA gene is a part of 30S small subunit of prokaryotic ribosomes. It contains approximately 1,500 base pairs (Blaut et al., 2002). This gene has been commonly used for phylogenetic identification it helps the scientists to differentiate between different bacterial species in different environments, such as human gut (Blaut et al., 2002; Favier et al., 2002). This widespread use of 16S rRNA as phylogenetic and molecular marker is due to that it is present in all bacteria; it can be directly sequenced, and it has high degree of functional and evolutionary homology (Woese, 1987). Additionally, relationships between 16S rRNAs can be applied in evolutionary relationships, taking into account the lack of artifacts of cross gene transfer.

Nine highly conserved and scattered hyper variable regions (V1 – V9) are detected when of 16S rRNA genere analyzed that include signatures of bacterial phylogenetic groups and species (Baker, Smith, & Cowan, 2003). This characteristic of this gene can be applied for oligonucleotide probes for hybridization engineering; which helps to differentiate bacteria at different levels of taxonomic hierarchy (Blaut et al., 2002). Additionally, the gene's conserved nucleotide sequence can be applied in the design of complimentary primer pairs for PCR amplification of a pool of bacterial 16S rRNA gene fragment (Baker et al., 2003). Designing universal primers is very useful for the amplification the entire bacterial pool of 16S rRNA gene .In addition, primers can be paired for the identification of bacterial species (Rudi, Skulberg, Larsen, & Jakobsen, 1997). The sequenced bacterial 16S rRNA gene fragment can be compared with the ever-expanding 16S rRNA sequence databases in Ribosome database project in order to identify the bacterial taxonomy.

4.2.2.9 Random Amplified Polymorphic DNA (RAPD)

This technique is commonly used (PCR) based technique used to develop DNA markers, which is a modification of the PCR. It is a single, short and random oligonucleotide primer. This primer has the ability to harden and highlight on multiple locations all over the genome. It is able to produce a band of amplification bits that are characteristics of the

template DNA. It has multiple applications for example gene mapping, population genetics, molecular evolutionary genetics. The advantages of this technique speed, cost and efficiency of the technique. It can produce large numbers of markers in a short period, it also can be performed in a moderate laboratory and no prior knowledge is needed of the genome under study (Griffiths, et al., 1996). A disadvantage of RAPD cannot be used for identification of heterozygotes because these are dominant markers because polymorphisms are produced from the mutation and rearrangements within or between primers binding locations and their presence and absence of RAPD band is detected (Kumar & Gurusubramanian, 2011).

The principle of RAPD is that it is a single and short oligonucleotide primer. It can bind to many different loci. It is used to amplify arbitrary sequences from a composite in a DNA template. The length and size of both the primer and the target genome affects the amplified fragment generated by PCR. The principal of this technique depends on a particular DNA sequence that is complementary to the primer, occurs in a genome on the opposite DNA strand, in opposite alignment within a distance that is capable of be amplified by PCR. PCR amplifiable products (of up to 3.0 kb) can be separated on agarose gels (1.5-2.0%) and ethidium bromide staining is used to visualize the PCR amplifiable products. This technology uses 10 bases long synthetic oligonucleotides of random sequences as primers to amplify nanogram amounts of genomic DNA under low hardening temperatures (Kumar & Gurusubramanian, 2011).

4.2.2.10 *recA* gene sequence analysis

The recA gene is a sensitive molecule for determining intrageneric phylogenetic relationships, it can be used to study natural ecosystem on a large scale like human intestinal ecosystem (O'Sullivan, 2000). It helps for encoding the RecA protein. This protein has a critical role in the recombination and DNA repair in response to SOS (Roca and Cox, 1997). Bacterial phylogenetic relationships can be studies through the analysis of this protein (Eisen, 1995; Karlin et al., 1995).

4.2.2.11 Multiplex-PCR

This method uses one set of primers or more to allow the amplification of targeting a number of DNA regions simultaneously. This theory behind this method is that the amplification of more target areas allows for more reliability of the technique. The main

disadvantage of the technique is that prior sequence knowledge is needed and optimal reaction conditions are challenging to imitate (O'Sullivan, 2000).

4.2.2.12 Arbitrary primed (AP) PCR

This technique uses a single short primer, varying between 10-12 bases that are chosen arbitrarily. This allows the primer annealing to the DNA template. The strength of the reaction is reduced to allow the primer to bind to regions that shows closest homology. The DNA regions between the opposite strands can be amplified with primer binding sites which are within a few thousand bases. The discrimination of this technique increase when more products are amplified (O'Sullivan, 2000).

4.2.2.13 Triplet arbitrary primed (TAP) PCR

The basis of this technique is that unintended changes in reaction conditions lead to the low reproducibility of arbitrary priming. The identification of the amplicons that share usceptible to changes can be through deliberately introducing particular changes to the reactions in three identical reactions (O'Sullivan, 2000). All three reactions are performed paralleled at three different annealing temperatures (38°, 40° and 42° C) and comparing the banding patterns follows gel electrophoresis for each reaction. Bands that are considered in the fingerprint analysis are the ones present in at least two lanes and are considered resilient to small changes in reaction conditions.

4.2.2.14 Denaturing gradient gel electrophoresis/ in temperature gradient gel electroporesis (DGGE/TGGE)

DGGE method is used to study the microbial diversity in complex ecosystems. 16S rDNA fragments are amplified with PCR, using universal primers. Then polyacrylamide gels are used to separate the fragments of 16S rDNA. The gel contains a gradient of denaturing agent such as urea or formamide. This process helps in separating the amplicons, even the ones with similar lengths. Different amplicons with heteroduplex and different guanine and cytosine content are separating at different points in the denaturing gradient, which hinders migration of bands. The melting point of rRNA gene depends on its nucleotide sequence is the principle behind this method. Gene sequence starts melting at a different denaturation point. The conformation of the DNA molecule changes due to melting and hinders the migration within the gel (Fraher et al., 2012). A pattern of bands are resulted from this

melting process that represents the characteristic of the bacterial community and species are represented with a band.

The TGGE methods are similar to the DGGE, separating DNA fragments temperature gradient gels, universal primers like 16S rDNA is used allowing for the estimation of the bacterial content of a sample (Zoetendal et al., 1998). Primers are helpful for studying bacterial communities at higher resolution. Further hybridization used to study bacterial communities in depth and correct identification of bacterial groups. The advantage of method is that it allows for studying predominant bacteria that are difficult to culture (Zoetendal et al., 1998).

DGGE and TGGE are used to compare two different bacterial communities. It time efficient and allows for analysis of several sample at the same time. This method is semiquantitative because it allows for studying the intensity and abundance of bacteria visually. Due to the unreliable amplification dynamic does not allow for an exact comparison of bacterial communities. Another disadvantage is the bias introduced by the PCR process. Unless the probe hybridization is carried out it is not possible to perform direct phylogenetic identification. The difference between TGGE and DGGE is that the latter uses a linear temperature gradient, whereas the first uses denaturing gradient gel.

4.2.2.15 Gene cloning and sequencing

This method uses a bank of gene clones that have been generated by the PCR amplification process. The heterogeneous product of from the amplification of 16Sr RNA of a sample DNA is cloned into standard sequencing vectors then they can be sequenced and phylogenetically analyzed (O'Sullivan, 2000). Chain termination by dideoxynucleotides is the principle of this method. Sanger sequencing method is used to the sequencing of cloned full-length 16S rRNA gene amplicons. This method was used for studying the variability of intestinal microbiota and uncultured bacteria. And the disadvantages of this method is that it due to the generation of large number of clones it can be labor intensive (Fraher et al., 2012).

4.2.2.16 Real-time PCR or quantitative real time PCR (qPCR)

Herbel (2013) described these methods as the same based principle of PCR techniques. However it is different from PCR techniques by that it can count the reactions after each reaction cycle. SYBR® Green, TaqMan® labelled primers, or molecular beacons is used as markers. SYBR® Green is a fluorescent dye used for binding DNA and has the tendency to bind to double stranded DNA (dsDNA) (Castoldi et al., 2013), whereas, the

TaqMan® labeled primer fluoresce after binding with its DNA compliment. To detect a species of a strain a TaqMan® labelled primer is designed to be species-specific and to link to sequence internal side universal primers. Hairpins that are not fluorescent form molecular signal probes form and are in non hybridized state (Meng et al., 2012). These methods are used for the detection and quantification of a strain; it can be possible without using further post-PCR analyzes steps, when a strain's specific sequence is known.

qPCR can be used to study bacteria in complex bacterial communities using a quantitative approach (Miller et al., 2012). The advantages of this technique are that it can be performed quickly and is highly sensitive. It can be specific and accurate for the detection of minor populations of bacteria (Postollec et al., 2011). It allows for the detection and quantification of bacteria in microbial communities simultaneously (Sohier et al., 2012). It is also cost-effective. One disadvantage summerized by Rudi and Isaksen (2012) is that this method is not applicable for monitoring changes in intestinal microbiota due to its specificity in the detection of individual bacteria. It cannot detect novel species because it has been designed to detect known species. In the case of the unavailability of suitable strains it is challenging to generate the standard curve required for the interpretation.

4.2.2.17 Terminal-restriction fragment (T-RFLP) analysis

T-RFLP length of a known 16S rRNA gene sequences will used to detect sequences of similar lengths, therefore, the prediction of bacterial species can be based on their T-RFs. It allows for the detection or prediction of multiple species from the same T-RF length. Multiple restriction enzymes of digests obtained for analyzing to identify bacterial species by extracting DNA and after PCR amplification the DNA then the purification and digestion with restriction enzymes is performed (Matsumoto et al., 2005). The advantage of this analysis is that it used for comparing complex bacterial communities. It can be performed it is time and cost effective. The disadvantage of this method is that it does not allow for phylogenetic studies. However, this limitation can be overcome by combining this method with 16S rRNA clone library.

4.3 Sequencing

Although it is considered the gold standard method for the identification of taxonomy of bacteria down to the level of species, full-length 16S rRNA gene (>1,500 base pairs long) information is needed that is sequenced from only a clone library insert. Then the resulted sequences are compared with a database to identify the microorganism by sequencing the full-length 16S rRNA gene. In order to delineate the taxonomic rank of a species a sequence deviation range of 0.5–1% is used in addition to a 97% cut-off point to define operational taxonomic units.16S rRNA amplicon are used for direct sequencing or by removing of the bands from a gel and then reamplified bands removed by PCR. Sanger sequencing is commonly used and then developed to the next-generation sequencing technologies for more time and cost efficiency. A brief description, advantages and disadvantages are listed in table 3, 4 and 5.

