Journal of Advances in Microbiology

7(2): 1-19, 2017; Article no.JAMB.37951
ISSN: 2456-7116

An Overview of Cultural, Molecular and Metagenomic Techniques in Description of Microbial Diversity

U. O. Edet[1,2*], S. P. Antai[2], A. A. Brooks[2], A. D. Asitok[2], O. Enya[3] and F. H. Japhet[1]

[1]Department of Microbiology, Faculty of Natural and Applied Sciences, Obong University, Obong Ntak, Etim Ekpo LGA, Akwa Ibom State, Nigeria.
[2]Department of Microbiology, Faculty of Biological Sciences, University of Calabar, Calabar, Cross River State, Nigeria.
[3]School of Environmental and Life Sciences, University of Salford, Manchester, United Kingdom.

Authors' contributions

This work was carried out in collaboration between all authors. Authors UOE, SPA, AAB, ADA and OE planned and designed the study. Authors UOE and FHJ wrote the first draft of the manuscript. Authors UOE, OE and FHJ managed the literature searches. All authors read and approved the final manuscript.

Article Information

DOI: 10.9734/JAMB/2017/37951
Editor(s):
(1) Ana Claudia Correia Coelho, Department of Veterinary Sciences, University of Trás-os-Montes and Alto Douro, Portugal.
Reviewers:
(1) Viviane Nakano, University of Sao Paulo, Brazil.
(2) Aqleem Abbas, The University of Agriculture Peshawar, Pakistan.
(3) Manal M. A. El-Naggar, National Institute of Oceanography and Fisheries, Egypt.
Complete Peer review History: http://www.sciencedomain.org/review-history/22367

Review Article

Received 3rd November 2017
Accepted 13th December 2017
Published 19th December 2017

ABSTRACT

Traditional microbiology based on the culture of microorganisms has been the mainstay of microbiology for over a century and has contributed immensely to what we now know about the harmful and beneficial roles of microorganisms. However, it has a number of limitations that has hampered the full utilization of the non-culturable majority in various ecosystems. Metagenomics is a technique that bypasses the need to culture microorganisms from various samples. Unarguably, it is one of the few powerful techniques that have revolutionized every aspect of molecular biology, microbiology, microbial ecology and even beyond. Although still in its infancy, metagenomics have far reaching applications already in medicine and health care, biotechnology, environmental microbiology, bioprospecting for new products of biological origin, discovery of functional genes,

**Corresponding author: E-mail: uwemedet27@gmail.com;*

evolution, to mention just a few. The various steps involved in metagenomics from sampling to sequencing have all been shown to affect the overall diversity of the sampled ecosystem. An ideal methodology should be one that produces a fair representation of the community sampled. The rapid progress in sequencing technologies has forced its cost down per sample. This has lead to a drastic increase in the number of research projects using sequencing technologies, and databases that can store, analyse and allow retrieval of metagenomics data. With this pace of growth, we will certainly see many putative genes, proteins and pathways come alive, and also witness a number of paradigm shifts in some key concepts in traditional microbiology like taxonomy and microbial physiology.

Keywords: Metagenomics; databases; cultural methods; molecular methods; microbial diversity.

1. INTRODUCTION

Microbiology is traditionally defined as the study of microorganisms that are too small to be seen with the unaided eye. This definition implies that microscopy is an inseparable aspect of microbiology. However, despite improvements in microscopy over the last century, much of what we see under the microscope is still limited to a large extent by what we can culture in the laboratory. The culturable majority has brought about a lot of benefits to humans in all areas of life. Without microbes, life would have been very difficult or almost non-existent when we look at the array of beneficial roles that these microorganisms play in our lives, industries, biotechnology, health and environment. They are intimately involved in biogeochemical cycling of nutrients like carbon, nitrogen and sulphur, where they unlock and make these nutrients available for re-use [1,2]. In human and animal guts, they are intricately linked with food digestion, degradation of toxins, and fighting off disease causing pathogens from our bodies [1,3]. Incessant spillages of xenobiotics into the environment are degraded via their outstanding ability to utilize these complex compounds as sources of carbon or energy, and in the process transform them into less harmful compounds [4].

Traditional microbiology tries to tie these functions and roles played by microbes in the environment, nutrition, health, and disease to just about one microbe or a few groups that are culturable. In reality, these events are usually not unimicrobial (one) or oligomicrobial (few) but polymicrobial [5]. Sadly, the routine cultural and microscopy techniques often used to study and characterize isolates in traditional microbiology does not allow for the desired polymicrobial identification of the microbial consortium. This has limited the full exploitation of the "unculturable majority", and seemed impossible for sometime [6].

Studies have shown that cultural methods can only accounts for < 1% of the overall microbial diversity in an environment [7]. Although many brilliant cultural media have been devised to increase the diversity of microbes, the "great plate anomaly" is still valid [8]. The anomaly tries to account for the fact that these unexploited majorities cannot be grown under laboratory conditions as some may be non-viable or viable but not culturable. Given the inabilities of the culture based methods, as expected a number of advancements have been recorded in culturable methods. However, they are still plagued with problems.

Coined by Handelsman and colleagues in 1982 in the United State of America, metagenomics offers a great insight into the unculturable world of various environments. Metagenomics is a culture independent technique that allow for the much desired polymicrobial and direct exploitation of microbes from various environments without the need to culture them. Metagenomics studies have soared tremendously over the last two decades in particular [9]. Together with the polymerase chain reaction; they are arguably two of the most powerful techniques of the 20th Century. It allows for the estimation of microbial diversities in various ecosystems, and exploitation of the resulting diversity for the recovery of novel genes, enzymes, pathways and various metabolic products [10]. The aim of this review paper is to provide a survey of cultural, molecular and metagenomics techniques used in characterizing microbial diversity, examine the basic steps, strategies and various applications of metagenomics with some emphasis on its environmental applications. The review ends with some of the databases used to analyse metagenomic data and its future prospects.

2. DIVERSITY OF VARIOUS ECO-SYSTEMS

Microbes represent the most diverse group of organisms on earth and make up 60% of the earth biomass. Prokaryotes dominate the biosphere with a population of about 4-6 x 10^{30} cells which is about 2-3 times in order of magnitude more than all the plants and animals cells combined together [11]. Biological diversity or biodiversity is defined as the collective variation at all levels of biological organization, from the genetic variations within populations and species, species within communities, to communities that compose an ecosystem [11,12,13]. Soil environment host the most diverse microbial community. It is estimated that the earth has a bacterial population of 4-6 x 10^{30} out of which 2.6 x 10^{29} is in the soil [11]. Water covers about 71% of the earth's surface with most of these waters found in the oceans and seas. Sogin [14] estimated that the oceans contain about 3.6 x10^{29} microbial cells. The intestinal tract remains the most densely populated part of the human body with microbial counts reaching $10^{12\text{-}14}$ bacterial cells [3,15,16]. Sludge environment are also microbiologically diverse and dynamic. Much of these diversities had remained elusive with traditional microbiology. A number of metagenomic studies have been carried out over the last 10 years on these ecosystems and they have revealed the enormous diversities in these environments [8,11,16,17,18].

3. ECOLOGICAL METHODS IN TRADITIO-NAL AND MODERN MICROBIOLOGY

In order to describe the diversities of microbes in various environments, a number of techniques have emerged. Passively, they are classified into cultural and molecular methods. This classification is already biased because we now have stable isotope probing and the "omics" based techniques. A better classification would be one that would separate metagenomics, metaproteomics, protogenomics and metatranscriptomics into the "omics" group and stable isotope probing into the non-molecular methods.

4. CULTURAL METHODS IN MICROBIAL ECOLOGY

Cultural methods of investigating the ecology of natural and impacted environments have been the mainstay of microbiology for over a century now. These techniques are still important and have led to the description of a number of habitats but are extremely biased in their ability to capture the microbial diversity of these environments. Cultural methods rely heavily on cultural characteristics, morphology, microscopy, Gram reaction, biochemical and physiological reactions to characterise microbial community diversity. The main challenge here has been that about 90-99% of these organisms cannot be cultured in the laboratory. However, cultivation allows for the isolation of microbes of interest for further studies [19]. Some of the commonly used cultural methods in microbial diversity description are based on plate counts, sole source of carbon utilization and analysis of fatty acids.

5. PLATE COUNTS

Plate counts are routinely used in traditional microbiology to describe diversity of bacteria and fungi using appropriate media. Following serial dilutions and plating, viable counts or total aerobic or total heterotrophic counts are obtained. These isolates are further subjected to purification via sub-culturing. Once purified and pure culture obtained, they are then subjected to Gram staining, microscopy and a battery of biochemical tests to help identify the isolates [20]. The reactions of bacterial isolates to these tests are used to place them in taxonomic groups in accordance to the Bergey's manual of determinative bacteriology [21]. Despite availability of commercial identification kits, selective media and advances in microscopy, cultural methods have a number of disadvantages. These include time consuming for slow growing microbes, prone to errors and contamination and could be very expensive. Furthermore, they only capture the culturable minority.

