Chapter

The Role of Metagenomic Approaches in the Analysis of Microbial Community in Extreme Environment

Ahmed M. Shuikan, Rakan M. Alshuwaykan and Ibrahim A. Arif

Abstract

Metagenomic is a promising technique that has many applications in different fields. In fact, metagenomics is the ideal culture-independent technique that unravels the microbial composition and biodiversity in the sample, which helps scientists to study and understand how this biodiversity is affected by continuously changing conditions in the environment and how this microbial community interacts with each other. In the past, the microbial composition in extreme environments was undiscovered due to the difficulty of isolation, culturing, and identification of microbes living there. However, nowadays after the development and combination of metagenomic and next-generation sequencing techniques, it became more easy to study the microbial composition in extreme environments without culturing. In this chapter, the use of metagenomic techniques to study the microbial biodiversity in different extreme environments are discussed. In addition, different NGS platforms are discussed in terms of principles, advantages, and limitations.

Keywords: metagenomics, extreme environment, salinity soil, microbial taxonomy, biodiversity, sequencing approaches

1. Introduction

Microorganisms are crucial to the control of many activities, including the recycling of essential elements and nutrients, the dynamics of the biogeochemical cycles (such as carbon, nitrogen, and oxygen cycles), and the development of soil structure. To increase our knowledge of microbial diversity, activity, and interactions with different ecosystem components, it is essential to have an understanding of the structure, function, and activities of microbes [1]. Prokaryotic taxonomy is the term used to describe the traditional classification of the diversity of microorganisms based on the ideas of classification, naming, and characterization. Based on their morphology (form), growth environments, bacterial pathogenicity, and early investigations introduced the ideas of genera and species into the taxonomy of bacteria [2].

Later, the classification scheme developed by the American Society for Microbiology included the biochemical and physiological characteristics of bacteria. DNA–DNA hybridization (DDH) methods have been used as the de facto method for prokaryotic classification based on genomic similarity since the 1960s [3].

As a way to distinguish between different bacterial species and research bacterial phylogeny, 16S rRNA sequence similarity emerged as one of the most used methods [4]. Later, following the discovery of the polymerase chain reaction (PCR), molecular methods to study prokaryotic classification were developed, including denaturing gradient gel electrophoresis (DGGE), temperature gradient gel electrophoresis (TGGE), restriction fragment length polymorphism (RFLP), terminal restriction fragment length polymorphism (TRFLP), ribosomal intergenic spacer analysis (RISA), and real-time PCR (quantitative PCR) [5]. The field of microbial taxonomy has been transformed most significantly by high-throughput DNA sequencing technologies, such as metagenomics shotgun sequencing.

When researching soil microbiology, genomics is especially crucial. According to different researches, one gram of soil may contain more different microorganisms than that have been grown in laboratory to date [6]. Therefore, metagenomics appears to be the best culture-independent method for revealing soil biodiversity and researching how this biodiversity is impacted by constantly changing environmental factors. The first step in the approach for metagenomics research is to choose an appropriate ecological or biological setting that supports a diverse range of microbial communities with potential biotechnological and therapeutic uses. Extreme conditions are mostly what draw metagenomic researchers to certain environments. These include highly alkaline or acidic pH conditions, high metal concentrations, pressures, or radiation, as well as settings with high salinity or extreme temperatures [7].

Beginning with the isolation of genomic DNA from the soil sample that represents the entire population, a DNA library is built from the isolated DNA, and the DNA library is then screened for a target sequence (**Figure 1**). It is crucial to choose a DNA extraction procedure that will generate enough DNA. The diversity of the entire microbial community in the target environment. One of the most difficult steps in the metagenomic analysis is still this one. Depending on the type of soil analyzed, the chemical and physical features of soils are quite diverse and complex, which will make it challenging to design a reference method for DNA extraction from soils. Additionally, a variety of chemicals found in soils co-extract genomic DNA and have an inhibiting influence on the subsequent processing of extracted DNA. Humic and fulvic acids are two examples [8]. As a result, for each kind of soil, optimization and comparison of various extraction methods are typically necessary [9–11].