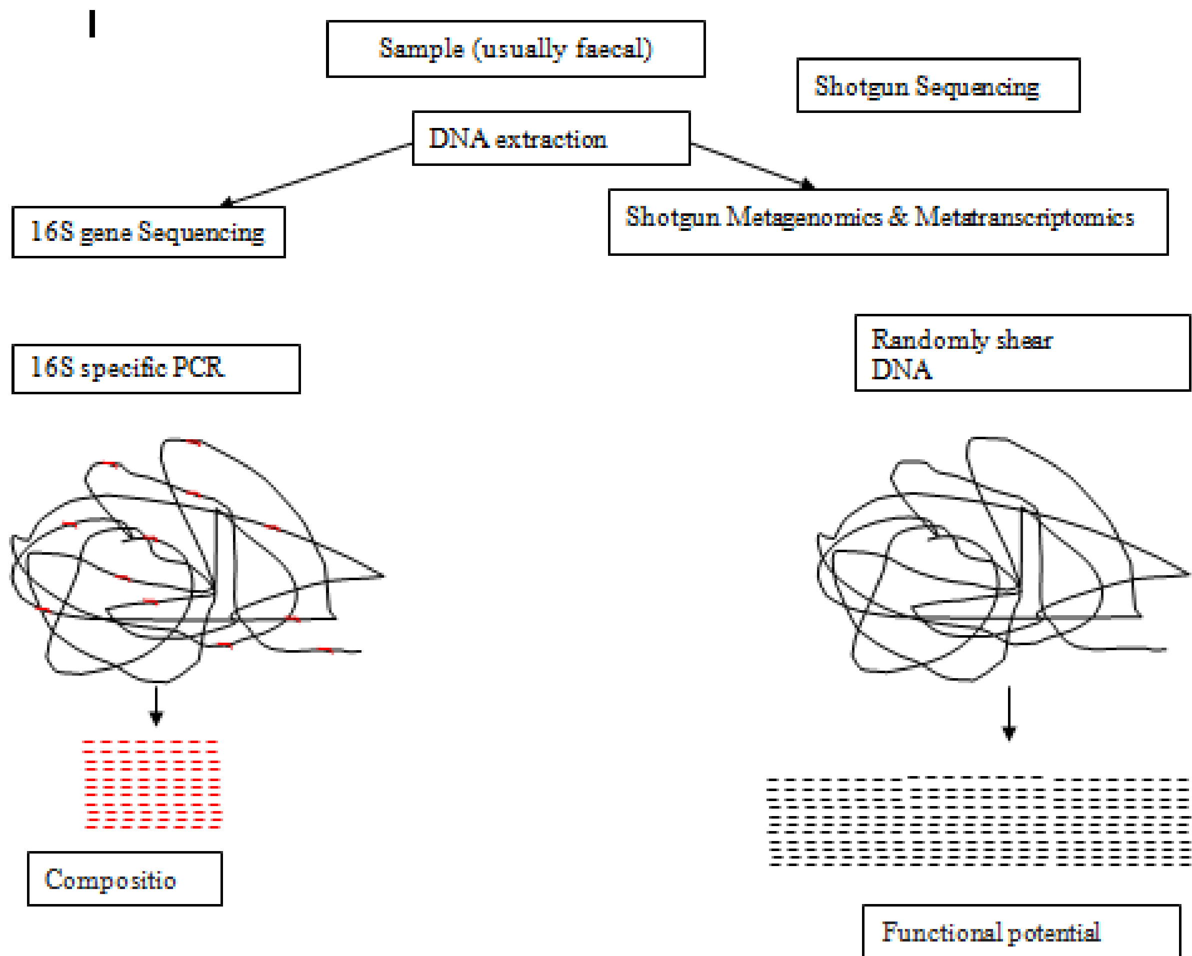


Figure 5: sequencing by both methods: Sanger and NGS

Table 3: Next Generation Sequencing platforms. (Loman, N. J. et al., 2012).

machine (manufacturer)	chemistry	Advantages	Disadvantages
454 GS FLX+(Roche)	Pyrosequencing	• read length is long • high throughput • sensitve • analysis of multiple sample in the same time • no cloning bias	• in hands-genes • high reagent costs, • high error rate in homopolymers • Short sequencing reads • Extensive bioinformatics required
HiSeq 2000/2500 (Illumina)	Reversible terminator	• effective • steadily improving read lengths • massive throughput • minimal hands - on time	• long run time • short read lengths • upgrade is under development
5500xl SOLiD (Life Technologies)	Ligation	• low error rate • massive throughput	• very short lengths • long run time
PacBio RS (Pacific Bioseince)	real-time sequencing	• sample easy to prepare • reagent are low in costs • very long read length	• error rates are high • system is highly costly • installation is difficult
Ion torrent (Life Technologies)	proton detection	• short run time • flexible chip reagents	• instrument under development

Table 4: Comparison of Sanger sequencing method with some next generation sequencing technologies (Kircher & Kelso, 2010) cited from (Naseribafrouei, 2013).

	Throughput	Length	Quality	Costs
Sanger	6Mb/day	800nt	10-4 – 10-5	~500$/Mb
454/Roche	750Mb/day	400nt	10-3 – 10-4	~20$/Mb
Illumina	5,000Mb/day	100nt	10-2 – 10-3	~0.5$/Mb
SOLiD	5,000Mb/day	50nt	10-2 – 10-3	~0.5$/Mb
Helicos	5,000Mb/day	32nt	10-2	<0.5$/Mb

Table 5: Comparison of currently available next-generation sequencing technologies (Shokralla et al., 2012)

Category	Platform	Read length (bp)	Max. number of reads / run	Sequencing output / run	Run time
PCR-based NGS technologies	Roche 454 GS FLX	400–500	$1 * 10^{6}$	<= 500 Mb	10 h
	Roche 454 GS FLX+	600–800	$1 * 10^{6}$	<=700 Mb	23 h
	Roche 454 GS Junior	400–450	$1 * 10^{5}$	35 Mb	10 h
	Illumina HiSeq 2000	100–200	$6 * 10^{9}$	<=540–600 Gb	11 d
	Illumina HiSeq 1000	100–200	$3 * 10^{9}$	<=270–300 Gb	8.5 d
	Illumina GAIIx	50–75	$6.4 * 10^{8}$	<=95 Gb	7.5–14.5 d
	Illumina MiSeq	100–150	$7 * 10^{6}$	<=1–2 Gb	19–27 h
	AB SOLiD 5500 system	35–75	$2,4 * 10^{9}$	100 Gb	4 d
	AB SOLiD 5500 xl system	35–75	$6 * 10^{9}$	250 Gb	7–8 d
	Ion Torrent -314 chip	100–200	$1 * 10^{6}$	>=10 Mb	3.5 h
	Ion Torrent -316 chip	100–200	$6 * 10^{6}$	>=100 Mb	4.7 h
	Ion Torrent -318 chip	100–200	$11 * 10^{6}$	>=1 Gb	5.5 h
SMS technologies	Helicos HeliScope	30–35	$1 * 10^{9}$	20–28 Gb	<=1 d
	Pacific Biosciences system	>= 1500	$50 * 10^{3}$	60–75 Mb	0.5 h

4.3.1 Sanger sequencing

Over the last 30 years, Sanger sequencing has been a gold standard for DNA sequencing techniques. Briefly this approach is performed in the following steps: after the DNA purification, and then labeled by chain termination method using dye-labeled dideoxynucleotides (ddNTPs), then capillary electrophoresis, finally detected through fluorescence detection. Lately, this method has been developed to allow for longer sequencing reads to approximately 800 bases, and lower error rate (700 bp or less), and larger insert sizes. Sanger sequencing has been contributing to scientific advances in many areas such as studying intestinal microbiota. In environments with lower diversity this method can produce up to complete genome sequencing. Sanger method is associated with the following disadvantages: analyzing large number of cloned genes in large number of samples requires technical difficulties and high cost (e.g. overall per gigabase is 400,000 USD), therefore, subdominant bacteria are challenging to be revealed (Fraher et al., 2012).

4.3.2 Shotgun sequencing

It is developed by the twice Nobel Prize winner Fredrick Sanger in 1970. This method involves fragmenting the entire genome into a series of short sequences and then cloned; each fragment is sequenced and then reassembled. Then the reassembled sequenced fraqments are analyzed for the gaps, single strands and double strands to generate the master sequence. The disadvantage of this method is that it is only used for sequencing microorganisms and lesser complex genomes. Another disadvantage is that it may lead to errors when overlapping fragments are repeated (Brown, 2006).

4.3.3 Next Generation Sequencing Platforms Capable of RNA Sequencing:

The NGS are ground breaking technologies in analyzing the intestinal microbiota. It consists of the following techniques and will be discussed in terms their principles, advantages and disadvantages in the following sections of this paper:

- Roch 454 GS (Pyrosequencing)
- Illumina HiSeq™ 2000 (Official Service Provider)
- SOLiD v4 (Official Service Provider)
- Ion Torrent PGM™
- Helicos Biosciences HeliSope

- Pacific Bioscience SMRT DNA sequencer

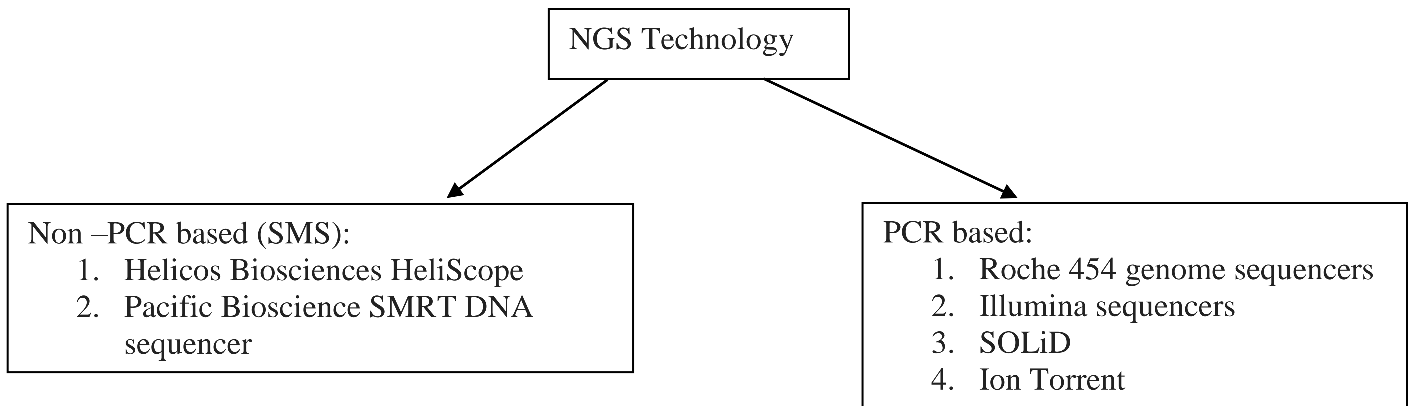


Figure 6: This digram shows the different NGS platforms technologies discussed in this paper.

4.3.3.1 Roche 454 genome sequencers (pyrosequencing technology)

It was introduced NGS technology in 2005. It utilizes real-time sequencing-by-synthesis pyrosequencing technology. The nucleotides are combined by DNA polymerase, which results in the release of a pyrophosphate molecule. As a result a light will be produced by the action of the enzyme luciferase, which has initiated a series of downstream reactions. The volume of the light generated from the reactions is directly proportional to the number of the combined nucleotides (Margulies et al., 2005). The 454 pyrosequencing includes the restriction of the library fragments on Sepharose or styrofoam beads, which on their surfaces carry oligonucleotides complementary to the 454-specific adapter sequences ligated or PCR-generated onto both ends of the fragmented library. Oil micro-reactors contain PCR ingredients, therefore the emulsion of the PCR thermal cycling into individual water is used to amplify the library fragments. The amplification of each library fragment is done on the surface of one bead in a single micro-reactor. This amplification process generates same fragments in billions of copies. The amplified beads are then retrieved from emulsion oil then they are enriched to keep only the amplified beads. The enriched beads are then prepared as single-stranded and strengthened to a specific sequencing primer.

These beads are then arranged into a picotiter plate (PTP) that has the capacity of one million wells per plate, each which can only hold one amplified DNA bead. Four layers of processed beads are deposited into the PTP. By centrifugation, all the layers are deposited from bottom to top, diluted pyrosequencing enzyme beads, DNA amplified beads, pyrosequencing enzyme beads and, then, PPiase enzyme beads. Later the PTP is sequenced in

the 454 GS pyrosequencing instrument. Throughout the sequencing steps a flow of nucleotide solutions (T, C, A and G) regularly introduced. A CCD camera is used to record the light the emitted from each bead accompanied with the flow of the nucleotide solutions. To generate up to 800-bp sequencing reads, the GS FLX+ system provides 200 nucleotide flow cycles. 454-pyrosequencing-analysis software is used to process the generated raw signals and then screened by several quality filters to remove poor-quality sequences (Mardis, 2008a).

