

## APPLICATION OF METAGENOMICS IN ECOLOGY: A BRIEF OVERVIEW

Luiza Silvia Chiriac<sup>1,2\*</sup>, Dumitru T. Murariu<sup>1,2</sup>

<sup>1</sup> Institute of Biology Bucharest, Romanian Academy, Splaiul Independenței 296, Bucharest, Romania

<sup>2</sup> Doctoral School of Biology, Faculty of Biology, University of Bucharest, Splaiul Independenței 91-95, Bucharest, Romania



### Abstract

*Starting from the fact that in a microbial community even the concept of species is poorly defined and often debated, but closely related organisms tend to share a substantial amount of genomic sequences, together with a large number of physiological and biochemical properties, it was developed metagenomics. The discipline of metagenomics, defined as the genomic analysis of all microorganisms in a given niche environment, has evolved as an effort to find out more about the microbial diversity of natural environments, such as soil, seawater and the gastrointestinal tract of vertebrates and invertebrates. The purpose of this paper was to bring in front the discipline of metagenomics which will be used in many scientific areas in the future. Our paper represents a brief review of the literature available on the internet regarding the definition, description of metagenomics but also its possibilities of its application in ecology. Moreover, this paper contains the description of MG-RAST software as bioinformatics method that is suitable for usage of metagenomics in ecological studies.*

*Keywords: ecology, genomic sequences, metagenomics, MG-RAST*

### 1. INTRODUCTION

Metagenomics, a term first invented by Handelsman in 1998, is a study of microbial populations in a particular habitat based on their genetic material. DNA is extracted directly from samples taken from the environment (Lorenz and Schleper, 2002; Steele and Streit, 2005; Parks and Beiko, 2010). Initially, the term was used to describe a specific discipline in genetics that deals with the mapping, sequencing and analysis of the genome (complete set of genes and chromosomes in an organism) (Xu, 2006). Metagenomics can represent "beyond the genome" (Gilbert and Dupont, 2011). While many scientists use genomics in this structural sense, an increasing number of researchers have expanded their use to include functional analyzes of the entire genome. These functional analytical aspects include transcripts of whole genome RNA (called transcriptomics), proteins (proteomics), and metabolites (metabolomics) (Xu, 2006). These areas of study can improve the understanding of the organization and functionality of biological systems. They can also track the molecular changes that occurred during development, under various conditions (physiological, pathological or influenced by environmental changes) (Esposito et al., 2016).

The basic definition of metagenomics is the analysis of genomic DNA from an entire community. This definition separates metagenomics from genomics, which is the analysis of genomic DNA in an individual organism or cell). Etymologically "meta", in Greek means "beyond", and "genomics"

means study the entire DNA content of an organism (Xu, 2006; Gilbert and Dupont, 2011). Metagenomics, therefore, it simply means studying several genomes at a time (Singh et al., 2020). In addition, various combinations of "-omic" terms have recently become fashionable. For example, the discipline that uses genomics methods to analyze natural ecological communities' metagenomics, ecological genomics, community genomics, and environmental genomics has been named (Xu, 2006).

## 2. MATERIALS AND METHODS

As the paper is a brief review of the literature on the application of metagenomics in ecology, the method used was the critical analysis of the literature. The total number of articles analysed was 46. The search engine for articles was Google Academic and Research Gate. Articles with free access were used or in some cases where the article was not free, I requested the work of the authors.

