We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,300

Open access books available

172,000

International authors and editors

190M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Genomic Techniques Used to Investigate the Human Gut Microbiota

Akhlesh P. Singh

Abstract

The human gut is the complex microbial ecosystem comprises more than 100 trillion microbes also known as microbiota. The gut microbiota does not only include about 400–500 types of bacterial strains, but it also contains archaea, bacteriophage, fungi, and protozoa species. In order to complete the characterization of the gut microbial community, we need the help of many culture-dependent and culture-independent genomic technologies. Recently, next-generation sequencing (NGS), mediated metagenomics that rely on 16S rRNA gene amplification, and whole-genome sequencing (WGS) have provided us deep knowledge related to important interactions such as host-microbiota and microbe-microbe interactions under various perturbation inside the gut. But, we still lack complete knowledge related to unique gene products encoded by gut meta-genome. Hence, it required the application of high-throughput “omics-based” methods to support metagenomics. Currently, a combination of high-throughput culturing and microfluidics assays is providing a new method to characterize non-amenable bacterial strains from the gut environment. The recent additions of artificial intelligence and deep learning to the area of microbiome studies have enhanced the capability of identification of thousand microbes simultaneously. Given above, it is necessary to apply new genome editing tools that can be used to design the personalized microflora which can be used to cure lifestyle-related diseases.

Keywords: culturomics, gut microbiota, human microbiome, metagenomics, metaproteomics, metabolomics, microfluidics, “multi-omics”, personalized diet

1. Introduction

In the beginning of the twenty-first century, the human genome was sequenced. The main aim of this gigantic scientific effort was to identify all genes present in the human genome, also considered as the “blueprint of human life.” Since then, most of the efforts are focused on the identification of all genes and annotate their functions which are responsible for genetic variation prevailed in human physiology and its association with diseases [1]. Currently, many experiments have proved that the gut microbes are more responsible than host genetics in the development of life style-related diseases. Hence, it becomes essential to investigate the crucial roles played by gut microbes in health and diseases. The human gut is a complex microbial ecosystem which is comprised of approximately 100 trillion microbes collectively known as “gut microbiota” [1]. It does not only include about 400–500 types of bacterial species but

also contains archaea, bacteriophage, fungi, and protozoa species [2]. According to a rough estimation, the human gut microbiome contains almost 3.3 million genes which are 150 times more than total human genes present in the human genome. Currently, gut microflora is also considered as “gold mines” because of its commercial value in the area of biopharmaceuticals and bioactive products. In order to complete the characterization of the gut microbial community and its mysteries, we need the help of many traditional and modern genomic technologies developed in due course of time. However, the study of the human microbiome is relatively a newly emerging area in the area of human biology, thus called the “forgotten organ” in the human body.

The study of the human microbiome was started with the help of reductionist approaches such as identification and characterization of a single bacterial strain by using culture media and microscopes. Initially, only culturable bacteria could only be identified and phylogenetically classified. It is well known that more than 40% of gut microbes cannot grow outside the natural environment. Hence, both culture-dependent and culture-independent analytical methods are applied that have improved our knowledge related to human gut microbiota. Recently, next-generation sequencing (NGS) has revolutionized all areas of biological sciences including the human gut microbiome. This also supports the most traditional metagenomic technique based on 16sRNA gene amplification via polymerase chain reaction (PCR) and whole-genome sequencing (WGS) also. However, both culture-dependent and culture-independent techniques have provided the snapshot of the gut microbial community, but they are still hazy in respect of host-microbiota and microbe-microbe interactions that make stable conditions of gut microbial communities under the influence of various perturbations such as environmental factors, diets, and drugs. In the last 20 years, it becomes apparent that gut microbes add in the metabolism and contribute to strengthening the host's immune system. The human gut microbiota constitutes a metagenome that encodes an intricate network of genes, proteins, and metabolites. In order to functionally characterize human microbiome, it requires applications of many supplementary high-throughput “omics-based” methods, e.g., metaproteomics, metatranscriptomics, and metabolomics.

Recently, several labs the world over have adopted new emerging technologies to support metagenomics consequently; it amasses the terabits data in various genomic databases. To retrieve meaningful information from a large amount of multi-omics data, the application of a high level of computational and bioinformatics knowledge is required. In view of the recent explosion of data in every field, machine learning and deep learning come forward for the rescue of scientists. Therefore, different algorithms have been created, tested, and applied to huge microbiome data to identify the results of numerous microbial strains. But the next aim of all plethora of technologies is to unravel the significant contribution of gut microbiota to human biochemistry and physiology, and ultimately, this knowledge can be translated to improve human health and reduce lifestyle-related pandemic prevailed worldwide. In view of the above facts, the current chapter describes a set of analytical methods that are used to dig deep into the human gut microbial community. These methods are exploited in phylogenetic classification and functional characterization of gut microbiota.

2. A brief history of the human microbiome study

The field of the human microbiome is closely associated with microbiology; hence, its study was started in the seventeenth century. Antonie van Leeuwenhoek, who is also considered the father of microbiology, discovered oral microbes by using a simple microscope and called them “animalcules” in 1676. In the 1800s, Robert Koch developed the investigation technique for anthrax. The pioneering

work of Pasteur, Koch, Escherich, and Kendall founded a strong base of microbiome research; hence, they are able to identify and count a large number of bacterial strains. In 1907, Metchnikoff proposed that lactic bacteria can ward off against harmful or putrefying bacteria from the gut [2]. Joshua Lederberg for the first time used the term “microbiome” for gut microbial community, and its relationship with the host. The microbial community can be defined as “the set of organisms (in this case, microorganisms) coexisting in the same space and time” [3].

In the beginning, only culturable bacterial species were studied, but there are a large number of microbes that are not grown inside the lab environment. That was revealed when the number of microorganisms observed by the microscope did not match with a number of microorganisms that grow on the media plate [4]. In 1970, Carl Woese suggested that ribosomal RNA genes can be used as molecular markers for bacterial classification [5]. Thus, scientists have developed the culture-independent technique based on amplification of 16S rRNA gene by PCR method and its sequencing by Sanger method. These strategies are used to classify gut microorganisms phylogenetically and then annotate their functions in a particular natural microbial ecosystem [6]. It has revolutionized the field of “microbiome research.” Other culture-independent techniques, which significantly influence the taxonomic research, were the PCR, rRNA gene cloning and sequencing, fluorescence in situ hybridization (FISH), denaturing gradient gel electrophoresis (DGGE) and temperature gradient gel electrophoresis (TGGE), restriction fragment length polymorphism (RFLP), and terminal restriction fragment length polymorphism (T-RFLP). But these techniques could not reveal the metabolic and ecological functions of microorganisms. In order to ascertain the function of individual bacterial strain in the gut ecosystem, germ-free mouse models were also developed.

But due to cumbersome and time-consuming methods of traditional metagenomic techniques, new methods based on NGS have taken over the central stage to investigate the microbial communities [7]. Currently, sequencing-based techniques are used to classify numerous uncultivable microbes. Most recently, mass spectroscopy (MS) and one of its variants, i.e., matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF)-based “omics”-based high-throughput methods, have been applied to functional characterization of microbial communities [8]. These sophisticated technologies have amassed a huge amount of genomic data that needs to be annotated by computer-based systems biology approaches. The systems biology will provide a holistic picture of the microbial community inside the human gut. Currently, seven major groups such as *Firmicutes*, *Bacteroidetes*, *Proteobacteria*, *Fusobacteria*, *Verrucomicrobia*, *Cyanobacteria*, and *Actinobacteria* which constitute a major chunk of gut microbes have been recognized, but out of these two phyla, namely, *Firmicutes* and *Bacteroidetes*, include most of the gut bacteria species [9].