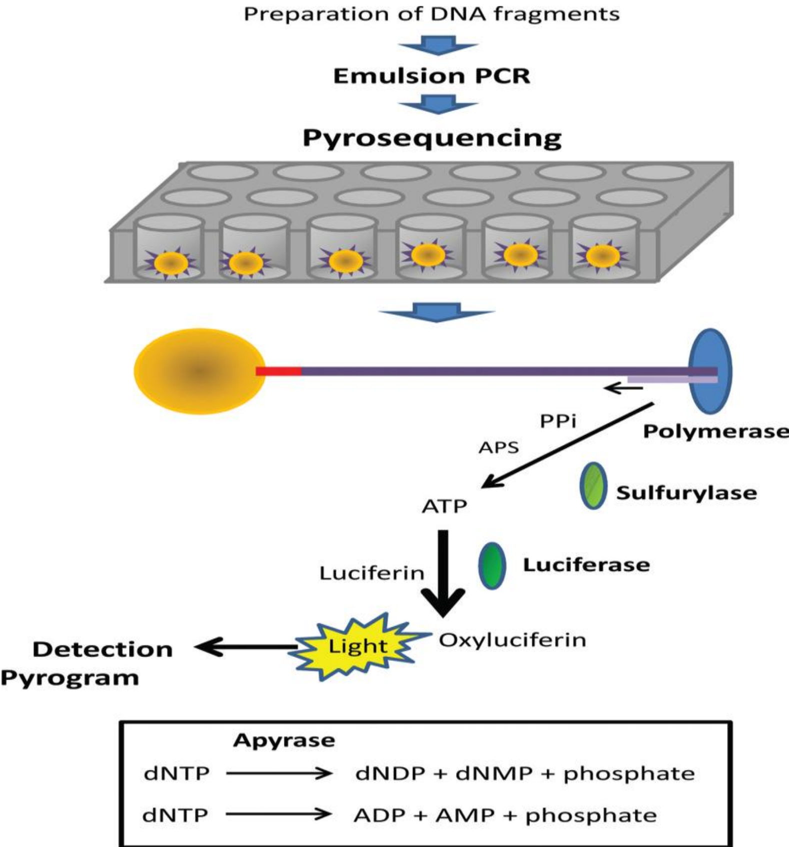


Figure 7: principles of pyrosequencing Rastogi, G., & Sani, R. K. (2011).

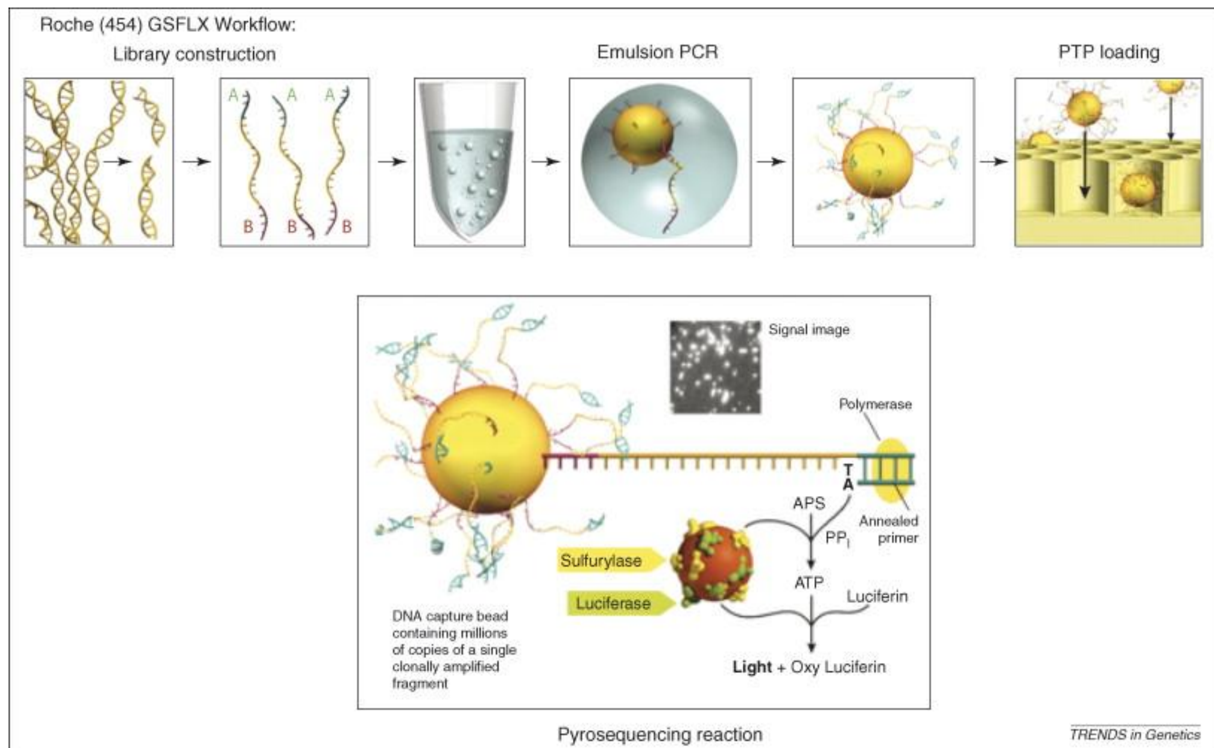


Figure 8. Shows Roche 454 workflow which is discussed in detail in the text.

4.3.3.2 Illumina sequencers

It was formerly known as Solexa, it was introduced 2007. Due to its high capacity it is mainly used in resequencing applications, including human and model organism genomic projects. Similar to Roche 454 genome sequencers it utilizes a sequencing-by-synthesis approach but it is combined with bridge amplification on the surface of a flow cell, each divided into eight separate lanes. The library fragments linked to oligos through adapters, which are attached to the interior surfaces of the flow cells. Active heating and cooling steps are done to help in the DNA fragment-to-oligo hybridization on the flow cell. The process of generating millions of clusters of library fragments is through the incubation and with the amplification reactants and an iso- thermal polymerase. The sequencing step includes supplying the cluster with polymerase and four fluorescent nucleotides that are labeled separately with chemically inactivated OH.

This step helps as a blocking step ensuring the incorporation of a single base per flow cycle. After the incorporation of each nucleotide, then stimulation and then imaging step occurs to isolate the incorporated nucleotide in each cluster. To remove the fluorescent group, a chemical deblocking treatment step to allow the flow of the incorporated nucleotide for the next flow cycle. Quality filtering step occurs after the computation of the sequence of each

cluster to separate the low-quality reads (Shendure & Ji, 2008). The four commercially available versions are: the HiSeq 2000, HiSeq 1000 and Genome Analyzer IIx. They have up to 600, 300 and 95 Gb sequencing outputs, respectively. The newest HiSeq2500 platform can generate up to 120 Gb of data in 27 h. This version allows researchers to sequence an entire genome in 24 h. Figure 9 illustrate amplification processes used in SOLiD and Illumina systems.

4.3.3.3 Applied Biosystems SOLiD sequencer (Life Technologies)

Applied Biosystems (Life Technologies) introduced the SOLiD sequencer in 2007. It is different from Roche 454 genome sequencers and Illumina sequencers that it is based on sequencing- by-oligo ligation technology. Oligo adapter-linked DNA fragments with matching oligos settled on a 1-mm magnetic beads' surface. An emulsion PCR is used to individually amplify the beads. Then the beads attached to the surface of a specially treated glass slide, which then placed in a sequencer on a fluidics cassette. Then the solidification of a suitable sequencing primer to connect the SOLiD-specific adapters to the library fragments. Four semi-degenerate 8-mer fluorescent oligos combined with DNA ligase are added in a regular manner in the instrument. Then the oligio mixed with the DNA fragment sequence next to the universal primer.

Therefore, the DNA-ligase seals the phosphate backbone. Then the ligated 8-mer oligo are identified by a fluorescent readout, corresponding to one of the four available bases. Then a chemical cleavage occurs between the fifth and sixth bases of the 8mer oligo, to continue the ligation round the fluorescent group will be removed. The hybridization of an n-1, n-2, n-3 and n-4 positioned universal primer starts the second sequencing round, including rounds of ligation-mediated sequencing occurs. The five universal primers generate florescence to be decoded with a two-base calling processing software. A single run can process two slides, while sequencing reagents is applied on the first slide, and the second slide is being image (Mardis, 2008b). Two versions of Applied Biosystems SOLiD sequencers are available, the 5500 system and the 5500xl system, with up to 100- and 250-Gb sequencing capacity, respectively. Figure 9 illustrate amplification processes used in SOLiD and Illumina systems.

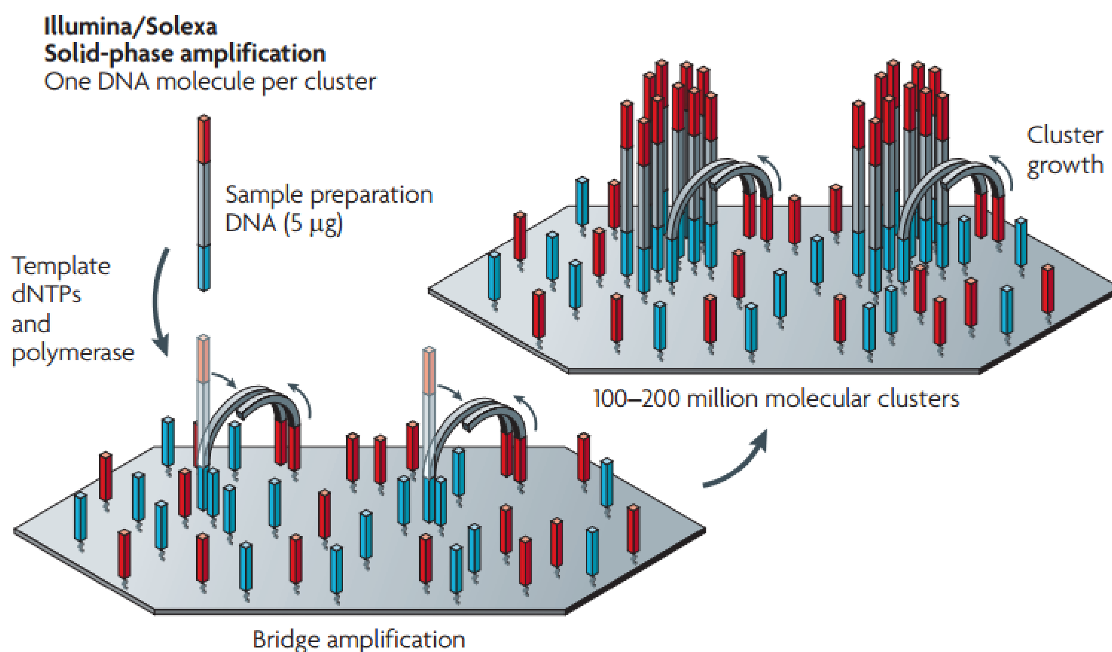


Figure 9. Shows Illumine / SOLiD sample preparation which is discussed in detail in the text (Metzker, 2010).

