GMR
The Original

METAGENOMIC ANALYSIS OF MICROBIAL COMMUNITIES FOR BIOREMEDIATION OF CONTAMINATED ENVIRONMENTS

Dr. Saji Sivan S[1], Valli Nachiyar C[2], Nisha Boopathy[3], Tharani Munusamy[4], Dr. Premal C. Patel[5], Dr Latha Narayan[6], Dr. Megha Ojha[7]

[1]Associate Professor, Vitsol, Vit University, Chennai Campus.
[2]Professor, Department Of Research, Meenakshi Academy Of Higher Education And Research
[3]Associate Professor, Community Medicine, Saveetha Medical College, Saveetha Institute Of Medical And Technical Sciences
[4]Department Of Research, Meenakshi Academy Of Higher Education And Research
[5]Associate Professor, Department Of Information Technology, College Of Technology, Gujarat, India.
[6]Teaching Faculty, Department Of Zoology, Jnana Bharathi Campus, Bangalore University, Bangalore-560056, Karnataka, India
[7]Associate Professor Of Law, Klef College Of Law, Kl (Deemed-To-Be) University, Vaddeswaram, Guntur,

ABSTRACT

With the advent of fast industrialization, the extensive agricultural use of cultured land and urban growth, environmental pollution has become the top priority worldwide, and it is extremely dangerous to the ecological state and human health. Traditional remediation solutions, such as physical and chemical processes, tend to be costly, consume a lot of energy, and can also produce secondary pollutants, which is why more acceptable and environmentally-friendly solutions are required. Microbial bioremediation has become one of the highly effective methods that leverage the metabolic diversity within microorganisms to turn, degrade, or immobilise a broad variety of environmental pollutants. More recently in metagenomics, improvement in this area has come in with the ability to study microbial communities at a comprehensive, culture-independent level that can give a deeper insight on the cornerstone of metabolic capability, functional genes, and taxonomic diversity of microbial communities. This review summarises the recent advances of metagenomic technologies, and their role in determining the key genes and enzymatic pathways that are used in the hydrocarbon, heavy metal, and other toxic compounds degradation. It also emphasises the significance of higher-level bioinformatics tools with regard to both the sequence analysis and the functional annotation as well as pathway reconstruction. Though this has made significant advances, issues including incomplete assembly of genome, constraints of full experimental validation and a large level of computation remain. The field of the future research needs to be integrating multi-omics approach and artificial intelligence-based analytics to improve the accuracy, scalability, and effectiveness of bioremediation systems based on metagenomics.

KEYWORDS: Metagenomics, Bioremediation, Microbial Communities, Functional Genes, Environmental Genomics, Pollutant Degradation

1. INTRODUCTION

The 21st-century world has placed environmental pollution among the highest priority global issues due to the blistering industrialization, urban growth, and extensive agricultural activity (Das and Chandran, 2011; Varjani, 2017). This has caused very high environmental degradation and human health risks because of the excessive colonialization of soil and water bodies by the release of heavy metals, hydrocarbons, pesticides, and industrial effluents (Ghosal et al., 2016). Such pollutants tend to be persistent, bioaccumulative, and toxic and are therefore difficult and of high importance to remove them out of the environment (Goswami et al., 2018). Moreover, the pollution of such natural resources like groundwater or agricultural soil also has the direct effect on food security, biodiversity, and the stability of the overall ecosystem (Fierer, 2017). Physical excavation, chemical treatment, and thermal degradation are also conventional remediation methods that have been widely used to solve environmental pollution (Tripathi and Singh, 2022). Nevertheless, it is necessary to relate these methods to the high cost of operation, energy consumption, and a possible formation of secondary pollutants (Mir et al., 2022). Moreover, several of the products cannot be applied in large-scale applications or in situ, which prevents their use in the long-term restoration of the environment (Goswami et al., 2018). Such constraints have motivated the process of seeking more sustainable, cost effective and environmental benign remediation solutions (Varjani, 2017). Microbial bioremediation is also a promising option, where microorganisms are used as natural metabolic processes to degrade, transform, or immobilise a broad spectrum of environmental pollutants (Das & Chandran, 2011; Ghosal et al., 2016). Heterogeneous microbial communities have specialised enzymes and biochemical pathways that facilitate the degradation of complex organic substances and inactivation of dangerous substances (Varjani, 2017). The efficacy and flexibility of bioremediation processes in the conditions of different environmental conditions are additionally complemented by the mutual impact of microbial populations (Fierer, 2017). Recently, a revolution in the scientific method of environmental microbiology has occurred following the introduction of the field of metagenomics concerned with culture-independent studies of microbial communities (Thomas et al., 2012; Wooley et al., 2010). Older microbiological methods have the drawback of not being able