## 3. RESULTS AND DISCUSSIONS

Genomics completes the genetic structure of an organism by sequencing a large number of base pairs of its DNA. Metagenomics involves sampling the genome sequences of a community of organisms living in a common environment (Hugenholtz and Tyson, 2008; Narasingarao et al., 2012; Ayling et al., 2020). Metagenomics is responsible for the impressive advances in the ecology, evolution and diversity of microorganisms in recent years, and many research institutes and laboratories are currently involved in this field (Thomas et al., 2012). The discipline of metagenomics, defined as the genomic analysis of all microorganisms in a given niche environment, has evolved as an effort to find out more about the microbial diversity of natural environments (Sleator et al., 2008). Advances in DNA sequencing technology have led to a huge increase in biodiversity information at all levels, from genes to ecosystems, and have revolutionized biodiversity studies (Macher et al., 2019). Metagenomics has become a powerful tool for the study of soil microbial diversity, population dynamics and ecophysiology of microorganisms in natural ecosystems. Metagenomics also provides genetic information of soil microorganisms (Rondon et al., 2000; Grob et al., 2015; Kouzuma, 2018). Metagenomics is also defined as the analysis of the genomic sequences of an entire community, directly from a sample (Coyotzi et al., 2016; Coutinho et al., 2018). Metagenomics, the reconstruction of the genome of microorganisms that are difficult to cultivate is an innovative method for access to biodiversity resources (Zeyaulah et al., 2009). Current estimates say that over 99% of microorganisms in natural environments cannot be easily grown in the laboratory so they are not accessible to biotechnology or basic research. Metagenomics means that the microbial diversity that cannot be grown in the laboratory can be analyzed for new therapeutic, biotechnological and sustainable agricultural applications (Carbonetto et al., 2014; Miliute et al., 2015; Gupta et al., 2018; Kirubakaran et al., 2020; Singh et al., 2020). It is a new, increasingly sophisticated field that refers to the direct isolation of DNA from a particular habitat, followed by the cloning (in a surrogate host, such as *Escherichia coli*), of the entire genome of the entire microbial populations. The resulting library is then analyzed for functions and sequences of interest (Shizuya et al., 1992; Gillespie et al., 2002; Knietsch et al., 2003; Sleator et al., 2008, Gupta et al., 2018). Recent molecular methods used to identify desired genes are direct extraction of genomic DNA, preparation of metagenomic libraries, and sequencing of environmental samples (Henne et al., 1999; Bakshi et al., 2020). Metagenomics can be divided into sequence-based and function-based analyzes of microorganisms. Sequence-based approaches involve screening clones for well-conserved 16S rRNA genes for identification purposes and then

sequencing the entire clone to identify other genes of interest or large-scale sequencing of the entire metagenome to look for phylogenetic anchors in the reconstituted genomes (Ward et al., 1990; Sleator et al., 2008; Gupta et al., 2018). Strategically metagenomics involves the consideration of three categories of data: (i) environmental (metadata); (ii) the community of organisms that environment supports; and (iii) the genes that the organisms contain (Amann et al., 1995; Li, 2015; Quince et al., 2017). The great discoveries in the ecology of microorganisms consist in changing the approach from studies on isolated species in laboratory experiments with controlled conditions to studies of communities of organisms in field conditions. Bioinformatics techniques for managing and analysing both new data types and increased volumes of data transform our understanding of life and its interdependencies. These data sets, together with bioinformatics, improve our understanding about the diversity of microorganisms and their ecological characteristics. Although it has been known for a long time while microorganisms are ubiquitous and play a crucial role in the energy circuit, nutrients, in agriculture, disease and health, our understanding of these organisms was limited to traditional approaches to observation (phenotype). With the development of molecular biology and DNA sequencing technologies was used approach based on sequencing methods to assess species diversity (Wilkening et al., 2009; Li, 2015; Gregor et al., 2016). Regarding the development of a sustainable agriculture program that often uses rhizobacteria that promote plant growth (PGPR), it requires the correct identification and characterization of important microorganisms from an agricultural point of view (Goel et al., 2017). It is impossible to comprehensively understand the functioning of any ecosystem without knowing which organisms are part of it. Thus, the determination of the taxonomic composition of microbial communities in situ is the basis of microbial ecology studies (Coutinho et al., 2018). One of the main goals of ecology is to link the identity of different microorganisms in a habitat with the processes they carry out in that environment (Steele and Streit, 2005). Metagenomics can also fill the lack of scientific knowledge between genetics and ecology. It indicates that the genes of a single microorganism are connected to the genes of other members of the community (Goel et al., 2017). Soil is probably the most challenging of all natural environments for microbiologists in regarding the size of the microbial community and the diversity of the species. The number of distinct prokaryotic genomes has been estimated to range from 2,000 to 18,000 genomes per gram of soil. These figures could be underestimated because the genomes that represent rare and unknown species could have been excluded from these analyses (Rolf, 2005; Granjou et al., 2019; Olson et al., 2019). Soils abound in life, but only in the last two decades have begun genomic approaches to reveal the secret life of soil. Since the soil is no longer understood as an "abiotic" part, it has become a network dense genetics containing various life forms, most of which remain not yet identified. The approaches have begun to be widely applied in the last two decades genomics to reveal soil secrets. The soil is now known as the largest biodiversity reservoir on Earth, with a gram of soil containing up to one million different organisms, most of which remain unidentified and uncharacterized. Optimism about microorganisms is based on the recent development of metagenomics (Edge et al., 2020). Lack of technologies to support visualization and quantitative measurements of the microenvironment soil have greatly hindered the development of soil science research. In the 1960s, most microbial studies have focused on the cultivation of a particular bacterial strain or species isolated in the laboratory. Even today, researchers can grow only about 5% of the strains existing bacteria in the soil under laboratory conditions. However, in the 1990s, the new molecular methods have made it possible to characterize microbial communities and associations by analyzing DNA extracted directly from environmental samples - including soil, water and other materials, such as feces (Grossart et al.,