6. SOLE CARBON SOURCE UTILIZATION (SCSU)

Also known as community level physiological profiling (CLPP). It is principled on the ability of microbes to utilize an array of compounds as sole source of carbon for their metabolism. SCSU is commonly used in traditional microbiology to access the microbial functional diversity based on sole carbon source utilized. It allows for monitoring bacteria communities abilities and rates of utilization of these carbon sources. Commercially available kits such as the Biolog capable of examining about 95 carbon sources are now available and they utilize colour

change from the reduction of tetrazolium violet as an indicator of utilization. They are adaptable to multivariate statistics that are used to analyse the results. Some pioneer studies exist that have used SCSU [22,23,24]. A critique by Preston-Mafham et al. [25] and a review by Fakruddin and Mannan [13] identified some challenges of SCSU including sensitivity to inoculum density, captures only carbon sources utilisers, and the viable and culturable fractions.

7. FATTY ACID ANALYSIS

Bacteria do have a cell membrane composed mainly of proteins and fatty acids (lipids) [26]. This has been exploited in the identification and description of community diversity. Unlike nucleic acids, fatty acids are very stable and not associated with plasmids or mutations [13]. Those with two carbons to twenty-four (C_2 – C_{24}) are very conserved across microbes and useful in community diversity profiling. Variant of this technique include fatty acid methyl ester (FAME) analysis that uses certain unique fatty acids to distinguish major taxonomic groups within a community and allows fatty acid to be directly extracted from soil or sediment, methylated and analyzed by gas chromatography [19,27]. Another variant is the phospholipid fatty acid analysis (PFLA). Fatty acid analysis is sensitive to external factors [13].

8. MOLECULAR METHODS

The introduction of molecular methods has led to a new era in microbial ecology. These molecular methods have shown that the age long cultural methods are highly biased in capturing the biodiversities of various environments. Given the limitations of cultural methods often used in tradition microbiology, there now exist a number of molecular methods that bypass the need to culture microbes directly and are used to describe diversity in extreme and non-extreme environments, and also monitor changes in microbial communities. Molecular methods of analysis of biodiversities are based on direct isolation and analysis of biomolecules such as proteins and nucleic acids. They include genetic fingerprinting, metagenomics, metaproteomics, metatranscriptomics, and proteomics. Others include but not limited to single stranded confirmation polymorphism (SSCP), denaturing or temperature gradient gel electrophoresis (DGGE), Fluorescent in-situ hybridization (FISH), G+C content, restriction fragment length polymorphism (RFLP) and microarray. These molecular methods are well reviewed in a number of papers [13,19,28]. However, metagenomics remain the gold standard and the most used of all the molecular techniques. Since other molecular methods used in diversity and functional genes analysis are equally important and worth mentioning, thus the next few lines are used to give a general overview of some these methods.

9. RATIO OF GUANINE PLUS CYTOSINE

Genomic guanine and cytosine content (G+C) of eubacteria and other microbes differ, and is related to their phylogeny. These differences are exploited in the study of microbial diversity. Basically, they vary by 3-5% in their guanine and cytosine content. Thus, employing fractionation of whole community DNA followed by density gradient centrifugation based on G+C content, diversity can be described [29,30]. As a technique, it is quantitative, gives a rough estimate of diversity, and is not affected by polymerase chain reaction bias [19]. The technique have been used to study microbial community changes following various agricultural management practices coupled with denaturing and temperature gradient gel electrophoresis (DGGE) and amplified ribosomal DNA restriction analysis (ARDRA) [30].

10. DNA-DNA HYBRIDIZATION

As the name implies, this technique is principled on the ability of separated DNA to hybridize or re-associate with one another based on the complementarities of their bases. In this technique, DNA from pure or environmental samples is first extracted, purified, denatured and incubated under conditions that would allow them to anneal. The degree of hybridization is a measure of correlation and viz-a-viz the diversity of the community. Basically, the higher the rate of association, the lower the sequence diversity of the community and vice versa [11,28,31].

11. DNA MICROARRAY

The basic principle behind DNA microarray technology is to immobilize known DNA sequences referred to as probes in micrometer-sized spots on a solid surface (microarray) and specifically hybridize a complementary sequence

of the analyte or target DNA. An entire microbial genome can be represented in a single array, making it possible to carry out a genome-wide analysis of such microbe [32]. As a high throughput technique, it is also used to screen environmental samples for diversity [28]. Panicker et al. [33] identified two major types of DNA microarray namely: The oligonucleotides based and the PCR product based arrays. Irrespective of the type, the components of an array remain basically similar: array fabrication, probe preparation, hybridization and data analysis. However, there are a number of fundamental differences regarding its applications in prokaryotes and eukaryotes. For example, total (ribonucleic acid) RNA is usually labelled for bacterial array experiments while poly RNA is usually labelled for eukaryotic arrays [32].

12. RIBOSOMAL INTERGENIC SPACER ANALYSIS (RISA)

Carlse Woese [34] pioneering work on taxonomy has shown that all organisms on earth can be classified using the small subunit ribosomal RNA. RISA is a molecular technique that employs the polymerase chain amplification of a portion of the intergenic spacer region (ISR) which is in between the 16S and the 23S subunits of the ribosomal subunits. Specifically, it uses primers annealing to the conserved regions in the ribosomal subunit genes [35]. A variant called automated ribosomal intergenic spacer analysis (ARISA) exists that detects the ISR fragment using a detector and a forward primer that is fluorescently labelled [28]. It is capable of generating highly reproducible bacteria community profiles and the automated version allows for analysis of more than one sample at a time. However, like microarray and other PCR based techniques, it is subject to PCR biases and requires large quantities of DNA [13].

13. DENATURING OR TEMPERATURE GRADIENT GEL ELECTROPHORESIS (DGGE OR TTGE)

First introduced by Muyzer et al. [36], it is a versatile genetic fingerprinting technique in microbial ecology. In this technique, extracted DNA from environmental sample is used to obtain a mixture of PCR products using primers specific for a molecular marker. Commonly used is the 16S rRNA gene. The mixture is then separated using a polyacrylamide gel containing a linear gradient of DNA denaturants [36]. In the temperature gradient, the principle remains the same. The only difference is that instead of a chemical denaturant, a temperature gradient is employed and sequence variation amongst the various amplicons determines their melting points. Generated profiles using universal primers are usually very complex and this challenge seems to be overcome using group specific primers that target specific physiological and or phylogenetic groups [37]. These techniques, can handle large amounts of samples at a time, is reliable, reproducible and fast. However, it is subject to PCR bias, issues of co-migration and poor resolution of non-dominant species are common [13].

14. SINGLE STRAND CONFORMATION POLYMORPHISM (SSCP)

This technique is similar in a number of ways to denaturing gradient gel electrophoresis by using electrophoresis and PCR. However, it differs from it by using a non-denaturing polyacrylamide gel [13,38]. In SSCP, PCR products from extracted and amplified DNA from environmental samples are denatured to produce single stranded DNA fragments that are then separated by electrophoresis. Separation on gel in this technique is based on the ability of the ssDNA to uniquely fold leading to differences in mobilities. It does not require a GC clamped primer, gel gradients or elaborate gel set ups. It works well for small fragments preferably 150 to 400 bp and is very simple and reliable. Like DGGE, it is subject to PCR biases. However, a major limitation of this technique appears to be the ability of some DNA strand to form multiple stable conformations [13].

15. RESTRICTION FRAGMENT LENGTH POLYMORPHISM (RFLP) OR TERMINAL RESTRICTION FRAGMENT LENGTH POLYMORPHISM (T-RFLP)

This molecular technique is similar to amplified ribosomal DNA restriction analysis but differ in the use of 5' fluorescently labelled primer during the polymerase chain reaction. As a method of studying microbial diversity, it relies on DNA polymorphisms, single nucleotides changes in sequences. The resulting 16S rDNA fragments from PCR are digested with restriction

enzymes and electrophoresed in agarose or acrylamide gels. RFLP has been used to estimate the prokaryotic diversity in hypersaline ponds [39].

16. FLUORESCENT *In-situ* HYBRIDIZATION (FISH)

FISH is a versatile molecular technique that allows at source (*in situ*) phylogenetic identification and enumeration of individual microbial cells using whole cell hybridization to oligonucleotides (18-30 nucleotides long) probes [40]. The probes usually contain a 5' end fluorescently labelled to a dye which allows for its detection following hybridization by epifluorescent microscope. The strength of the signal is correlated to the cellular rRNA contents and growth rates. These parameters can be used to estimate the metabolism dynamics of the community under study. Some of the disadvantages of FISH include low signal intensity, background noises, and poor accessibility of targets. However, variants such as catalyzed reporter deposition (CARD) FISH, flow cytometry coupled FISH, and FISH coupled to ion mass spectrophotometry give better resolutions [28].