to cultivate a large percentage of microbes in the environment in the laboratory (Simon & Daniel, 2011). Metagenomics has overcome this and it by itself analyses the extracted genetic material in the form of environmental samples and in so doing it gives detailed information on the diversity of the microbes found, the functional genes and metabolic pathways involved in pollutant degradation (Suenaga, 2012; Thomas et al., 2012). Such a method has greatly broadened the knowledge of known complex microbial ecosystems and their functional capacity in the bioremediation process (Fierer, 2017). Besides, the recent technological improvements in the high-throughput sequencing systems and bioinformatical software have made it easier to conduct large-scale metagenomic research, thus providing a precision in identities of genes, enzymes and pathways that degrade contaminants (Shokralla et al., 2012; Li et al., 2015). The new opportunities have led to designing specific and effective bioremediation strategies, as well as the integration of metagenomics and other forms of omics to provide a systems-level insight of the environmental processes (Knight et al., 2018; Kim et al., 2010). This review will set out to critically and succinctly review the metagenomic method in the study of contaminated environment microbial communities, and specifically discuss the role of such methods in bioremediation. It also analyses modern-day technological developments, essential functional genes, bioinformatics tools and quantitative models as well as discusses the current issues and perspectives of future research in this fast developing field.

2. LITERATURE REVIEW

Bioremediation as an area has developed a lot within recent decades to a new metagenomic based method, which makes analysis of the microbial community complete instead of a culture based approach (Simon and Daniel, 2011). Primary early bioremediation studies were based on culture dependent approaches to isolate and characterise microorganisms that could be used to degrade environmental pollutants (Das & Chandran, 2011). Although these methods offered useful information about the metabolism and activities of the microbes, it was only naturally crippled by the reality that too large a majority of the environment microorganisms cannot be cultured under laboratory standard conditions (Wooley et al., 2010). This weakness limited the knowledge of diversity related to microbes and impeded the recognition of new functional genes in degradation of pollutants (Suenaga, 2012). The development of metagenomics has facilitated a shift in the paradigm of environmental microbiology to provide culture-independent information regarding the microbe communities (Thomas et al., 2012). Metagenomics can be used to provide a more detailed characterization of microbial diversity, functional potential and metabolic pathways by extracting and sequencing genetic material of direct interest (environmental samples) which is then transformed into a format that can be analysed (Shokralla et al., 2012). The initial work in the field was concerned with 16S rRNA gene sequence, which gave taxonomic data on the populations of microbes (Langille et al., 2013). Nevertheless, these methods were not able to exhaustively clarify the functional capabilities thus leading to the consideration of more sophisticated methods (Simon and Daniel, 2011). Over the last several years (20152025), with great progress in high-throughput sequencing technologies and analysis software, the use of metagenomics in a bioremediation use has gained momentum (Li et al., 2015; Luo et al., 2015). Shotgun metagenomics has made possible the whole-genome study of microbial communities, which offers comprehensive data on both their taxonomic properties and on their functional genes (Bhattacharjee et al., 2017). Genome-resolved metagenomics expands on this possibility further by assembling genomes that are metagenome-assembled (MAGs), which can be used to gain a comprehensive view of metabolic pathways and gene networks (Luo et al., 2015). Also the application of functional metagenomics has aided the discovery of actively expressed genes and enzymes engaged in the degradation of pollutants that provides a direct correlation between genetic capability and environmental role (Suenaga, 2012). Table 1 provides a comparative description of key metagenomic methods applied in bioremediation projects, and key features of the methodologies, their advantages, and disadvantages. Amplicon sequencing techniques provide an excellent tool to use in diversity profiling of microorganisms because they are cost-effective, but they can offer limited functional data (Langille et al., 2013). Comparatively, shotgun metagenomics provides detailed genome information at a high cost (Li et al., 2015). Functional and genome-resolved metagenomics are thought to be able to give more insight into active genes and metabolic processes, yet can present experimental and analytical complexities (Bhattacharjee et al., 2017). Moreover, multi-omics methodologies, such as metatranscriptomics and proteomics, have become a potent tool of getting system-level insights on the processes in microorganisms, though they are more expensive and technically challenging (Kim et al., 2010; Knight et al., 2018).

