Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

HHS Public Access

Author manuscript

Postdoc J. Author manuscript; available in PMC 2016 December 12.

Published in final edited form as:

Postdoc J. 2016 November ; 4(11): 3–17. doi:10.14304/SURYA.JPR.V4N11.2.

Microbial Ecology: Where are we now?

Lisa A. Boughner, Ph.D.[#1] and **Pallavi Singh, Ph.D.**[#†,2]

[1] Center for Microbial Ecology, Michigan State University, E. Lansing MI 48823

[2]Department of Microbiology and Molecular Genetics, Michigan State University, E. Lansing MI 48823

[#] These authors contributed equally to this work.

Abstract

Conventional microbiological methods have been readily taken over by newer molecular techniques due to the ease of use, reproducibility, sensitivity and speed of working with nucleic acids. These tools allow high throughput analysis of complex and diverse microbial communities, such as those in soil, freshwater, saltwater, or the microbiota living in collaboration with a host organism (plant, mouse, human, etc). For instance, these methods have been robustly used for characterizing the plant (rhizosphere), animal and human microbiome specifically the complex intestinal microbiota. The human body has been referred to as the Superorganism since microbial genes are more numerous than the number of human genes and are essential to the health of the host. In this review we provide an overview of the Next Generation tools currently available to study microbial ecology, along with their limitations and advantages.

Keywords

Microbial ecology; 16S rRNA; NextGen tools; High throughput sequencing

Introduction

Molecular techniques for analysis of microbial ecology and characterization have taken over conventional analytical tools, which were laborious and time consuming. These new approaches provide better insight into the complex interactions taking place within indigenous microbiota of various environments, a few of which include: host and host-pathogens, plants and soil microbiota, aquatic microbes with surrounding water. With the advent of Next-Generation sequencing (NGS) technologies, it has become even faster and more economical to comprehensively evaluate the complexity of microbiota in an environment. This review focuses on these molecular approaches for understanding complex microbial communities and their implications.

Microbiology really began in the late 17th century with Robert Hooke and Antoni van Leeuwenhoek, who were the first to individually publish on the original observations of single celled organisms, or as van Leeuwenhoek referred to them “little animals” (Gest

[†] pallavis@msu.edu.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

2004). Traditionally, microbial populations were solely understood through culture based techniques, of which less than 1% of microbes are culturable, and those that are require numerous physiological and biochemical tests for full characterization. These techniques were very time consuming, laborious, as well as requiring prior knowledge of the organisms of interest to selectively and successfully culture from a complex microbial network.

There are several limitations to culture based techniques, most important of which is a biases towards readily cultivatable organisms, completely overlooking non-culturable bacteria (i.e. 'the great plate count anomaly'), the largest proportion of microbes (Staley and Konopka 1985). To some extent immunological based methods for characterization and identification of bacteria, including enzyme linked immunosorbent assay (ELISA) requiring specific antibodies, can help circumvent these problems but are still limited in assessing functionality and have not been extensively used due to the need for specific antigens/antibodies (Ricke and Pillai 1999). Consequently, focus was turned towards development of molecular approaches for understanding these complex communities, eliminating the need for culturing beforehand.

Dawn of a new era began with the innovation of NGS technologies, which revolutionized the existing gold standard techniques of microbial community analysis. NGS enables the generation of massive sequencing results in parallel utilizing phylogenetic markers, such as 16S rRNA gene, inexpensively and in considerably less time. The generated high throughput results not only enable comprehensive analysis of community microbiota from various environments (for instance, the soil, intestine, fresh/saltwater etc.) but also insights into their interactions with the environmental system. There are endless possibilities for applying NGS technologies. Sequencing also finds its utility in profiling the microbial community of many complex environments, and has been extensively reviewed for performing such projects (Hamady and Knight 2009). Herein we include a brief overview of the molecular techniques readily used in environmental microbiological investigations, along with their various advantages and limitations. We initially begin with molecular approaches on bacterial characterization (16S rRNA, DGGE, T-RFLP) and a section on typing bacterial pathogens. Followed by the various NGS technologies currently used (targeted amplification, metagenomics, RNA-seq, proteomics, Tn-seq). Included in each section are selected examples where the individual research groups aim at identifying key species (or species specific genes) responsible for a dynamic state within a niche. The tools reviewed here are being readily implemented and provide a comprehensive overview of microbial ecology today.

Unravelling Microbial Ecology

Current trends in understanding complex microbial communities are focused on molecular approaches that not only provide unbiased, comprehensive information but are also less labor intensive and can be accomplished in a shorter duration of time. There are several molecular approaches that were developed for classification of complex microbial communities and provide a clearer understanding of the microbe's ecology and interactions when compared to conventional culturing methods.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Several recombinant DNA techniques based on 16S rRNA [originally proposed by Carl Woese (1987)], chaperonin-60, *rpoB* sequences etc., for unbiased characterization of mixed microbial communities have been developed (Hill et al. 2002; Theron and Cloete 2000) and extensively used. For instance, a study on cohort of nonpregnant, healthy women to assess their vaginal microflora based on chaperonin-60 sequencing revealed that the dominant species found there were *Lactobacillus* spp (Hill et al. 2005). Also, a study on the intestinal tract of market age broilers based on 16S rRNA gene sequencing revealed a shift towards higher ratio of Firmicutes to Bacteroidetes upon administration of Penicillin (Singh et al. 2013b).

16S rRNA gene sequencing

Microbial community analysis based on 16S rRNA gene sequencing is a rapid and reliable tool that is being readily used for comprehensive analysis of microbial communities within all kinds of niches, be it deep sea sub-surface (Kimura et al. 2007) or an estuary (Bernhard et al. 2005). 16S rRNA is an excellent phylogenetic marker that is conserved across all prokaryotic species and is more resistant to mutations as compared to other conserved genes, however some regions called hot spots are different in all species (Clarridge III 2004). An important note: while it is unlikely for 16S genes to be horizontally transferred (horizontal gene transfer, HGT), such a thing is not impossible (Kitahara and Miyazaki 2013; Tian et al. 2015). Universal microbial primers (for both bacteria and archaea) have been designed to complement the constant while amplifying the variable regions of the ~1500bp 16S rRNA genes, which may then be sequenced for phylogenetic analysis.