3. Methodology for human gut microbiome studies

Initially, culture and biochemical typing were the standard methods to identify any new bacterial species. To know more about the human gut microbes’ diversity, its compositions, and relationships with various diseases, thus, many other techniques are also developed. The evolution of various methods applied to investigate the human gut microbiota is described above. Recently, significant advancements have been made in the area of sequencing-based genome technologies including metatranscriptomics, proteomics, and metagenomics, which are further supported by culturomics and computational biology for studies of human gut microbiome research. These techniques are rapid and, hence, provided a huge wealth of genomic

data related to uncultured microorganisms. This helped us in the identification of new microbe species inside the gastrointestinal tract. But there are many important issues associated with the accurate and proper investigation of a gut ecosystem like sample preparation, storage, and handling from the human as well as animal subjects. In the current chapter, total techniques under three major headings (1) culture-dependent methods, (2) culture-independent genomic technologies, and (3) latest techniques are described (**Figure 1**).

3.1 Phylogenetic analysis of microbial community

3.1.1 Culture-dependent methods

In the last century, most of the microbiome studies were based on culture-based methods. Almost for the last 300 years, this approach mainly relied on main identification features like colony features, bacterial growth, and selection of some biochemical typing and microscopic investigation of culturable microbes in lab condition. In the 1980s, large numbers of gram-negative bacterial species were identified from the fecal samples [10]. Later on, many species have been identified and phylogenetically classified by using fermentation profiling or *in vitro* requirements of bacterial species. It has contributed enormously to the identification of microbial agents and given birth to a new branch, i.e., microbial ecology [11]. The culture-based method is still considered as the gold standard protocol for the identification of new species and provided a deep understanding related to the microbial world. They are a cheap and most credible method of bacterial identification. But they could not be proven completely effective against anaerobic and not amenable bacterial species. It is already given that more than 30% of bacterial species cannot be grown outside from their habitat. Moreover, gut microbiota not only includes the bacteria but also consists of bacteriophages, archaea, fungal species, and single-celled eukaryotes. Hence, we need more wide investigative approaches to cover all the microbial agents involved and contribute to the stable form of gut microbiota.

Figure 1. Summary of techniques used to phylogenetic classification and functional characterization of the human gut microbiome.

3.1.2 Culturomics

The significance of culture-dependent methods cannot be undermined for the identification of microbes from the gut microbial community. Therefore, microbiologists have rediscovered and focused once again to revive culture-based methods by adding many sophisticated instrumentations and suitable growth media. This has allowed growing most of the unculturable bacteria that were earlier thought to be impossible in a lab environment. Hence, it will allow to know more about the functional aspects of gut microbiome that include its composition, microbial gene expression, metabolic pathways and host-bacteria relationships [12]. Actually, diverse types of favorable growing and incubation conditions are required to grow unculturable microbes that are provided by the new culturomics procedures. Currently, more than 50% of bacterial species that were earlier identified by classical 16S rRNA metagenomics could be re-identified with the help of culturomics. Simultaneously, it will also allow isolating hundreds of new bacterial species in the gut microbial ecosystem in the near future [13].

The culturomics is a multistep protocol that includes sample preparations and their diversification under different growth conditions that promote the growth of fastidious bacteria but, simultaneously, also cease the growth of few microbes. The targeted samples are subjected to further MALDI-TOF mass spectroscopy-based investigations such as a comparison of newly obtained protein spectra in recent protein databases. If the applied method could fail to establish the identification of bacteria, then the sample is processed for NGS-based 16sRNA metagenomic methods. Based on the 16sRNA gene sequencing, various toxicogenomics principles are applied to classified new species in phyla or family. Culturomics is quite an effective growth strategy particularly in microbes that are involved in mechanistic networks or intricate host-microbiome interactions. More recently, many culture techniques, for example, gel microdroplets, microculture, and microbial chips, provide very diverse growth conditions; hence, a large number of unknown microbes are able to grow [14]. Although new methods are quite helpful in the identification of new microbial species, e.g., from gut microbial ecosystem, these are also used to study human vaginal and urinary microbiota. Currently, almost 2671 new species have been identified by using culturomics ranging from commensals to pathogens, for example, 31 new bacterial species that belong to *Synergistetes* or *Deinococcus-Thermus* phyla. But, there are certain demerits like nonavailability of suitable culture media and growth conditions that allow the growth of uncultured bacteria in an artificial environment [15]. Moreover, certain bacteria grow in a highly intrigue environment inside the human gut because several microbes use common metabolites as a food and live in symbiotic and mutual interrelationship inside the gut environment.

3.1.3 Microfluidics assays

Microfluidics systems or cell on-chip offers a specific microenvironment for biochemical reactions. Microfluidics comprises numerous microchannels enshrined on the glass or polymer surface such as polydimethylsiloxane [16]. These channels are linked to each other that are based on principles of mixing, pumping, sorting, or offering biochemical environment; hence it can produce a suitable environment for microbial reactions. Recently, great advances have been made in this area; consequently, high-throughput screening, multiplexing, and automation of biochemical reactions could be achieved [17]. Microfluidics technique is also applied in the studies of gut microbiota; hence, some scientists called it gut-on-chip. With microchips, many uncultured microbes are identified because it provides specific growth environment and nutrition required for these bacterial growths, for example,

microfluidics-based model (human-microbial cross talk (HuMiX)). The HuMiX provide gastrointestinal-like environment for the co-growth of human epithelial cell and obligate anaerobe *Bacteroides caccae* cells [18]. Recently developed iChip containing multiple microchambers which are further divided into hundreds of miniature multiple cells has been used to grow bacteria. This technique mainly acts by providing a selective supply of nutrients to an inoculated single bacterial cell on-chip. Another chip-based method l-tip also acts on the same principles as iChip, but it allows bacterial cell multiplication in a gel and supplies required nutrients which are essential for growth [19]. Microfluidics is the combination of gel-based methods and sophisticated instruments, for example, first we grow a single bacterial cell, then amplify its genome, and, finally, sequence its genome that helps in identifying new species [20]. Recently, TM7, bacterium, and *Sulcia muelleri* could be identified which produced very unique metabolites. By using the same method, 34 various bacterial strains are identified and phylogenetically classified.

3.2 Culture-independent methods

3.2.1 Sample collection and standardization methods

The sample preparation is a very crucial and important step of any microbial or biochemical analysis that determines the accuracy and efficacy of any simple or sophisticated analytical technique. In the human microbiome studies, there are two major types of samples, namely, stool and mucosal biopsy. However, the mucosal biopsy sample must be preferred, but their availability and handling are not easy. Ideally, stool samples must be used in conjunction with the mucosal samples [21, 22]. Several proofs of investigation have shown that there are great ambiguities prevailed between the presence of microbiota in mucosal and stool samples. Sample collection and their storage conditions also influence the final results in terms of the genetic composition of gut microbes. It has been noticed that the populations of the two most abundant gut microbial species such as *Firmicutes* to *Bacteroidetes* are affected with storage temperature in the fecal sample [23]. The sample processing methods are also held responsible for the variations in results. Hence, different consortiums associated with large-scale investigation of the gut microbiome have suggested that we must adopt the standard and calibrated protocols for sample processing [24]. Therefore, many kits are developed, for example, Qiagen QIAamp DNA Stool Mini Kit (QIAG) has significantly improved the DNA extraction and reproducibility of results from fecal samples. Moreover, researchers have also recommended other methods, namely, phenol/chloroform (PHEC), chaotropic (CHAO), and THSTI. Their comparative efficacies and performance were analyzed in terms of the final yield of DNA [26]. Currently, one more DNA/RNA Extraction Kit (TS), i.e., TianLong Stool, is also used, which mainly acts on mechanical shearing and bead beating method. It also provided good reproducible results. However, comparative studies reflect that the TS kit offers a higher quantity of nucleic acids than the other extraction kits. Conclusively, we can say that, standard protocols that are available in the form of kits that save our time and efforts of researchers [25].