Table 1. Comparative Trends in Metagenomic Approaches for Bioremediation Studies

Approach Type	Methodology Focus	Typical Environment	Strengths	Limitations
Amplicon-Based Analysis	16S rRNA sequencing	Soil, water	Cost-effective, diversity profiling	Limited functional insight
Shotgun Metagenomics	Whole-genome sequencing	Marine, industrial waste	High-resolution data	High computational cost
Functional Metagenomics	Gene screening	Contaminated soil	Detects active genes	Experimental complexity
Genome-Resolved Metagenomics	MAG reconstruction	Complex ecosystems	Pathway identification	Partial genome recovery

Multi-Omics Integration	Integrated omics	Advanced systems	System-level insight	Expensive

In spite of this development, there are still a number of research gaps in the implementation of metagenomics in bioremediation. The absence of functional validation of the predicted genes and pathways is considered as one of the key challenges because most studies are heavily based on computational inference and not necessarily supported with experimental validation. Moreover, the majority of metagenomic work is carried out in controlled laboratory settings, and it is not well applied to large-scale environmental systems. Moreover, the combination of metagenomics and the newly developed technologies, related to artificial intelligence and multi-omics, is in its initial phases, which points to the necessity of larger-scale and more comprehensive tools. These challenges are critical issues that need to be resolved to increase metagenomics-based bioremediation to practical and sustainable uses of the environment.

3. Metagenomic Approaches and Microbial Community Dynamics

3.1 Types of Metagenomics

Metagenomics has become a potent instrument in the study of microbial communities in polluted habitats, allowing both taxonomic diversity and functional potential to be studied comprehensively without having to grow any bacteria. Metagenomic methods can be broadly categorised as amplicon sequencing, shotgun metagenomics and functional metagenomics depending on the sequencing strategy, and the analytical and/or functional goals. Amplicon sequencing is among the most commonly used methods in the profiling of the microbial community composition. This technique is normally aimed at preserved genetic markers like the 16S rRNA gene in bacteria and archaea in order to determine and categorise microorganisms found in environmental samples. The use of amplicon based studies is cheap and they offer important information on the diversity of microbes within the community, structure and the relative abundance of microorganisms. They however have a restriction to the capability of exposing operational genes and metabolic tracts in pollutant degradation. Shotgun metagenomics, by contrast, requires the random next-generation sequencing of each and every fragment of genetic material recovered off the environment. In this way, a complete characterization of taxonomic and functional traits of microbial communities is possible. Shotgun metagenomics is an excellent tool to determine the metabolic pathways, gene networks, and ecological interactions, which is why it is particularly well adapted to analyse complex bioremediation processes through the study of whole genomes. Although the method has some merits, it has a high cost in terms of computing resources as well as advanced bioinformatics systems to process and interpret the results. Functional metagenomics puts emphasis on the identification of active expressed genes and enzymes in a defined manner of biochemical degradation, especially the degradation of pollutants. It is a method of early cloning of environmental DNA into appropriate hosts and selection of functional activities, like enzymatic breakdown of contaminants. Functional metagenomics helps in sealing the divide that exists between the theoretical and practical possibilities of a biological process, leading to the identification of novel enzymes and metabolic pathways. It is however, experimentally intensive and is usually constrained by gene expression efficiency in heterologous systems. Combined, these metagenomic methods have resulted in complementary understanding of microbial communities and their functional potential which is the basis of enhanced bioremediation schemes.

3.2 Metagenomic Workflow

Bioremediation through metagenomic analysis is done in a form of a systematic chain of actions that make it possible to identify the diversity of microbes and their functional genes that help in the degradation of the pollutants. This is initiated by the environmental sampling in which soils, sludge, or water samples of contaminated areas are taken. This is then followed by DNA extraction whereby total environmental DNA (eDNA) is extracted out of the sampled collected, which is the genetic material of all the microorganisms present. The DNA is extracted and followed by high-throughput sequencing on a large scale using ultra-high-tech instruments like the Illumina or the Nanopore systems. Such sequencing methods produce genomic data in masses, which are then subjected to the bioinformatics analysis. It involves quality control, assembling of sequences, annotation and taxonomic and functional profiling to determine the metabolic capacity and species of the microbes. Lastly, the workflow facilitates the functional identification of genes, in which particular genes and pathways are defined during the process of pollutant degradation, e.g. the genes of hydrocarbon-degradation, the genes of heavy metal resistance, etc. This information is critical in the realisation of the microbial role in bioremediation and in creating effective strategies of environmental remedies. This metagenomic core flow as utilised in bioremediation is depicted in Figure 1 which is an overview of the important steps undertaken in the course of the environmental sampling to functional gene discovery.