The genomic DNA that was extracted from the target environment is then used to create a DNA library. This is accomplished by first dividing the isolated DNA into pieces of the proper sizes to facilitate cloning. Either mechanical shearing or restriction enzyme digestion is used to accomplish this. These procedures produce fragments of DNA that are then cloned into the appropriate cloning vector. Short-insert genomic libraries are created by inserting small DNA fragments into plasmid vectors. Large inserts can be cloned onto BAC or cosmid vectors, which can carry inserts larger than 40 Kb, or fosmid or cosmid vectors, which can handle inserts up to 40 Kb in size [12].

1.1 Metagenomic methods

The term "metagenomics" refers to a collection of methods and techniques used to analyze the entire genomes of microorganisms residing in a given environment

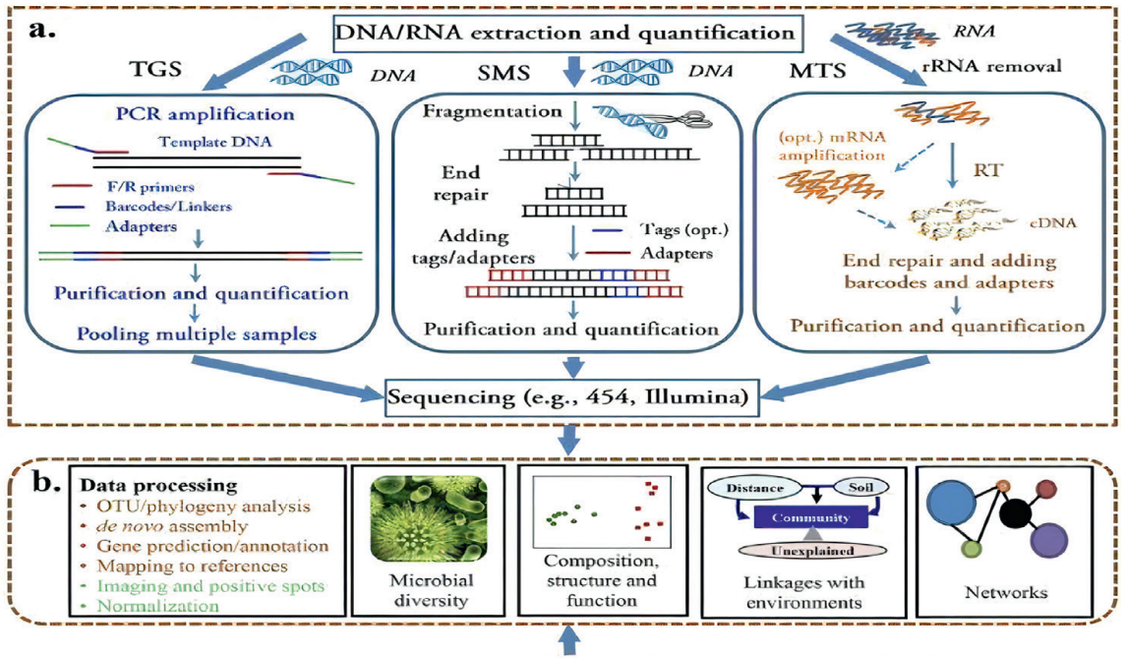

Figure 1.
Metagenomic protocol. (a) the first step of metagenomic analysis is DNA extraction by using different DNA extraction techniques and (b) the second step of metagenomic analysis by using software data analysis to analyze the huge data collected from samples.

without regard to culture [13]. It has a wide range of beneficial applications and shows great promise in both medical and environmental microbiology. Metagenomics is most frequently used in environmental microbiology to research the total genomic DNA analysis that has been used to study the diversity of microbial communities in various settings, how various microorganisms interact with one another there, and how these communities adapt to changes in the physical and chemical characteristics of these environments [14]. Additionally, metagenomics offers the chance to find and isolate new enzymes with industrial uses from harsh habitats where uncultivable extremophiles reside. Functional metagenomics enables the isolation of genes encoding extremozymes, enzymes capable of catalyzing in severe settings, or genes that will improve understanding of the mechanisms that make such organisms resistant to extreme environmental conditions in such situations [7].