4.3.3.4 Life Technologies Ion Torrent

Life Technologies introduced the Ion Personal Genome Machine (PGM) in 2010. It uses the postlight sequencing technology. A nucleotide is incorporated into a strand of DNA by the polymerase. This releases the hydrogen ion concentration, which allows for its detection. The high-density array of micro-machined wells is used to perform this process, simultaneously and in large numbers. The well carries DNA templates from the library. The change in hydrogen ion concentration due to the nucleotides incorporation is detected by an ion-sensitive layer under the wells and a proprietary ion sensor (Rothberg et al., 2011). The Ion Torrent platform uses 314, 316, or 318 ion chips that can produce up to 10 Mb, 100 Mb or 1 Gb, respectively, per the required sequencing coverage. The new generation of Ion semi-conductor sequencers: the Ion Proton bench top sequencer, that is available in two versions: Ion Proton I chip with 165 million wells (about 100-fold more than the Ion 314 chip); and Ion Proton II chip with 660 million wells (about 1000-fold more than the Ion 314 chip). These newer versions use CMOS semiconductor chip technology to decode these chemical changes into digital data.

4.3.3.5 Single-molecule DNA-sequencing technologies

4.3.3.5.1 Helicos Biosciences HeliScope

It was introduced in 2008; it was the first single-molecule sequencing (SMS) system available. It uses the sequencing-by-synthesis on a single DNA molecule technique (Harris et al., 2007), which constitutes the bases for the library construction. It doesn't require an amplification step. A combination of DNA polymerase and one of the four fluorescently labelled nucleotides is added in a regular fashion, during the sequencing cycles. These results in the extension of DNA that is template-dependent based on the flowed nucleotide. In order to capture the incorporated nucleotide's fluorescence, they are modified to stop the polymerase extension and then they are recorded with a highly sensitive CCD camera connected to a fluorescent microscope. Then the unincorporated nucleotides are washed away in addition to the residue of the previous cycle. Another cycle of single-base extension follows the chemically cleaving and removing of fluorescent labels on the extended strands (Zhang et al., 2011). This technique can produce approximately 1 billion sequence reads. Figure 10 is an illustration for single molecule primer immobilization.

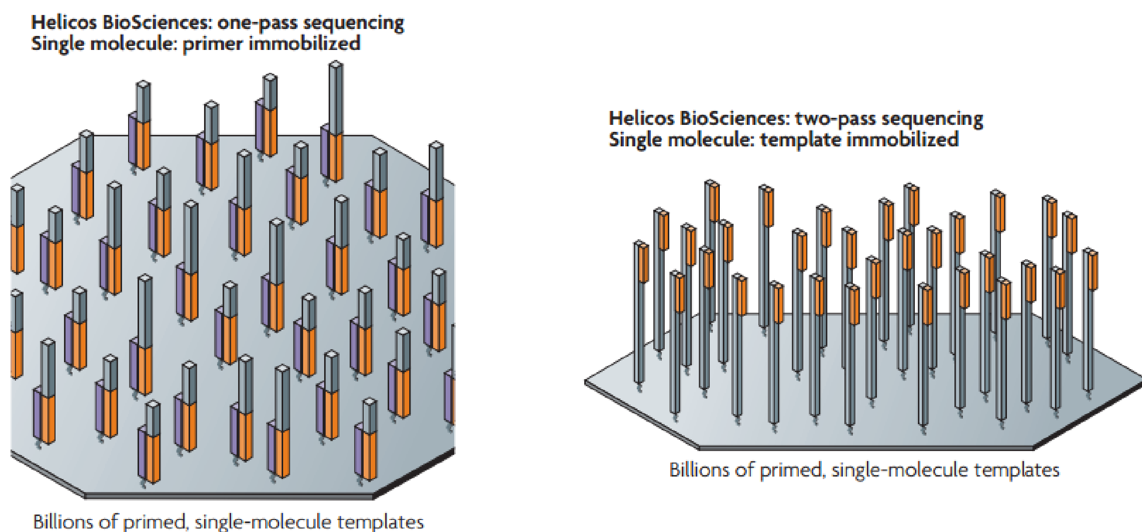


Figure 10. Shows single molecule Helicos BioScience immobilization principle, discussed in detail in the text (Metzker, 2010).

4.3.3.5.2 Pacific Biosciences SMRT DNA sequencer

Pacific Biosciences introduced it in 2010. It uses a single-molecule real-time fluorescence-based, SMS platform (Korlach et al., 2010). It is a single- molecule sequencing-by-synthesis approach, which doesn't require an amplification step for sample preparation.

For real-time observation of DNA polymerization, this technique uses a nano-structure called a Zero Mode Waveguide (ZMW) .Tens of thousands of sub wavelengths with, ten nanometer diameter holes worked perforating a thin metal film by supported by a transparent substrate. While sequencing, matched DNA strands are synthesized from the single-stranded template through the DNA polymerase, which are placed at the lower level of the waveguides. There are four multicolored nucleotides that are phosphor-linked, are used in the platform. The terminal phosphate group holds the fluorescence label and not the nucleotide base; this helps to release the fluorescence moiety with the nucleotide incorporation (Flusberg et al. 2010), which distinguishes this technology, is different from other technologies. Washing step between each nucleotide flow is avoided using this technology, which reduces the time needed for the nucleotide incorporation this also improves sequence quality. This approach utilizes the natural capacity of DNA polymerase for the sequencing ten or more nucleotides in a second in several thousand parallel ZMWs (Zhou et al., 2010). Figure 11 illustrate the principle of polymerase immobilization of threads.

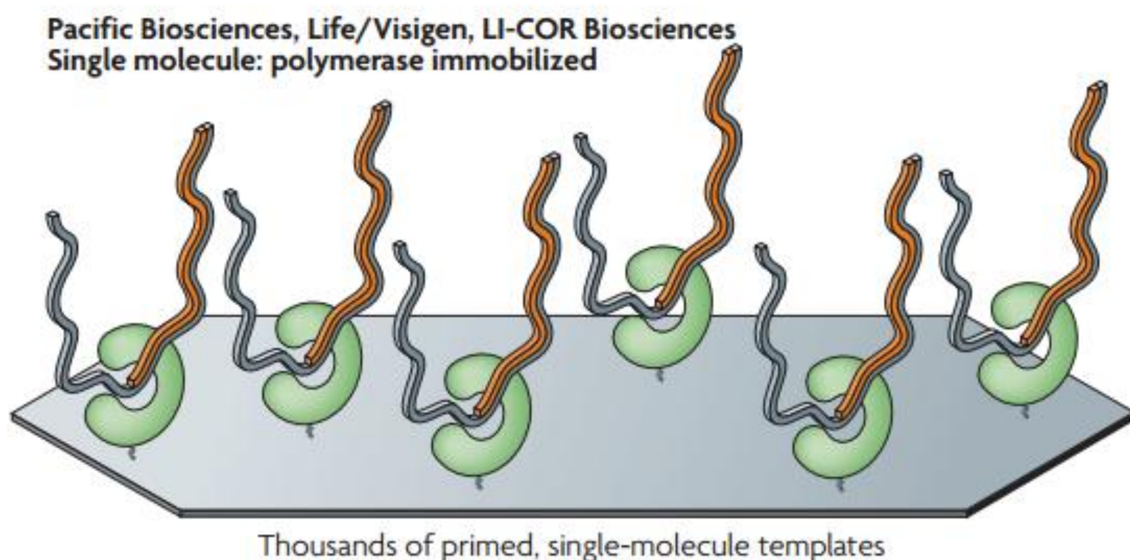


Figure 11. Shows Pacific Bioscience immobilization principle, discussed in detail in the text (Metzker, 2010).

4.4 Advantages and disadvantages of NGS Platforms

There are several advantages related to the use of 454 pyrosequencing platform, one is its high reading length (600-800 bp) and its relative short running time. In a single run, pyrocequencing sequences 500 million bases with accuracy rate more than 99%. It is superior to Sanger sequencing, because it allows for read of 2000 times increase in throughput that

helps in detecting the abundance of minor bacteria. These technologies do not require an additional chemical deblocking step for allowing DNA extension through the action of the DNA polymerase, thus reducing the chances of the two causes of dephasing, which are the termination of chains prematurely and nonsimultaneous extension (Zhou et al., 2010). This platform is commonly used for the analysis of environmental DNA for ecological applications due to its capacity for generating longer sequences with higher flexibility with accurate annotation of reads in ecological applications for nonmodel organisms. Unknown bacteria can be detected with these methods in addition phylogenetic identification of microbiota.

Where as the disadvantages of 454 pyrosequencing platform can be summarized in the following: the challenges concerning the reading of homopolymer regions due to the termination of the division to stop the extension run. The error type associated to this platform is insertion–deletion instead of substitution, which can be limiting in the analysis of environmental DNA, unique haplotypes indicating rare biota can be the results of these sequence errors. These errors have been reduced and can be identified and separated using computational tools (Quince et al., 2009). Cost associated with reagents per megabase sequencing output during the use of this platform is another disadvantage (Claesson et al., 2010), it can reach up to 20,000 USD per gigabase. There are reading biases associated with the 454 pyrosequencing, artificial replicate of sequences affect the estimation of gene abundance, therefore knowledge and understanding the replicas is crucial so they could be filtered out later. Another bias related to 454 pyrosequencing intensity of light may indicate the true number of nucleotides positions.

4.4.1 Illumina and SOLiD systems

The advantage is of these systems are that the sequencing of the homopolymer regions are relatively accurate because the process of the nucleotide detection is carried out one at a time. The chemical deblocking step is carried out prior to the next nucleotide incorporation or in the prior to further ligation in Illumina system and SOLiD systems respectively. High output per run is the second advantage of these two systems compared to 454 pyrosequencing. However, one disadvantage of these two systems is that their relative short-read length due to the decay of optical signals and dephasing, which limits the application of these technologies in analysis carried out without reference sequence used to align, assign and annotate the generated short sequences. In both systems, error rate is accumulative per longer sequencing reads (Zhou et al., 2010).

4.4.2 PCR-based NGS systems

There is bias introduced during amplification process, which can affect the results in the following two stages: the first bias is introduced during amplicon library preparation step. Temperature is an important factor for primer hardening and binding (Polz & Cavanaugh, 1998). Investigating the annealing temperature using denaturing gradient gel electrophoretic analysis in order to reduce PCR bias of primer sets. Another bias can be reduced at lower temperatures when achieving a specific amplification. Bias can also be strongly associated to the number of replication cycles, which can be reduced by lowering the number of cycles low Another strategy to reduce the amplification bias can be through using PCR cyclers with a fast ramping rate by identifying the fastest ramping rate from the denaturation step to the annealing step; this strategy can increase the chances of the formation of heteroduplex when PCR reaches the plateau phase. PCR bias can be reduced using effectively using concentrations with high templates, selecting primers wisely, and lower cycle numbers, lower hardening temperatures and mixed replicate reaction preparations (Lim et al., 2010). Bias can be introduced later during the library amplification step prior to sequencing through emulsion PCR or bridge PCR. Schuster (2008) found the despite the use of universal probes with bias-free amplification quality, it can exaggerate biased amplification in the original amplicon library preparation. Single-molecule non-PCR sequencing technologies do not use template amplification step, which helped in eliminating amplification bias.

4.4.3 Improvements of NGS technologies

Sequence capture is the newest strategy to improve the sequencing results and reduce or eliminate the limitations and bias associated with PCR platforms. This strategy allows the PCR platforms to analyze large numbers of DNA target sequences. This strategy involves two hybridization-based methods through oligonucleotide probes: first, is immobilized to a solid array ‘Capture arrays’; second is solution ‘Baits’, to capture the sequencing targets (Lee et al., 2011). Hybridization probes (60– 120bp) are designed to capture target regions across the genome, whereas the second process involves washing to remove unspecific hybrids and then targeted DNA is eluted for sequencing. The sequencing from the second process has shown to enhance uniformity and specificity; while the first strategy is costly since it requires a hybridization station. These processes have allowed to efficiently using the NGS for population genetic analyses of ancient DNA samples (Horn, 2012) and for applications in

environmental DNA research (Adey et al., 2010; Barzon et al., 2011; Jones et al., 2011; Faircloth et al., 2012). Modifying sample preparation protocols (e.g. library construction) are required as an enhancement strategy especially for Illumina platforms (Caruccio., 2011).