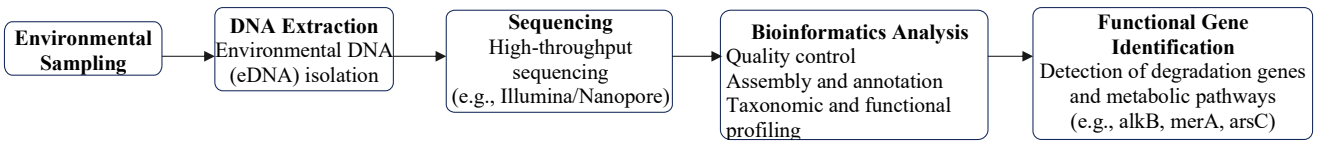

Figure 1. Metagenomic workflow for bioremediation of contaminated environments

3.3 Microbial Community of Contaminated Environments.

The microbial communities are very critical in natural attenuation and bioremediation of contaminated systems. The composition, diversity and functional potential of these communities is quite different according to the environment type there and the character of the pollutants there. These microbial dynamics can be used to understand how best the metagenomics-based bioremediation approaches can be optimised. Soil microbiomes have been considered to be one of the most diverse and functionally abundant microbial communities, especially found in polluting environments such as agricultural lands and industrial spill areas. Microorganisms in the soils have a large diversity of metabolic activities which allows the breakdown of organic pollutants, e.g. hydrocarbons, pesticides, and xenobiotic substances. The occurrence of contaminants usually causes a selective enrichment of populations of polluting based microbes, thus changing the indigenous system of the community, and functional gene expressional profiles. Complex microbial communities also exist in the waters of aquatic ecosystems (both fresh water and marine) that assist in the breakdown of pollutants and the cycling of nutrients. Bacteria which decompose hydrocarbon molecules like Alcanivorax and Marinobacter features prominently in oil contaminated sea environments to decompose petroleum compounds. In the same way, wastewater systems promote microbial consortia that have an ability to degenerate organic material and decontaminate problematic chemicals, which are crucial aspects of planned bioremediation steps. Specialised microbial communities adapted to extreme conditions are found in industrial environments and extreme settings like in mining sites, landfills, high-salinity or high-temperature environments. These extremophiles have got special metabolic processes which allow them to tolerate and convert toxic elements including heavy metals and industrial chemicals. They are good candidates of metagenomic discovery and possible biotechnological development due to their robustness and functional diversity. Besides personal microbial potentials, there is a significant role of microbial interactions in improving the action of bioremediation. Microbial consortia can also deimate complex pollutants more efficiently than single species because of the existence of synergistic relationships, including metabolic cooperation and cross-feeding. The interactions would help in stabilizing and adapting the microbial communities in different environmental conditions. Altogether, understanding microbial communities of various contaminated sites can be of critical importance in their ecological services and functionality potential, which makes it possible to create more efficient and specific bioremediation measures.

4. Functional Metagenomics and Bioinformatics Analysis

4.1 Functional Genes in Bioremediation

The expression and presence of certain functional genes that encode an enzyme with the ability to degrade or transform environmental pollutants largely determine microbial bioremediation. Metagenomic studies have facilitated the detection of these genes in environmental samples and this has led to better understanding of the metabolic capabilities of microbial assemblies that take part in bioremediation events. The most significant of them are hydrocarbon degradation genes that are of significant importance in the degradation of petroleum-based contaminants. Examples of such genes include the alkB genes that encode the alkane monooxygenases responsible to catalyse the first oxidation of alkanes to alcohols, which is a major process in hydrocarbon degradation pathways. Likewise, nah genes are also implicated in the breakdown of the aromatic hydrocarbons like naphthalene to convert them into intermediate substances that can then proceed to be broken down by microbial enzymes. Besides the degradation of organic pollutants, the microorganisms also have specialised genes, which make them resistant to heavy metals which are very toxic. MerA and arsC are heavy metal resistance genes required during the process of detoxification. Mercuric reductase, which is produced by merA gene, pumps the toxic amount of mercury ions to elemental mercury and more harmless, whereas arsenate reductase, produced by arsC gene, pumps the arsenate to arsenite, after which they can be further eliminated or sequestered by the microbial systems. Table 2 summarises some of the important functional genes in the bioremediation process, the pollutants that they degrade and their biological roles.