17. STABLE ISOTOPE PROBING

Abbreviated as SIP, it is a wonder technique that has been coupled to metagenomics in bioremediation studies. It monitors an incorporated stable isotopes (^{13}C, ^{15}N, ^{18}O and ^{2}H) given to microbial community of interest so as to monitor a certain metabolic process [28,41]. The labelled molecule which could be DNA or RNA or other biomolecules are then separated using various biochemical means and the biochemical property of the community established using molecular means such as 16S rRNA clone libraries. Like FISH, it is very versatile and has been coupled to FISH and Raman microscopy to enable the simultaneous resolution of taxonomic identities and activity of microbial communities [42].

The most commonly used in the determination of functional genes has been DNA-SIP as it is a lot easier to detect functional genes in ^{13}C-DNA compared to mRNA. It is therefore not surprising that SIP has already been coupled to metagenomics, offering an opportunity to easily describe a functional gene of interest whose diversity is low from a vast array of functional genes that may be present in an ecosystem. This is interesting because it allows the full capture of microbial groups of low abundance. It is proposed that SIP coupled to metagenomics could be of great potential benefit to bioremediation because it can assess a sample and determine if in the first place, the desired functional genes and microbes capable of metabolizing the pollutants are present or not. Furthermore, it could provide an understanding of the mechanisms and pathways underlying biostimulation and phytoremediation. However, SIP has some limitations and these include high cost of labelled substrate, labour intensive and low throughput [41].

18. METAGENOMICS

First coined by Jo Handelsman and colleagues at the University of Wisconsin in the United States of America in 1982, metagenomics is defined as the collective study and analysis of the entire genome of a microbial community other than that of an individual microorganism bypassing the need to culture the microorganisms [43]. It is important to note that early forms of microbial community phylogenetic analysis existed [44]. Etymologically, the word metagenomics simply means beyond (In Greek “meta” means “transcendent”) individual organism [1]. A much simpler definition is given by De Mandal et al. [15] where they defined metagenomics as the study of genetic materials retrieved directly from environmental samples. It is also called ecological, community or environmental genomics [43]. Irrespective of the definition, certain key words make up metagenomics and these include sample, DNA extraction and amplification, sequencing and analysis. It is principled on the differences that exist in nucleotide sequences that drive the functional and structural differences that exist between these organisms. There are basically two approaches to metagenomics and these are sequence based metagenomics (that answers the question who is there?) and the function based metagenomics (which answers the question what are they doing?). The flowchart below shows the basic steps involved in metagenomics. This is followed by a brief description of the basic steps involved in metagenomics.

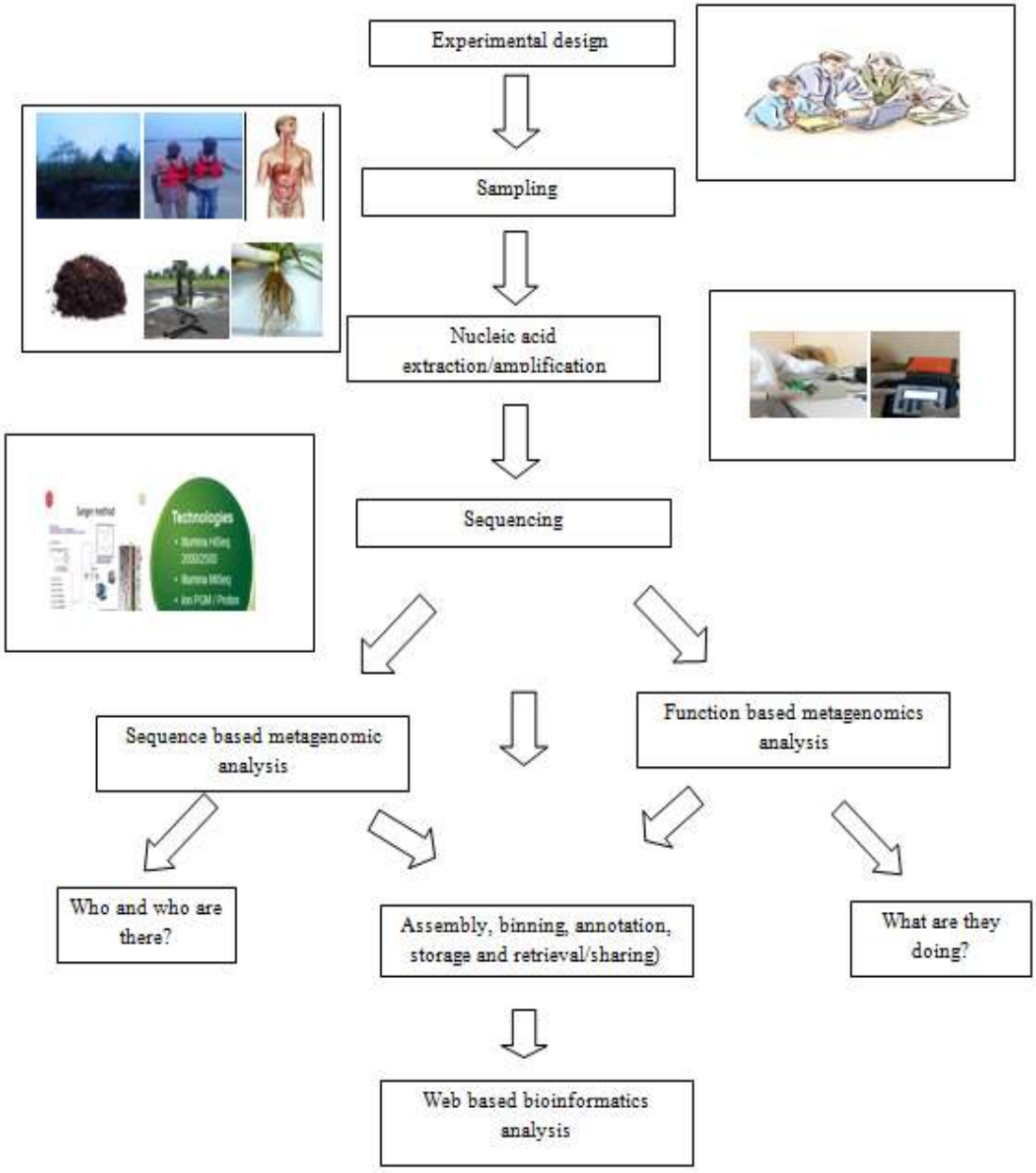

Fig. 1. Basic steps involved in metagenomics

19. EXPERIMENTAL DESIGN AND SAMPLING

Experimental design is a very important step in carrying out a metagenomics study just like every other study. It is the stage where pertinent questions about sampling, DNA extraction, amplification and metagenomics analysis will be asked and answered before venturing into the actual study. Without a good design and proper planning, time and resource wastage can ensue [45,46,47]. Key decisions such as whether replicate sample collection or single sample collection is enough, whether to shallow sample or deep sample, protocol for DNA extraction/amplification, sequencing technology to adopt, strategy for interpretation and metadata collection are reached [45]. Sampling is a crucial step in metagenomic studies that affects the quality and quantity of the final results

[10]. It covers strategies for sample collection, storage and transportation to the laboratory for further analysis. Samples collected for metagenomics studies vary and can be divided basically into environmental and non-environmental samples. Environmental samples can be described as heterologous and homogenous [10]. Environmental samples could be pristine or non pristine samples. They include water (oceans, seas, lakes, ponds, streams, rivers, estuaries, and lagoons), various soils, sludge/waste waters, sediments, animal feaces, effluents (such as palm oil mill, cassava mill, textile and rubber mill waste water or effluents) and assays. Samples can also be obtained from extreme environments such as hot springs, acid mines and deserts. Non-environmental samples include the guts and skin of human, animals, and clinical samples such as blood, faeces, etc. Others include biofilms. The best sampling methodology should be the one that will eliminate completely the entry of exogenous microbes and possible contaminants that could interfere with the quality of the community nucleic acid. It is also important that the sampling methodology adopted in the collection of the samples be representative enough to capture the structural diversity and/or functional gene diversity sought [1,48]. Furthermore, the storage time should be reduced to the barest minimum time possible in order to reduce the effect of time on communities .Storage conditions (-20 and -80^{o}C) should be optimum enough to preserve the microbial community.