2020). Microbial ecology examines the diversity and activity of microorganisms. In the last 20 years, the application of genomics tools has revolutionized microbial ecological studies and drastically expanded our view of the microbial world as previously valued. Indeed, microorganisms are still most often considered from an anthropocentric perspective, with a focus on the relatively few species that cause human disease. Microbial ecology examines the diversity of microorganisms and how they interact with each other and with their environment. As a result, microbial ecologists have traditionally focused on two areas of study: (i) microbial diversity (isolation, identification and quantification of microorganisms in different habitats); and (ii) microbial activity (how their activities contribute to microbial diversity and the biogeochemical cycle) (Xu, 2006). Microbial communities are essential components of ecosystems through their contribution to the dynamics of the food web and to biogeochemical processes. Microbial diversity is immense and a general challenge is to understand how the metabolism and interactions of organisms shape the dynamics of the microbial community and ecosystem-wide biogeochemical transformations. Metagenomic approaches have developed rapidly and have proven to be strong in linking microbial community dynamics to biogeochemical processes. Microorganisms are essential components of ecosystems, providing more ecosystem services. Metagenomic approaches have been widely adopted by scientists and are now commonly used in studies of various habitats. About 20 years after the term was invented, metagenomics can now be considered a mature technology, where huge sets of data and an extensive set of analysis tools are accessible to many researchers. The challenge of integrating metagenomics into existing ecological and biogeochemical frameworks and advancing beyond the descriptive science of "what's in that ecosystem?" Remains a challenge. and "what are they capable of doing?" to a better predictive understanding based on mechanisms. The need for such an understanding has never been more relevant in the light of ongoing global climate change and massive man-made changes in ecosystems (Nesme et al., 2016; Grossart et al., 2020). Metagenomics provides the first general information about coexisting (sympatric) populations, because each sequence read is derived from a different individual in a given community. For example, archaeological populations of data from acid mine drainage have been used to show that genetic recombination occurs at a much higher frequency than previously predicted and is the primary evolutionary force shaping these populations (Hugenholtz and Tyson, 2008). One of the main areas of metagenomic research from the beginning it was applied in the discovery of new biocatalysts. A wide range of biocatalysts were obtained from metagenomic libraries. Some examples of biocatalysts derived from soil metagenome include esterases, nitrile hydratases, alcohol reductases, amidases, cellulases, amylases, branching enzymes 1,4 - glucan and pectate liazele (Steele and Streit, 2005). Metagenomics provides a relatively unbiased view not only of community structure (species richness and distribution), but also of the functional (metabolic) potential of a community (Hugenholtz and Tyson, 2008).

### **Bioinformatics method**

MG-RAST is a portal that includes free open access data for processing, analysis, sharing and dissemination of metagenomic data sets. The system contains over 200,000 data sets and is constantly updated (Glass et al., 2010; Wilke et al., 2016, Wilke et al., 2017). The MG-RAST network performs quality control, grouping and annotation based on similarity on nucleic acid sequence data sets using a number of tools bioinformatics. Bioinformatics methods that allow statistical comparisons of libraries built are needed to determine if the differences in the libraries are either artifacts of the sampling and construction of the library, or are caused by changes in composition community. Programs like LIBSHUFF, which has been used for comparison 16S