Quantitative polymerase chain reaction (qPCR) has also been readily used for complementation and verification of 16S rRNA gene sequencing data, as studies have shown sequencing biases and variability based on the protocol used (Hang et al. 2014). Recent advancements in sequencing platforms, from obsolete 454 pyrosequencing to short read based Illumina, has made use of 16S rRNA gene sequencing. More recently Illumina based sequencing for microbial community analysis has been the robust and economical sequencing platform of choice. The Human Microbiome Project Consortium studied the complex microbial communities from various sites of the human body, including the gut, skin and vagina, making use of the Illumina GAIIx platform (The Human Microbiome Project Consortium 2012). Unfortunately, 16S rRNA gene sequencing is not infallible and sometimes exhibits low resolution due to sequence similarities even when exhibiting different phenotypes. An alternative has been suggested by utilizing 16S rDNA sequencing in combination with DNA-DNA hybridization techniques, like microarray (Janda and Abbott 2007). Several open source software tools have been consequently developed to efficiently analyze the massive amount of data being generated. The most commonly used software tools include Mothur (Schloss et al. 2009) and QIIME (Caporaso et al. 2010) which include steps for quality filtering the data sets to purge ambiguous and sequencing artifacts. Chimera are a major sequencing artifact, and both Mothur and QIIME include a chimera checker (Edgar et al. 2011), to eliminate such issues from the sequencing data before clustering and phylogenetic analyses. Phylogenetic classification is carried out with the help of well-maintained and periodically updated 16S rRNA sequence databases, including the green genes (http://greengenes.lbl.gov/) and the Ribosomal Database Project (Cole et al.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

2014). Further downstream statistical analysis tools are also freely available from various sources, which include but are not limited to: LEfSe for identifying differentially abundant microbial features (Segata et al. 2011), Picrust for metagenomic prediction (Langille et al. 2013) and R tools for generating heat maps and graphs.

Denaturing Gradient Gel Electrophoresis (DGGE)

DGGE is an important tool for studying complex microbial communities as well as their shift over time (Muyzer and Smalla 1998). Community profiling by DGGE is recommended as an important clinical tool that depends on the differences between the genetic sequences of various bacterial communities, thereby generating different banding patterns based on denaturation and electrophoretic mobility of the denatured DNA (Ariefdjohan et al. 2010). DNA sequences amplified by PCR are chemically denatured as the amplified product runs on polyacrylamide gel with increasing concentrations of denaturants. In addition to being less laborious, DGGE is also considered superior to sequencing with cloning techniques, and has been demonstrated by denaturation analysis of 16S rRNA genes as a useful diagnostic tool for microorganisms including bacteria, fungi and yeast (Muyzer et al. 1993). Even though DGGE has been shown incapable of detecting two or three base variations, the addition of a GC clamp increased the methods ability to detect most single base substitutions for DNA fragments ranging from 25-500bp (Myers et al. 1985). Soil microbial community DGGE analysis using 16S rRNA gene amplification products was found to be efficient in resolving and characterizing a bacterial community (Nakatsu et al. 2000). In another study, DGGE was implemented on a large scale to elucidate marine microbial ecology and its spatial and temporal variations. Significant difference in bacterial community compositions were identified within a relatively small geographic distance (Singh and Ramaiah 2011).

Terminal Restriction Length Polymorphism (T-RFLP)

T-RFLP allows for reproducible profiling of a microbial community, as well as its changing dynamics in response to environmental factors. Terminal end fluorescent labeling of 16S rRNA PCR products are digested by a 4bp cutter enzyme and run on an automated sequencer, leading to differentiating peaks that can then be analyzed and generate a phylogenetic pattern. Extensive examination of gut microbial composition revealed a heterogeneity between bacterial populations of ceca and ileum in broiler chickens by T-RFLP analysis of 16S rRNA genes (Gong et al. 2002). Similar analyses on human intestinal microbiota exhibited approximately 10 phylogenetic bacterial groups (Nagashima et al. 2006). Limitations of this technique include generation of ‘pseudo-T-RFs’ (terminal restriction fragments), generated from undesired enzymatic digestion of secondary single stranded molecular structures. Formation of ‘pseudo-T-RFs’ result in additional peaks that may lead to over estimation of microbial diversity and the need to generate clone libraries due to inaccessibility of sequences (Egert and Friedrich 2003).

Typing of Bacterial Pathogens

The need to characterize and differentiate bacterial pathogens has resulted in various identification schemes to understand their ecology and epidemiology. Typing systems are evaluated in terms of reliability, reproducibility, discriminatory power, as well as data

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

interpretation. The method of choice also depends upon ease of use, time taken and cost involved (Foxman et al. 2005). Perhaps the most important application of identification is during an infectious outbreak by providing insight on the spread of infection (Pitt 1994). Conventional typing methods are phenotypic based and include serotyping, biotyping, bacteriocin typing, phage typing. Even though these gold standard techniques are laborious and time consuming, they have their benefits and can be used depending on the level of discrimination desired, as a standalone technique or in combination with molecular techniques. Moreover, the advent of DNA based techniques were based on the need to overcome shortcomings of conventional phenotypic typing methods.

Genotyping techniques based on total DNA sequences can be divided into two categories, one that is based on enzyme digestion patterns [ex. restriction enzyme analysis (REA) or Pulse Field Gel Electrophoresis (PFGE)] and a second category based on southern hybridization technique (ex. Ribotyping). Ribotyping was shown to have a higher discriminatory power over a combination of conventional methods, including biochemical tests, serotyping, phage-typing and LPS profiling in *Enterobacter cloacae* [a nosocomial pathogen, isolated from patients (Garaizar et al. 1991)].

PCR based techniques include amplified fragment length polymorphisms, random amplified polymorphic DNA (RAPD), PCR-RFLP and sequencing based on multi locus sequence typing (MLST). Menon et al. (2003) identified RAPD as a quick typing technique with high reproducibility and ease of use, based on an electrophoretic pattern generated for *Pseudomonas aeruginosa* in contrast to a set of conventional methods. Given that the discriminatory power of RAPD could not be compared due to a lack of correlation, Menon et al. (2003) suggest an advantage to combine more than one technique for typing purposes. This was supported by a study conducted to evaluate PCR based molecular typing techniques like RAPD, PCR fingerprinting and REA to assess the genetic diversity in *Saccharomyces cerevisiae*, and indicated the need to combine more than one molecular typing strategy to achieve a greater degree of discrimination (Couto et al. 1996). *Salmonella* species were shown to exhibit higher discrimination with more than one technique, like that of a PFGE pattern generated by XbaI and antibiotic susceptibility tests (performed for typing and source tracking) (Nayak and Stewart-King 2008). As well as a combination of enterobacterial repetitive intergenic consensus (ERIC) fingerprinting and RAPD (Lim et al. 2005).