3.2.2 Metagenomic analysis of microbial community

In order to overcome the drawbacks of traditional culture-based protocols, microbiologists have developed several advanced culture-independent methods to know the composition of gut microbiota. In this series, metagenomics was the first technique by which 80% of uncultured microbes are phylogenetically identified.

This culture-independent technique for microbial growth has revolutionized the area of human microflora investigations in the last two decades.

The classical techniques of metagenomics rely on the 16S ribosomal RNA (16S rRNA) gene. The 70S ribosome is the major component of prokaryotic cells and involved in protein synthesis which is highly conserved processes in all bacterial cells. The major function of 16S rRNA is the regulation of protein synthesis. During protein synthesis process, 3' end of 16S RNA combines with the ribosomal proteins S1 and S21 involved in activation and initiation of protein synthesis. Although 16S rRNA is highly conserved among microbial species, it also contains few hypervariable regions that offer phylogenetic linkage; hence, it is also proven to be helpful in the classification of enormous microbial diversities that prevailed on earth [26]. With the development of DNA sequencing methods, 16S rRNA gene amplicons are isolated and sequenced; hence, it now becomes the most successful and prevalent culture-independent method for taxonomic classification of microorganisms. After the availability of PCR-based cloning and 16S rRNAs, gene sequencing has revolutionized the area of taxonomic classification of uncultured bacterial strains in the last two decades [27].

The metagenomic protocols include the extraction of nucleic acid from the sample followed by PCR amplification of species-specific 1500-bp-long whole 16S ribosomal RNA genes [28]. It also contains highly hypervariable regions (the V4–V5 region out of nine short hypervariable regions from V1 to V9). PCR-based amplification is carried out by using universal and specific primers, and after that, physical separation of DNA fragments are carried out on electrophoresis gels [29].

Initially, 16S ribosomal RNA gene amplification was based on cloning in a suitable host, e.g., *Escherichia coli*, and then sequencing by Sanger sequencing method. After availability of PCR based cloning of 16S ribosomal RNA gene and then, sequencing of clones (amplicons) by using any DNA sequencing method. These methods have tremendously enhanced phylogenetically the identification of the gut microbiota [30]. At that time, the pace and cost of sequencing were the great impediments that could be overcome by the advent of NGS. It is now well known that PCR-mediated protocols used for characterization of microbial diversity have certain demerits. These are attributed to PCR-based amplification of 16S rRNA gene, which is a multi-step process that introduced several ambiguities into the final results, and it became more error prone due to the PCR-based sequencing method, e.g., pyrosequencing [33]. Generally gene-specific amplifications are primer based which must be appropriate for all major taxa. Furthermore, the amplified DNA fragments can harbor mutations because of the nonspecific binding of PCR primers to template DNA strands [31].

Recently, next-generation DNA sequencing has made metagenomic and whole-genome sequencing metagenomic methods more rapid and highly sophisticated. The latest sequencing methods such as 454 pyrosequencing, Illumina, SOLiD, Ion Torrent, and single-molecule real-time (SMRT) circular consensus sequencing equipment from Pacific Biosciences [32] and Oxford Nanopore have provided more pace and deep analytic power to the analyzed gut microbiome [33]. More recently, the application of Oxford Nanopore in gut microbe analysis can overcome the abovementioned PCR-based limitations such as PCR temperature, cloning, and long and deep sequencing by MinION™ nanopore sequencing technologies.

3.2.3 Real-time PCR

It is well known that PCR is a nonquantitative technique, but its variant, real-time PCR also known as quantitative PCR (qPCR), is used for microbiome analysis particularly for phylogenetic analysis. It can be used quantitatively and semiquantitatively depending upon the applications; qPCR can quantify the amount of DNA

in the stool or gut mucosa samples. In this technique, fluorescent probes or dye molecules are used that intercalate between the double strand of DNA molecules or 16 s RNA amplicons. These probes send a strong signal, and its intensity is directly proportional to the amount of DNA sample present. Sometimes sequence-specific oligonucleotide probes are linked with molecular markers or complementary DNA sequence [34]. The primers designing is a crucial step in the RT-PCR technique; therefore, primers must be specific for all bacterial phyla or taxa or species present in a sample [35]. Real-time PCR has been used to investigate the state of the ecological environment in normal and obese persons [36]. Quantitative PCR technique is also used solely or in combination with other gel and non-gel-based techniques. This combination of protocols is used to understand the functional microbial diversity of gut microbiota in the patient of age and effect of antibiotics on gut microbes [37], for example, DGGE and qPCR.

Real-time PCR-based methods are suitable for the prediction of accurate phylogenetic analysis. The appropriate primers provide great help to know the composition of a microbial community and microbial load. The protocol is simple to complex, and all chemicals and consumables are easily available in laboratories. But, this is also suffering due to PCR biases, which percolate at each step of the protocol. Quantitative PCR cannot be used to detect new bacterial strains in the gut microbiota without prior information of primers or probe.

3.2.4 Genetic fingerprinting of gut microbiota

There are many culture-independent methods which mainly rely on gel-based separation and hybridization of 16sRNA sequences with the probe, for example, T-RFLP, DGGE, TGGE, and a combination of FISH and flow cytometry [38]. These methods are also known as fingerprinting methods have been used to investigate microbial diversity. In the last two decades, fingerprinting methods have offered more information related to the composition of gut microflora. This group of techniques does not provide information about the phylogenetic compositions of the gut ecosystem. But the disturbance in the composition of gut microbiome, which is also known as “gut dysbioses,” caused by various environmental perturbations, including foreign bacterial species and antibiotics, could be investigated in the case of humans [39].

3.2.5 Denaturing gradient gel electrophoresis

It is the most widely used method built on the separation of 16S rRNA gene amplicons on polyacrylamide gel electrophoresis from the complex mixture of DNA fragments that have the same length but different nucleotides sequence [40]. The electrophoretic separation of DNA fragments is influenced by the gel gradient generally produced by denaturant agents, for example, urea and/or formamide. Actually, when the current passes through the electrophoresis gel, 16S rRNA gene amplicons/DNA fragments get separated at various positions on gel according to their molecular weight in linear order, and it continues till their complete denaturation. Consequently, a heterogeneous mixture of DNA sequences is separated in the form of bands on the gel due to their compositions and denatured gradient present in the gel. DGGE is a semiquantitative technique and practiced in the comparison of two different types of microbial communities, i.e., from a healthy or diseased person. The technique is fast and can be used for the separation of multiple samples in single experiments [41]. The main disadvantage of DGGE is that the final results are influenced by PCR-originated bias and not suitable for direct identification of new strains without the availability of a compatible probe.

3.2.6 Temperature gradient gel electrophoresis

It is well known that the DNA sequence influenced the value of the melting temperature (T_m) of a fragment. The high GC content is mainly responsible for high T_m , while the high AT content, for lesser T_m . That can be attributed to the fact that base pairing between G and C contains three hydrogen bonds, while A and T form two hydrogen bonds. Therefore, GC base pairing is more stable than AT in a DNA fragment. In the case of TGGE, denaturant agents are replaced with a temperature gradient. The final results of TGGE protocol mainly depend on amplicon stability and melting behavior, which are determined by GC content. Therefore, when current is passed through the slab gel, intact DNA strands get separated under the influence of temperature gradient inside the gel, but simultaneously, their movements are halted. Consequently, a banding pattern is produced under the influence of the temperature gradient; it is also known as fingerprinting or TGGE [42]. The technique of TGGE is fast and semiquantitative, but like DGGE, its results are also influenced by PCR predispositions. TGGE is not suitable for direct identification of microbes and phylogenetic analysis in absence of sequence-based suitable probes or appropriate hybridization processes.