Table 2. Functional Genes and Their Roles in Bioremediation

Gene	Pollutant Type	Function	Representative Microorganisms
alkB	Hydrocarbons	Alkane degradation via oxidation	*Pseudomonas*, *Alcanivorax*
nahA	Aromatic hydrocarbons	Naphthalene degradation	*Rhodococcus*, *Pseudomonas*
merA	Mercury	Reduction of Hg^{2+} to elemental Hg	*Bacillus*, *Pseudomonas*
arsC	Arsenic	Arsenate reduction and detoxification	*E. coli*, *Staphylococcus*

4.2 Metabolic Pathways

There are essentially complex metabolic pathways in which microbial bioremediation is fundamentally driven through transformations, degradations or detoxification of environmental pollutants. The pathways are facilitated by an array of enzymes using functional genes that enable such microorganisms to either use contaminants as a source of energy or carbon. Metagenomic research has formed an immense contribution to the study of these metabolic processes pointing to the variety and benefit of microbial enzyme systems in polluted systems. Enzymatic degradation is one of the leading process involved in the bioremediation, and through this, the microorganisms produce particular enzymes, which catalyse the breakage of the complex pollutant into less toxic and simpler substances. Oxygenase and monooxygenase enzymes, as an example, are essential in the breakdown of hydrocarbons, in the process of adding oxygen atoms to the non-reactive organic compound, which in turn makes the organic compound more reactive and hence allows further breakage. Equally, aromatic rings cleavages facilitate the degradation of persistent organic pollutants including polycyclic aromatic hydrocarbons (PAHs) and require dioxylases. These enzyme related processes are commonly involved in series of metabolic processes, which end up in the transformation of pollutants into intermediate substances leading to the tricarboxylic acid (TCA) cycle. Besides enzyme degradation, redox reactions are key to the deltoxification of inorganic contaminants, especially those of the heavy metals. Reduction and oxidation of the metal ions by microorganisms can change the oxidation state of the metal ions and hence affect the solubility, mobility and toxicity of the metal. As an illustration, the reduction of toxic hexavalent chromium (Cr^{6+}), to the less toxic trivalent form (Cr^{3+}), or mercury ions (Hg^{2+}) to elemental mercury (Hg^{0}), are major detoxification processes catalysed by microbial enzymes. These redox reactions do not only decrease toxication of the environment, they also aid the elimination or fixation of toxins. Moreover, microbial metabolic processes also tend to work cooperatively as they are found in microbial consortia whereby the intermediate metabolites generated by one microbe are further metabolic by others. This beneficial symbiotic metabolism gives the system better efficiency in pollutant destruction especially in the lacking and heterogenous environment. These correlated pathways have been revealed through metagenomic studies, and so has the overall contribution of microbial communities to the process of bioremediation. Altogether, redox-based transformations and enzymatic degradation are the main basis of the key microbial metabolism in bioremediation and can serve crucial information on the creation of effective and sustainable environmental remediation strategies.

4.3 Bioinformatics Tools

Metagenomic data analysis is based on a networked bioinformatics processing relying in transforming the raw sequencing data into biologically significant insights. The raw data used in the process are raw sequencing reads (FASTQ files) produced out of high-throughput sequencing systems. The reads used are as the input in downstream computing. The initial stage entails quality control, during which the sequencing artefacts, as well as low-quality reads, are eliminated with the aid of trimming and filtering of datasets by the help of FastQC and Trimmomatic. This makes the following analyses reliable and accurate. After quality control, the filtered reads are then assembled in which overlapping reads are joined together to produce longer contiggering sequences (contigs) of their assemblers like MEGAHIT or SPAdes. This will be necessary in reassembling genomic fragments in complex microbial communities. These assembled sequences are annotated by incorporating gene prediction and mapping onto various databases, such as KEGG, NCBI and UniProt. It is a step that can be used in identifying functional genes and their biological roles. Lastly, the functional pathway mapping is undertaken in order to connect annotated genes to the metabolic and degradation pathways taking place in bioremediation processes. Figure 2 shows the general bioinformatics pipeline, which consists of raw sequencing reads, up to functional pathway identification.