The importance of combing typing methods has been stressed by many researchers to enhance organismal discrimination, as demonstrated for 1) *Salmonella* Newport where a combination of MLST, PFGE and antimicrobial susceptibility (Harbottle et al. 2006) and 2) *Salmonella* Typhimurium where a combination of MLST, PFGE and Rep-PCR was evident (Foley et al. 2006) as well as 3) *Candida albicans* (Chen et al. 2005).

Furthermore, Whole Genome Sequencing (WGS) provides more characterization resolution wherein genome wide information of a bacterial pathogen is made available. WGS has been implemented in conjunction with epidemiological data to unravel transmission source. For instance, a study conducted on carbapenem-resistant *Klebsiella pneumonia* linked

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

dissemination of this pathogen to a single patient providing insight into nosocomial infections based on NGS technologies (Snitkin et al. 2012).

Efficient strategies for NGS technologies

Development of NGS technologies has enabled a great number of genomic facilities with varied applications for large scale sequencing by drastically increasing the capacity of high throughput sequencing in a time and cost effective manner. To completely utilize the potential of NGS technologies, enrichment techniques capable of large scale multiplexing are required. Even though as compared to conventional sequencing technologies, the cost of large scale sequencing has drastically reduced (Stein 2010), it is still very expensive for WGS of complex organisms, such as humans. For this reason, emphasis is being placed on NGS that allows for selective amplification of targeted genes relevant to individual sequencing projects.

Targeted Amplicon Sequencing

A primer based method by which a gene or region of interest is sequenced. Target enrichment can be done in two ways, hybridization based and amplification based methods. Hybridization based methods rely on sequence-specific nucleic acid hybridization in both a solid phase built on microarray based enrichment (Summerer et al. 2009) and a solution phase including molecular inversion probes; both methods have been elaborately reviewed (Fan et al. 2006; Mamanova et al. 2010; Summerer 2009). Tewhey *et. al.* has developed a hybridization based sequence enrichment technique in solution phase, that can selectively capture and enrich for a 3.9Mb sequence of the human genome and has been reported 99% accurate for SNP (single nucleotide polymorphism) calling (Tewhey et al. 2009). Amplification based methods like nested patch PCR for multiplexing (Varley and Mitra 2008), circular selectors (Dahl et al. 2005) and megaplex PCR (Meuzelaar et al. 2007), have been developed to efficiently and selectively capture target DNA, using either primers or probes each of which have been demonstrated to have unique NGS application. Target enrichment techniques are evaluated in terms of multiplexity, specificity, uniformity, drop-off rate and reproducibility. The above-mentioned techniques were all evaluated using these key parameters (Ting et al. 2009). Senapathy et al. developed an enrichment technique based on fixed-randomization that binds multiple genomic sites with similar sequence complementarity, called Functional Genomic Fingerprinting (FGF) and has been used to enrich functional genomic regions in humans (Senapathy et al. 2010). One important aspect of multiplexing by target enrichment is normalization of the amplicon pool, where many samples are assessed and barcoded for increased performance and efficiency, as well as generation of a complete microbial profile (Harris et al. 2010). Similarly, in another study, enhanced bacterial characterization technique (MLST-seq) for foodborne pathogen utilizing *Salmonella* as a model system, was demonstrated. Novel target enrichment strategies that selectively amplified housekeeping and virulence genes of *Salmonella* for NGS sequencing in a high-throughput manner were efficiently developed (Singh et al. 2011; Singh et al. 2012, 2013a. Comprehensive reviews of several other applications in the fields of functional genomics (Olena and Marra 2008), clinical diagnostics (Bosch and Grody 2008), ancient DNA (Knapp and Hofreiter 2010), and transcriptomic characterization (Wall et al. 2009)

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

have been assessed. Numerous other applications could be performed, like studying host-pathogen interactions or carrying out genomic studies for drug discoveries.

A major advantage of NGS is the elimination of vector based sequencing that is likely to be biased, laborious, time consuming and would not generate a complete profile of the microbial community. One such study conducted on cecal microbiome of chickens, demonstrated a role of mobile DNA elements in microbial horizontal gene transfer and evolution of a functional microbiome conferring resistance to zoonotic organisms in the host (Qu et al. 2008). Phylogenetic markers like *cpn60* are also being explored and utilized for metagenomic analysis (Schellenberg et al. 2009).

Other areas of research and interest include transcriptome analysis, cancer genetics, SNP analysis, mutation discoveries etc. A few limitations to this technology exist since it cannot be applied to research requiring minimal sequencing. Relatively speaking, NGS is an inexpensive technology when comparing its data output with Sanger's sequencing. Another limitation of Illumina technology is the generation of shorter reads with an average of ~300bp, this technology is now being replaced by new improved kits, however, still requires bidirectional sequencing to cover amplicons over 600bp in length.

Metagenomics

A method by which all the DNA extracted from an environmental sample is sequenced, allowing for a comprehensive analysis of microbial communities. Metagenomics, or shotgun sequencing, is dependent upon the method of DNA obtainment. A poor method of DNA removal will lead to false interpretations of the microbiota in, and the functional profile of, an environment. DNA may be extracted through either indirect or direct approaches. Indirect DNA extractions consist of removing cells from an environmental sample, followed by extracting DNA from the separated cells; and direct DNA extractions, involve *in situ* cell lysis and subsequent removal of DNA from within an environmental sample (Williamson et al. 2011). There are many approaches to choose from for direct DNA extraction, it is important to consider sample characteristics before determining the ideal method. For example, there are many substances in soil that bind nucleic acids or inhibit downstream applications; such as humic acids (inhibitors of PCR) and clay (binds nucleic acids) (Nair et al. 2014; Yankson and Steck 2009). Regardless, extracted DNA is fragmented to about 150bp in size, these fragments are then end labeled with barcodes for downstream sample identification, enabling multi-sample sequencing. Fully sequencing the DNA within a sample allows researchers to examine the taxonomic diversity of all microbes, not just the culturable. However, the degree to which the diversity can be determined, depends on the depth of sequencing, i.e. how much sequencing data is obtained from a sample (Sims et al. 2014). The more abundant an organism in the environment, the more DNA extracted and the more sequencing data obtained. Greater sequencing depth is necessary to obtain information on the least abundant microorganisms. An important note is that the DNA obtained from a sample will depend on the method of extraction, as well as the sample characteristics from which the DNA is extracted (Sagova-Mareckova et al. 2008).