3.2.7 Terminal restriction fragment length polymorphism assay

RFLP is a classical molecular biological technique used for genetic fingerprinting in the case of animals and plant samples. Its variant T-RFLP is applied to compare the microbial communities and the microbial diversities of gut microbiota. In the process of T-RFLP technique, 16sRNA gene amplicons are isolated from different stool samples and then amplified by PCR. Next, 16sRNA gene amplicons are cut by using different types of restriction enzymes that produced restriction fragments of varying lengths following the isolation of the electrophoresis gel. So that due to different length/M. wt, restriction fragments move to different distances on gel, thus producing a banding pattern. Being fluorescent, each terminal fragment can be identified, whereby each band represents an individual species in the gut community. T-RFLP is used in the comparison of two ecological communities [43]; it is a fast and cheap technique, but not suitable for direct phylogenetic analysis of bacterial strains. Moreover, incompatibility between primer and target genomic DNA influences the T-RFLP results [44]; therefore, it can underrepresent the crucial species, for example, *Lactobacillus* and *Actinobacteria*.

3.2.8 Probe hybridization-based methods

Probe hybridization techniques are mainly used for species identification and their quantification in particular samples. These methods depend on the complementarity between specific oligonucleotide probes and specific target DNA sequences in the bacterial genome. Two major techniques, namely, FISH and DNA microarrays, are included in this class of probe hybridization-based methods which are mainly used in phylogenetic identification and quantification of species living in the microbial ecosystem.

3.2.8.1 Fluorescence in situ hybridization

Basically, FISH is a cytogenetic technique that is applied to pinpoint a specific DNA sequence on the chromosomal landscape by using a suitable fluorescent probe. But, it is also widely used in gut microbiome studies, also known as bacterial

FISH. In the studies of microbial communities, the 16S rRNA gene amplicons are prepared and denatured in a solution. After that, both fluorescent probe and DNA strands are also added in the hybridization solution. In order to allow maximum hybridization process, some cross-linking agents like aldehyde or any precipitating agent (methanol) are also added and incubated in the reaction mixture and kept at 65–75°C for 12 h [45]. After ensuring that the hybridization process is completed, the intensity of fluorescence is measured by using suitable laser available fitted in the flow cytometry instrument. The combination of FISH and flow cytometry is a sort of high-throughput method used in the genome comparison of two different species in the gut sample [46]. The FISH technique is efficiently applied to compare two types of microbial communities such as breast- and formula-fed newborns, and two different species *Bifidobacterium* and *Atopobium* are identified [47]. The merits of this method are that it is semiquantitative and rapid. Due to the availability of diverse probes for specific phyla or species, FISH can be widely used in microbiome studies. But the technique completely failed to identify de novo identification of a bacterial strain. Some researchers have used FISH to estimate the time of sample stability and change in their species compositions with the passage of time and storage conditions.

3.2.8.2 DNA microarrays

DNA microarray technology or DNA chip method is widely applied to learn more about the microbial ecosystem, particularly in gut microbiota. The component of the DNA microarray is a small chip containing a large number of microscopic spots on a solid surface which are used to immobilize fluorescent probes. DNA spots hold pico-level DNA, which is sufficient for hybridization process of a small part of a gene or its regulatory element with cDNA already immobilized on a DNA chip under suitable reaction environments. The microarray protocol includes the following: firstly, the 16S rRNA amplicon or extracted DNA from the samples is processed to make them fluorescent. Secondly, oligonucleotide probes are spotted and immobilized on the surface of the microarray chip [48]. Finally, hybridization is allowed between 16S rRNA amplicons and fluorescent probes. The fluorescence intensity after complete hybridization is quantified by using a laser. The microarray can identify the expression of hundreds of genes in a single experiment. The effect of *C. difficile* infection and its successful cure by fecal microbiota transplantation (FMT) is studied by microarray [49]. This method is quite fast and rapid and offers a high-throughput method for phylogenetic analysis of gut microbiota. It requires a very small amount of DNA for accurate analysis. The most noticed demerit of a microarray experiment is the possibility of cross hybridization, i.e., binding of multiple oligonucleotide probes to a single DNA fragment. In the absence of the probe, a microarray cannot identify a new bacterial species.

4. Functional analysis of the microbial community

4.1 Next-generation sequencing-based methods

Before the advent of NGS, the Sanger sequencing method was the only protocol available to read DNA sequence or full-length 16sRNA gene amplicons. Sanger method was based on the DNA replication process and capillary electrophoresis. In this procedure, all components required for DNA synthesis, i.e., enzyme DNA polymerase, primers for 16sRNA gene, four types of deoxynucleotides (dATP, dGTP,

dCTP, dTTP), and four types of fluorescent chain terminators (dideoxynucleotides: ddATP, ddGTP, ddCTP, ddTTP), are added to single-stranded template DNA and initiate the DNA synthesis process. Consequently, new DNA fragments of various lengths are synthesized with corresponding fluorescent chain terminators which stop further elongation of strands. Hence, randomly terminated DNA fragments are produced that are isolated with capillary gel electrophoresis. On the slab gel, four types of fluorescent dideoxynucleotides fragments can be read by a suitable laser scanning method on the basis of light emitted by them [50]. Therefore, a nucleotide sequence of 16sRNA gene amplicon can be inferred that can be searched in a large number of databases. There are many databases used for the 16sRNA gene amplicon, for example, GenBank and ribosomal RNA gene bank.

Sanger sequencing method not only supports the traditional metagenomic experiments but also supplemented to DGGE, TGGE, and T-RFLP methods as well as whole-genome sequencing metagenomics. The protocol includes the combination of gel and DNA sequencing based methods. In this, the isolated DNA bands from DGGE, TGGE, and T-RFLP gels are removed and sequenced by Sanger's sequencing methods. But in the case of scarcity of DNA, in a particular band, it can be further amplified by PCR and then sequenced. Sanger's sequencing method is most suitable to quantify and carry out phylogenetic identifications of the gut microbiota. Sanger method belongs to first-generation sequencing (FGS) technology, being the most important tool, and is also used for first human genome sequencing. This method is still considered as the gold standard method for long-read sequencing up to 500 nucleotides which are highly essential for genome assemblies. The main disadvantage associated with the Sanger method is its high cost and time-consuming nature.

In the beginning of the twenty-first century, many high-throughput methods of DNA sequencing were developed, for example, pyrosequencing which is a PCR-based massively parallel sequencing platform like Roche/454 pyrosequencing exploited for investigation of gut microbiota. It provided huge genomic data related to human microbiome analyses. Pyrosequencing technique is cheap and high-throughput and requires a small amount of DNA, but short read is a major limitation of the method and unsuitable for comparisons between species within the genus and bioinformatics analysis [51]. Parallely, other next-generation sequencing platforms for DNA sequencing are also developed such as Illumina, SOLiD, Ion Torrent, and single-molecule real-time circular consensus sequencing equipment from Pacific Biosciences and Oxford Nanopore [52]. These technologies have to make microbiome analysis very fast and easy and amass the genomic data for phylogenetic analysis. NGS has provided great speed and accuracy to culture-independent methods used for the study of the functional diversity of microflora. Recently, MinION™ nanopore sequencing technologies used PCR-independent methods; hence, this is free from PCR-based cloning biases, such as amplification temperatures and biased primers sequences. Simultaneously, nanopore sequencing methods offer long reads, which are more suitable for genome assemblies. The above said NGS methods are applied to sequenced cloned amplicons or total community DNA [53]. These methods allow us to investigate gut microbiota qualitatively as well as quantitatively which is influenced by various perturbations, e.g., environmental factors, perturbation, and diets.