20. DNA AND AMPLIFICATION

A thorough DNA extraction is very important in metagenomics and the protocol varies depending on the sample from which the DNA is to be extracted from. DNA extraction is well reviewed and evaluated by a number of authors [10,46,48,49,50]. Extracted DNA should be of enough quantity and high quality for subsequent library production and sequencing [46]. DNA extraction procedures have come a long way. The pioneering method of DNA extraction of Torsvik [51] has undergone a lot of transformation and some recent available procedures require less time, less quantity of samples and are commercialized using kits. Thus far, there is no universally acceptable method of DNA extraction. Extensive studies have been carried out on soil DNA extraction and they highlight the need to use multiple extraction methods for comparism [46]. Studies have shown that extraction of nucleic acids using kits work differently on different soil samples [52]. Furthermore, none is free of bias [53]. However, the basic steps as described by Felczykowska et al. [10] can be divided into indirect and direct methods. The steps usually involve separation of cells via (shaking, centrifugation and filtration), cell lysis (enzymes: lysozymes, proteinase K, RNase A, and achromopeptodase; mechanical: beading, vortexing and sonication, thermal means: freezing- thawing and boiling, and chemical means: SDS (sodium dodecyl sulphate), phenol and CTAB- cetyltrimethyl ammonium bromide), and purification. Some of the solvents commonly used include ethanol, isopropanol, chloroform and polyvinyl polypyrrolidone (PVPP). An ideal extraction method should have the following properties. First, it should extract the community DNA in such a way that it is representative of the whole community. Second, it should be easy to replicate and with reasonable speed and accuracy. Third, it should be such that the employed physical means do not disrupt the nucleic acid. Fourth, it should be such that the yields of nucleic acid are high. Fifth, contaminations from exogenous compounds such as humic acid, clay, metals and proteins should be kept near elimination as possible [10,54]. Finally, it should eliminate excessive fragmentation of the resulting genetic material [10].

21. POLYMERASE CHAIN REACTION (PCR)

Invented in early 80s by Dr Kary Mullis and colleagues, this Noble prize winning invention popularly called polymerase chain reaction (PCR) together with metagenomics are two key events invented in the last quarter of the twentieth century that have shaped and revolutionized molecular biology, microbiology, microbial ecology and biotechnology. PCR is an exponentially progressing *in vitro* synthesis of a defined target of DNA sequence. The technique is very versatile, robust, sensitive, rapid and simple to use. Although the basic steps of PCR is still the same - denaturation, annealing and extension [55], using mock microbial communities, it was observed that PCR and DNA extraction are potential sources of bias in next generation sequencing (NGS) metagenomic studies, and are capable of distorting results [53]. Earlier, Wu et al. [56] demonstrated the effects of polymerase, template dilution and cycle number on 16S rRNA diversity analysis. Their findings support much earlier study and it has been

suggested that the cycle number be kept as minimum as possible preferably 30-35 as higher cycle numbers lead to accumulation of mutations [57]. Hong et al. [58], using marine tidal flat diversity as their model, showed that even with unlimited sampling and sequencing efforts, a single combination of PCR primers and extraction protocol only enabled half of the diversity to be described. Pinto and Raskin [59], using mock bacterial and archaeal communities concluded that the greatest obstacles to correctly evaluate community structures are errors in the estimation of mean relative abundance of each detected operational taxonomic units (OTU) that arise from biases due to multi template PCR reactions.

22. SEQUENCING TECHNOLOGIES

Next generation sequencing (NGS) technology is gradually becoming very popular and taking the centre stage from the classical Sanger di-deoxy sequencing technology in not just metagenomics studies but other studies [46,60]. However, Sanger sequencing is still regarded as the gold standard for sequencing due to its low error rates, large insert size of >30kb for use in bacterial artificial chromosomes (BACs) and fosmids, and long read lengths >700kb. However, its major drawback include the high cost per gigabase, inability to do more than one sample at a time, and requires a lot of DNA to start [46]. Other limitations include the need for gels or polymers for sieving labelled DNA fragments, automation and de novo assembly to generate longer genomics contigs challenges [60]. To avert these challenges, a number of NGS technologies have emerged. These include the 454/Roche, the Illumina/Solexa systems, applied Biosystem SOLiD, Hydrogen Ion (pH) based Ion Torrent, Pacific biosciences (PacBio) (single molecule sequencing –SMRT), and Oxford nanopore technology. Of these technologies, the most used remains the 454/Roche and Illiumina/Solexa technologies [46]. These technologies have been excellently reviewed by several authors [46,60,61,62,63]. We present here the key features of some of these technologies especially those commonly used in metagenomic studies including their merits and demerits.

The 454/Roche pyrosequencing systems use emulsion polymerase reaction (ePCR) to clonally amplify DNA fragments attached to a microscopic bead and then pyrosequenced individually [46]. Pyrosequencing is based on sequence-by-synthesis of the nucleotides using DNA polymerase and on detection of the released pyrophosphate (PPi), a by-product of the DNA synthesis. Subsequently, the PPi is then quantitatively converted into adenotriphosphate (ATP) by ATP sufurylase in the presence of adenosine 5' phosphosulfate (APS). The generated ATP then drives light production using luciferase [60]. It has a number of advantages over Sanger sequencing. It is carried out in real-time, can be automated for large-scale screening, offers multiplexing and allows for up to 12 samples to be analyzed in a single run of approximately 500Mbp. Furthermore, it has been adapted to tens of nanogram of DNA for sequencing single-end libraries. Some of the demerits include artificial replicate sequences that affect estimation of gene abundance, and also the difficulties in correlating the light intensity to actual nucleotide position [46].

The Illumina/Solexa unlike 454/Roche technology performs solid surface PCR and does not use emulsion like the latter. It immobilizes random DNA fragments onto a solid support and then performs PCR that produces a cluster of identical fragments that are then sequenced with reversible terminators using sequencing by synthesis process. It has high error rates at the tail ends of its reads, cheaper and can work with a few nanograms. However, it is limited by small read lengths (36 bp) and longer run time. The run time appears to be improved by the Miseq instrument [46].

Oxford nanopore (MinION) technology is a technology pioneered by David Deamer and colleagues. Beginning from the 1990s, the technology has evolved into a competitive and portable technology. Unlike other NGS technologies, it can detect base modification in native DNA with accuracies up to 92-98%, it can achieve real time targeted sequencing, portable, and requires simple library preparation [64].

Other sequencing technologies exist that are not very popular with metagenomics sequencing but look like they will be good alternatives in the future especially with bias and cost reduction. Pacific Biosciences (PacBio) sequencing technology is based on a single molecule sequencing (SMRT) with real time detection in zero-mode wave guide wells. It is not popular with metagenomics given it 85% accuracy of single reads, it has been used to study direct DNA methylation [46,65]. However, it can get really long sequences of 5,000 to 8,000 bp on

the average. PCR bias is eliminated as it does not require amplification before sequencing. However, it is more expensive than the aforementioned techniques. Others include the Ion Torrent and the hydrogen ion (pH) based, Helicos HeliScope and Applied Biosystems SoLiD [63].

23. ASSEMBLY, BINNING AND ANNOTATION

Assembly of the short read fragments is necessary to obtain large contigs where genomic recovery of uncultured organisms is desired. Furthermore, without assembly longer and more complex genetic elements cannot be analysed and also study of repeat classes. There are two major approaches for assembly in metagenomics and these are: *de novo* and reference-based assemblies. Binning is the process of sorting DNA sequences into groups that be used to describe an individual organism or a genome. Several algorithms exist that have been used in binning and these are based on two identified facts. First, in compositional binning, nucleotides have conserved guanine-cytosine (GC) regions that will be reflected in the fragments of the genome. Second, genes of interest could be used to bin the sequence by referring to such genes in reference databases. Annotation is simply the assignment of function(s) to genes of interest. This can be done using two approaches namely feature prediction and functional annotation [46].

24. DATA BASES FOR METAGENOME

The works of Venter et al. [17] that generated the Sargasso Sea data set and the Global Ocean Survey that stretched from the Gulf of Maine all the way to the Eastern tropical South Pacific Ocean generated huge amounts of data [66]. Given the huge datasets that emanates from metagenomic studies, a number of databases are now available for storage, retrieval and analysis of metagenome data. Six of such databases have been identified by Mineta and Gojoboru [67].

The iMicrobe (http://imicrobe.us) database holds in addition to metagenomic datasets, genomics and transcriptomics data sets and also offer web based computerized environment that enables metagenomic research. The next metagenomic data base is the VIROME. As the name implies, it is a collection of DNA and RNA that make up the viral community. In this context, it refers to environmental viral data (http://virome.dbi.udel.edu/). Using open reading frames, it performs classification of viruses [68].

MetaGenomics Rapid Annotation using Subsystem Technology (MR-RAST), is an automated user friendly SEED environment that allows users to do metagenomic analyses and its server is built on a modified version of the rapid annotation using subsystem technology (RAST). The server can provide amongst other annotation of sequence fragment, phylogenetic classification, functional classification of samples, and comparison between multiple metagenome and also allows for an initial metabolic reconstruction and comparism of metagenome. The web site is hosted at (http://www.mcs.anl.gov/project/mg-rast-metagenomics-rast-server) while the analytical site is hosted at http://metagenomics.anl.gov/.