Tagged or bar-coded DNA templates for sequencing have allowed reading millions of DNA sequences in parallel to be applicable for large-scale biodiversity analyses of environmental samples and possesses large potential for many applications (Binladen et al., 2007). Speeding the ecological studies by allowing multiplying several target gene markers of single bulk samples or multiplying of a single marker from multiple samples (Xu et al., 2012). Tags can be designed considering the use of sequencing chemistry to decrease the chances of uncertainties due to potential sequencing error, for instance tags should not start or end with the same nucleotide as same as the sequencing chemistry adaptor ends or as the amplification primer starts; allowing for two or less identical successive nucleotides within the unique tag Human gut microbiome is one of the fields that has benefited from the tag-encoded strategy as one of other benefiting fields , which can carried out within a reasonable cost (Sun et al., 2011).

4.5 Application of NGS for Analyzing DNA Intestinal Microbiota

The capacity of sequencing in massive amounts has been a critical advancement for the sequencing of environmental samples of ecology and biodiversity research in recent years. The NGS technologies have made it possible to analyze environmentally originated samples from various ecosystems such as freshwater, marine, soil, and gut microbiota. The use of the massive sequencing through NGS platforms made possible to observe the slim changes in community structure that can occur caused by anthropogenic or natural environmental changes which is feasible using the traditional sequencing tools such as Sanger sequencing (Shokralla, 2012). Diet analysis and gut microbiota analysis and its ecology is one of the applications of the NGS technology including the analyses of herbivore diet from gut contents using the plastid trnL sequence the effect of diet on the gut microbiome of mice using 16S rDNA amplicons the diet of bats was conducted using short COI amplicons (Shokralla, 2012). NGS application to diet analysis enables for a comprehensive relationship of the diet of sympatric ambiguous species by enabling species-level identification of dietary components (Razgour et al., 2011). In addition to 454 pyrosequencing, Illumina platforms have been utilized for analyzing microbial community diversity using short fragments of 16S rDNA.

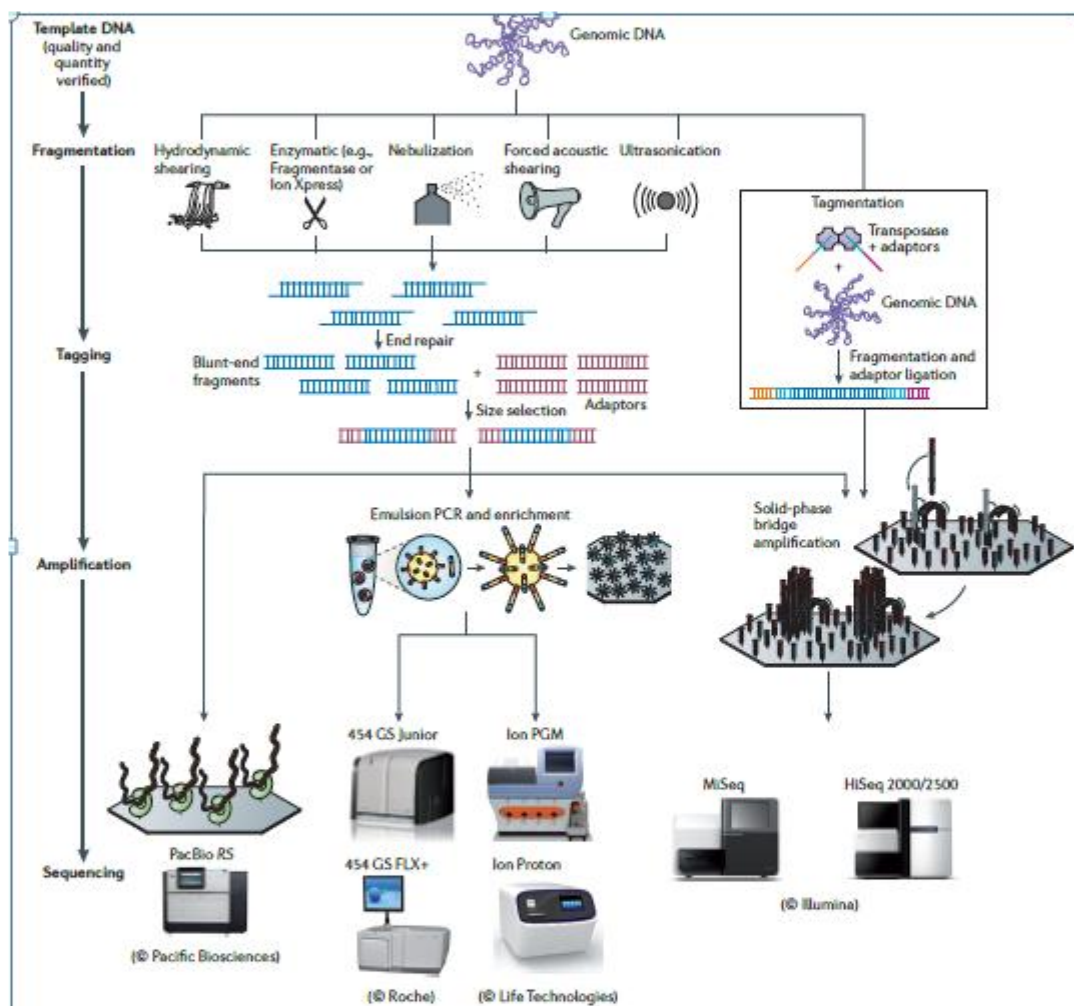


Figure 12: represents different platforms of NGS.

4.6 Tools for functional studies of intestinal microbiota

There are new approaches to study microbiota through its molecular function. These techniques includes fingerprinting techniques of 16S rRNA gene amplicons, DNA sequencing of 16S rRNA gene clones, FISH, flow cytometry, DNA microarrays, and high throughput sequencing with 16S rRNA genes as the target .Literature shows that the gut microbiota composition varies among individuals due to host genotypes, age, health status, and diets. Whith the use of methods targeting 16S rRNA genes has provided insights into the functional characteristics of the microbiota and their influences on host health. To fully understand the contribution of a particular bacterial group to host health, the development of advanced tools is essential. Their applications in functional studies are discussed separately in the following sections (Gong et al., 2012).

4.6.1 Stable isotope probing (SIP)

It is used to connect the identity of microbes to their function in the community. Commercially prepared substrates (i.e., starch) are used by enriching it in a stable isotope, such as 13C. Then the identification of active microbes is conducted by the selective recovery and analysis of isotope-labeled cellular components. The advantage of this technique is it can be used to explain metabolic activities and identify new potential pathways in microbial communities. It can be used to identify the bacterium or bacterial group responsible for the metabolic activity. It can be used for in vitro (Egert et al., 2007) and in vivo studies (Bombach et al., 2010). However, the main disadvantage of this technique is the low resolution of the density-gradient ultracentrifugation. This can only produce nucleic acids with large differences in the degree of isotope incorporation (Kovatcheva-Datchary et al., 2009).

4.6.2 'Omics'

'Omics' refers to a range of techniques with high throughput sequencing that allows for a large-scale analysis of microbial communities. These techniques include metagenomics, metatranscriptomics, metaproteomics, and metabolomics. These techniques are used to analyze the DNA, mRNA, proteins and metabolites of the gut microbiota then generally analyzing the complex ecosystem functions of the gut microbiota (Gong et al., 2012). A brief description, advantages and disadvantages of the 'omics' groups are listed in table 6 and figure 13

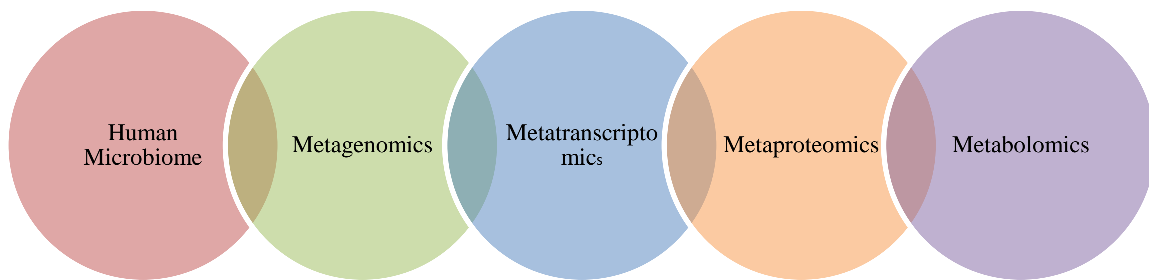


Figure 13: The overlap among the"omics" platforms.

Metagenomics provides DNA information; metatranscriptomics provides RNA Information; metaproteomics provides protein information; metabolomics provides information of metabolites of the small-molecule.

Table 6: Culturing independent techniques 'omics' Lepage et al., (2013)

Techniques description: Meta-omics	Principle	Disadvantages	Advantages
Metagenomics	Studying genomes collectively in high resolution microbiota profiling, for genes contents from uncultivated microorganism	microorganism expressed functions are not produced	• Comprehensive sequence information • Information of functional contributions of the microorganisms • No cloning of sequencing of specific genes required
Metatranscriptomics	Messenger RNA/cDNA sequencing for high resolution gene expression profiling, differential microbial gene in expression various physiological/environmental conditions	• Low stability of bacterial mRNA, • representatively unknown/multiple purification steps needed, • Lack of unique protocol	• High throughput • High sensitivity • Quantification • Characterization of known and unknown microorganism • Information of microbial interaction
Metaproteomics	Proteins/Peptides are analyzed for high resolution protein monitoring and profiling, differential microbial proteins production under various physiological/environmental conditions	• many unknown proteins in databases, • heterogeneous stability, • 'no unique protocol	• Locating and monitoring new functional genes
Metabolomics	Metabolites are analyzed for microbial host metabolic profiling	• many unknown metabolites in databases, • strict identification of compound labor intensive, • no unique protocol, • combination of host and microbial molecules	• easy to perform on every low amount of material faces/serum/urine • time efficient • impact of microorganisms in health and disease

4.6.2.1 Metagenomics

This method uses a comprehensive approach to analyze the structure (composition) and function of gut microbiota by sequencing information from the combined genomes of the microbiota (microbiome) without the need for isolating them. The advantage of this method is that it has a high throughput and capacity to recognize new functional genes. However, the limitation is that is cannot predict genes expressed or not and how different conditions can regulate it. One more disadvantage is that it cannot distinguish DNA from dead cells and DNA of live cells (Gong et al., 2012).

4.6.2.2 Metatranscriptomics

Metatranscriptomics is one of the new techniques which depend on the high-throughput sequencing RNA isolated from complex microbial populations. Metatransciptomics can be used to obtain functional information into the gut microbiota and the effect of changes in the host and diet on community-wide alterations in gene expression. Its advantages including high throughput, high sensitivity, quantization, the ability to characterize both known and unknown gene transcripts, comparison of results from different laboratories .

These techniques include RNA-Sequencing (RNA-Seq) and cDNA amplified fragment length polymorphism (cDNA-AFLP). Despite the above stated advantages of the Metatransciptomics techniques, they can be limited in some aspects. For example RNA-Seq does not distinguish de novo transcription and posttranscriptional events; because it identifies the number of RNA present and it can have bioinformatic implications. Cost is another disadvantage of this type of methods. Metatranscriptomic analysis is cDNA amplified fragment length polymorphism (cDNA-AFLP) has high levels of reproducibility, sensitivity, and specificity. Its advantages also include that it can help sequencing or studying any unknown genome or set of genomes without prior sequence knowledge, and allows the detection of lowly expressed genes (Gong et al., 2012).