NGS is not only useful in phylogenetic classification but also helps in the functional analysis of microbial communities. Therefore, several supplementary technologies also emerged which can differentiate between microbial species in an ecosystem. But it requires analysis of different molecular signatures like DNA, RNAs, proteins, and metabolites generally produced by microbial communities. NGS provided the basic foundation for many omics-based methods, for example,

metatranscriptomics, metaproteomics, and metabolomics, which have helped us in the functional analysis of metagenome represented by a whole microbial community [54]. These methods offered a huge amount of genomic data stored in different databases that can be integrated with the help of bioinformatics tools.

4.1.1 Metatranscriptomics

In fact, transcriptomics is the analysis of the whole gamut of RNA molecules expressed by a particular cell. There are many RNA molecules including mRNA, rRNA, tRNA, and other noncoding RNA transcribed in a microbial ecosystem which play an important role in the gene expression or metagenome expression in the case of the microbial community. Traditionally, the transcriptomics analysis is carried out by measuring the level of RNA expression by using cDNA-based microarray chip. To study microbial communities, thousands of fluorescent probes were required to be immobilized on the microarray chip surface. Actually, metatranscriptomics is the studies of RNA molecules encoded by a metagenome present in a local ecosystem, for example, gut microbiota. Recently, metatranscriptome is studied with the help of the RNA-seq method; this technique is extremely suitable to confirm the gene expression of complete metagenome in the sample which provides the basic data for proteomics and metabolomics [55]. Metatranscriptomics is highly sensitive methods which can even differentiate between dead and live bacterial cell present in a sample. The major drawback of the method is its high cost and it requires great care during the design and execution of experiments because of the momentary stability of mRNA and its contaminations. There are several demerits associated with this method, for example, less amount of mRNA in bacteria, and hence, it creates an experimental problem. Recently, metatranscriptomics methods have been used to identify the pathway of carbohydrate metabolism and energy extraction and physiological functions regulated by a metagenome [56].

4.1.2 Metaproteomics

The proteome is the complete protein complement expressed by a cell or tissue at a particular moment, and the study of the proteome is known as “proteomics.” The metaproteomics or community proteomics is the variant of proteomics in the sense that it is the protein complement expressed by a metagenome from a microbial community. Currently, a small number of reports are available on gut community meta-proteomics that is attributed to the small amount of proteins available in the sample, and its detection makes it further a less applied method in comparison to metagenomics and metatranscriptome. There are still lacking standardized protocols related to protein extraction and its downstream processing. The detection of low abundant proteins in the sample is still a challenge. Moreover, its high cost, time-consuming, and labor-intensive nature further restricted its applications. But many labs have applied metaproteomics in the study of functional analysis of host-microbiome interactions and proteins expressed by gut metagenome. There are two types of proteomics methods, i.e., gel-dependent and gel-independent methods. First, the category of protocols includes the combination of 2D gel electrophoresis, mass spectroscopy, and various bioinformatics tools. Second, categories, namely, shotgun proteomics, mainly depend on most expensive and more sophisticated instruments like two-dimensional liquid chromatography (LC) coupled with nano-spray tandem mass spectrometry (nano 2D LC–MS/MS) and powerful bioinformatics data analysis pipeline. Both types of technologies have provided large-scale protein analysis data in the case of the human gut proteome [57]. Currently,

metaproteomics methods are applied to analyze the effect of dietary components, e.g., resistant starch on protein expression, enzymes, and composition of microbes involved in starch metabolism inside the gut. This technique is useful to investigate the ratio of two important bacterial species *Firmicutes* to *Bacteroidetes* inside gut microbiota [58].

4.1.3 Metabolomics

Metabolites are the final outcome of the gene expression process; they are highly unique in the case of the gut microbiota. Large numbers of metabolites are produced by gut microbiota, which can act as pharmaceutical agents or bioactive products. The metabolomics is a high-throughput omics-based method that mainly deals with the identification and quantification of total metabolites produced in a cell, tissue, and organ which are also called the metabolome. The “meta-metabolome” is the whole complement of metabolites and is produced by a specific microbial community. The analysis of meta-metabolomics requires a set of very sophisticated tools and techniques like matrix-assisted laser desorption/ionization time-of-flight, secondary ion mass spectrometry (SIMS), and Fourier transform ion cyclotron resonance MS that are used for metabolome analysis [59]. The complete annotation of the metabolome produced by a metagenome will help us to understand the physiology and functionality of a microbial community. Inside the human gut, fermentation of short-chain fatty acid is carried out by specific bacteria and produced many types of metabolites that participate in host metabolism and influence the physiology of both host-microbial communities inside the gut. The metabolome analysis offered the investigation of functional gene products in a sample that is helpful in functional analysis of microbes present a microbial niche. Currently, many unique metabolites are identified that are produced by gut microbiota.

4.1.4 Bioinformatics and multi-omics data integration

In the last two decades, bioinformatics has provided much needed help to annotate the complex genome sequences and metagenomic data. The microbial bioinformatics offers help to understand microbial agents of the microbial ecosystem and their mutual and host-microbes interactions. Recently, community-based bioinformatics platforms and pipelines are developed like Mothur and QIIME which help in downstreaming of high-throughput genome sequencing data of variable regions of bacterial 16S ribosomal genes or amplicons. These platforms also help in data analysis and visualization of gut microbiome composition. The high-throughput method like shotgun sequencing and WGS metagenomics produced a huge amount of data, and its annotation is a great challenge in the field of microbiome analysis [60].

In order to know the functions of a particular microbial community, it requires integrating data from other studies such as metatranscriptomics sequencing, metagenomics, metatranscriptomics, metaproteomics, metabolomics, and other techniques. The integration of data provides holistic knowledge of a gut community in terms of its structure and functions [61]. For example, any perturbation such as antibiotics or heavy metal toxicities leads to the change in gut microbial community that can be studied at the level of metabolite production and protein expression. Multi-omics data integration is the uphill task and requires a highly advanced level of computational skill, but current few tools have been developed, e.g., XCMS is a new web-based tool that integrates transcriptome, proteome, and metabolome data [62]. The new systems-level integration can also provide valuable insights, especially when they are combined with community surveys and metagenomics (**Table 1**).