Upon extracting the DNA and sequencing the samples, one must then analyze the data. The obtained sequencing reads are sorted according to the barcodes, poor quality reads are

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

removed, then the remaining quality reads are assembled, annotated, and analyzed for taxonomic and/or functional diversity (i.e. who's there and/or what are they doing?) (Albertsen et al. 2013; Sharpton 2014; Thomas et al. 2012). Each stage is computationally laborious and there are many programs available to assist in each of the steps (Howe et al. 2014; Miller et al. 2010; Wang et al. 2015). Fierer et al. (2012) investigated the impact of nitrogen (N) fertilization on terrestrial ecosystems utilizing various approaches, including shotgun sequencing, and indicate that anthropogenic inputs of N are favoring copiotrophic (fast growing) microbes.

Transcriptomics

In addition to metagenomics, there are other similar ‘omic's to assist in broadly examining the taxonomic and/or functional diversity of environments. Transcriptomics, or RNA sequencing (RNA-seq), allows for the examination of RNA present at the time of extraction (i.e. which genes were considered necessary at that time?) (Wang et al. 2009). As noted above for DNA extraction, the method of nucleic acid removal will inform the results. Additional consideration is required when working with RNA since it is a short-lived molecule, and not to mention that RNases are everywhere and are difficult to avoid (Dineen et al. 2010; Saleh-Lakha et al. 2005). Also, during the extraction process, all RNA is obtained but usually, only mRNA is of interest, therefore before beginning with sequence analysis of prokaryotic RNA, mRNA (no poly-A tail) needs to be ‘selected’ above other RNA molecules (tRNA, rRNA) (Giannoukos et al. 2012).

Once the additional considerations when working with RNA are accounted for, the downstream analysis are quite similar to those discussed for metagenomics. For example, RNA-seq was recently used to examine the differential expression of genes from *Sphingomonas wittichii* RW1 [debatably the only organism in isolation that is able to utilize the toxin dibenzo-*p*-dioxin (DD) as its sole carbon and energy source] in the presence of either DD, dibenzofuran (DF; a similar molecule with lower toxicity), or succinate (SUC; negative control carbon source) with and without clay particles (a prominent soil component) (Chai et al. 2016). Such investigations can assist in understanding why dibenzo-*p*-dioxin is so resistant to biodegradation, whereas dibenzofuran biodegradation is relatively common (toxicity, molecular structure, or something else entirely?).

Proteomics

Similar to transcriptomics, there is another method for examining the functional profile of a microbial community or organism of interest. However, instead of looking at which genes were transcribed, proteomics examines the proteins that have actually been produced (Simon and Daniel 2011). An important distinction, since not all amplified mRNA leads to synthesized protein (Ingolia 2014). Similar to metagenomics and transcriptomics, the method of extracting proteins from a sample requires due consideration (Keiblinger et al. 2012; Wang et al. 2006). One of the main issues is to successfully recover a representative profile from samples with various characteristics, the challenges of soil proteomics is reviewed by Renella et al. (2014). The purified protein samples are then run on a mass spectrophotometer to determine the sequences of proteins present (Armengaud 2013). An alternative is to run the purified protein samples on two-dimensional polyacrylamide gel

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

electrophoresis (2D-PAGE), but this method will only provide a protein profile for comparison between samples, not the specific proteins present. Toueille et al. (2012) used proteomics to comparatively analyze the nucleoid associated proteins (NAPs) from *Deinococcus* spp, with the goal of identifying whether novel NAPs play a role in conferring radiation stability to *Deinococcus* organisms.

Transposon Sequencing

Transposon sequencing (Tn-seq), is a method that can be utilized to examine gene functions and/or genes of essentiality within an organism under varying growth conditions. Before beginning a Tn-seq experiment, there are numerous considerations to be made, much like with any method (Kwon et al. 2016). Briefly, Tn-seq involves a mutating transposable element that is introduced into a dense culture of an organism, ideally resulting in a collection of mutants where under such growth conditions, all non-essential genes are interrupted (i.e. transposon library for that organism) (Opijnen et al. 2009). Subsequently, the transposon library can then be grown up under various conditions of interest, as well as the starting/negative control condition (which provides the background essential genes). The DNA is then extracted and sequenced (targeting the region where the transposable elements were inserted), genes with a transposable element are not essential for that organism's growth under that condition (i.e. the cell grew without that gene). Through Tn-seq analyses, Pechter et al. (2016) examined the essential aerobic heterotrophic genes of *Rhodopseudomonas palustris*, an organism that is able to fluctuate between anaerobic photoautotrophic growth and aerobic heterotrophy. Obtaining a better understanding of the metabolic capabilities of organisms such as *R. palustris* is a most worthwhile endeavor. A tool such as Tn-seq is useful when working to understand the metabolic diversity of microbes: which genes enable an organism to metabolize a compound, but not others? What are the functional differences between two similar organisms with differing metabolic capabilities? Is the ability conferred by metabolic pathway enzymes, or through reduced stress/toxicity affects, or something else entirely? Tn-seq is an increasingly utilized method to address these and many other questions.

Conclusion

Current approaches to microbial ecology have come a long way in the past few centuries, enabling a broader and more thorough investigation into the small, single celled organisms on which we all depend. NGS technologies have been utilized for high throughput screening and have not only become more efficient and accessible tools for studying complex microbial communities, but also provide better insight into the interactions taking place within complex microbial communities. NGS technologies have aided in determining the core factors responsible for the dynamic nature of these complex communities and their interaction with host and native environments. One can only wonder what new advances, and wondrous discoveries will come tomorrow.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Acknowledgements

Salary support for Dr. Pallavi Singh was provided NIH Enterics Research Investigational Network Cooperative Research Center at Michigan State University (U19AI090872) and in part by the USDA NIFA (2011-67005-30004) awarded to Dr. Shannon D. Manning, MSU foundation Associate Professor.