4.6.2.3 Metaproteomics

It refers to as community of proteomics, function-based approach to identify microbial functions in the community. The advantages of metaproteomics are that microbial protein

expression levels and the identification of new functional genes can be monitored directly. However, the challenges of the use of this technique include uneven species distribution, purification and detection capabilities for microbial proteins that have a low abundance, and the large genetic heterogeneity of proteins (Zoetendal et al., 2008). It has also shown potential to link the composition and activities of gut microbiota with the functions of the microbiome. In addition enzymal intestinal activity and abundance can be linked to their origins of phylogenic depending on protein using this method.

4.6.2.4 Metabolomics

Metabolomics is another method used to study the function of intestinal microbiota by surveying of their metabolic profiles and host metabolic profiles with mass spectroscopy, nuclear magnetic resonance (NMR), or other analytical methods. This technique helps with simultaneously analyzing multiple small metabolites in a given sample. Metabolomics has been used to analyze the effect of antibiotics or nutrition on the intestinal microbiota. It has been also used to analyze the metabolite profiles in feces with various compounds, including short-chain fatty acids (SCFA, e.g. butyrate), organic acids (e.g. succinate), amino acids, uracil, trimethylamine, ethanol, glycerol, glucose, phenolic acids, cholate, and lipid components. It has been also used to investigate the effect of the murine intestinal microbiota on blood metabolites, viewing that the intestinal microbiota has a profound and systemic impact on host metabolism. However one limitation of metabolomics is the difficulty to analyze all the metabolites present in a sample due to the complexity of the intestinal microbiota (Gong et al., 2012).

4.7 Insertion Sequencing (INSeq)

It is a technique than merges genomewide transposon mutation and large numbers of parallel sequencing to study the genetical function of the microbiota. It “uses transposons with an identifiable DNA “bar code” to introduce mutations into tens of thousands of bacteria. These transposon-mutated strains are introduced into the guts of various kinds of animals such as germ-free (GF) and gnotobiotic animals. After a given time period for establishment of these bacteria, genomic DNA from the gut microbiota is extracted, digested by type IIs restriction enzyme Mmel, and separated by PAGE. Transposon-sized fragments are recovered from gels and sequencing adapters are appended by ligation” (Gong & Yang, 2012, p. 922). Then several cycles of PCR is conducted with adaptors that are specific to the transposon and

adaptors. Then parallel sequencing is performed that the abundance of gene mutant transposon indicates the ubandance or corresponding microbiota. Amplicons are sequenced using a massively parallel (Goodman & Gordon, 2010; Goodman et al., 2009).

4.8 Animal models

Animal models use to study the biological complexities of the interactions among host; diet, and microbiota, groups of animal models have been developed to study the dynamic, ecological diversity and functions of the gut microbiota. One of the advantages of the GF animal models is that it can provide a simplified experimental system to study specific members of the gut microbiota. However, the disadvantage of the use of germ free animals that it responses exhibited by it cannot reflect true changes in the physiology of normal animals. The gut microbiota is essential for the proper gut development (Gong & Yang, 2012).

4.9 Data analyses and bioinformatics

In order to study, analyze and understand the complexity of the intestinal microbiota. The first step for the analysis of the microbiota is to determine the evolutionary relations among the microorganisms in the intestine. DNA sequence-based (culture-independent) both alignment-based and alignment independent methods are used to determine the evolutionary relatedness.

In the alignment-based analyses, the homologous positions in the gene sequence are identified through a multiple sequence alignment against databases such as ARB and the Ribosomal Database Project II (RDPII) .The most popular approach for making alignments is the CLUSTAL online software and databases such as NCBI and multivariate statistical analysis are utilized for this process. One of the advantages of the alignment-based approach is that it is the most accurate method to provide a very detailed map of the phylogenetic relations, but it’s not suitable for analysis of large sets of data. However alternatives used for large datasets can be costly. Currently, the most used method for alignment independent analysis is principal component analysis (Gong et al., 2012).

4.10 Microbial culturomics

It is a new approach to analyze biodiversity applied to human intestinal microbiota. It was estimated that 80% of bacteria detected by molecular techniques were not culturable. Culturomics is a technique that combines over 200 types of culturing with rapid identification tools such as mass spectrometry tools (MALDI-TOF) (Gorlas et al., 2012; Lagier et al., 2012). With this technique, researchers were able to detect bacteria, 80% of which could not be detected by molecular techniques. This technique was developed by Prof. Raoult from Aix-Marseilles University in early 2010s. Using this method allowed for the detection and isolation of new viruses, one that is the largest known virus *Senegalvirus* (Gorlas et al., 2012; Lagier et al., 2012). Also it allowed for the detection of 31 new bacteria, one which has the largest genome such as *Microvirga massiliensis* with a genome of 9.35 Mb. New Archea were identified and isolated using this method, one of which the largest known Archea *Methanomassiliicoccus luminyensis*, 2.6 Mb. Compared with metagenomic techniques, culturomics techniques was able to identify 85% of 340 cultivated species in gut microbiota. Whereas researchers found that this technique can be time consuming (Gorlas et al., 2012; Lagier et al., 2012).

5 Conclusions

As it has been described above that the human intestinal microbiota is composted of large number of microorganisms, in which bacteria are the predominate one. The population of the microbial community is very high but also has wide range of diversity and complexity. These large numbers of population with wide range of diversity have significant influence on the nutritional, physiological and immunological status of human. This significant influence of the microbiota on the host cells for more study and understanding of the composition, abundance and their interaction, for better benefit.

To study the microbiota it is important to have versatile and reliable methods and techniques. Previously the study of the microbiota was depending on the conventional cultivation method, which has been found to be slow, laborious and time consuming. From the large and diversified microbiota population, there are large numbers of uncultivable and abdominal microorganisms but viable microorganisms that could not be studied by this method, because their growth requirement is not known.

These conventional methods did not allow for comprehensive studying and did not allow for studying the environmental aspect of intestinal human microbita such as the

microorganism/microorganism/host interactions. Recently, culture-independent molecular techniques have been developed. Generally these methods are rapid because of the automation nature and a very precise than the conventional one. As it has been described above, these molecular techniques are not only help to identify and characterize the microbiota but methods such as 16S r RNA sequencing techniques are also helpful for the classification and finding the phylogenetic relationship of the micrbiota. Although these molecular technologies also were helpful in epidemiological implications, but still were not helpful in the study of the comprehensive environmental study of the intestinal microbiota.

Most of the culture-independent molecular techniques to study the microbiota depend on Polymerase Chain Reaction (PCR) procedures. PCR is an extremely powerful technique used to amplify any specific piece of DNA of microbiota. One advantage of PCR is that it is highly precise, sensitive procedure and less time consuming. The DNA of interest can be amplified with the DNA from just one cell, thus, very small amounts of starting material can be used. However, it can be susceptible to much bias if not handled properly. Small contamination during the process can give biased results as well as many inhibitors can affect the PCR amplification.

Molecular techniques such as DGGE/TGGE and T-RFLP are very helpful techniques to monitor the microbiota population shift and gives rapid comparative analysis however their limitations are that they are subjected to PCR biases and their semi-quantitative identification requires clone library. FISH, and Dot-blot hybridization molecular techniques are also useful in the detection and enumeration of the microbiota population, however their disadvantage is that they required sequence information and the procedure is laborious at species level. Anyhow, qualitative PCR and diversity microarray techniques detect and estimate the relative abundance of the microbiota in specific samples. Their limitations are that they are expensive at early stages of development and the quantitative PCR is labor intensive.

Then the molecular approach was introduced and has helped researchers for in-depth study of intestinal microbiota and has been contributing in the field in several aspects. For example, DGGE is useful to assess the diversity of intestinal microbiota rapidly and can be considered a first descriptive step of an analysis process and qPCR is a qualitative and can be used for a higher profile analysis of the intestinal microbiota; FISH is a technique that can be a useful tool that can be designed to target specific species.

The need for the environmental, ecological, and phylogenic study and analysis of intestinal microbiota has prompted researchers to investigate newer platforms. These new platforms ranging from the 'omics' to culturomics have contributed in the discovery of new

virus, bacteria, and Archea among which has the largest ever discovered DNA sequences. Although, the focus of studying intestinal microbiota has been explorative and depended upon deep sequencing and although the tools, techniques and platforms of studying microbiota are an ongoing developing process the newer platforms of NGS are promising to further deepen our understanding of human intestinal microbiota.

NGS techniques both PCR-based and non-PCR-based are depending on genetic markers, require DNA-libraries, allowed for simultaneous and longer read length and high throughput (>several thousand to tens of millions of sequencing) with a comparatively lower cost. The platforms include 454 pyrosequencers, Illumine sequencers, SOLiD sequencers, Ion torrent, and SMS technologies the difference among them lies in the cost and runtime. All these qualities of NGS have put its techniques in the front run of the analysis of human intestinal microbiota. Although known for their relative accuracy due to the non use of cloning, the NGS platforms include the multification of genes which can introduce PCR related biases. Other limitations are still undesirable read length, accuracy, cost, labor, and the amplification step related challenges. Culturomics, the newest trend in the field, is a collective of 200 techniques that can put the process of study intestinal mucrobiota in a new era.

Cost and depth of an investigation are the main factors to decide which tools, techniques and platforms has to be used most of NGS platforms have high throughput and have environmental, ecological, and phylogenic implications. Each of the NGS platforms can be used for a specific objective. Generally speaking as it can be concluded from the present survey of literature regarding the study of microbiota composition that the molecular methods are better in many aspects and more accurate than the classical one, while the different recent molecular techniques and platforms each has an advantage and disadvantage which seem difficult to judge which one to be recommended. Furthermore, it is revealed that each of a specific molecular method or platform has its specific target to be reached.

13. References:

Amann R, Ludwig W. 2000. Ribosomal RNA-targeted nucleic acid probes for studies in microbial ecology. FEMS Microbiol Rev 24:555-565.

Amann, R. I., Ludwig, W., & Schleifer, K. H. (1995). Phylogenetic identification and in situ detection of individual microbial cells without cultivation. Microbiological reviews, 59(1), 143-169.

Amor, K. B., Vaughan, E. E., & de Vos, W. M. (2007). Advanced molecular tools for the identification of lactic acid bacteria. *The Journal of nutrition*, *137*(3), 741S-747S.

Applied Math (n.d.). *Pulsed-field gel electrophoresis (PFGE) typing*. Retrieved 1 March, 2014 from

Baker, G. C., Smith, J. J., & Cowan, D. A. (2003). Review and re-analysis of domain-specific 16S primers. Journal of Microbiological Methods, 55(3), 541-555.

Baker, G. C., Smith, J. J., & Cowan, D. A. (2003). Review and re-analysis of domain-specific 16S primers. Journal of Microbiological Methods, 55(3), 541-555.

Becker, S., Oelschlaeger, T. A., Wullaert, A., Pasparakis, M., Wehkamp, J., Stange, E. F., & Gersemann, M. (2013). Bacteria Regulate Intestinal Epithelial Cell Differentiation Factors Both In Vitro and In Vivo. PloS one, 8(2), e55620.

Bik, E. M. (2009). Composition and function of the human-associated microbiota. *Nutrition reviews*, *67*(s2), S164-S171.

Blaut, M., Collins, M. D., Welling, G. W., Dore, J., Van Loo, J., & De Vos, W. (2002). Molecular biological methods for studying the gut microbiota: the EU human gut flora project. British journal of nutrition, 87(S2), S203-S211.