Technique	Basis of techniques	Advantage	Disadvantage
Method for phylogenetic classifications			
<i>Culture-dependent methods</i>			
Culture of bacteria and microscopic studies	Colony features, microscopic and biochemical studies	Low cost	Not suitable for microbiota studies
Culturomics	Culture of microbes and MALDI-TOF mass spectroscopy-based investigations	Appropriate for uncultured microbes	Extremely costly
Microfluidics assays	Microchips based on biochemical reactions	Co-culture of microbes	Need high technical knowledge
<i>Culture-independent</i>			
Quantitative PCR	Fluorescent dyes bind with 16S rRNA gene and quantification of DNA	Highly suitable for phylogenetic classification	Not suitable to identify new bacterial species and biased due to PCR steps
DGGE/TGGE	Separation of 16S rRNA amplicons on electrophoresis based on DNA denaturants and temperature gradients	Fast, less-expensive, and semiquantitative	Results also affected by PCR biases
T-RFLP	Fragmentation of 16S rRNA amplicons by one or more restriction enzymes followed by electrophoretic isolation	Fast and semiquantitative	Not suitable for phylogenetic identification, results are affected with PCR biases
FISH (fluorescence in situ hybridization)	The 16S rRNA amplicon-specific fluorescent probes and flow cytometry	Used for phylogenetic analysis	Unable to identify a new bacterial species. Free from PCR-based bias
DNA microarray	Fluorescent probes immobilized on DNA chip hybridized with 16S rRNA gene. Fluorescence intensity is measured by laser	Phylogenetic identification is possible, high-throughput method, a semiquantitative method which is very fast	Possibility of cross hybridization, PCR biases, detect low-level species in gut microbiota
Cloning of 16sRNA gene (<i>classical metagenomics</i>)	Amplification of 16sRNA gene by PCR- and Sanger-based sequencing by capillary electrophoresis	Highly suitable for phylogenetic classifications and microbiota composition	Affected with the PCR/cloning bias, time-consuming, and extremely expensive
Direct sequencing of 16sRNA amplicon (<i>modern metagenomics</i>)	Sequencing of 16S rRNA amplicons by fast NGS methods, e.g., 454 pyrosequencing, Illumina, SOLiD, single-molecule real-time, Pacific Biosciences and nanopore sequencing methods	Cheap, fast, suitable for phylogenetic identification of unknown microbes	PCR biases, expensive, laborious, and computer intensive
Whole-genome sequencing of bacterial species	Sequencing of the whole genome by NGS-based methods	Suitable for phylogenetic identification of new species	Expensive and computer intensive

Technique	Basis of techniques	Advantage	Disadvantage
Shotgun cloning of microbiome genome/ metagenome	Random shearing of genome. Then assemble genomes on the basis of overlapping sequences by bioinformatics methods	Useful for phylogenetic identification of new species and suitable for microbiome studies	Method is costly and not suitable for phylogenetic classification of a new bacterial species
Method for functional analysis			
<i>Multi-omics methods</i>			
Metatranscriptomics	Sequenced RNA molecules encoded by a metagenome through NGS-based RNA-seq methods	Can identify the metabolism encoded by metagenome	Expensive and requires technical knowledge to conduct experiments
Metaproteomics	Detection of all proteins encoded by metagenome by applying nano 2D LC-MS/MS	Can identify the unique proteins and enzymes encoded by metagenome	Difficult to protein extraction and its downstream processing
Metabolomics	Detection of all metabolites encoded by metagenome using MALDI-TOF, SIMS, and Fourier transform ion cyclotron resonance MS	The method can be used to identify noble metabolites and metabolic pathways imparted by a microbial community	Highly expensive and sophisticated, lack of standard protocols so far
Bioinformatics	Various web-based data analysis pipelines/ platforms are developed QIIME and XCMS	Integration of omics-based data, it provides holistic knowledge about gut microbiome	Need high level of computational skill

Table 1.
Summary of various techniques used for phylogenetic classification and functional characterization of the human gut microbiome.

5. New advancements

5.1 Machine learning

The advancements made in the area of NGS also coincide with machine learning in the last two decades. Machine learning, a branch of artificial intelligence, is based on computational and statistical principles and is recently applied to various fields of genomics including microbiome genomics. Machine learning deals with the development and testing of algorithms to identify, classify, and forecast patterns that emerged from a huge data set [63]. The gut microbial community is comprised of trillion of microbes which further affected various types of factors such as diet, drugs, age, environment, and even lifestyles. To extract the information from such an intricate system cannot be carried out by humans but rather require machine intervention. The machine learning methods such as deep learning and neural network are used to predict severity and susceptibility gingivitis on the basis of the oral microbiome. The two most important machine learning algorithms, random forest and SourceTracker, are applied to know the effect of antibiotics on the genomic and metagenomic studies [64]. In the near future, machine learning can be used to know the host-trait prediction.

5.2 Genome editing/synthetic biology of microbial community

Genome sequencing data of thousands of bacteria are now available in various databases. Currently, many types of genome editing tools are available to manipulate the genome of animals and plants including microbial genomes. Many scientists have exploited these tools in the manipulation of gut microbiota so that desirable genetic changes can be brought into the metagenomes. The most widely used genome editing tool CRISPR-Cas systems also called clustered regularly interspaced short palindromic repeats (CRISPRs) and CRISPR-associated (Cas) proteins are present in the microbes which are mainly responsible for adaptive immunity for prokaryote cells. CRISPR-Cas systems comprise combinations of short DNA sequences called spacers that guide Cas proteins to cleave foreign DNA. So far, CRISPR-Cas systems are the most widely studied and applied method used for genetic manipulation. There are several types of a spacer or genome editing CRISPR-Cas systems, for example, Cas9, CasX, and CRISPR-CasY, that can be used to manipulate genomic content of gut microorganisms. Class 2 CRISPR-Cas systems are streamlined versions, in which a single RNA-bound Cas protein recognizes and cleaves target sequences. Actually, components of Class 2 CRISPR-Cas systems are studied, and assembly from its components in vitro system has revolutionized the field of synthetic biology.

The gut microbiota also comprises a microorganism, for example, single-cell eukaryotes, bacteria, fungus, and bacteriophages. They live in the gut in a very harmonious manner with trillions of bacteria in a natural environment, hence, well adapted to the local environment. Therefore, researchers are embarking on the idea that gut symbionts can be potential agents or vectors for genetic manipulation of gut microbial communities. The new genome editing tools are used to genetically reprogram gut communities under synthetic biology [65]. CRISPR-Cas systems have been exploited to modification of gene expression, change of the production of metabolites, biocatalyst, and protein production that can act as better microbiome modulators. Moreover, genome editing tools will prove extremely helpful in the functional characterization of gut microbiota. Current genome editing tools have offered opportunities in the investigation of intricate relationships between members of the microbiome and host and have opened new avenues for the development of pharmaceutical agents that target the microbiome. But still many demerits are also linked with genome editing tools including their off-targets and inability to introduce exogenous DNA into the metagenome [66]. Moreover, many bacteria particularly unculturable are naturally ill-adapted to transformation methods such as electroporation, conjugation, or transduction in lab conditions.

6. Summary and future prospectus

The gut microbiome is an unexploited huge wealth of microbes that synthesized the valuable and unique metabolites to be used for pharmaceutical industries and the preparation of functional foods. Additionally, metabolites produced by the gut microbiome also contribute in maintaining the health and immunity of the host. In order to exploit microbiome's wealth, we need to apply appropriate and suitable analytical techniques in a highly systematic manner to dig out unique biomolecules. The gut microbial community contains trillions of microbes that make it highly complex. It carried out thousands of metabolic and biochemical reactions in the natural environment. Hence, investigating gut microbiota requires new culturomics methods because of a large number of microbes not able to grow in an artificial environment. Currently, data generated by high-throughput sequencing

contain a wealth of information and must be analyzed by using advanced tools and techniques of bioinformatics and microbiology techniques.

Now the picture of the human gut microbiome is available but still hazy in terms of how microbes impact their host and other microbes living in the gut microbial community. The NGS has revolutionized every field of biological sciences including human microbiome research. It not only sequenced thousands of genome of microorganisms but also helped to emerge many supplementary technologies which are very significant in the functional investigation of the microbial community. Therefore, the advent of modern “omics-based” high-throughput methodologies will help in the identification and characterization of previously unknown microbial strains and modulation mechanisms of the gut ecosystem. But the huge data generation by the omics-based methodologies is a great challenge which needs to be dealt with the development of new bioinformatics tools and techniques. Simultaneously, methods of big data analysis also need to be designed like machine learning and deep learning that will certainly help us in the study of microbial communities.