E.coli, a microbe that has been extensively studied and is simple to work with in the lab, is typically used to create DNA libraries. Shuttle vectors are employed to transfer the libraries into the appropriate host in the event that the genes contained in DNA inserts need to be expressed in other microorganisms [15].

Finally, a screening test is used to look for a gene that performs a certain function, and the gene product is then functionally examined. There are two distinct metagenomics approaches that are frequently applied in studies. The first one focuses on using marker genes, such as the ribosomal genes 16S rRNA [16] and 18S rRNA, to research the makeup of microbial communities in certain environments or a specific protein-coding gene of medical or industrial value [17, 18]. Targeted metagenomics is the name given to this tactic. Shotgun metagenomics is the second method. In this method, high-throughput next-generation sequencing is used to get broad coverage of genomic DNA sequences in order to evaluate the overall taxonomic structure or functional potential of microbial communities [19].

Recent times have seen a significant increase in the use of various next-generation sequencing (NGS) platforms, each with its own advantages and disadvantages, for the taxonomic profiling, characterization, and analysis of microbial communities. High-throughput, short read sequences, and relatively declining costs are all characteristics of metagenomic samples. These platforms are beneficial in that they do not require DNA fragment cloning [20]. Recent improvements in NGS technology have been made to accommodate a wide range of applications, costs, and capabilities [21]. The 454 life sciences (Roche) and Illumina systems (Solexa) platforms are the most often utilized ones [22]. The 454-sequencing technology, which was the first next-generation technology to be commercially accessible, is based on the pyrosequencing method. It offers analysis with a high-throughput at a reasonable cost [23]. This method sequencing procedure involves the insertion of nucleotides into the capture of the released pyrophosphate, which undergoes an enzyme process to produce light, and allows for the detection of the developing chain. In order to assign a distinct nucleotide to each nucleotide incorporation event, different nucleotides are successively added. The light signals are finally transformed into sequencing data. The DNA fragments are fixed on beads in a water–oil emulsion before being amplified in a 454 pyrosequencer [24]. The thermophilic cellulose-degrading microbial species isolated from the hot springs in Xiamen, China, have been studied using pyrosequencing. It was also widely used to analyze the diversity of microbes in a variety of habitats, such as different soil environments [25–28] and marine environments.

Reversible terminator nucleotides that are fluorescently tagged are a key component of Illumina sequencing technology. The terminator nucleotides are coupled with blocking groups that may be removed from the nitrogen base in a single step, unlike deoxynucleotides, which are chemically changed to stop further DNA synthesis as is the case with sanger sequencing. On a chip with attached primers, DNA synthesis takes to happen. Following each cycle of synthesis, a laser is used to excite the dyes attached to each nucleotide. This is followed by scanning of the integrated bases. The blocking group and the dye must first be eliminated by a chemical reaction in order for the subsequent synthesis cycle to start. Multiple settings, including freshwater sponges, the gastrointestinal system, soils, and marine environments, were effectively studied using the Illumina sequencing technology.

In addition to the methods already described, metagenomics research also makes use of recently developed sequencing technology. These include the single-molecule real-time (SMRT) DNA sequencing from Pacific Biosciences, Ion Torrent's semiconductor sequencing, and Applied Biosystems' SOLiD 5500 W Series [22]. Technologies that are more advanced and cutting-edge are being created, and they could be very helpful in metagenomics research. Oxford nanopore technologies are actively working on strand sequencing technologies, which make it possible to sequence whole DNA strands as they travel through a protein nanopore [29]. One of the most exciting new technologies in the genomics era is Irys technology, created by BioNano genomics. The analysis of the enormous number of sequence data that are produced from the screening phase in metagenomics is the component of the procedure that is the most difficult built-in library many bioinformatics tools have grown over time to assist in analyzing the metagenomic data and comparing it to online databases.