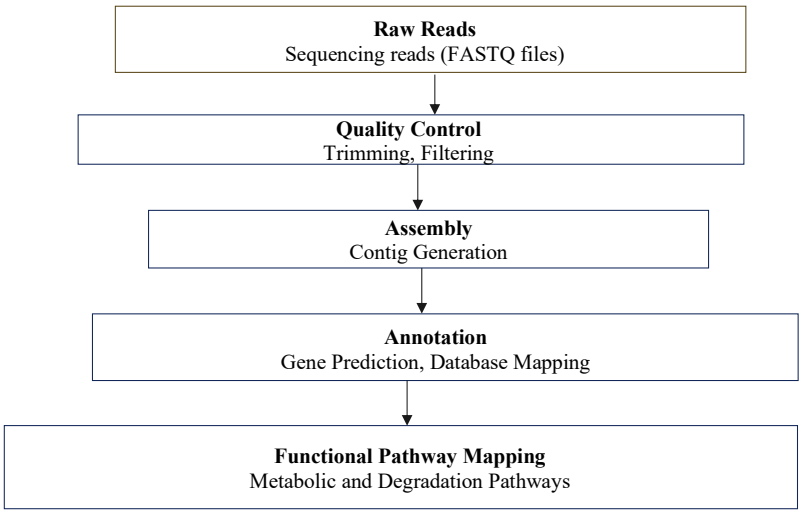

Figure 2. Bioinformatics pipeline for metagenomic data analysis

5. Quantitative Models in Bioremediation

Quantitative models are very important in the study of microbial diversities and forecasting the dynamics of pollutant degradation during bioremediation processes. Using these models is the mathematical explanations of the microbial community structure, and the efficiency of contaminants removal with time course. The Shannon Diversity index is one of the most popular and commonly used diversity metrics of the microbes in a metagenome study that determines the richness and evenness of species within a given micro community. It is expressed as:

$$H' = -\sum_{i=1}^{S} p_i \, lnp_i$$

where p_i represents the relative abundance of the i^{th}species, and S denotes the total number of species in the community. The higher the Shannon index, the more diverse the microbes, and the higher the functional stability and resilience of bioremediation systems are related to its high value. Besides the evaluation of diversity, the kinetics of pollutant degradation are also evaluated by the use of mathematical models. One of the widely used models is the first-order biodegradation model that takes into account that the degradation rate is constant depending on the concentration of the pollutant. This model is given by:

$$C_t = C_0 e^{-kt}$$

where C_t is the pollutant concentration at time t, C_0 is the initial concentration, and k is the degradation rate constant. The model is very popular in estimating degradation efficiency and also in comparing the activity of various bioremediation strategies. The combination of these quantitative models would be a useful addition in understanding the dynamics of microbial community behaviour and pollutant transformation processes to assist in designing and optimization of efficient bioremediation processes.

6. Metagenomics uses in Bioremediation.

It has been observed that metagenomics has become a revolutionary technique in environmental biotechnology and facilitated the discovery of functional genes and microbial communities engaging in the degradation of various pollutants. Its use in bioremediation has greatly enhanced the effectiveness and selectivity and scalability of the environmental cleanup procedures. Among the most notable of them is the use in oil spill cleanup, where metagenomic analysis is used to identify hydrocarbon-degrading microorganisms and the related genes that cause the degradation of petroleum compounds. Some microbial species, including Alcanivorax and Pseudomonas, are commonly overgrown in oil-contaminated habitats and with help of metagenomics, their metabolic routes can be characterised to implement them to best bioremediation approaches. Metagenomics aids in identifying the resistance genes and metabolic interactions in the process of detecting the transformation and immobilisation of toxic metals like mercury, arsenic, and chromium during the process of heavy metal detoxification. Such observations can be used to create specialised microbial consortia that will help to lower the levels of metal toxicity in polluted environments. Among the wastewater treatment processes, Metagenomics also plays a crucial role because complex microbial communities are involved in the degradation of organic pollutants and nutrient extraction. Metagenomics can be used to maximise the efficiency of treatment and increase the functionality of the system through the study of microbial diversity and functional genes. Moreover, in the cleanup process of industrial waste, a diverse of metagenomic methods can be used to identify microorganisms that can be used to break down dangerous chemicals and xenobiotic compounds. This assists the construction of effective bioremediation processes of industrial effluent and contaminated location. Combination of metagenomics and high-performance computation, engineered microbial systems, and environmental monitoring technologies can also be revealed in Figure 3 that provides a detailed scheme of metagenomics-mediated bioremediation. The figure illustrates the relationship between the microbial communities, metagenomic analysis, bioinformatics-guided information, and the processes of the pollutant degradation resulting in regeneration of the environment.