The European Bioinformatics Institute (EBI) metagenomics is a freely available data analysis hub that is hosted at https://www.ebi.ac.uk/metagenomics/ and is used for analysis and archiving of metagenomic and meta-transcriptomic data sets from important projects like the Ocean sampling day and Tara oceans oceanographic projects. This platform support analysis of meta data derived from NGS platforms such as Roche 454, Ion Torrent and Illumina [69]. Others include the IMG/M, MetaGenome Analyzer DataBase (MeganDB) and CAMERA. It is also important to mention the cluster of orthologous genes (COG) and Kyoto encyclopedia of genes and genomes (KEGG) data bases as they also allow for functional gene diversity analysis. Much more detailed information about metagenomics analytical tools and databases is given by Kim et al. [70].

25. APPLICATIONS OF METAGENOMICS

The entire earth is endowed with microbial communities that may hold the very keys to the enormous challenges facing mankind in the environment, health and medicine, agriculture and even industry [1]. It is therefore not surprising that metagenomics is already being applied in various environments to seek answers to this challenges. Some of the applications of metagenomics are discussed below.

26. MUSHROOM COMPOST

Mushroom is a very rich source of nutrients for human and animal consumption. Several studies

have shown that both edible commercially cultivated and wild mushroom are very rich in nutrients [71,72]. Mushroom play important roles in the ecosystems and one of such is decomposition and nutrient recycling in the environment. Mushroom compost is a very complex and microbiologically diverse ecosystem that is swarming with bacteria, fungi and actinomycetes with one group succeeding the other during compositing [73,74,75]. The study of mushroom diversity and succession activity as compared to that of other ecosystems is still dominated by cultural methods. In their flow chart, Argwal et al. [74] listed a number of culture dependent and independent methodologies commonly used to study mushroom compost including their merits and demerits. In all cases, the common demerit of the culture independent techniques was the inability to capture the unculturable fraction of the compost ecosystem. Methodologies now available for whole community studies of bacterial diversities in mushroom compost include metagenomics, metaproteomics, proteomics, metatranscipto-mics, whole genome sequencing, analysis of guanine plus cytosine contents and metabolomics [73].

27. MICROBIAL POPULATION DYNAMICS DURING BIOREMEDIATION AND OTHER WASTER TREATMENTS

Bioremediation is a low cost green clean up technology that is capable of removing pollutants or at least transform them to an innocuous state [4,76]. Bioremediation whether as applied to petroleum sludge, crude oil and other effluents such as waste water is a polymicrobial and complex event that can be performed *in situ* and *ex situ.* Furthermore, they are polymicrobial events [77,78,79]. As a technique, it is not completely fool-proof and some of its limitations include nutrient availability and non-optimization of the process. Furthermore, the non-usage of molecular techniques to capture the polymicrobial events viz-a-viz the microbial diversity in earlier bioremediation studies was also a major set-back in the optimization of bioremediation processes. Other challenges include the fact that less than 10% of these organisms are actually culturable and the poor performance of culturable species with the non-culturable ones especially when they are used *in-situ.* Apart from process optimization, the use of molecular identification is important for a number of reasons. In their study, metagenomics analysis of sediments impacted with aviation fuel revealed a total of 718 species compared to the only 6 species revealed by cultural method [8]. Secondly, metagenomics can also reveal the functional genes with which these organisms use in the degradation of the pollutants such as alkane monogenase, napthlene dioxygenase, pyrene dioxygenase and biphenyl dioxygenase genes. Molecular methods can also be applied to monitor the community changes that occur during bioremediation using 16s amplicon sequencing and also community changes and stability. Furthermore, it can help provide insight into some of the pollutant pathways that are still elusive or even help explain better those already known. Finally, microbial signature resulting from application of metagenomics over time can be useful as to fingerprint pollutants [80].

28. DIAGNOSIS OF BACTERIAL, VIRAL AND PARASITIC INFECTIONS

Traditionally diagnostic microbiology is based on three approaches namely: culture and microscopy, serology and molecular techniques. Some of these techniques especially those based on the first two were developed well over a hundred years ago and are very tedious, prone to errors and contaminations, complex workflows, and tedious. Most importantly, the cultural methods are only applicable to those organisms that are culturable. In instances where the causative microorganism is not culturable, diagnosis could be misleading. Although, it is well known that molecular techniques such as metagenomics have a number of advantages over the routinely used cultural and serological approaches often used in clinical microbiology, its impact is still not fully felt in diagnostic microbiology. Techniques already in use such as immunoassays and nucleic acid detection are target specific and are not able to detect untargeted agents. Viruses for instance are quite difficult to grow in the laboratory and the application of metagenomics is already been used in a number of viral outbreaks and other cases, and can be applied to just any clinical sample [81].

The application of metagenomics is still in its infancy in parasitology compared to bacteriology. The vector associated with visceral leishamaniasis, *Lutzomyia longipalpis* taxa has been analyzed using cDNA metagenomics. Future applications will certainly include evolution of parasites and insights into the interaction of parasites on gut ecosystems of lower and higher animals which will allow for a better

understanding of their physiology especially the obligate intracellular parasites [81].

In clinical bacteriology, the application of metagenomics has shown great promise in the isolation and characterization of target and non-target microbes directly from samples such as feaces. In a classical study aimed at exploring the diagnostic potentials of metagenomics in stool samples during an outbreak of Shiga-toxigenic *Esherichia coli* also provided a genome wide coverage of the strain genome and that of other non- target pathogens [82].

29. APPLICATION TO HUMAN MICRO-BIOTA

The collection of microorganisms on the human body is called microbiota or normal flora of the body. The different body parts have their characteristics normal flora and those that are transient, and there is a relationship between these organisms and possible infections they can cause [3]. The intestinal tract remains the most densely populated part of the human body with microbial counts reaching $10^{12\text{-}14}$ bacterial cells [3,15]. Such a huge number of cells can only be accessible by culture independent methods. The major phyla include the Actinobacteria, Bacteriodes, Firmucutes, Fusobacteria and Proteobacteria in addition to the Archae, Viruses and Eukaryotes [84]. These microbes play important roles such as development of innate immune system and other beneficial effects on the hosts [3]. The gut is a nutrient rich environment and other roles as rightly observed by Kumar et al. [83], include digestion of polysaccharides such as cellulose, vitamin synthesis and bile acid metabolism. The gut genome is very unstable unlike those of the human body as it can be affected by drugs, diets and even probiotics [16]. The deployment of next generation sequencing to the study of the human gut microbiota has given some insights already into the roles played by microbes in disease and health of humans as seen with type-2 diabetes and inflammatory bowel disease [16].

30. STUDY OF MARINE BIODIVERSITY

The marine ecosystem is very important to a number of industries including the transport, petroleum, food, and biotechnology and microbiology industries. As highlighted by Titilade and Olalekan [85], the importance of marine genomics to life include areas such as clinical diagnostics, agrobiotechnology, marine biology, environmental biotechnology and pharmacogenomics. It is also important in bioprospecting of commercially viable biological products from the marine environment [86]. The marine microbial ecology and their functional role are well reviewed in Alexopoulos et al. [87]. Understanding and preserving ecological diversity is still a concern and challenge despite increasing knowledge of endangered species [88]. Marine ecosystems are among the largest aquatic ecosystems and it includes salt marshes, sea grasses, mangrove forests, oceans, intertidal ecology, estuaries, lagoons, coral reefs, deep seas and its floors [19]. As expected from the large volume of land mass occupied by the marine environment, it is ecologically diverse with all forms of lives ranging from mammals to microbes. Microbes are involved in decomposition, photosynthesis and biogeochemical cycles. Cultural independent methods have identified bacteria, fungi, viruses, actinomycetes and halophiles. Majority of the marine bacteria include members of the order Pseudomonadales, Hyphomicrobia, Chalamydobacetria, Actinomycetes, Cytophagas, Beggiatoas and Spirochetes. The dominant fungi include Myxomycetes, Ascomycetes, Phycomycetes and Basidomycetes. Marine actinomycetes group is dominated by the Streptomyces [19]. Das et al. [89], identified some 32 genera based on chemical composition to include Actinomyces, Actinopolyspora, Micromonospora, Micropolyspora, Nocardia, Rhodococcus, Streptomyces, Streptosporangium and Streptoerticillium. The bacteriophages dominate the virus group and it is estimated that at least 107 viruses exist in one millimeter of sea water [90]. Despite the well known biodiversity and their importance to humans, their biodiversity was poorly described until the arrival of culture-independent techniques. Combination of traditional and molecular techniques was very powerful in the assessment of functional and characteristics of marine and fresh water from India [19]. Using 454 pyrosequencing to analyse marine community that colonize and degrade insoluble polysaccharide *in situ* in the Irish sea, functional gene analysis showed dominance by members of the Gammaproteobacteria and Bacteriodes [91]. A number of extremozymes such as esterase, laccase, merduric reductase, lipase, glycoside and hydrolases have already been identified using sequence based and functional based strategies [92]. Using comparative metagenomics, Eiler et al. [93] have developed an hypothesis based on their findings that states that oceanic dwelling

microbes invest more in the elaboration of functional genes needed for amino acid metabolism and that strategies of carbohydrate metabolism differs significantly amongst fresh water and marine dwellers. Elsewhere using comparative metagenomics and network analyses, a scalable approach has been developed that can test and generate ecological hypotheses that could help answer some ecological questions concerning microbial communities in nature [94]. Ganest et al. [95], using short gun approach and 16S rRNA, to study marine oxygen zones showed that size fraction was a stronger predictor of communities than depth. For more details on recent applications of metagenomics to ocean microbiology, the review of Kerkhof and Goodman [96] is recommended.