rRNA gene libraries could be useful for this purpose after further development. To take full advantage of the enormous diversity of soil microorganisms, a combination of Sequence-based approaches of different types of libraries should be used to test soil metagenome. Recently, a third efficiency screening strategy was introduced which is called substrate - induced gene expression cloning (SIGEX) for identification and retrieval of genes encoding catabolic pathways (Rolf, 2005). Approaches to Molecular ecology is advancing rapidly in understanding the microbial communities involved in synthesis and degradation of hydrophobic organic substances involved with major consequences for applications in climate change, environmental pollution, human health and biotechnology. Metagenomics allows researchers to inventory microbial genes in different environments to understand their genetic potential. Metagenomics allows researchers to access functional and metabolic diversity of microbial communities. Since 2008, MG-RAST serves as a repository for metagenomic data sets and as a provider of analyzes. At the moment, the system has analyzed and hosts over 130,000 data sets. Over the years, MG-RAST has undergone a significant number of revisions to accommodate the dramatic increase in the size of the data set, new data types and the wider adoption of systems among the research community. The MG-RAST system provides a web user interface rich covering all aspects of metagenome analysis, from data loading to order analysis. The web interface can also be used to discover data. Metagenomic data sets can be easily selected individually or based on filters such as technology (including reading length), quality, sample type and keyword, with filtering dynamics of results based on similarity to known reference proteins or taxonomy. The results can be displayed in familiar formats, including bar charts, trees which incorporates information about the abundance, maps or analyzes of the components main, or exported in tabular form. Raw or processed data can be recovered by via download pages (Glass and Meyer, 2015). The MG-RAST website is rich in functionality and offers a lot of different options. The site from <http://metagenomics.anl.gov> has five main pages and one main page: download - lists all publicly available data for download. The data is structured in projects. Page navigation - allows interactive navigation of all data sets and is powered by metadata. Search page - allows identification, taxonomy and searches based on functions against all public data. Analysis page - allows analysis in-depth comparisons between data sets. Upload page - allows users to provide their samples and metadata to MG-RAST. Home (presentation overview) - provides an overview for each individual data set (Meyer et al., 2008; Glass and Meyer, 2015).

#### 4. CONCLUSIONS

The genome of the dominant species can be completely reconstructed from environmental samples using random sequencing. Having the complete or almost complete genome of a population dominant we obtain the inventory of genes for the body this allowing the determination its key metabolic potential. A key feature of genomes obtained from the environment is that they are composite samples of the population from which they were obtained and comprise the genetic microheterogeneity present in that population. Metagenomics allows scientists to investigate prokaryotic ecology more fully and unlock the vast biotechnological potential of the prokaryotic population.

In conclusion, metagenomics is a mix of genomics, bioinformatics and biological systems that can be used to study the genomes of many organisms at the same time.

## 5. ACKNOWLEDGEMENTS

I thank my colleagues Dr. Onete Marilena and Dr. Manu Minodora for supervising the realization of this work. The results were obtained within: the doctoral project (Doctoral School of Biology, Faculty of Biology, University of Bucharest), project RO1567-IBB01 / 2021 and PED project PN-III-CERC-CO-PED-2-2019 (RESET).