1.2 Metagenomics to identify microbial diversity in extreme environments

Both industrial microbiologists and ecologists are interested in the study of microbial communities in extreme conditions, such as saline soil or other saline

habitats, such as saline waterways or saline sediments [30]. Understanding the impacts of salinity on soil ecosystems can aid in understanding how the structure of the microbial community changes in response to changes in salinity. However, due to their benefits over mesophilic enzymes in commercial applications, industrial microbiologists are more interested in isolating halophilic bacteria or organisms that like to survive in salty environments, or their enzymes.

Vera-Gargallo [31] used metagenomics to explore the microbial communities in two hypersaline soils in the Odiel saltmarshes of Spain. They then matched their findings to the information found in the databases for the 13 saline aquatic habitats. In contrast to hypersaline aquatic habitats, they discovered that hypersaline soils had. more diversified microbial communities that include non-halophilic organisms. In order to understand how salt impacts microbial dispersion in saline environments, more research has recently been focused on saline soils. Using high-throughput sequencing technology, it was recently shown that bacterial communities were more abundant in surface soils (0–10 cm) than in deep soils in the saline soils of Qarhan Sale Lake, China (15–30 cm). This was linked to the fact that underground soils had low oxygen content in other investigations as well [32]. *Proteobacteria, Bacteroidetes,* and *Gemmatimonadetes* were discovered to be the three phyla with the greatest abundance in this study. *Proteobacteria* contain the *Alphaproteobacteria*, *Deltaproteobacteria,* and *Betaproteobacteria* and were the next most prevalent classes after *Gammaproteobacteria* [33]. Previous research [34] revealed similar findings. Nevertheless, there was no discernible variation in the microbial community between the surface and deep saline soils.

The structure and metabolic processes of microbial communities are being impacted by soil salinity, according to more recent studies. According to a study by Chen [35], the phyla *Planctomycetes* and *Bacteroidetes* of bacteria were observed to decline under continuous irrigation with saline water, while *proteobacteria, Actinobacteria,* and *Chloroflexi* grew. They also discovered that irrigation with salt water has a significant impact on the metabolic activities of soil microorganisms. The salinity of irrigation was observed to boost soil bacterial richness.

Uncertainty surrounds the exact method by which salinity alters the composition of microbial communities. Morrissey [36] hypothesized that microbes' preferences for saline soils are influenced by their evolutionary history. They discovered that most *proteobacteria* favored saltwater, while many *proteobacteria* chose freshwater in their study of wetland soils. There was a connection between the quantity of microorganisms that responded to salinity and phylogenetically grouped salinity preferences. Analysis of the 16S rRNA gene sequences of bacteria isolated from soil in Italy's A horizon that has varied salt concentrations. The findings demonstrated that variations in soil salt concentration cause differences in the composition of bacterial communities. *Proteobacteria* > *Actinobacteria* > *Acidobacteria* > *Verrucomicrobia* > *Gemmatimonodates* > *Firmicutes* > *Chloroflexi* > *Bacteriodes* > *Chlorobi* were the most prevalent bacterial species in this soil in terms of abundance. However, they discovered that various levels of some taxonomic groups were present in the soil that had varying salt values. Nevertheless, several bacterial species did not have salt levels, which were present in equal amounts at all of the investigated sites [37].

The same bacterial groups that were found in the prior research were also found in large numbers in saline soils [38]. Functional genomics has been effectively used to separate the genetic components of osmoadaptation from the isolated metagenome library. In order to better understand the mechanisms by which the microbial population responds to salinity, they were able to locate salt-tolerant genes. Through genetic

engineering, numerous genes have been cloned to enhance salt tolerance in crops, such as *P5CS, DREB1A,* and *AtNHX1*.