31. DRINKING WATER SYSTEMS

Life arguably began in water billions of years ago. Sources of drinking otherwise called potable water varies from country to country. In resource poor settings, potable water in some rural areas could be water obtained directly from ponds, streams, rivers, estuaries and seas without any prior treatment. In cities of some developing and developed countries, potable water is obtained from simple to complex treatment plants and drinking water distribution systems (DWDS). The chlorination step in water treatment is meant to eliminate microbes that may be present in water after the initial steps. Microbial ecology of DWDS systems were commonly done using bulk water sampling with cultural methods such as total heterotrophic counts, total coliform counts using membrane filter and multiple fermentation tubes [97]. Gomez-Alvarez et al. [98], while carrying out pyrosequencing, obtained metagenome data from free-chlorine and monchloramine treated samples that showed mostly proteins of bacterial descent in addition to eukaryotic, archaeal and viral proteins. The complexity as argued by the authors was comparable to that of the distal gut but less than that of the waste water and Sargasso sea. Furthermore, genes associated with multiple disinfectant mechanisms, antibiotics resistance and virulence were also observed. The bacterial species included *Mycobacterium, Legionella* and *Cyanobacterium.* Eukayotes identified also included Amoeba, Ciliate and Slime mold. Specifically, *Mycobacterium* (Actinobacteria), *Acidovorax* (Betaproteo-bacteria), *Burkholderia* (Betaproteobacteria), *Pseudomonas* (Gammaproteobacteria), and *Dechloromonas* (Betaproteobacteria) were dominant in the CHM water, while *Caulobacter* (Alphaproteobacteria), *Rhodopseudomonas* (Alphaproteobacteria), *Synechococcus* (Cyano-bacteria) and *Bradyrhizobium* (Alphaproteo-bacteria) were the most abundant members in chlorinated water. Some of the molecular techniques commonly available for microbial ecology studies and regulatory standards compliance is well reviewed by Douterelo et al. [97]. They include those based on nucleic acid extractions which include DGGE/TGGE, ARDRA, ARISA, SSCP and metagenomics. Those used to assess microbial activity include RT-PCR, microarrays, SIP, MAR-FISH, proteomics and metabolomics. One can only but imagine the diversity that would exist in some of the “potable” being used by dweller of some resource poor settings and the public health implications associated with drinking such waters.

32. STUDY OF SOIL MICROBIAL DIVERSITY

Soil is considered a very complex environment that is regarded as a reservoir of microbial diversity [99,100]. This complexity is understood better when the composition of soil is examined a little closely. Soil is made up of 1-5%, 25%, 25% and 50% of organic matter, water, air and minerals, respectively. It is even amazing to think that soil organisms including microorganism make up just 0.04% of the soil organic matter. The microbial groups in the soil include bacteria, fungi, cyanobacteria, algae, protozoa and viruses. A little fraction of the soil bacteria and fungi have been exploited using traditional microbiology but same cannot be said for other groups of microbes in the soil as they are much more difficult to cultivate than former. Metagenomics has a number of potential applications in soil. The biotechnological and industrial potentials in this diversity can better be tapped using metagenomics. Secondly, it could help answer some of the ecological questions posed by Sutherland et al. [101] and as well shed some light into nutrients cycling in the soil, xenobiotic transformation and the pathways degrading microbes use to transform them. In addition, it will increase our knowledge of the microbial diversity of the soil environment. Soil environment is very heterogeneous in terms of its growth condition, matrix substances and microbial distribution across its different horizons. Metagenomic studies have shown that a gram of soil contains about 3,000 to 11,000 genomes [11] with great biotechnological and industrial implications. Some of the applications of soil

metagenomics as highlighted by Ghazanfar et al. [99] include production of antibiotics, oxidoreductase / dehydrogenase, amidase, vitamins biosynthesis, polysaccharide degrading / modifying enzymes, amylolytic and lipolytic genes.

33. CHALLENGES AND WAY FORWARD OF METAGENOMICS

Metagenomics is still in its infancy and like every other new technology, it has its own challenges. These challenges are becoming clearer with the passage of time and popularity of technique. Some of the challenges as highlighted by the Committee on Metagenomics headed by Handelsman include lack of interdisciplinary collaboration, methodological challenges, data analysis, data archiving to mention a few. Other challenges include the inability to cultivate most of the microbes from environmental samples that metagenomics can be applied to, DNA extraction methods are often not very effective, isolated DNA sometimes get contaminated with various compounds and screening methods varies. It is unlikely that a single methodology will be adopted for DNA extraction from soil, sediment, water, and other samples used for metagenomics. Impediments to usage of metagenomics and other molecular methods include high cost and inadequate bioinformatics data analysis skills amongst most researchers could affect the ability to maximize the multivariate data coming from such analyses.

34. CONCLUSION

Given the enormous applications of metagenomics, there is a need to establish a global metagenomics initiative that would bring together larger, medium and scale metagenomic projects around the globe with a view to making maximum use of the resulting datasets. This will certainly foster and rapidly expand the applications of metagenomics. Traditional microbiology will certainly get better because with the identity of previously uncultivated microbes becoming available, there is a high tendency that we would certainly be seeing 'smarter' media that would cater for such microbes and this will certainly add value to the traditional microbiology.

ACKNOWLEDGEMENT

We acknowledge the financial support by the Office of the Governor of Cross River State, Nigeria that made this review possible.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. A report by Committee on Metagenomics: Challenges and functional applications. The New Science of Metagenomics: Revealing the Secrets of Our Microbial Planet; The National Academies Press, Washington, DC, USA; 2007.
2. Handelsman J. Metagenomics: Application of genomics to uncultured microorganisms. Microbiology and Molecular Biology Reviews. 2004;669–685.
3. Prescott LM, Harley JP, Klein DN. Microbiology. 4th Edition, The McGraw-Hill Companies Inc., New York; 1999.
4. Chikere CB, Okpokwasili GC, Chikere BO. Monitoring of microbial hydrocarbon remediation in the soil. Biotechnology. 2011;1:117–138.
5. Oliveira NM, Niehus R, Foster KR. Evolutionary limits to cooperation microbial communities. Proceedings of the National Academy of Science. 2014;111(50): 17941-17946.
6. Sabree ZL. Rondon MR, Handelsman J. Metagenomics. Genetic, Genomics. 2009;623-631.
7. Tyson GW, Banfield JF. Cultivating the uncultivated: A community genomics perspective. Trends in Microbiology. 2005; 13(9):411-415.
8. Udotong IR, Uko MP, Udotong JIR. Microbial Diversity of a Remote Aviation Fuel Contaminated Sediment of a Lentic Ecosystem in Ibeno, Nigeria. Journal of Environmental and Analytical Toxicology. 2015;5(320):1-7.
9. Escobar-Zepeda A, de Leon AVP, Sanchez-Flores A. The road to metagenomics: From microbiology to DNA sequencing technologies and bioinformatics. Frontiers in Genetics. 2015; 6:348.
10. Felczykowska A, Krajewska AK, Zielinska S, Los JM. Sampling, metadata and DNA extraction- important steps in metagenomics studies. Acta Biochimica Polonica. 2015;62(1):151-160.
11. Torsvik V, Ovreas L. Microbial diversity and function in soil: From genes that ecosystems. Curr Opin Microbiol. 2002;5: 240–245.