## 6. REFERENCES

- Amann, R. I., Ludwig, W., & Schleifer, K. H. (1995). Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiological reviews*, 59(1), 143-169.
- Ayling, M., Clark, M. D., & Leggett, R. M. (2020). New approaches for metagenome assembly with short reads. *Briefings in bioinformatics*, 21(2), 584-594.
- Bakshi, A., Moin, M., & Madhav, M. S. (2020). Metagenomics in Agriculture: State-of-the-Art. In *Metagenomics: Techniques, Applications, Challenges and Opportunities* (pp. 167-187). Springer, Singapore.
- Carbonetto, B., Rascovan, N., Álvarez, R., Mentaberry, A., & Vázquez, M. P. (2014). Structure, composition and metagenomic profile of soil microbiomes associated to agricultural land use and tillage systems in Argentine Pampas. *PLoS one*, 9(6), e99949.
- Coutinho, F. H., Gregoracci, G. B., Walter, J. M., Thompson, C. C., & Thompson, F. L. (2018). Metagenomics sheds light on the ecology of marine microbes and their viruses. *Trends in microbiology*, 26(11), 955-965.
- Coyotzi, S., Pratscher, J., Murrell, J. C., & Neufeld, J. D. (2016). Targeted metagenomics of active microbial populations with stable-isotope probing. *Current opinion in biotechnology*, 41, 1-8.
- Edge, T. A., Baird, D. J., Bilodeau, G., Gagné, N., Greer, C., Konkin, D., ... & Macklin, J. (2020). The Ecobiomics project: Advancing metagenomics assessment of soil health and freshwater quality in Canada. *Science of The Total Environment*, 710, 135906.
- Esposito, A., Colantuono, C., Ruggieri, V., & Chiusano, M. L. (2016). Bioinformatics for agriculture in the Next-Generation sequencing era. *Chemical and Biological Technologies in Agriculture*, 3(1), 1-12.
- Gilbert, J. A., & Dupont, C. L. (2011). Microbial metagenomics: beyond the genome. *Annual review of marine science*, 3, 347-371.
- Gillespie, D. E., Brady, S. F., Bettermann, A. D., Cianciotto, N. P., Liles, M. R., Rondon, M. R., ... & Handelsman, J. (2002). Isolation of antibiotics turbomycin A and B from a metagenomic library of soil microbial DNA. *Applied and environmental microbiology*, 68(9), 4301-4306.
- Glass, E. M., & Meyer, F. (2015). MG-RAST, a metagenomics service for the analysis of microbial community structure and function. In *Hydrocarbon and Lipid Microbiology Protocols* (pp. 69-87). Springer, Berlin, Heidelberg.
- Glass, E. M., Wilkening, J., Wilke, A., Antonopoulos, D., & Meyer, F. (2010). Using the metagenomics RAST server (MG-RAST) for analyzing shotgun metagenomes. *Cold Spring Harbor Protocols*, 2010(1), pdb-prot5368.
- Goel, R., Suyal, D. C., Dash, B., & Soni, R. (2017). Soil metagenomics: a tool for sustainable agriculture. In *Mining of microbial wealth and metagenomics* (pp. 217-225). Springer, Singapore.
- Granjou, C., & Phillips, C. (2019). Living and labouring soils: Metagenomic ecology and a new agricultural revolution?. *BioSocieties*, 14(3), 393-415.
- Gregor, I., Schönhuth, A., & McHardy, A. C. (2016). Snowball: strain aware gene assembly of metagenomes. *Bioinformatics*, 32(17), i649-i657.
- Grob, C., Taubert, M., Howat, A. M., Burns, O. J., Dixon, J. L., Richnow, H. H., ... & Murrell, J. C. (2015). Combining metagenomics with metaproteomics and stable isotope probing reveals metabolic pathways used by a naturally occurring marine methylotroph. *Environmental microbiology*, 17(10), 4007-4018.
- Grossart, H. P., Massana, R., McMahon, K. D., & Walsh, D. A. (2020). Linking metagenomics to aquatic microbial ecology and biogeochemical cycles. *Limnology and Oceanography*, 65, S2-S20.
- Gupta, N., Vats, S., & Bhargava, P. (2018). Sustainable agriculture: role of metagenomics and metabolomics in exploring the soil microbiota. In *In Silico Approach for Sustainable Agriculture* (pp. 183-199). Springer, Singapore.
- Henne, A., Daniel, R., Schmitz, R. A., & Gottschalk, G. (1999). Construction of Environmental DNA Libraries in *Escherichia coli* and Screening for the Presence of Genes Conferring Utilization of 4-Hydroxybutyrate. *Applied and Environmental Microbiology*, 65(9), 3901-3907.