References

Albertsen M, et al. Genome sequences of rare, uncultured bacteria obtained by differential coverage binning of multiple metagenomes. Nature Biotechnology. 2013; 31(6):533. http://dx.doi.org/10.1038/nbt.2579 PMid:23707974.

The Human Microbiome Project Consortium. Structure, function and diversity of the healthy human microbiome. Nature. 2012; 486(7402):207–14. http://dx.doi.org/10.1038/nature11234 PMid:22699609 PMCid:PMC3564958. [PubMed: 22699609]

Ariefdjohan MW, Savaiano DA, Nakatsu CH. Comparison of DNA extraction kits for PCR-DGGE analysis of human intestinal microbial communities from fecal specimens. Nutrition Journal. 2010; 9(23) http://dx.doi.org/10.1186/1475-2891-9-23.

Armengaud J. Microbiology and proteomics, getting the best of both worlds! Environ Microbiol. 2013; 15(1):12–23. http://dx.doi.org/10.1111/j.1462-2920.2012.02811.x PMid:22708953. [PubMed: 22708953]

Bernhard AE, et al. Microbial community dynamics based on 16S rRNA gene profiles in a Pacific Northwest estuary and its tributaries. FEMS Microbiology Ecology. 2005; 52:115–28. http://dx.doi.org/10.1016/j.femsec.2004.10.016 PMid:16329898. [PubMed: 16329898]

Bosch JR, ten and Grody WW. Keeping Up With the Next Generation: Massively Parallel Sequencing in Clinical Diagnostics. Journal of Molecular Diagnostics. 2008; 10(6):484–92. http://dx.doi.org/10.2353/jmoldx.2008.080027 PMid:18832462 PMCid:PMC2570630. [PubMed: 18832462]

Caporaso JG, et al. QIIME allows analysis of high-throughput community sequencing data. Nat Methods. 2010; 7(5):335–6. http://dx.doi.org/10.1038/nmeth.f.303 PMid:20383131 PMCid:PMC3156573. [PubMed: 20383131]

Chai B, et al. Sphingomonas wittichii Strain RW1 Genome-Wide Gene Expression Shifts in Response to Dioxins and Clay. PLoS ONE. 2016; 11(6):e0157008. http://dx.doi.org/10.1371/journal.pone.0157008 PMid:27309357 PMCid:PMC4911050. [PubMed: 27309357]

Chen KW, et al. Comparison of four molecular typing methods to assess genetic relatedness of Candida albicans clinical isolates in Taiwan. Journal of Medical Microbiology. 2005; 54:249–58. http://dx.doi.org/10.1099/jmm.0.45829-0 PMid:15713608. [PubMed: 15713608]

Clarridge III JE. Impact of 16S rRNA Gene Sequence Analysis for identification of Bacteria on Clinical Microbiology and Infectious diseases. Clinical Microbiology Reviews. 2004; 17(4):840–62. http://dx.doi.org/10.1128/CMR.17.4.840-862.2004 PMid:15489351 PMCid:PMC523561. [PubMed: 15489351]

Cole JR, et al. Ribosomal Database Project: data and tools for high throughput rRNA analysis. Nucleic Acids Res. 2014; 42(Database issue):D633–42. http://dx.doi.org/10.1093/nar/gkt1244 PMid:24288368 PMCid:PMC3965039. [PubMed: 24288368]

Couto MMB, et al. Evaluation of Molecular Typing Techniques To Assign Genetic Diversity among Saccharomyces cerevisiae Strains. APPLIED AND ENVIRONMENTAL MICROBIOLOGY. 1996; 62(1):41–46. [PubMed: 8572710]

Dahl F, et al. Multiplex amplification enables by selective circularization of large sets of genomic DNA fragments. Nucleic Acids Research. 2005; 33(8) http://dx.doi.org/10.1093/nar/gni070.

Dineen SM, et al. Evaluation of commercial RNA extraction kits for the isolation of viral MS2 RNA from soil. Journal of Virological Methods. 2010; 168(1–2):44–50. http://dx.doi.org/10.1016/j.jviromet.2010.04.014 PMid:20417664. [PubMed: 20417664]

Edgar RC, et al. UCHIME improves sensitivity and speed of chimera detection. Bioinformatics. 2011; 27(16):2194–200. http://dx.doi.org/10.1093/bioinformatics/btr381 PMid:21700674 PMCid:PMC3150044. [PubMed: 21700674]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Egert M, Friedrich MW. Formation of Pseudo-Terminal Restriction Fragments, a PCR-Related Bias Affecting Terminal Restriction Fragment Length Polymorphism Analysis of Microbial Community Structure. Applied and Environmental Microbiology. 2003; 69(5):2555–62. http://dx.doi.org/10.1128/AEM.69.5.2555-2562.2003 PMCid:PMC154551. [PubMed: 12732521]

Fan JB, Chee MS, Gunderson KL. Highly parallel genomic assays. Nature Reviews Genetics. 2006; 7 http://dx.doi.org/10.1038/nrg1901.

Fierer N, et al. Comparative metagenomic, phylogenetic and physiological analyses of soil microbial communities across nitrogen gradients. Isme Journal. 2012; 6(5):1007–17. http://dx.doi.org/10.1038/ismej.2011.159 PMid:22134642 PMCid:PMC3329107. [PubMed: 22134642]

Foley SL, et al. Comparison of Subtyping Methods for Differentiating Salmonella enterica Serovar Typhimurium Isolates Obtained from Food Animal Sources. Journal of Clinical Microbiology. 2006; 44(10) http://dx.doi.org/10.1128/JCM.00745-06.

Foxman B, et al. Choosing an appropriate bacterial typing technique for epidemiologic studies. Epidemiologic Perspectives & Innovations. 2005; 2(10) PMid:16309556 PMCid:PMC1308839.