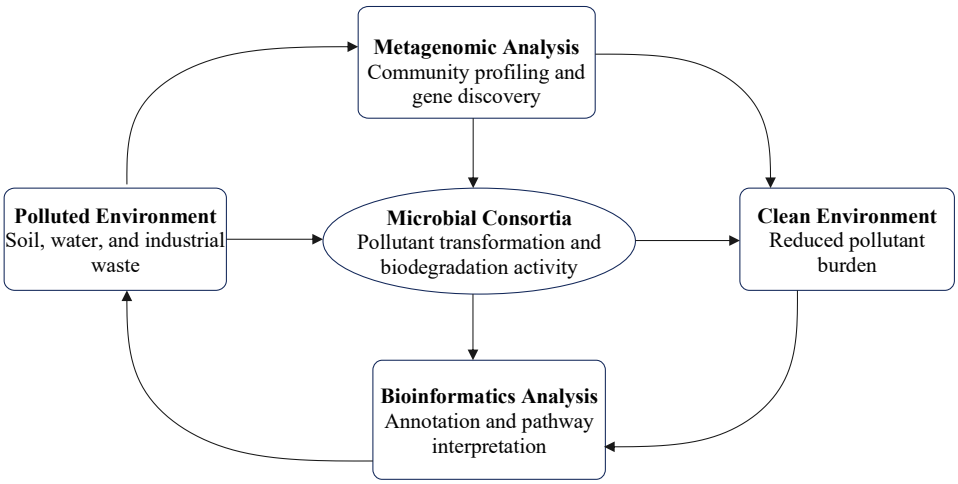

Figure 3. Integrated framework of metagenomics-based bioremediation

7. FUTURE DIRECTIONS AND CHALLENGES

Although metagenomics has greatly improved in the context of bioremediation, a number of difficulties are still seen to limit the full-scale application of this technique in the system. The complexity of microbial ecosystems is one of the major threats with various and dynamic microbial communities that respond to different environmental conditions. This complication places it hard to realistically conduct microbial operations and foresee their action in real-life bioremediation conditions. The other significant constraint is incompleteness in genome reconstruction especially in complex systems where the metagenomic information may lead to assembly fragments. The assembly of entire genomes of metagenomes (MAGs) is still not yet effectively constructed because of the limitations imposed by sequencing and presence of closely related species of microbes. This limits the proper determination of functional genes and metabolic pathways. Moreover, a restriction on annotation is a serious constraint in the metagenomic analysis. Considerable percentage of the genes detected in the environment samples are yet to be characterised because of the uncharacterized or biassed reference databases. This creates loopholes in functional interpretation and constrained explanation of role that microbes play in degrading contaminants. Another significant challenge is the high cost of computations involved in processing of metagenomic data. Larger sequencing datasets consume large amounts of computational time, storage environment, and more powerful tools to analyse, which can be costly and unavailable in every research project. In the future, future research directions would work on incorporating future technologies in metagenomics to reduce these shortcomings. Pattern recognition, fabrication of genes, and pathway analysis may be improved with the integration of the artificial intelligence (AI) and machine learning. Additionally, the multi-omics, such as metatranscriptomics, proteomics, and metabolomics, can give a more holistic view of the microbial dynamics of activity and its functions. Synthetic biology also has the potential to be used to design microbial consortia with more powerful bioremediational properties. All in all, technological innovation and interdisciplinary solutions addressing these challenges will be critical in the context of implementing metagenomics-based bioremediation into scalable and sustainable solutions to the environment.

8. CONCLUSION

Metagenomics has become a revolutionary methodology in environmental biotechnology with the ability to provide a global view of the structure and functional capability of stochastic communities of microbes in the polluted environment at a gene level. Metagenomics can help identify relevant functional genes and metabolic pathways during which pollutants are degraded which in turn optimises the efficacy, sensitivity and versatility of bioremediation efforts. The culture-neutral approach has also increased the range of known unculturable microorganisms, and thus creating new enzymatic systems and biodegradation pathways that have not been available using traditional methods. Combination of metagenomics and highly developed bioinformatics tools has also enhanced the capacity to characterise complex microbial ecosystems and anticipate their functional part in the restoration of the environment. Through such integrative strategies, it becomes possible to produce specific, economically viable and ecologically sustainable remediation/clean up strategies to an extensive variety of pollutants, encompassing hydrocarbons, heavy metals, and industrial waste products. However, in spite of the progress, a number of problems exist, especially with regard to the functional verification of the predicted genes, incomplete genome reconstruction and weaknesses in the currently accessible annotation databases. Subsequent studies must then be directed towards assembling metagenomics with technologies that are multi-omic, like metatranscriptomics and metabolomics, in order to have a more detailed picture of microbial action. Furthermore, the integration of machine learning and artificial intelligence will also play an essential role in working with big data and enhancing the accuracy of predictions. Altogether, metagenomics has a tremendous potential in enhancing sustainable environmental management and dealing with the issues of global pollutions by developing innovative and efficient bioremediation plans.