12. Singh S, Padavani D, Leslie RA, Chiku T, Banerjee R. Relative contributions of cystathionine (beta) synthesis and (gamma)-cystethionase to H2S biogenesis via alternative transulfuration reactions. Journal of Biological Chemistry. 2009; 284(33):22457-22466.
13. Fakruddin M, Mannan KSB. Methods for analyzing diversity of microbial communities in natural environments. Ceylon Journal of Science (Bio. Sci.). 2013;42(1):19-33.
14. Sogin ML, Morrison HG, Huber JA, Welch DM, Huse SM, Neal PR, Arrieta JM, Herndl GJ. Microbial diversity in the deep sea and the underexplored ''rare biosphere''. Proc. Natl. Acad. Sci. 2006; 103:12115–12120.
15. De Mandal S, Sanga Z, Nachimuthu SK. Metagenomic analysis of bacterial community composition among the cave sediments of Indo-Burman biodiversity hotspot region. Policies and Procedures Pre Print. 2014;2:631.
16. Tremaroli V, Bäckhed F. Functional interactions between the gut microbiota and host metabolism. Nature. 2012; 489(7415):242–249
17. Venter JC, Remington K, Heidelberg JF, Halpern AL, Rusch D, et al Environmental genome shotgun sequencing of the Sargasso Sea. Science. 2004;304:66–74.
18. Novak O, Henykova E, Sairanen I, Kowalczyk M, Pospisil T, Ljung K.. Tissue specific profiling of the Arabidopsis thaliana auxin metabolome. The Plant Journal. 2012;72(3):523-536.
19. Alone AG. Functional and molecular characterization of marine and freshwater microorganisms. An MSc Thesis submitted to the University of Agricultural Sciences, Dharwad; 2012.
20. Antai SP, Asitok AD, Ekpenyong M. Principle of petroleum microbiology: A practical approach. Nature Printers, Calabar; 2015.
21. Holt JG, Krieg NP, Sneath PHA, Staley J, Williams ST. Bergey's manual of determination bacteriology. 9th Edn. Williams and Wilkins Publishers, Baltimore; 1994.
22. Zhang Y, Zheng Y, Hu J, Du N, Chen F. Functional Diversity of the Microbial Community in Healthy Subjects and Periodontitis Patients Based on Sole Carbon Source Utilization. PLOS One. 2014;9(3):e91977.
23. Garland JL, Mills AL. Classification and characterization of heterotrophic microbial communities on the basis of patterns of community-level sole- carbon-source utilization. Applied and Environmental Microbiology. 1991;57(8):2351-2359.
24. Snell JJS, Lapage SP. Carbon Source Utilization Tests as an Aid to the Classification of Non-fermenting Gram-negative Bacteria. Journal of General Microbiology. 1972;749-20.
25. Preston-Mafham J, Boddy L, Randerson PF. Analysis of microbial community functional diversity using sole-carbon-source utilisation profiles - a critique. FEMS Microbiology Ecology. 2002;42: 1-14.
26. Lores M, Gomez-brandon M, Dominguez J. Tracking down microbial communities via fatty acids analysis: Analytical strategy for solid organic samples. Current Research, technology and Educatiuon Topics in Applied Microbiology and Microbial Biotechnology. A. mandez-Vilas (Ed). 2010;1-7.
27. Ibekwe AM, Kennedy AC. Fatty acid methyl ester (FAME) profiles as a tool to investigate community structure of two agricultural soils. Plant Soil. 1999;206: 151-161.
28. Rastogi G, Sani RK. Molecular techniques to assess microbial community structure, function, and dynamics in the environment. Microbes and Microbial Technology. 2011; 1:29–57.
29. Muto A, Osawa S. The guanine and cytosine content of genomic DNA and bacterial evolution. Proc Natl. Acad. Sci. USA. 1987;84(1):166-9.
30. Tiedje JM, Asuming-Brempong S, Nusslein K, Marsh TL, Flynn SJ. Opening the black box of soil microbial diversity. Appl. Soil Ecol. 1999;13:109–122.
31. Theron J, Cloete TE. Molecular techniques for determining microbial diversity and community structure in natural environments. Crit. Rev. Microbiol. 2000; 26:37–57.
32. Nsofor CA. DNA microarray and their applications in medical microbiology. Biotechnology and Molecular Biology Reviews. 2014;9(1):1-11.
33. Panicker G, Call DR, Krug MJ, Bej AK. Detection of Pathogenic *Vibrio* spp. in shellfish by using multiplex PCR and

DNA microarrays. Applied Environmental Microbiology. 2004;5:7436-7444.

34. Woese CR. Bacterial evolution. Microbiological Reviews. 1987;51(2):221-271.
35. Fisher MM, Triplett EW. Automated approach for ribosomal intergenic spacer analysis of microbial diversity and its application to freshwater bacterial communities. Appl. Environ. Microbiol. 1999;65:4630–4636.
36. Muyzer G, de Waal EC, Uitterlinden AG. Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes encoding for 16S rRNA. Appl. Environ. Microbiology. 1993;59:695-700.
37. Mühling M, Woolven-Allen J, Murrell JC, Joint I. Improved group-specific PCR primers for denaturing gradient gel electrophoresis analysis of the genetic diversity of complex microbial communities. ISME. Journal. 2008;2:379–392.
38. Schwieger F, Tebbe CC. A new approach to utilize PCR–single-strand- conformation polymorphism for 16S rRNA gene-based microbial community analysis. Applied and Environmental Microbiology. 1998; 4870–4876.
39. Martinez-Murcia AJ, Acinas SG, Rodriguez-Valera F. Evaluation of prokaryotic diversity by restrictase digestion of 16S rDNA directly amplified from hypersaline environments. FEMS Microbiol. Ecol. 1995;17:247-256.
40. Amann RI, Ludwig W, Schleifer KH. Phylogenetic identification and *in situ* detection of individual microbial cells without cultivation. Microbiol. Rev. 1995; 59:143-169.
41. Uhlik O, Leewis MC, Strejcek M, Musilova L, Mackova M, Leigh MB, Macek T. Stable isotope probing in the metagenomics era: A bridge towards improved bioremediation. Biotechnolgy Adv. 2013;31(2):154–165.
42. Huang WE, Stoecker K, Griffiths R, Newbold L, Daims H, Whiteley AS, Wagner M. Raman–FISH: Combining stable-isotope Raman spectroscopy and fluorescence *in situ* hybridization for the single cell analysis of identity and function. Environ. Microbiol. 2007;9:1878–1889.
43. Handelsman JMR, Rondon SF, Brady JC, Goodman RM. Molecular biological access to the chemistry of unknown soil microbes: A new frontier for natural products. Chem. Biol. 1998;5:R245–R249.
44. Pace NR, Stahl DA, Lane DJ, Olsen GJ. Analyzing natural microbial populations by rRNA sequences. ASM News. 1985;51:4–12.
45. Knight R, Jansson J, Field D, Fierer N, Desai N, Fuhrman JA, Hugeholtz P, van der Lelie D, Meyer F, Stevens R, Bailey MJ, Gordon JI, Kowalchuk GA, Gilbert JA. Unlocking the potential of metagenomics through replicated experimental design. Nature Biotechnology. 2012;30(6):513-520.
46. Thomas T, Gilberts J, Meyer F. Metagenomics - a guide from sampling to data analysis. Microbial Informatics and Experimentation. 2012;2(3):1-12.
47. Wooley JC, Godzik A, Friedberg I. A primer on metagenomics. PLoS Comput. Biol. 2010;6(2):e1000667.
Available:https://doi.org/10.1371/journal.pcbi.1000667
48. Hu Y, Liu Z, Yan J, Qi X, Li J, Zhong S, Yu J, Liu Q. A de velopes DNA extraction method for different soil samples. J. Basic Microbiol. 2010;50:401–407.
49. Delmont TO, Robe P, Clark I, Simonet P, Vogel TM. Metagenomic comparison of direct and indirect soil DNA extraction approaches. J. Microbiol. Methods. 2011; 86(3):397-400.
50. Yuan S, Cohen DB, Ravel J, Abdo Z, Forney LJ. Evaluation of Methods for the Extraction and Purification of DNA from the Human Microbiome. PLoS ONE. 2012; 7(3):e33865.
DOI: 10.1371/journal.pone.0033865
51. Torsvik V. Isolation of bacterial DNA from soil. Soil Biol Biochem. 1980;12:15–21.
52. Lim NYN, Roco CA, frostegard A. Transparaent DNA/RNA co-extraction workflow protocol suitable for inhibitor-Rich Environmnet Samples that focuses on complete DNA removal for transcritptomic analyses. Frontiers in Mcirobiology. 2016;7:1588.
DOI: 10.3389/fmicb.2016.01588
53. Brooks JP, Edwards DJ, Harwich Jr MD, Rivera MC, Fettweis JM, Serrano MG, Reris RA, Sheth NU, Huang B, Girerd P. Vaginal microbiome consortium members, strauss III, JF, Jefferson KK, Buck GA. The truth about metagenomics:

quantifying and counteracting bias in 16S rRNA studies. BMC. Microbiology. 2015; 15:66.