Garaizar J, Kaufmann ME, Pitt TL. Comparison of Ribotyping with Conventional Methods for the Type Identification of Enterobacter cloacae. Journal of Clinical Microbiology. 1991; 29(7):1303–07. PMid:1715880 PMCid:PMC270105. [PubMed: 1715880]

Gest H. The discovery of microorganisms by Robert Hooke and Antoni Van Leeuwenhoek, fellows of the Royal Society. Notes and Records of the Royal Society of London. 2004; 58(2):187–201. http://dx.doi.org/10.1098/rsnr.2004.0055 PMid:15209075. [PubMed: 15209075]

Giannoukos G, et al. Efficient and robust RNA-seq process for cultured bacteria and complex community transcriptomes. Genome Biology. 2012; 13(3):r23. http://dx.doi.org/10.1186/gb-2012-13-3-r23 PMid:22455878 PMCid:PMC3439974. [PubMed: 22455878]

Gong J, et al. Molecular analysis of bacterial populations in the ileum of broiler chickens and comparison with bacteria in the cecum. FEMS Microbiology Ecology. 2002; 41:171–79. http://dx.doi.org/10.1111/j.1574-6941.2002.tb00978.x PMid:19709251. [PubMed: 19709251]

Hamady M, Knight R. Microbial community profiling for human microbiome projects: Tools, techniques, and challenges. Genome Res. 2009; 19:1141–52. http://dx.doi.org/10.1101/gr.085464.108 PMid:19383763 PMCid:PMC3776646. [PubMed: 19383763]

Hang J, et al. 16S rRNA gene pyrosequencing of reference and clinical samples and investigation of the temperature stability of microbiome profiles. Microbiome. 2014; 2:31. http://dx.doi.org/10.1186/2049-2618-2-31 PMid:25228989 PMCid:PMC4165438. [PubMed: 25228989]

Harbottle H, et al. Comparison of Multilocus Sequence Typing, Pulsed-Field Gel Electrophoresis, and Antimicrobial Susceptibility Typing for Characterization of Salmonella enterica Serotype Newport Isolates. Journal of Clinical Microbiology. 2006; 44(7):2449–57. http://dx.doi.org/10.1128/JCM.00019-06 PMid:16825363 PMCid:PMC1489510. [PubMed: 16825363]

Harris JK, et al. Comparison of Normalization Methods for Construction of Large, Multiplex Amplicon Pools for Next-Generation Sequencing. Applied and Environmental Microbiology. 2010; 76(12):3863–68. http://dx.doi.org/10.1128/AEM.02585-09 PMid:20418443 PMCid:PMC2893486. [PubMed: 20418443]

Hill JE, et al. Characterization of vaginal microflora of healthy, nonpregnant women by chaperonin-60 sequence-based methods. Am J Obstet Gynecol. 2005; 193(3 Pt 1):682–92. http://dx.doi.org/10.1016/j.ajog.2005.02.094 PMid:16150261. [PubMed: 16150261]

Hill JE, et al. Extensive Profiling of a Complex Microbial Community by High-Throughput Sequencing. Applied and Environmental Microbiology. 2002; 68(6):3055–66. http://dx.doi.org/10.1128/AEM.68.6.3055-3066.2002 PMid:12039767 PMCid:PMC123911. [PubMed: 12039767]

Howe, AC., et al. Tackling soil diversity with the assembly of large, complex metagenomes. Proceedings of the National Academy of Sciences. 2014. http://dx.doi.org/10.1073/pnas.1402564111

Ingolia NT. Ribosome profiling: new views of translation, from single codons to genome scale APPLICATIONS OF NEXT-GENERATION SEQUENCING - INNOVATION. Nature Reviews Genetics. 2014; 15(3):205–13. http://dx.doi.org/10.1038/nrg3645 PMid:24468696.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Janda JM, Abbott SL. 16S rRNA Gene Sequencing for Bacterial Identification in the Diagnostic Laboratory: Pluses, Perils, and Pitfalls. Journal of Clinical Microbiology. 2007; 45(9):2761–64. http://dx.doi.org/10.1128/JCM.01228-07 PMid:17626177 PMCid:PMC2045242. [PubMed: 17626177]

Keiblinger KM, et al. Soil metaproteomics – Comparative evaluation of protein extraction protocols. Soil Biology & Biochemistry. 2012; 54(15-10):14–24. http://dx.doi.org/10.1016/j.soilbio.2012.05.014 PMid:23125465 PMCid:PMC3413887. [PubMed: 23125465]

Kimura H, et al. Selective Phylogenetic Analysis Targeting 16S rRNA Genes of Hyperthermophilic Archaea in the Deep-Subsurface Hot Biosphere. Applied and Environmental Microbiology. 2007; 73(7):2110–17. http://dx.doi.org/10.1128/AEM.02800-06 PMid:17277228 PMCid:PMC1855651. [PubMed: 17277228]

Knapp M, Hofreiter M. Next Generation Sequencing of Ancient DNA: Requirements, Strategies and Perspectives. Genes. 2010; 1:227–43. http://dx.doi.org/10.3390/genes1020227 PMid:24710043 PMCid:PMC3954087. [PubMed: 24710043]

Kwon YM, Ricke SC, Mandal RK. Transposon sequencing: methods and expanding applications. Applied Microbiology and Biotechnology. 2016; 100(1):31–43. http://dx.doi.org/10.1007/s00253-015-7037-8 PMid:26476650. [PubMed: 26476650]

Langille MGI, et al. Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. Nature Biotechnology. 2013; 31(9):814. http://dx.doi.org/10.1038/nbt.2676 PMid:23975157 PMCid:PMC3819121.

Lim H, et al. Comparison of four molecular typing methods for the differentiation of Salmonella spp. International Journal of Food Microbiology. 2005; 105:411–18. http://dx.doi.org/10.1016/j.ijfoodmicro.2005.03.019 PMid:16139380. [PubMed: 16139380]

Mamanova L, et al. Target-enrichment strategies for next-generation sequencing. Nature Methods. 2010; 7(2) http://dx.doi.org/10.1038/nmeth.1419 PMCid:PMC2861772.

Menon PK, et al. Random Amplification of Polymorphic DNA Based Typing of Pseudomonas aeruginosa. Medical journal Armed Forces India. 2003; 59(1):25–28. http://dx.doi.org/10.1016/S0377-1237(03)80099-0.