By using metagenomic analysis, several more severe settings have been examined. Organic carbon sources are scarce in alkaline environments, such as soda lakes and low-saline-alkaline habitats. Understanding the diversity of microbes present in these harsh conditions, according to scientists, may help us better comprehend the early and rare forms of life [39]. Five novel species of the candidate phyla radiation were abundantly found in the hypersaline soda lake, according to a recent metagenomic investigation [40]. To investigate the microbial diversity in contaminated soils, metagenomics is a potent investigative tool. Heavy metal contamination of the soil at mining sites can be detrimental. The most prevalent bacteria in these conditions were determined to be *Solirubrobacter*, *Geobacter, Edaphobacter, and Pseudomonas* [41]. Other environments that metagenomics is probably being used to investigate acidic, low-oxygen, and volcanic settings are among the methods [42–44].

2. Conclusion

Due to the presence of non-cultivable microorganisms, traditional culturing techniques, such as isolation, do not reflect the real structure of microbial community when studying extreme environments. However, the emergence and development of metagenomic techniques boosted the identification and profiling of the microbial community in extreme environments and allowed to overcome non-cultivable microorganisms issues. Nowadays, by using metagenomic approaches scientists are able to profile microbial community and investigate the biodiversity in a given environment and also able to study and understand how these microorganisms have adapted to such environments. Resulting in, many different secondary metabolites have been discovered that are produced by microorganisms in order to adapt to extreme environments. The 16S rRNA gene is the most suitable gene to study microbial biodiversity and establish a phylogenetic relationship among unculturable and novel microorganisms.

DOI: http://dx.doi.org/10.5772/intechopen.108050

Author details

Ahmed M. Shuikan[1*], Rakan M. Alshuwaykan[2] and Ibrahim A. Arif[1]

1 Botany and Microbiology Department, College of Science, King Saud University, Riyadh, Saudi Arabia

2 Suergery Department, King Khalid Civilian Hospital, Tabuk, Saudi Arabia

*Address all correspondence to: ahmedshuikan@gmail.com

References

[1] Quince C, Curtis TP, Sloan WT. The rational exploration of microbial diversity. The ISME Journal. 2008;**2**(10):997

[2] Drews G. The roots of microbiology and the influence of Ferdinand Cohn on microbiology of the 19th century. FEMS Microbiology Reviews. 2000;**24**(3):225-249

[3] Grimont PA. Use of DNA reassociation in bacterial classification. Canadian Journal of Microbiology. 1988;**34**(4):541-546

[4] Lane DJ, Pace B, Olsen GJ, Stahl DA, Sogin ML, Pace NR. Rapid determination of 16S ribosomal RNA sequences for phylogenetic analyses. Proceedings of the National Academy of Sciences. 1985;**82**(20):6955-6959

[5] Stefanis C, Alexopoulos A, Voidarou C, Vavias S, E. & Bezirtzoglou. Principal methods for isolation and identification of soil microbial communities. Folia Microbiologia (Praha). 2013;**58**(1):61-68

[6] Torsvik V, Øvreås L. Microbial diversity and function in soil: From genes to ecosystems. Current Opinion in Microbiology. 2002;**5**(3):240-245

[7] Mirete S, Morgante V, González JE. Functional metagenomics of extreme environments. Current Opinion in Biotechnology. 2016;**38**:143-149

[8] Young JM, Rawlence NJ, Weyrich LS, Cooper A. Limitations and recommendations for successful DNA extraction from forensic soil samples: A review. Science & Justice. 2014;**54**(3):238-244

[9] Finley SJ, Lorenco N, Mulle J, Robertson BK, Javan GT. Assessment of microbial DNA extraction methods of cadaver soil samples for criminal investigations. Australian Journal of Forensic Sciences. 2016;**48**(3):265-272

[10] Lim NY, Roco CA, Frostegård Å. Transparent DNA/RNA Coextraction workflow protocol suitable for inhibitor-rich environmental samples that focuses on complete DNA removal for Transcriptomic analyses. Frontiers in Microbiology. 2016;**7**:1588