The availability of cheap and sufficient raw data has opened new avenues. In the near future, gut microbiota can be used as biomarkers and can be personalized to microflora on the line of personalized diet and personalized genomics. Moreover, the recent development of genomic editing tools can manipulate the microbial community under the techniques of synthetic biology. Hence we can cure lifestyle-related diseases such as obesity, cancer, and diabetes by positive manipulation in the composition of gut microbiota.

Acknowledgements

I would like to thank to the Department of Biochemistry, GGSDS College (Panjab University), India, for providing me all the facilities and support for preparation of this manuscript.

Conflict of interest

The authors declare no conflict of interest.

Author details

Akhlesh P. Singh
Genomics and Proteomics Lab, Department of Biochemistry, GGSDS College,
Panjab University, Chandigarh, India

*Address all correspondence to: akhlesh@ggdsd.ac.in; akhlesh@gmail.com

IntechOpen

© 2020 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

References

- [1] Grice EA, Segre JA. The human microbiome: Our second genome. *Annual Review of Genomics and Human Genetics*. 2012;**13**:151-170
- [2] Savage DC. Microbial ecology of the gastrointestinal tract. *Annual Review of Microbiology*. 1977;**31**(1):107-133
- [3] Steinhoff U. Who controls the crowd? New findings and old questions about the intestinal microflora. *Immunology Letters*. 2005;**99**(1):12-16
- [4] Staley JT, Konopka A. Measurement of in situ activities of nonphotosynthetic microorganisms in aquatic and terrestrial habitats. *Annual review of microbiology*. Oct 1985;**39**(1):321-346
- [5] Woese CR, Fox GE. Phylogenetic structure of the prokaryotic domain: The primary kingdoms. *Proceedings of the National Academy of Sciences*. 1977;**74**(11):5088-5090
- [6] Staley JT, Konopka A. Measurement of in situ activities of nonphotosynthetic microorganisms in aquatic and terrestrial habitats. *Annual Review of Microbiology*. 1985;**39**(1):321-346
- [7] Woese CR, Fox GE. Phylogenetic structure of the prokaryotic domain: The primary kingdoms. *Proceedings of the National Academy of Sciences*. 1977;**74**(11):5088-5090
- [8] Zoetendal EG, Vaughan EE, De Vos WM. A microbial world within us. *Molecular Microbiology*. 2006;**59**(6):1639-1650
- [9] Amann RI, Ludwig W, Schleifer KH. Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiological Reviews*. 1995;**59**:143-169
- [10] Savage DC. Microbial biota of the human intestine: A tribute to some pioneering scientists. *Current Issues in Intestinal Microbiology*. 2001;**2**:1-15
- [11] Backhed F, Ley RE, Sonnenburg JL, Peterson DA, Gordon JI. Host-bacterial mutualism in the human intestine. *Science*. 2005;**307**(5717):1915-1920
- [12] Blevins SM, Bronze MS. Robert Koch and the 'golden age' of bacteriology. *International Journal of Infectious Diseases*. 2010;**14**(9):e744-e751
- [13] Ackert L. Sergei Vinogradskii and the Cycle of Life: From the Thermodynamics of Life to Ecological Microbiology, 1850-1950. Vol. 34. Netherlands: Springer Science & Business Media; 2012. p. 194
- [14] Goodman AL et al. Extensive personal human gut microbiota culture collections characterized and manipulated in gnotobiotic mice. *Proceedings of the National Academy of Sciences of the United States of America*. 2011;**108**:6252-6257
- [15] Lagier JC, Dubourg G, Million M, Cadoret F, Bilen M, Fenollar F, et al. Culturing the human microbiota and culturomics. *Nature Reviews Microbiology*. 2018;**16**(9):540-550
- [16] Ma L, Kim J, Hatzenpichler R, Karymov MA, Hubert N, Hanan IM, et al. Gene-targeted microfluidic cultivation validated by isolation of a gut bacterium listed in human microbiome project's most wanted taxa. *Proceedings of the National Academy of Sciences*. 2014;**111**(27):9768-9773
- [17] Shah P, Fritz JV, Glaab E, Desai MS, Greenhalgh K, Frachet A, et al. A microfluidics-based in vitro model of the gastrointestinal human-microbe interface. *Nature Communications*. 2016;**7**:11535

- [18] Nichols D, Cahoon N, Trakhtenberg EM, Pham L, Mehta A, et al. Use of ichip for high throughput in situ cultivation of “uncultivable” microbial species. *Applied and Environmental Microbiology*. 2010;**76**:2445-245
- [19] Jung D, Seo EY, Epstein SS, Joung Y, Han J, Parfenova VV, et al. Application of a new cultivation technology, I-tip, for studying microbial diversity in freshwater sponges of Lake Baikal, Russia. *FEMS Microbiology Ecology*. 2014;**90**(2):417-423
- [20] Arnold JW, Roach J, Azcarate-Peril MA. Emerging technologies for gut microbiome research. *Trends in Microbiology*. 2016;**24**(11):887-901
- [21] Zoetendal EG, von Wright A, Vilpponen-Salmela T, Ben-Amor K, Akkermans AD, de Vos WM. Mucosa-associated bacteria in the human gastrointestinal tract are uniformly distributed along the colon and differ from the community recovered from feces. *Applied and Environmental Microbiology*. 2002;**68**(7):3401-3407
- [22] Tong M, Jacobs JP, McHardy IH, Braun J. Sampling of intestinal microbiota and targeted amplification of bacterial 16S rRNA genes for microbial ecologic analysis. *Current Protocols in Immunology*. 2014;**107**(1):7-41
- [23] Stinson LF, Keelan JA, Payne MS. Comparison of meconium DNA extraction methods for use in microbiome studies. *Frontiers in Microbiology*. 2018;**9**:270
- [24] Bag S, Saha B, Mehta O, Anbumani D, Kumar N, Dayal M, et al. An improved method for high quality metagenomics DNA extraction from human and environmental samples. *Scientific Reports*. 2016;**6**:26775
- [25] Lim MY, Song EJ, Kim SH, Lee J, Nam YD. Comparison of DNA extraction methods for human gut microbial community profiling. *Systematic and Applied Microbiology*. 2018;**41**(2):151-157
- [26] Eckburg PB, Bik EM, Bernstein CN, Purdom E, Dethlefsen L, Sargent M, et al. Diversity of the human intestinal microbial flora. *Science*. 2005;**308**(5728):1635-1638
- [27] Kolbert CP, Persing DH. Ribosomal DNA sequencing as a tool for identification of bacterial pathogens. *Current Opinion in Microbiology*. 1999;**2**(3):299-305
- [28] Olsen GJ, Lane DJ, Giovannoni SJ, Pace NR, Stahl DA. Microbial ecology and evolution: A ribosomal RNA approach. *Annual Reviews in Microbiology*. 1986;**40**(1):337-365
- [29] Lane DJ, Pace B, Olsen GJ, Stahl DA, Sogin ML, Pace NR. Rapid determination of 16S ribosomal RNA sequences for phylogenetic analyses. *Proceedings of the National Academy of Sciences*. 1985;**82**(20):6955-6959
- [30] Rajendhran J, Gunasekaran P. Microbial phylogeny and diversity: Small subunit ribosomal RNA sequence analysis and beyond. *Microbiological Research*. 2011;**166**(2):99-110
- [31] Andoh A, Benno Y, Kanauchi O, Fujiyama Y. Recent advances in molecular approaches to gut microbiota in inflammatory bowel disease. *Current Pharmaceutical Design*. 2009;**15**(18):2066-2073
- [32] Lu T, Stroot PG, Oerther DB. Reverse transcription of 16S rRNA to monitor ribosome-synthesizing bacterial populations in the environment. *Applied and Environmental Microbiology*. 2009;**75**(13):4589-4598
- [33] Nicholls SM, Quick JC, Tang S, Loman NJ. Ultra-deep, long-read