Meuzelaar LS, et al. MegaPlex PCR: a strategy for multiplex amplification. Nature Methods. 2007; 4(10):835–37. http://dx.doi.org/10.1038/nmeth1091 PMid:17873887. [PubMed: 17873887]

Miller JR, Koren S, Sutton G. Assembly algorithms for next-generation sequencing data. Genomics. 2010; 95(6):315–27. http://dx.doi.org/10.1016/j.ygeno.2010.03.001 PMid:20211242 PMCid:PMC2874646. [PubMed: 20211242]

Muyzer G, Waal EC. De, Uitierlinden AG. Profiling of Complex Microbial Populations by Denaturing Gradient Gel Electrophoresis Analysis of Polymerase Chain Reaction-Amplified Genes Coding for 16S rRNA. Applied And Environmental Microbiology. 1993; 59(3):695–700. PMid:7683183 PMCid:PMC202176. [PubMed: 7683183]

Muyzer G, Smalla K. Application of denaturing gradient gel electrophoresis (DGGE) and temperature gradient gel electrophoresis (TGGE) in microbial ecology. Antonie van Leeuwenhoek. 1998; 73:127–41. http://dx.doi.org/10.1023/A:1000669317571 PMid:9602286. [PubMed: 9602286]

Myers RM, et al. Nearly all single base substitutions In DNA fragments joined to a GC-clamp can be detected by denaturing gradient gel electrophoresis. Nucleic Acids Research. 1985; 13(9):3131–45. http://dx.doi.org/10.1093/nar/13.9.3131 PMid:4000972 PMCid:PMC341225. [PubMed: 4000972]

Nagashima K, et al. Phylogenetic Analysis of 16S Ribosomal RNA Gene Sequences from Human Fecal Microbiota and Improved Utility of Terminal Restriction Fragment Length Polymorphism Profiling. Bioscience and Microflora. 2006; 25(3):99–107. http://dx.doi.org/10.12938/bifidus.25.99.

Nair HP, Vincent H, Bhat SG. Evaluation of five in situ lysis protocols for PCR amenable metagenomic DNA from mangrove soils. Biotechnology Reports. 2014; 4(0):134–38. http://dx.doi.org/10.1016/j.btre.2014.09.008.

Nakatsu CH, Torsvik V, Ovreas L. Soil community analysis using DGGE of 16S rDNA polymerase chain reaction products. Soil Science Society of America Journal. 2000; 64(4):1382–88. http://dx.doi.org/10.2136/sssaj2000.6441382x.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Nayak R, Stewart-King T. Molecular Epidemiological Analysis and Microbial Source Tracking of Salmonella enterica Serovars in a Preharvest Turkey Production Environment. Foodborne Pathogens and Disease. 2008; 5(2):115–26. http://dx.doi.org/10.1089/fpd.2007.0029 PMid: 18361686. [PubMed: 18361686]

Olena M, Marra MA. Applications of next-generation sequencing technologies in functional genomics. Genomics. 2008; 92:255–64. http://dx.doi.org/10.1016/j.ygeno.2008.07.001 PMid:18703132. [PubMed: 18703132]

Pechter KB, et al. Essential Genome of the Metabolically Versatile Alphaproteobacterium Rhodopseudomonas palustris. Journal of Bacteriology. 2016; 198(5):867–76. http://dx.doi.org/10.1128/JB.00771-15 PMid:26712940 PMCid:PMC4810602.

Pitt TL. Bacterial typing systems: the way ahead. J. Med. Microbiol. 1994; 40:1–2. http://dx.doi.org/10.1099/00222615-40-1-1 PMid:8289208. [PubMed: 8289208]

Qu A, et al. Comparative Metagenomics Reveals Host Specific Metavirulomes and Horizontal Gene Transfer Elements in the Chicken Cecum Microbiome. PLoS ONE. 2008; 3(8) http://dx.doi.org/10.1371/journal.pone.0002945.

Renella G, et al. Environmental proteomics: A long march in the pedosphere. Soil Biology and Biochemistry. 2014; 69:34–37. http://dx.doi.org/10.1016/j.soilbio.2013.10.035.

Ricke SC, Pillai SD. Conventional and Molecular Methods for Understanding Probiotic Bacteria Functionality in Gastrointestinal Tracts. Critical Reviews in Microbiology. 1999; 25(1):19–38. http://dx.doi.org/10.1080/10408419991299176 PMid:10342098. [PubMed: 10342098]

Sagova-Mareckova M, et al. Innovative Methods for Soil DNA Purification Tested in Soils with Widely Differing Characteristics. Applied and Environmental Microbiology. 2008; 74(9):2902–07. http://dx.doi.org/10.1128/AEM.02161-07 PMid:18344341 PMCid:PMC2394906. [PubMed: 18344341]

Saleh-Lakha S, et al. Microbial gene expression in soil: methods, applications and challenges. Journal of Microbiological Methods. 2005; 63(1):1–19. http://dx.doi.org/10.1016/j.mimet.2005.03.007 PMid:15939495. [PubMed: 15939495]

Schellenberg J, et al. Pyrosequencing of the Chaperonin-60 Universal Target as a Tool for Determining Microbial Community Composition. Applied and Environmental Microbiology. 2009; 75(9):2889–98. http://dx.doi.org/10.1128/AEM.01640-08 PMid:19270139 PMCid:PMC2681704. [PubMed: 19270139]

Schloss PD, et al. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. Appl Environ Microbiol. 2009; 75(23):7537–41. http://dx.doi.org/10.1128/AEM.01541-09 PMid:19801464 PMCid:PMC2786419. [PubMed: 19801464]

Segata N, et al. Metagenomic biomarker discovery and explanation. Genome Biol. 2011; 12(6):R60. http://dx.doi.org/10.1186/gb-2011-12-6-r60 PMid:21702898 PMCid:PMC3218848. [PubMed: 21702898]

Senapathy P, et al. Targeted Genome-Wide Enrichment of Functional Regions. PLoS ONE. 2010; 5(6) http://dx.doi.org/10.1371/journal.pone.0011138 PMid:20585402 PMCid:PMC2886846.

Sharpton TJ. An introduction to the analysis of shotgun metagenomic data. Frontiers in plant science. 2014; 5 http://dx.doi.org/10.3389/fpls.2014.00209.

Simon C, Daniel R. Metagenomic Analyses: Past and Future Trends. Applied and Environmental Microbiology. 2011; 77(4):1153–61. http://dx.doi.org/10.1128/AEM.02345-10 PMid:21169428 PMCid:PMC3067235. [PubMed: 21169428]

Sims D, et al. Sequencing depth and coverage: key considerations in genomic analyses. Nat Rev Genet. 2014; 15(2):121–32. http://dx.doi.org/10.1038/nrg3642 PMid:24434847. [PubMed: 24434847]

Singh P, Nayak R, Kwon YM. Target-Enrichment Through Amplification of Hairpin-Ligated Universal Targets for Next-Generation Sequencing Analysis. High-Throughput Next Generation Sequencing: Methods and Application. 2011; 733:267–78. http://dx.doi.org/10.1007/978-1-61779-089-8_19 PMid:21431777.