[11] Mazziotti M, Henry S, Laval P, Bonnefoy A, Falla J. Comparison of two bacterial DNA extraction methods from non-polluted and polluted soils. Folia Microbiologica. 2018;**63**(1):85-92

[12] Simon C, Daniel R. Construction of small-insert and large-insert metagenomic libraries. In: Metagenomics. New York, NY: Humana Press; 2017. pp. 1-12

[13] Daniel R. The metagenomics of soil. Nature Reviews Microbiology. 2005;**3**(6):470

[14] Delmont TO, Robe P, Cecillon S, Clark IM, Constancias F, Simonet P, et al. Accessing the soil metagenome for studies of microbial diversity. Applied and Environmental Microbiology. 2011;**77**(4):1315-1324

[15] Lam KN, Cheng J, Engel K, Neufeld JD, Charles TC. Current and future resources for functional metagenomics. Frontiers in Microbiology. 2015;**6**:1196

[16] Païssé S, Valle C, Servant F, Courtney M, Burcelin R, Amar J, et al. Comprehensive description of blood microbiome from healthy donors assessed by 16 S targeted

metagenomic sequencing. Transfusion. 2016;**56**(5):1138-1147

[17] Nurdiani D, Ito M, Maruyama T, Terahara T, Mori T, Ugawa S, et al. Analysis of bacterial xylose isomerase gene diversity using gene- targeted metagenomics. Journal of Bioscience and Bioengineering. 2015;**120**(2):174-180

[18] Lanza VF, Baquero F, Martínez JL, Ramos R, Gonzalez B, Andremont A, et al. In-depth resistome analysis by targeted metagenomics. Microbiome. 2018;**6**(1):11

[19] Quince C, Walker AW, Simpson JT, Loman NJ, Segata N. Shotgun metagenomics, from sampling to analysis. Nature Biotechnology. 2017;**35**(9):833

[20] Ansorge WJ. Next-generation DNA sequencing techniques. New Biotechnology. 2009;**25**(4):195-203

[21] Scholz L, Chain PS. Next generation sequencing and bioinformatic bottlenecks: The current state of metagenomic data analysis. Current Opinion in Biotechnology. 2012;**23**(1):9-15

[22] Oulas A, Pavloudi C, Polymenakou P, Pavlopoulos GA, Papanikolaou N, Kotoulas G, et al. Metagenomics: Tools and Insights for Analyzing next-generation sequencing data derived from biodiversity studies. Bioinformatics and Biology Insights. 2015;**9**:BBI-S12462

[23] Ronaghi M. Pyrosequencing sheds light on DNA sequencing. Genome Research. 2001;**11**(1):3-11

[24] Goodwin S, McPherson JD, McCombie WR. Coming of age: Ten years of next-generation sequencing technologies. Nature Reviews Genetics. 2016;**17**(6):333

[25] Egge E, Bittner L, Andersen T, Audic S, de Vargas C, Edvardsen B. 454 pyrosequencing to describe microbial eukaryotic community composition, diversity and relative abundance: A test for marine haptophytes. PLoS One. 2013;**8**(9):e74371

[26] Fan W, Huo G, Li X, Yang L, Duan C, Wang T, et al. Diversity of the intestinal microbiota in different patterns of feeding infants by Illumina high-throughput sequencing. World Journal of Microbiology and Biotechnology. 2013;**29**(12):2365-2372

[27] Hong C, Si Y, Xing Y, Li Y. Illumina MiSeq sequencing investigation on the contrasting soil bacterial community structures in different iron mining areas. Environmental Science and Pollution Research. 2015;**22**(14):10788-10799

[28] Lentini V, Gugliandolo C, Bunk B, Overmann J, Maugeri TL. Diversity of prokaryotic community at a shallow marine hydrothermal site elucidated by Illumina sequencing technology. Current Microbiology. 2014;**69**(4):457-466

[29] Jain M, Olsen HE, Paten B, Akeson M. The Oxford Nanopore MinION: Delivery of nanopore sequencing to the genomics community. Genome Biology. 2016;**17**(1):239