- nanopore sequencing of mock microbial community standards. *Gigascience*. 2019;**8**(5):giz043
- [34] Riddle MS, Connor BA. The traveling microbiome. *Current Infectious Disease Reports*. 2016;**18**(9):29
- [35] Carey CM, Kirk JL, Ojha S, Kostrzynska M. Current and future uses of real-time polymerase chain reaction and microarrays in the study of intestinal microbiota, and probiotic use and effectiveness. *Canadian Journal of Microbiology*. 2007;**53**(5):537-550
- [36] Kieler IN, Mølbak L, Hansen LL, Hermann-Bank ML, Bjornvad CR. Overweight and the feline gut microbiome—A pilot study. *Journal of Animal Physiology and Animal Nutrition*. 2016;**100**(3):478-484
- [37] Zwielehner J, Liszt K, Handschur M, Lassl C, Lapin A, Haslberger AG. Combined PCR-DGGE fingerprinting and quantitative-PCR indicates shifts in fecal population sizes and diversity of *Bacteroides*, *bifidobacteria* and *Clostridium* cluster IV in institutionalized elderly. *Experimental Gerontology*. 2009;**44**(6-7):440-446
- [38] Feng XW, Ding WP, Xiong LY, Guo L, Sun JM, Xiao P. Recent advancements in intestinal microbiota analyses: A review for non-microbiologists. *Current Medical Science*. 2018;**38**(6):949-961
- [39] Lange K, Buerger M, Stallmach A, Bruns T. Effects of antibiotics on gut microbiota. *Digestive Diseases*. 2016;**34**(3):260-268
- [40] Muyzer G. DGGE/TGGE a method for identifying genes from natural ecosystems. *Current Opinion in Microbiology*. 1999;**2**(3):317-322
- [41] Noor SO, Ridgway K, Scovell L, Kemsley EK, Lund EK, Jamieson C, et al. Ulcerative colitis and irritable bowel patients exhibit distinct abnormalities of the gut microbiota. *BMC Gastroenterology*. 2010;**10**(1):134
- [42] Fischer SG, Lerman LS. Separation of random fragments of DNA according to properties of their sequences. *Proceedings of the National Academy of Sciences*. 1980;**77**(8):4420-4424
- [43] Li F, Hullar MA, Lampe JW. Optimization of terminal restriction fragment polymorphism (TRFLP) analysis of human gut microbiota. *Journal of Microbiological Methods*. 2007;**68**(2):303-311
- [44] Polz MF, Cavanaugh CM. Bias in template-to-product ratios in multitemplate PCR. *Applied and Environmental Microbiology*. 1998;**64**(10):3724-3730
- [45] Lukumbuzya M, Schmid M, Pjevac P, Daims H. A multicolor fluorescence in situ hybridization approach using an extended set of fluorophores to visualize microorganisms. *Frontiers in Microbiology*. 2019;**10**:1383
- [46] Swidsinski A, Loening-Baucke V, Vanechoutte M, Doerffel Y. Active Crohn's disease and ulcerative colitis can be specifically diagnosed and monitored based on the biostructure of the fecal flora. *Inflammatory Bowel Diseases*. 2008;**14**(2):147-161
- [47] Bezirtzoglou E, Tsiotsias A, Welling GW. Microbiota profile in feces of breast-and formula-fed newborns by using fluorescence in situ hybridization (FISH). *Anaerobe*. 2011;**17**(6):478-482
- [48] Ingber DE. Reverse engineering human pathophysiology with organs-on-chips. *Cell*. 2016;**164**(6):1105-1109
- [49] Shankar V, Hamilton MJ, Khoruts A, Kilburn A, Unno T, Paliy O, et al. Species and genus level resolution analysis of gut microbiota in

Clostridium difficile patients following fecal microbiota transplantation. Microbiome. 2014;2(1):13

[50] Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. Proceedings of the National Academy of Sciences. 1977;74(12):5463-5467

[51] Rhodes AN, Urbance JW, Youga H, Corlew-Newman H, Reddy CA, Klug MJ, et al. Identification of bacterial isolates obtained from intestinal contents associated with 12,000-year-old mastodon remains. Applied and Environmental Microbiology. 1998;64(2):651-658

[52] Wagner J, Coupland P, Browne HP, Lawley TD, Francis SC, Parkhill J. Evaluation of PacBio sequencing for full-length bacterial 16S rRNA gene classification. BMC Microbiology. 2016;16(1):274

[53] Loman NJ, Pallen MJ. Twenty years of bacterial genome sequencing. Nature Reviews Microbiology. 2015;13(12):787-794

[54] Petrizz BA, Franco OL. Metaproteomics as a complementary approach to gut microbiota in health and disease. Frontiers in Chemistry. 2017;5:4

[55] Franzosa EA, Morgan XC, Segata N, Waldron L, Reyes J, Earl AM, et al. Relating the metatranscriptome and metagenome of the human gut. Proceedings of the National Academy of Sciences. 2014;111(22):E2329-E2338

[56] Gosalbes MJ, Durbán A, Pignatelli M, Abellan JJ, Jiménez-Hernández N, Pérez-Cobas AE, et al. Metatranscriptomic approach to analyze the functional human gut microbiota. PLoS One. 2011;6(3):e17447

[57] Xiong W, Abraham PE, Li Z, Pan C, Hettich RL. Microbial

metaproteomics for characterizing the range of metabolic functions and activities of human gut microbiota. Proteomics. 2015;15(20):3424-3438

[58] Maier TV, Lucio M, Lee LH, VerBerkmoes NC, Brislawn CJ, Bernhardt J, et al. Impact of dietary resistant starch on the human gut microbiome, metaproteome, and metabolome. MBio. 2017;8(5):e01343-e01341

[59] Scholz M, Ward DV, Pasolli E, Tolio T, Zolfo M, Asnicar F, et al. Strain-level microbial epidemiology and population genomics from shotgun metagenomics. Nature Methods. 2016;13(5):435-438

[60] Leone V, Gibbons SM, Martinez K, Hutchison AL, Huang EY, Cham CM, et al. Effects of diurnal variation of gut microbes and high-fat feeding on host circadian clock function and metabolism. Cell Host & Microbe. 2015;17(5):681-689

[61] Roume H, Muller EE, Cordes T, Renaut J, Hiller K, Wilmes P. A biomolecular isolation framework for eco-systems biology. The ISME Journal. 2013;7(1):110

[62] Huan T, Forsberg EM, Rinehart D, Johnson CH, Ivanisevic J, Benton HP, et al. Systems biology guided by XCMS online metabolomics. Nature Methods. 2017;14(5):461-462

[63] Tarca AL, Carey VJ, Chen XW, Romero R, Drăghici S. Machine learning and its applications to biology. PLoS Computational Biology. 2007;3(6):e116

[64] Rahman SF, Olm MR, Morowitz MJ, Banfield JF. Machine learning leveraging genomes from metagenomes identifies influential antibiotic resistance genes in the infant gut microbiome. mSystems. 2018;3(1):e00123-e00117

[65] Ramachandran G, Bikard D.
Editing the microbiome the
CRISPR way. Philosophical
Transactions of the Royal Society B.
2019;**374**(1772):20180103

[66] Burstein D, Harrington LB,
Strutt SC, Probst AJ, Anantharaman K,
Thomas BC, et al. New CRISPR—Cas
systems from uncultivated microbes.
Nature. 2017;**542**(7640):237-241