Bombach, P., Chatzinotas, A., Neu, T. R., Kästner, M., Lueders, T., & Vogt, C. (2010). Enrichment and characterization of a sulfate- reducing toluene- degrading microbial consortium by combining in situ microcosms and stable isotope probing techniques. FEMS microbiology ecology, 71(2), 237-246.

Brown, T. A. (2006). *Genomes*. Wiley-Liss, Oxford.

Cani, P. D. (2013). Gut microbiota and obesity: lessons from the microbiome. Briefings in functional genomics.

Cano-Gómez, A., Goulden, E. F., Owens, L., & Høj, L. (2010). Vibrio owensii sp. nov., isolated from cultured crustaceans in Australia. FEMS microbiology letters, 302(2), 175-181.

Castoldi M et al(2013). Expression profiling of MicroRNAs by quantitative real-time PCR. In *PCR technology: current innovations*. Edited by Nolan T, Bustin SA. Boca Raton: CRC Press, Taylor & Francis Group

Charteris WP, Kelly PM, Morelli L, Collins JK.(1997). Selective detection, enumeration and identification of potentially probiotic *Lactobacillus* and *Bifidobacterium* species in mixed bacterial populations. *Int.J Food Microbiol;35*:1-27.

Claesson M, Wang Q, O'Sullivan O et al. (2010) Comparison of two next-generation sequencing technologies for resolving highly complex microbiota composition using tandem variable 16S rRNA gene regions. Nucleic Acids Research, 38,1–13.

Clarridge, J. E. (2004). Impact of 16S rRNA gene sequence analysis for identification of bacteria on clinical microbiology and infectious diseases. *Clinical microbiology reviews*, *17*(4), 840-862.

Collado, M. C., Isolauri, E., Laitinen, K., & Salminen, S. (2008). Distinct composition of gut microbiota during pregnancy in overweight and normal-weight women. The American journal of clinical nutrition, 88(4), 894-899.

Collins MD , Gibson GR. (1999). Probiotics, prebiotics, and synbiotics: approaches for modulating the microbial ecology of the gut. *Am J Clin Nutr 69*: l052S-l057S.

Collins, S. M., & Bercik, P. (2009). The relationship between intestinal microbiota and the central nervous system in normal gastrointestinal function and disease. *Gastroenterology, 136*(6), 2003-2014.

Corthier, G., Muller, M. C., & L'Haridon, R. (1996). Selective enumeration of Bacteroides vulgatus and B. distasonis organisms in the predominant human fecal flora by using monoclonal antibodies. *Applied and environmental microbiology*, *62*(2), 735-738.

Dethlefsen L, McFall-Ngai M, Relman DA. (2007). An ecological and evolutionary perspective on human-microbe mutualism and disease. *Nature 2007; 449*:811-818.

Egert, M., De Graaf, A. A., Maathuis, A., De Waard, P., Plugge, C. M., Smidt, H., ... & Venema, K. (2007). Identification of glucose‐fermenting bacteria present in an in vitro model of the human intestine by RNA‐stable isotope probing. FEMS microbiology ecology, 60(1), 126-135.

Favier CF, Vaughan EE, de Vos WM, Akkermans ADL. 2002. Molecular monitoring of succession of bacterial communities in human neonates. *Appl Environ Microbiol 68*:219-226.

Festin, R., Björklund, B., & Tötterman, T. H. (1987). Detection of triple antibody-binding lymphocytes in standard single laser flow cytometry using colloidal gold, fluorescein and phycoerythrin as labels. Journal of immunological methods, 101(1), 23-28.

Flusberg BA, Webster DR, Lee JH et al. (2010) Direct detection of DNA methylation during single-molecule, real-time sequencing. Nature Methods, 7, 461–465.

Fraher, M. H., O'toole, P. W. & Quigley, E. M. 2012. Techniques used to characterize the gut microbiota: a guide for the clinician. *Nature Reviews Gastroenterology and Hepatology,* 9**,** 312-322. '

Franks AH, Harmsen HJ, Raangs GC, Jansen GJ, Schut F, Welling GW.(1998) .Variations of bacterial populations in human feces measured by fluorescent in situ hybridization with group-specific 16S rRNA-targeted oligonucleotide probes. *Appl Environ Microbiol ,64*:3336-3345.

Gong, J. and Yang, C.-B. (2012). "Advances in the methods for studying gut microbiota and their relevance to the research of dietary fiber functions.", *Food Research International, 48*(2), pp. 916-929. doi : 10.1016/j.foodres.2011.

Gorlas, A., Robert, C., Gimenez, G., Drancourt, M., & Raoult, D. (2012). Complete genome sequence of Methanomassiliicoccus luminyensis, the largest genome of a human-associated Archaea species. *Journal of bacteriology, 194*(17), 4745-4745.

Griffiths AJF, Miller JH, Suzuki DT, Lewontin RC & Gelbart WM (1996). An Introduction to Genetic Analysis, 6th edn. W.H. Freeman and Co., NY

Guinane, C. M., & Cotter, P. D. (2013). Role of the gut microbiota in health and chronic gastrointestinal disease: understanding a hidden metabolic organ. *Therapeutic Advances in* Gastroenterology, 6(4), 295-308.

Hajibabaei, M., Shokralla, S., Zhou, X., Singer, G. A., & Baird, D. J. (2011). Environmental barcoding: a next-generation sequencing approach for biomonitoring applications using river benthos. PLoS One, *6*(4), e17497.

Hamady, M., & Knight, R. (2009). Microbial community profiling for human microbiome projects: Tools, techniques, and challenges. *Genome research*, *19*(7), 1141-1152.

Harris T, Buzby P, Jarosz M et al. (2007) Optical train and method for TIRF single molecule detection and analysis. US patent application 20070070349.

Hopkins MJ, Sharp R, Macfarlane GT. 2001. Age and disease related changes in intestinal bacterial populations assessed by cell culture, 16S rRNA abundance, and community cellular fatty acid profiles. Gut 48:198-205.

Horn S (2012) Target enrichment via DNA hybridization capture. Methods in Molecular Biology, 840, 177–188.

Jansen, G. J., Wildeboer-Veloo, A. C. M., Tonk, R. H. J., Franks, A. H., & Welling, G. W. (1999). Development and validation of an automated, microscopy-based method for enumeration of groups of intestinal bacteria. Journal of microbiological methods, 37(3), 215-221.

Karlin S, Weinstock GM, Brendel V. 1995. Bacterial classifications derived from RecA protein sequence comparisons. J Bacteriol 177:6881-6893.

Kircher, M., & Kelso, J. (2010). High- throughput DNA sequencing–concepts and limitations. Bioessays, 32(6), 524-536.

Kleessen, Eugenia Bezirtzoglou, Jaana Mättö, B. (2000). Culture-based knowledge on biodiversity, development and stability of human gastrointestinal microflora. *Microbial ecology in health and disease*, *12*(2), 53-63.

Korlach J, Bjornson KB, Chaudhuri BP et al. (2010) Real-time DNA sequencing from single polymerase molecules. Methods in Enzymology, 472, 431–455.

Kovatcheva-Datchary, P., Egert, M., Maathuis, A., Rajilić-Stojanović, M., De Graaf, A. A., Smidt, H., ... & Venema, K. (2009). Linking phylogenetic identities of bacteria to starch fermentation in an in vitro model of the large intestine by RNA-based stable isotope probing. Environmental microbiology, 11(4), 914-926.

Kullen MJ, Brady LJ, O'Sullivan DJ. 1997. Evaluation of using a short region of the recA gene for rapid and sensitive speciation of dominant bifidobacteria in the human intestine. FEMS Microbiol Letts 1541377-383.

Kumar, N. S., & Gurusubramanian, G. (2011). Random amplified polymorphic DNA (RAPD) markers and its applications. Science vision, 116-124.

Kuokka, A. M., Saarela, M., & Mattila-Sandholm, T. (2002). Food, GI-tract Functionality and Human Health Cluster. *Microbial Ecology in Health and Disease*, *14*, 65-74.

Lagier, J. C., Armougom, F., Million, M., Hugon, P., Pagnier, I., Robert, C., ... & Raoult, D. (2012). Microbial culturomics: paradigm shift in the human gut microbiome study. *Clinical Microbiology and Infection*, *18*(12), 1185-1193.

Lagier, J. C., Million, M., Hugon, P., Armougom, F., & Raoult, D. (2012). Human gut microbiota: repertoire and variations. *Frontiers in Cellular & Infection Microbiology*, *2*.

Leblond-Bourget N, Philippe H, Mangin I, Decaris B. 1996. 16S rRNA and 16S to 23S internal transcribed spacer sequence analyses reveal inter- and intraspecific Bifdobacterium phylogeny. lnt J Syst Bacteriol 56:102-11 1.

Lee E-J, Pei L, Srivastava G et al. (2011) Targeted bisulfate sequencing by solution hybrid selection and massively parallel sequencing. Nucleic Acids Research, 39, e127.

Lepage, P., Leclerc, M. C., Joossens, M., Mondot, S., Blottière, H. M., Raes, J., ... & Doré, J. (2013). A metagenomic insight into our gut's microbiome. *Gut*, *62*(1), 146-158.

Levene, M. J., Korlach, J., Turner, S. W., Foquet, M., Craighead, H. G., & Webb, W. W. (2003). Zero-mode waveguides for single-molecule analysis at high concentrations. Science, 299(5607), 682-686.

Lim YW, Kim BK, Kim C et al. (2010) Assessment of soil fungal communities using pyrosequencing. Journal of Microbiology, 48, 284–289.

Lin C, Raskin L, Stahl DA. (1997). Microbial community structure in gastrointestinal tracts of domestic animals: comparative analyses using rRNA-targeted oligonucleotide probes FEMS *Microbiol Ecol;22*:281-294.

Loman, N. J., Constantinidou, C., Chan, J. Z., Halachev, M., Sergeant, M., Penn, C. W., ... & Pallen, M. J. (2012). High-throughput bacterial genome sequencing: an embarrassment of choice, a world of opportunity. *Nature Reviews Microbiology*.

Maidak BL, Cole JR, Lilbum TG, Parker Jr CT, Saxman PR, Farris RJ, Garrity GM, Olsen GJ, Schmidt TM, Tiedje JM. 2001] .The RDB-II (Ribosomal Database Project). Nucleic Acids Res 29:173-174.

Mardis ER (2008a) Next-generation DNA sequencing methods. Annual Review of Genomics and Human Genetics, 9, 387–402.

Mardis ER (2008b) The impact of next-generation sequencing technology on genetics. Trends in Genetics, 24, 133–141

Margulies M, Egholm M, Altman W et al. (2005) Genome sequencing in microfabricated high-density picolitre reactors.Nature, 437, 376–380.

Matsumoto, M., Sakamoto, M., Hayashi, H., & Benno, Y. (2005). Novel phylogenetic assignment database for terminal-restriction fragment length polymorphism analysis of human colonic microbiota. Journal of microbiological methods, 61(3), 305-319.

Maukonen, J. (2012). Characterization of the human predominant fecal microbiota-With special focus on the Clostridial clusters IV and XIVa.

McCartney, A. L. (2002). Application of molecular biological methods for studying probiotics and the gut flora. *British Journal of Nutrition*, *88*, S29-S37.

McCartney, A.L., W. Wenzhi, and G.W. Tannock. 1996. Molecular analysis of the composition of the bifidobacterial and Lactobacillus microflora of humans. Appl. Environ. Microbiol. 62: 4608-4613.

Meng HM, *et a.l.*(2012). Efficient fluorescence turn-on probe for zirconium via a target-triggered DNA molecular beacon strategy.*Anal Chem* , 84(5):2124-2128.