[30] Shuikan AM, Hozzein WN, Alzharani MM, Sandouka MN, Al Yousef SA, Alharbi SA, et al. Enhancement and identification of microbial secondary metabolites. In: Extremophilic Microbes and Metabolites - Diversity, Bioprespecting and Biotechnological Applications. London: IntechOpen; 2020

[31] Vera-Gargallo B, Ventosa A. Metagenomic insights into the phylogenetic and metabolic diversity of the prokaryotic community dwelling in Hypersaline soils from the Odiel saltmarshes (SW Spain). Genes. 2018;**9**(3):152

[32] Serkebaeva YM, Kim Y, Liesack W, Dedysh SN. Pyrosequencingbased assessment of the bacteria diversity in surface and subsurface peat layers of a northern wetland, with focus on poorly studied phyla and candidate divisions. PLoS One. 2013;**8**(5):e63994

[33] Xie K, Deng Y, Zhang S, Zhang W, Liu J, Xie Y, et al. Prokaryotic community distribution along an ecological gradient of salinity in surface and subsurface saline soils. Scientific Reports. 2017;7(1):13332

[34] Ma B, Gong J. A meta-analysis of the publicly available bacterial and archaeal sequence diversity in saline soils. World Journal of Microbiology and Biotechnology. 2013;**29**(12):2325-2334

[35] Chen L, Li C, Feng Q, Wei Y, Zheng H, Zhao Y, et al. Shifts in soil microbial metabolic activities and community structures along a salinity gradient of irrigation water in a typical arid region of China. Science of the Total Environment. 2017;**598**:64-70

[36] Morrissey EM, Franklin RB. Evolutionary history influences the salinity preference of bacterial taxa in wetland soils. Frontiers in Microbiology. 2015;**6**:1013

[37] Canfora L, Bacci G, Pinzari F, Papa GL, Dazzi C, Benedetti A. Salinity and bacterial diversity: To what extent does the concentration of salt affect the bacterial community in a saline soil? PLoS One. 2014;**9**(9):e106662

[38] Ahmed V, Verma MK, Gupta S, Mandhan V, Chauhan NS. Metagenomic profiling of soil microbes to mine salt stress tolerance genes. Frontiers in Microbiology. 2018;**9**:159

[39] Veríssimo A, Tiago I. Bacterial diversity in a nonsaline alkaline environment. In: Nelson KE, editor. Encyclopedia of Metagenomics: Environmental Metagenomics. Boston, MA: Springer; 2015. pp. 37-41

[40] Vavourakis CD, Andrei AS, Mehrshad M, Ghai R, Sorokin DY, Muyzer G. A metagenomics roadmap to the uncultured genome diversity in hypersaline soda lake sediments. Microbiome. 2018;**6**(1):168

[41] Hemmat-Jou MH, Safari-Sinegani AA, Mirzaie-Asl A, Tahmourespour A. Analysis of microbial communities in heavy metals-contaminated soils using the metagenomic approach. Ecotoxicology. 2018;**27**(9):1281-1291

[42] Korzhenkov AA, Toshchakov SV, Bargiela R, Gibbard H, Ferrer M, Teplyuk AV, et al. Archaea dominate the microbial community in an ecosystem with low-to-moderate temperature and extreme acidity. Microbiome. 2019;7(1):11

[43] Voorhies AA, Eisenlord SD, Marcus DN, Duhaime MB, Biddanda BA, Cavalcoli JD, et al. Ecological and genetic interactions between cyanobacteria and viruses in a low-oxygen mat community inferred through metagenomics and metatranscriptomics. Environmental Microbiology. 2016;**18**(2):358-371

[44] Riquelme C, Dapkevicius MDLE, Miller AZ, Charlop Z, Brady S, Mason C, et al. Biotechnological potential of Actinobacteria from Canadian and Azorean volcanic caves. Applied Microbiology and Biotechnology. 2017;**101**(2):843-857