54. Macrae A. The use of 16s rDNA methods in soil microbial ecology. Brazilian Journal of Microbiology. 2000;31:77-82.
55. Joshi M. Deshpande JD. Polymerase chain reaction: Methods, principles and application. International Journal of Biomedical Research. 2010;1(5):81-97.
56. Wu JY, Jiang XT, Jiang YX, Lu SY, Zou F, Zhou HW. Effects of polymerase, template dilution and cycle number on PCR based 16 S rRNA diversity analysis using the deep sequencing method. BMC Microbiology. 2010;10:255.
57. Acinas SG, Sarma-Rupavtarm R, Klepac-Ceraj V, Polz MF. PCR-induced sequence artifacts and bias: Insights from comparison of two 16 S rRNA clone libraries constructed from the same sample. Appl. Environ. Microbiol. 2005; 71(12):8966-8969.
58. Hong SH, Bunge J, Leslin C, Jeon S, Epstein S. (2009). Polymerase chain reaction primers miss half of rRNA microbial diversity. The International Society for Microbial Ecology. 2009;3: 1365–1373.
59. Pinto AJ, Raskin L. PCR biases distort bacterial and archaeal community structure in pyrosequencing datasets. PLoS ONE. 2012;7:43093.
60. Fakruddin and Chowdhury A. Pyrosequencing-an alternative to traditional Sanger sequencing. American Journal of Biochemistry and Biotechnology. 2012;8(1):14-20.
61. Bahassi EM, Stambrook PJ. Next-generation sequencing technologies: Breaking the sound barrier of human genetics. Mutagenesis. 2014;29(5):303–310.
62. Shendure J, Hanlee J. Next generation DNA sequencing. Nature Biotechnology. 2008;26:1135-1145.
63. Kicher M. Kelso J. High-throughput DNA sequencing – concepts and limitations. Bioessays. 2010;32:524–536.
64. Jain M, Olsen HE, Paten B, Akeson M. The Oxford Nanopore MinION: Delivery of nanopore sequencing to the genomics community. Genome Biology. 2016;17: 239.
65. Flusberg BA, Webster D, Lee J, Travers K, Olivares E, Clark Ta, Korlach J, Turner SW. Direct detection of DNA methylation during single-molecule, real-time sequencing. Nat. Methods. 2010;7(6): 461–465.
66. Rusch DB, Halpem Al, Sutton G, Heid elberg KB, Williamson S, Yooseph S, Wu D, et al. The sorcerer II global ocean sampling expedition: Northwest Atlantic through Eastern Tropical Pacific. PLoS Biol. 2007;5(3):e77.
67. Mineta K, Gojobori T. Databases of the marine metagenomics. Gene; 2015. DOI: 10.1016/j.gene.2015.10.035.
68. Wommack KE, Bhavsar J, Polson SW, Chen J, Dumas M, Srinivasiah S, Furman M, Jamindar S, Nasko DJ. VIROME: A standard operating procedure for analysis of viral metagenome sequences. Standards in Genomic Sciences. 2012;6: 427-439.
69. Alex M, Francois B, Guy C, Hubert D, Petra TH, Matthew F, Sebastien P, Simon P, Maxim S, Peter S, Robert DF. EBI metagenomics in 2016 - an expanding and evolving resource for the analysis and archiving of metagenomic data. Nucleic Acids Research. 2016;44:595-603.
70. Kim M, Lee KH, Yoon SW, Kim BS, Chun J, Yi H. Analytical tools and databases for metagenomics in the next-generation sequencing era. Genomics & Informatics. 2013;11(3):102-113.
71. Edet UO, Ebana RUB, Etok CA, Udoidiong VO. Nutrient profile and phytochemical analysis of commercially cultivated oyster mushroom in Calabar, South-South Nigeria. Advances in Research. 2016;7(3): 1-6.
72. Jonathan G, Ayaji A, Oku I, Donbebe W. Nutritive value of common wild edible mushrooms from Southern Nigeria. Global Journal of Biotechnology & Biochemistry. 2006;1(1):16-21.
73. Agarwal PK, Agrawal S, Shrivastava R. Modern molecular approaches for analyzing microbial diversity from mushroom compost ecosystem. Biotechnology. 2015;5:853-866.
74. Agrawal PK. Microbial ecology of compost ecosystem: With special reference to mushroom compost. Journal of Biological and Scientific Opinion. 2014;2(1):45–50.
75. Argrawal R, Schwer B, Shuman S. Structure function analysis and genetic interactions of the Luc7 subunit of the *Saccharomyces cerevisae* U1 snRNP. RNA. Genetics. 2016;22:1320-1328.

76. Vidali M. Bioremediation. An overview. Pure and Applied Chemistry. 2001;73(7): 163–1172.
77. Meerbergen K, Geel MV, Waud M, Willems KA, Dewil R, Impe JV, Appels L, Lievens B. Assessing the composition of microbial communities in textile wastewater treatment plants in comparison with wastewater treatment plants. Microbiology Open. 2017;6:e00413. Available:https:doi.org/10.1002/mbo3.413
78. Antai SP, Mgbomo E. Distribution of hydrocarbon utilizing bacteria in the Ahoada Oil-spilled areas. Microbios Letters. 1989;40:137-143.
79. Antai SP. Biodegradation of Bonny Light crude oil by Bacillus sp and *Pseudomonas* sp. Waste Management. 1990;10:61-64.
80. MacNaughton SJ, Stephen JR, Venosa AD, Davis GA, Chang YJ, White DC. Microbial population changes during bioremediation of an experimental oil spill. Applied and Environmental Microbiology. 1999;65:3566-3574.
81. Pallen MJ. Diagnostic metagenomics: potential applications to bacterial, viral and parasitic infections. Parasitology. 2014; 141:1856-1862.
82. Loman NJ, Constantinidou C, Christner M, Rohde H, Chan JZ, Quick J, Weir JC, Quince C, Smith GP, Betley JR, Aepfelbacher M. Pallen MJ. A culture independent sequence-based metagenomics approach to the investigation of an outbreak of Shiga-toxigenic *Escherichia coli* O104:H4. Journal of the American Medical Association. 2013;309:1502–1510.
83. Kumar M, Babaei P, Ji B, Nielsen J. Human gut microbiota and healthy aging: Recent developments and future prospective. Nutrition and Healthy Aging. 2016;4:3–16.
84. Ji B, Nielsen J. From next-generation sequencing to systematic modeling of the gut microbiome. Frontiers in Genetics. 2015;6(219):219.
85. Titilade PR, Olalekan EI. The importance of marine genomics to life. Journal of Ocean Research. 2015;3(1):1-13.
86. Kodzius R, Gojobari T. metagenomics as a source for bioprospecting. Marine Genomics. 2015;24:21-30.
87. Alexopoulos A, Plessas S, Bezirtzoglou E. Water microbial ecology- An overview. Encyclopedia of Life Sci. 2009;1-24.
88. Medlin LK, Groben R, Valentin K. Molecular tools for the study of marine microbial diversity. In: Doelle HW (ed) Encyclopedia of Life Support Sys-tem. Eolss Publisher, Oxford, UK; 2002.
89. Das S, Lyla PS, Khan SA. Marine microbial diversity and ecology: Importance and future perspectives. Curr. Sci. 2006; 90(10):1325-1335.
90. Wommack K, Colwell RR, Virioplankton: Viruses in aquatic ecosystem. Microbiol. Mol. Biol. Rev. 2000;64:69-114.
91. Edwards JL, Smith DL, Connolly J, McDonald JE, Cox MJ, Jojnt I, Edwards C, McCarthy AJ. Identification of carbohydrate metabolism genes in the metagenome of a marine biofilm community shown to be dominated by gammaproteobacteria and bacteroidetes. Genes. 2010;1:371-384.
92. Barone R, De Santi C, Palma EF, Tedesco P, Galati F, Visone M, Di Scala A, De PD. Marine metagenomics, a valuable tool for enzymes and bioactive compounds discovery. Frontiers in Marine Science. 2014;1:38.
93. Eiler A, Zaremba-Niedzwiedzka K, Martínez-García M, McMahon KD, Stepanauskas R, Andersson SGE, Bertilsson S. Productivity and salinity structuring of the microplankton revealed by comparative freshwater metagenomics. Environmental Micro-biology. 2014;16(9):2682-2698.
94. Hurwitz BL, Westveld AH, Brum JR, Sullivan MB. Modeling ecological drivers in marine viral communities using comparative metagenomics and network analyses. Proceedings of National Academy of Science. 2013;111(29): 10714-10719.
95. Ganesh S, Parris, DJ, Delong EF, Stewart FJ. Metagenomic analysis of size fractionated picoplankton in amrine oxygen minimum zone. The ISME Journal. 2014; 8:187–211.
96. Kerkhof LJ, Goodman RM. Ocean microbial metagenomics. Deep-Sea Research II; 2009. DOI: 10.1016/j.dsr2.2009.05.005.
97. Douterelo I, Boxall JB, Deines P, Sekar R, Fish KE, Biggs CA. Methodological approaches for studying the microbial ecology of drinking water distribution systems. Water Resources. 2014;65:134-156.

98. Gomez-Alvarez V, Revetta RP. Domingo JWS. Metagenomic analyses of drinking water receiving different disinfection treatments. Applied and Environmental Microbiology. 2012;78(17):6095-6102.
99. Ghazanfar S, Azim A, Ghazanfar MA, Anjum MI. Metagenomics and its application in soil microbial community studies: Biotechnological prospects. Journal of Animalia and Plants Science. 2010;6(2):611-622.
100. Rosello MR, Amann R. The species concept for prokaryotes. FEMS. Microboil. Rev. 2001;25:39–67.
101. Sutherland WJ. Identification of 100 fundamental ecological questions. Journal of Ecology. 2013;101:58–67.

Peer-review history:
The peer review history for this paper can be accessed here:
http://sciencedomain.org/review-history/22367