Metzker, M. L. (2010). Sequencing technologies—the next generation. *Nature Reviews Genetics*, *11*(1), 31-46.

Miller DM, Dudley EG, Roberts RF.(2012). Technical note: development of a quantitative PCR method for monitoring strain dynamics during yogurt manufacture.*J Dairy Sci* , 95(9)**:**4868-4872.

Miller W, Hayes V, Ratan A et al. (2011) Genetic diversity and population structure of the endangered marsupial Sarcophilus harrisii (Tasmanian Devil). Proceedings of the National Academy of Sciences, USA, 108, 12348–12353.

Moore LVH, Bourne D, Moore WEC.(1994). Comparative distribution and taxonomic value of cellular fatty acids in thirty-three genera of anaerobic gram-negative bacilli. Int J Syst Bact 44:338-347.

Namsolleck, P., Thiel, R., Lawson, P., Holmstrøm, K., Rajilic, M., Vaughan, E. E., ... & Blaut, M. (2004). Molecular methods for the analysis of gut microbiota. Microbial ecology in health and disease, 16(2-3), 71-85.

Naseribafrouei, A. (2013). Correlation Between Gut Microbiota and Depression.

Nikkilä, J., & de Vos, W. M. (2010). Advanced approaches to characterize the human intestinal microbiota by computational meta-analysis. Journal of clinical gastroenterology, 44, S2-S5.

Nocker, A., & Camper, A. K. (2009). Novel approaches toward preferential detection of viable cells using nucleic acid amplification techniques. FEMS microbiology letters, 291(2), 137-142.

O'Sullivan, D. J. (2000). Methods for analysis of the intestinal microflora. *Current issues in intestinal microbiology*, *1*(2), 39-50.

Paliy, O., & Agans, R. (2012). Application of phylogenetic microarrays to interrogation of human microbiota. *FEMS microbiology ecology*, *79*(1), 2-11.

Palmer, C., Bik, E. M., DiGiulio, D. B., Relman, D. A., & Brown, P. 0. (2007). Development of the Human Infant Intestinal Microbiota. Plos Biology, 5(7), 1 5 5 6- 1 5 73.

Phianphak, W., Rengpipat, S., Rukpratanporn, S., Longyant, S., Chaivisuthangkura, P., Sithigorngul, W., & Sithigorngul, P. (2005). Production of monoclonal antibodies for detection of Vibrio harveyi. *Diseases of aquatic organisms*, *63*(2), 161-168.

Polz MF, Cavanaugh CM (1998) Bias in template-to-product ratios in multitemplate PCR. Applied and Environmental Microbiology, 64, 3724–3730.

Pond, S. L. K., Posada, D., Gravenor, M. B., Woelk, C. H., & Frost, S. D. (2006). GARD: a genetic algorithm for recombination detection. *Bioinformatics*, *22*(24), 3096-3098.

Postollec, F., Falentin, H., Pavan, S., Combrisson, J., & Sohier, D. (2011). Recent advances in quantitative PCR (qPCR) applications in food microbiology. Food microbiology, 28(5), 848-861.

Qin, J., Li, R., Raes, J., Arumugam, M., Burgdorf, K. S., Manichanh, C., ... & Weissenbach, J. (2010). A human gut microbial gene catalogue established by metagenomic sequencing. *Nature*, *464*(7285), 59-65.

Quince C, Lanze´n A, Curtis TP et al. (2009) Accurate determination of microbial diversity from 454 pyrosequencing data. Nature Methods, 6, 639–641.

Rakoff-Nahoum, S., Paglino, J., Eslami-Varzaneh, F., Edberg, S., & Medzhitov, R. (2004). Recognition of commensal microflora by toll-like receptors is required for intestinal homeostasis. Cell, 118(2), 229-241.

Raskin L, Stromley JM, Rittman BE, Stahl DA. (1994). Group-specific 16S rRNA hybridization probes to describe natural communities of methanogens. *Appl Environ Microbiol* 6021232-1240

Rastogi, G., & Sani, R. K. (2011). Molecular techniques to assess microbial community structure, function, and dynamics in the environment. In Microbes and Microbial Technology (pp. 29-57). Springer New York.

Rautio, M. (2002). Indigenous Intestinal Microbiota and Disease (Doctoral dissertation, Faculty of Sciences, Department of Biosciences, University of Helsinki).

Razgour, O., Clare, E. L., Zeale, M. R., Hanmer, J., Schnell, I. B., Rasmussen, M., ... & Jones, G. (2011). High-throughput sequencing offers insight into mechanisms of resource partitioning in cryptic bat species. Ecology and evolution, 1(4), 556-570.

Roca Al, Cox MM. 1997. RecA protein: structure, function, and role in recombinational DNA repair. *Prog Nucleic Acid Res M01 Biol 56*:129-223.

Rogers, G. B., & Bruce, K. D. (2010). Next-Generation Sequencing in the Analysis of Human Microbiota. *Molecular Diagnosis & Therapy*, *14*(6), 343-350.

Rothberg J, Hinz W, Rearick T et al. (2011) An integrated semiconductor device enabling non-optical genome sequencing. Nature, 475, 348–352.

Rudi, K., & Isaksen, M. (2012). Gut Microbiota in Disease Diagnostics. *New Advances in the Basic and Clinical Gastroenterology*.

Rudi, K., Skulberg, O. M., Larsen, F., & Jakobsen, K. S. (1997). Strain characterization and classification of oxyphotobacteria in clone cultures on the basis of 16S rRNA sequences from the variable regions V6, V7, and V8. Applied and environmental microbiology, 63(7), 2593-2599.

Salipante, S. J., Sengupta, D. J., Rosenthal, C., Costa, G., Spangler, J., Sims, E. H., ... & Hoffman, N. G. (2013). Rapid 16S rRNA Next-Generation Sequencing of Polymicrobial Clinical Samples for Diagnosis of Complex Bacterial Infections. *PloS one, 8*(5), e65226.

Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors" *Proceedings of the National Academy of Sciences* 74:5463-5467.

Santos, S. R., & Ochman, H. (2004). Identification and phylogenetic sorting of bacterial lineages with universally conserved genes and proteins. Environmental Microbiology, 6(7), 754-759.

Sanz M, Lau L, Herrera D, Morillo JM, Silva A.(2004). Methods of detection of *Actinobacillus* actinomycetemcomitans, Porphyromonas gingivalis and Tannerella forsythensis in periodontal microbiology, with special emphasis on advanced molecular techniques: a review. J Clin Periodontol.;31:1034-1047.

Satake S, Clark N, Rimland D, Nolte FS, T enover FC. 1997. Detection on vancomycin- resistant enterococci in fecal samples by PCR. *J Clin Microbiol 35*:2325-2330.

Schuster S. C. (2008) Next-generation sequencing transforms today‘s biology. Nature Methods, 5(1) 16–18.

Searchhigh, L. (2012). Characterization of the human predominant fecal microbiota.

Sekirov, I., Russell, S. L., Antunes, L. C. M., & Finlay, B. B. (2010). Gut microbiota in health and disease. *Physiological Reviews*, *90*(3), 859-904.

Sghir A, Gramet G, Suau A, Rochet V, Pochart P, Dore J.(2000). Quantification of bacterial groups within human fecal flora by oligonucleotide probe hybridization. *Appl Environ Microbiol* 6622263-2266.

Shendure J, Ji H (2008) Next-generation DNA sequencing. Nature Biotechnology, 26, 1135–1145.

Shokralla, S., Spall, J. L., Gibson, J. F., & Hajibabaei, M. (2012). Next-generation sequencing technologies for environmental DNA research. *Molecular Ecology*, *21*(8), 1794-1805.

Sohier D et al. (2012). Polyphasic approach for quantitative analysis of obligately heterofermentative Lactobacillus species in cheese.*Food Microbiol* , 31(2):271-277.

Sun Y, Wolcott RD, Dowd SE (2011) Tag-encoded FLX amplicon pyrosequencing for the elucidation of microbial and functional gene diversity in any environment. Methods in Molecular Biology, 733, 129–141.

Suzuki MT, Giovannoni SJ. (1996). Bias caused by template annealing in the amplification of mixtures of 16S rRNA genes by PCR. *Appl Environ Microbiol 62*:625-630.

Tannock GW. (l999). Analysis of the intestin al microflora : a renaissance. *Antonie Leewenhoek 76*:265-278.

Ursell, L. K., Clemente, J. C., Rideout, J. R., Gevers, D., Caporaso, J. G., & Knight, R. (2012). The interpersonal and intrapersonal diversity of human-associated microbiota in key body sites. *Journal of Allergy and Clinical Immunology, 129*(5), 1204-1208.

Ursell, L. K., Metcalf, J. L., Parfrey, L. W., & Knight, R. (2012). Defining the human microbiome. *Nutrition reviews*, *70*(s1), S38-S44.

Vaughan EE, Schut F, Heiling HGHJ, Zoetendaal EG, de Vos WM, Akkermans ADL.(2000). A molecular view ofthe intestinal ecosystem. Curr Issues Intest Microbiol 1 : 1-12.

Wang, Z., Gerstein, M., & Snyder, M. (2009). RNA-Seq: a revolutionary tool for transcriptomics. *Nature Reviews Genetics*, *10*(1), 57-63.

Wilson KH, Blitchington RB. (1996). Human colonic biota studied by ribosomal DNA sequence analysis. *Appl Environ Microbiol;7*:2273-2278.

Woese CR.1987.Bacterial evolution.Microbial Rev 51:221-271.

Woo, P. C. Y., Lau, S. K. P., Teng, J. L. L., Tse, H., & Yuen, K. Y. (2008). Then and now: use of 16S rDNA gene sequencing for bacterial identification and discovery of novel bacteria in clinical microbiology laboratories. *Clinical Microbiology and Infection*, *14*(10), 908-934.

Wu, L., Thompson, D. K., Li, G., Hurt, R. A., Tiedje, J. M., & Zhou, J. (2001). Development and evaluation of functional gene arrays for detection of selected genes in the environment. Applied and environmental microbiology, 67(12), 5780-5790.

Xu M, Chen X, Qiu M et al. (2012) Bar-coded pyrosequencing reveals the responses of PDBE-degrading microbial com-munities to electron donor amendments. PLoS ONE, 7,e30439.

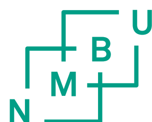
Zhang, W., Chen, J., Yang, Y., Tang, Y., Shang, J., & Shen, B. (2011). A practical comparison of de novo genome assembly software tools for next-generation sequencing technologies. PloS one, 6(3), e17915.

Zhou XG, Ren LF, Li YT et al. (2010) The next-generation sequencing technology: a technology review and future perspective. Science China. Life Sciences, 53, 44–57.

Zoetendal EG, Akkermans ADL, de Vos WM.(1998). Temperature gradient gel electrophoresis analysis of 16S rRNA from human fecal samples reveals stable and host- specific communities of active bacteria. *Appl Environ Microbiol 64*: 3854-3859.

Zoetendal, E. G., Collier, C. T., Koike, S., Mackie, R. I., & Gaskins, H. R. (2004). Molecular ecological analysis of the gastrointestinal microbiota: a review. *The Journal of nutrition*, *134*(2), 465-472.

Zoetendal, E. G., Rajilic-Stojanovic, M., & Vos, W. M. (2008). High-throughput diversity and functionality analysis of the gastrointestinal tract microbiota. *Gut, 57*,1605-1615.



Norges miljø- og
biovitenskapelige
universitet

Postboks 5003
NO-1432 Ås
67 23 00 00
www.nmbu.no