REFERENCES

1. Bhattacharjee, A. S., Wu, S., Lawson, C. E., Jetten, M. S., Kapoor, V., Domingo, J. W. S., ... & Goel, R. (2017). Whole-community metagenomics in two different anammox configurations: process performance and community structure. Environmental Science & Technology, 51(8), 4317–4327.
2. Das, N., & Chandran, P. (2011). Microbial degradation of petroleum hydrocarbon contaminants: an overview. Biotechnology Research International, 2011(1), 941810.
3. Fierer, N. (2017). Embracing the unknown: disentangling the complexities of the soil microbiome. Nature Reviews Microbiology, 15(10), 579–590.
4. Ghosal, D., Ghosh, S., Dutta, T. K., & Ahn, Y. (2016). Current state of knowledge in microbial degradation of polycyclic aromatic hydrocarbons (PAHs): a review. Frontiers in Microbiology, 7, 208111.
5. Goswami, M., Chakraborty, P., Mukherjee, K., Mitra, G., Bhattacharyya, P., Dey, S., & Tribedi, P. (2018). Bioaugmentation and biostimulation: a potential strategy for environmental remediation. Journal of Microbiology & Experimentation, 6(5), 223–231.
6. Hazen, T. C., Dubinsky, E. A., DeSantis, T. Z., Andersen, G. L., Piceno, Y. M., Singh, N., ... & Mason, O. U. (2010). Deep-sea oil plume enriches indigenous oil-degrading bacteria. Science, 330(6001), 204–208.
7. Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y., & Morishima, K. (2017). KEGG: new perspectives on genomes, pathways, diseases and drugs. Nucleic Acids Research, 45(D1), D353–D361.

8. Kim, J. S., Woo, J. H., Kim, J. T., Park, N. H., & Kim, C. G. (2010). Metaproteomics in microbial ecology. The Korean Journal of Microbiology, 46(1).
9. Knight, R., Vrbanac, A., Taylor, B. C., Aksenov, A., Callewaert, C., Debelius, J., ... & Dorrestein, P. C. (2018). Best practices for analysing microbiomes. Nature Reviews Microbiology, 16(7), 410–422.
10. Langille, M. G., Zaneveld, J., Caporaso, J. G., McDonald, D., Knights, D., Reyes, J. A., ... & Huttenhower, C. (2013). Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. Nature Biotechnology, 31(9), 814–821.
11. Li, D., Liu, C. M., Luo, R., Sadakane, K., & Lam, T. W. (2015). MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. Bioinformatics, 31(10), 1674–1676.
12. Luo, C., Knight, R., Siljander, H., Knip, M., Xavier, R. J., & Gevers, D. (2015). ConStrains identifies microbial strains in metagenomic datasets. Nature Biotechnology, 33(10), 1045–1052.
13. Mir, M. Y., Hamid, S., & Rohela, G. K. (2022). Bioaugmentation: A way out for remediation of polluted environments. In Environmental Biotechnology (pp. 137–156). Apple Academic Press.
14. Shokralla, S., Spall, J. L., Gibson, J. F., & Hajibabaei, M. (2012). Next-generation sequencing technologies for environmental DNA research. Molecular Ecology, 21(8), 1794–1805.
15. Simon, C., & Daniel, R. (2011). Metagenomic analyses: past and future trends. Applied and Environmental Microbiology, 77(4), 1153–1161.
16. Suenaga, H. (2012). Targeted metagenomics: a high-resolution metagenomics approach for specific gene clusters in complex microbial communities. Environmental Microbiology, 14(1), 13–22.
17. Thomas, T., Gilbert, J., & Meyer, F. (2012). Metagenomics—a guide from sampling to data analysis. Microbial Informatics and Experimentation, 2(1), 3.
18. Tripathi, M., & Singh, D. N. (Eds.). (2022). Bioremediation: Challenges and advancements. Bentham Science Publishers.
19. Varjani, S. J. (2017). Microbial degradation of petroleum hydrocarbons. Bioresource Technology, 223, 277–286.
20. Wooley, J. C., Godzik, A., & Friedberg, I. (2010). A primer on metagenomics. PLoS Computational Biology, 6(2), e1000667.