Singh P, et al. Multilocus sequence typing of Salmonella strains by high-throughput sequencing of selectively amplified target genes. J Microbiol Methods. 2012; 88(1):127–33. http://dx.doi.org/10.1016/j.mimet.2011.11.004 PMid:22108494. [PubMed: 22108494]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Singh P, et al. Massively parallel sequencing of enriched target amplicons for high-resolution genotyping of Salmonella serovars. Mol Cell Probes. 2013a; 27(2):80–5. http://dx.doi.org/10.1016/j.mcp.2012.11.004. [PubMed: 23201627]

Singh P, et al. Influence of penicillin on microbial diversity of the cecal microbiota in broiler chickens. Poult Sci. 2013b; 92(1):272–6. http://dx.doi.org/10.3382/ps.2012-02603 PMid:23243258. [PubMed: 23243258]

Singh SK, Ramaiah N. Denaturing gradient gel electrophoresis profiling of bacterial communities composition in Arabian Sea. J Environ Biol. 2011; 32(3):339–46. PMid:22167947. [PubMed: 22167947]

Snitkin ES, et al. Tracking a hospital outbreak of carbapenem-resistant Klebsiella pneumoniae with whole-genome sequencing. Sci Transl Med. 2012; 4(148):148ra116. https://dx.doi.org/10.1126/scitranslmed.3004129 PMid: 22914622.

Staley JT, Konopka Allan. Measurement of In Situ Activities of Nonphotosynthetic Microorganisms in Aquaticand Terrestrial Habitats. Ann. Rev. Microbiol. 1985; 39:321–46. http://dx.doi.org/10.1146/annurev.mi.39.100185.001541 PMid:3904603. [PubMed: 3904603]

Stein LD. The case for cloud computing in genome informatics. Genome Biology. 2010; 11(5):207. http://dx.doi.org/10.1186/gb-2010-11-5-207 PMid:20441614 PMCid:PMC2898083. [PubMed: 20441614]

Summerer D. Enabling technologies of genomic-scale sequence enrichment for targeted high-throughput sequencing. Genomics. 2009; 94:363–68. http://dx.doi.org/10.1016/j.ygeno.2009.08.012 PMid:19720138. [PubMed: 19720138]

Summerer D, et al. Microarray-based multicycle-enrichment of genomic subsets for targeted next-generation sequencing. Genome Res. 2009; 19:1616–21. http://dx.doi.org/10.1101/gr.091942.109 PMid:19638418 PMCid:PMC2752126. [PubMed: 19638418]

Tewhey R, et al. Enrichment of sequencing targets from the human genome by solution hybridization. Genome Biology. 2009; 10(10) http://dx.doi.org/10.1186/gb-2009-10-10-r116 PMid:19835619 PMCid:PMC2784331.

Theron J, Cloete TE. Molecular Techniques for Determining Microbial Diversity and Community Structure in Natural Environments. Critical Reviews in Microbiology. 2000; 26(1):37–57. http://dx.doi.org/10.1080/10408410091154174 PMid:10782339. [PubMed: 10782339]

Thomas T, Gilbert J, Meyer F. Metagenomics-a guide from sampling to data analysis. Microb Inform Exp. 2012; 2(3):1–12. http://dx.doi.org/10.1186/2042-5783-2-3. [PubMed: 22587916]

Ting N, et al. Selective gene amplification for High-Throughput Sequencing. Recent Patents on DNA & Gene Sequences. 2009; 3:29–38. http://dx.doi.org/10.2174/187221509787236183. [PubMed: 19149736]

Toueille, M., et al. A comparative proteomic approach to better define Deinococcus nucleoid specificities. Journal of proteomics. 2012. http://dx.doi.org/10.1016/j.jprot.2012.03.002 PMid: 22446890

van Opijnen T, Bodi KL, Camilli A. Tn-seq: high-throughput parallel sequencing for fitness and genetic interaction studies in microorganisms. Nature Methods. 2009; 6(10):767–U21. http://dx.doi.org/10.1038/nmeth.1377 PMid:19767758 PMCid:PMC2957483. [PubMed: 19767758]

Varley KE, Mitra RD. Nested Patch PCR enables highly multiplexed mutation discovery in candidate genes. Genome Research. 2008; 18:1844–50. http://dx.doi.org/10.1101/gr.078204.108 PMid: 18849522 PMCid:PMC2577855. [PubMed: 18849522]

Wall PK, et al. Comparison of next generation sequencing technologies for transcriptome characterization. BMC Genomics. 2009; 10(347) http://dx.doi.org/10.1186/1471-2164-10-347.

Wang Q, et al. Xander: employing a novel method for efficient gene-targeted metagenomic assembly. Microbiome. 2015; 3(1):32. http://dx.doi.org/10.1186/s40168-015-0093-6 PMid:26246894 PMCid:PMC4526283. [PubMed: 26246894]

Wang W, et al. A universal and rapid protocol for protein extraction from recalcitrant plant tissues for proteomic analysis. Electrophoresis. 2006; 27(13):2782–86. http://dx.doi.org/10.1002/elps.200500722 PMid:16732618. [PubMed: 16732618]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. Nature Reviews Genetics. 2009; 10(1):57–63. http://dx.doi.org/10.1038/nrg2484 PMid:19015660 PMCid:PMC2949280.

Williamson KE, et al. Optimizing the indirect extraction of prokaryotic DNA from soils. Soil Biology and Biochemistry. 2011; 43(4):736–48. http://dx.doi.org/10.1016/j.soilbio.2010.04.017.

Woese CR. Bacterial Evolution. Microbiological Reviews. 1987; 51(2):221–71. PMid:2439888 PMCid:PMC373105. [PubMed: 2439888]

Yankson KK, Steck TR. Strategy for Extracting DNA from Clay Soil and Detecting a Specific Target Sequence via Selective Enrichment and Real-Time (Quantitative) PCR Amplification. Applied and Environmental Microbiology. 2009; 75(18):6017–21. http://dx.doi.org/10.1128/AEM.00211-09 PMid:19633108 PMCid:PMC2747842. [PubMed: 19633108]