- Hugenholtz, P., & Tyson, G. W. (2008). Metagenomics. *Nature*, 455(7212), 481-483.
- Kirubakaran, R., ArulJothi, K. N., Revathi, S., Shameem, N., & Parray, J. A. (2020). Emerging priorities for microbial metagenome research. *Bioresource Technology Reports*, 100485.
- Knietsch, A., Waschkowitz, T., Bowien, S., Henne, A., & Daniel, R. (2003). Metagenomes of complex microbial consortia derived from different soils as sources for novel genes conferring formation of carbonyls from short-chain polyols on *Escherichia coli*. *Journal of molecular microbiology and biotechnology*, 5(1), 46-56.
- Kouzuma, A., Ishii, S. I., & Watanabe, K. (2018). Metagenomic insights into the ecology and physiology of microbes in bioelectrochemical systems. *Bioresource technology*, 255, 302-307.
- Li, L. (Ed.). (2015). *Infectious Microecology: Theory and Applications*. Springer.
- Lorenz, P., & Schleper, C. (2002). Metagenome—a challenging source of enzyme discovery. *Journal of Molecular Catalysis B: Enzymatic*, 19, 13-19.
- Macher, J. N., Speksnijder, A., Choo, L. Q., van Der Hoorn, B., & Renema, W. (2019). Uncovering bacterial and functional diversity in macroinvertebrate mitochondrial-metagenomic datasets by differential centrifugation. *Scientific reports*, 9(1), 1-9.
- Meyer, F., Paarmann, D., D'Souza, M., Olson, R., Glass, E. M., Kubal, M., ... & Edwards, R. A. (2008). The metagenomics RAST server—a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC bioinformatics*, 9(1), 1-8.
- Miliute, I., Buzaitė, O., Baniulis, D., & Stanys, V. (2015). Bacterial endophytes in agricultural crops and their role in stress tolerance: a review. *Zemdirbyste-Agriculture*, 102(4), 465-478.
- Narasimgarao, P., Podell, S., Ugalde, J. A., Brochier-Armanet, C., Emerson, J. B., Brocks, J. J., ... & Allen, E. E. (2012). De novo metagenomic assembly reveals abundant novel major lineage of Archaea in hypersaline microbial communities. *The ISME journal*, 6(1), 81-93.
- Nesme, J., Achouak, W., Agathos, S. N., Bailey, M., Baldrian, P., Brunel, D., ... & Simonet, P. (2016). Back to the future of soil metagenomics. *Frontiers in microbiology*, 7, 73.
- Olson, N. D., Treangen, T. J., Hill, C. M., Cepeda-Espinoza, V., Ghurye, J., Koren, S., & Pop, M. (2019). Metagenomic assembly through the lens of validation: recent advances in assessing and improving the quality of genomes assembled from metagenomes. *Briefings in bioinformatics*, 20(4), 1140-1150.
- Parks, D. H., & Beiko, R. G. (2010). Identifying biologically relevant differences between metagenomic communities. *Bioinformatics*, 26(6), 715-721.
- Quince, C., Walker, A. W., Simpson, J. T., Loman, N. J., & Segata, N. (2017). Shotgun metagenomics, from sampling to analysis. *Nature biotechnology*, 35(9), 833-844.
- Rolf, D. (2005). The metagenomics of soil. *Nature Reviews Microbiology*, 3(6), 470-478.
- Rondon, M. R., August, P. R., Bettermann, A. D., Brady, S. F., Grossman, T. H., Liles, M. R., ... & Goodman, R. M. (2000). Cloning the soil metagenome: a strategy for accessing the genetic and functional diversity of uncultured microorganisms. *Applied and environmental microbiology*, 66(6), 2541-2547.
- Shizuya, H., Birren, B., Kim, U. J., Mancino, V., Slepak, T., Tachiiri, Y., & Simon, M. (1992). Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in *Escherichia coli* using an F-factor-based vector. *Proceedings of the National Academy of Sciences*, 89(18), 8794-8797.
- Singh, H., Sharma, P., Kaur, R. P., Thakur, D., & Kaur, P. (2020). Computational Metagenomics: State-of-the-Art, Facts and Artifacts. In *Metagenomics: Techniques, Applications, Challenges and Opportunities* (pp. 199-227). Springer, Singapore.
- Sleator, R. D., Shortall, C., & Hill, C. (2008). Metagenomics. *Letters in applied microbiology*, 47(5), 361-366.
- Steele, H. L., & Streit, W. R. (2005). Metagenomics: advances in ecology and biotechnology. *FEMS microbiology letters*, 247(2), 105-111.
- Thomas, T., Gilbert, J., & Meyer, F. (2012). Metagenomics—a guide from sampling to data analysis. *Microbial informatics and experimentation*, 2(1), 1-12.
- Ward, D. M., Weller, R., & Bateson, M. M. (1990). 16S rRNA sequences reveal numerous uncultured microorganisms in a natural community. *Nature*, 345(6270), 63-65.
- Wilke, A., Bischof, J., Gerlach, W., Glass, E., Harrison, T., Keegan, K. P., ... & Meyer, F. (2016). The MG-RAST metagenomics database and portal in 2015. *Nucleic acids research*, 44(D1), D590-D594.
- Wilke, A., Gerlach, W., Harrison, T., Paczian, T., Trimble, W. L., & Meyer, F. (2017). *MG-RAST Manual for version 4, revision 3*.
- Wilkening, J., Wilke, A., Desai, N., & Meyer, F. (2009). Using clouds for metagenomics: a case study. In *2009 IEEE International Conference on Cluster Computing and Workshops* (pp. 1-6).

- Xu, J. (2006). Invited review: microbial ecology in the age of genomics and metagenomics: concepts, tools, and recent advances. *Molecular ecology*, 15(7), 1713-1731.
- Zeyauallah, M., Kamli, M. R., Islam, B., Atif, M., Benkhayal, F. A., Nehal, M., ... & Ali, A. (2009). Metagenomics-An advanced approach for noncultivable micro-organisms. *Biotechnology and Molecular Biology Reviews*, 4(3), 49-54.