Bioinformatics Toward Improving Bioremediation

27

Surbhi Khanna and Ashwani Kumar

Abstract

Environmental pollution has become a major issue of concern. With the rapid growth of industrialization, agricultural practices, and energy generation processes, the exploitation of natural resources has occurred. The result of which is the pollution of air and soil. To get rid of it, several practices are applied one of which is the bioremediation. This requires microbes, which have potential and enzymatic capability to undergo complete transformation or mineralization into harmless end products. To aid this process at a faster rate, bioinformatics has emerged as an advantageous approach. It helps in diversification and implementation of bioremediation in a productive way by employing the computational tools and software packages. This in silico approach of studying the bioremediation is very efficient by giving knowledge and understanding of the pathways and structural and functional aspects of microorganisms involved in biodegradation. Thus, this chapter gives the detailed complete idea of tools and software that bioinformatics provide toward improvement of bioremediation.

Keywords

Bioinformatics · Computational tools · Metabolic pathways · Chemical toxicity · Omic approaches · Databases

S. Khanna (✉)
Department of Computational Biology & Bioinformatics, University of Kerala, Thiruvananthapuram, Kerala, India

A. Kumar
Department of Botany & Biotechnology, University of Rajasthan, Jaipur, Rajasthan, India

27.1   Background

With the rapid industrialization, thousands of chemical compounds are produced causing air and soil pollution. Out of which, some are very toxic in nature and remain in the environment possessing a major threat to life of living organisms. Hence, therefore it is important to look for the techniques that can be employed for either the removal of these contaminants or to convert them into nonhazardous products that are eco-friendly to the environment. This is achieved by the use of enzymatic capabilities of microorganisms that break down toxic chemical compounds into end products or metabolites, which are not toxic anymore, and this whole process is known as bioremediation. Thus, such degradation is carried out by particular microbes, and to know more about it, the knowledge about the properties of these toxic chemicals such as classification, identification, environmental properties, toxicity, and distribution can enhance the biodegradation process. This technique has potential to restore the contaminated environment effectively with low cost and labor. But the information for the factors that control the growth and metabolism is still not known completely, making the implementation of it bit restricted.

Bioinformatics, which has now become the essential part of every life science-related research, has given new direction in view of bioremediation technique also. With the development of software packages and tools with the help of computational biology, bioinformatics has revolutionized the integration of it with bioremediation. In last few decades, branches of bioinformatics like genomics, proteomic, transcriptomics, or metabolomics have given a lot of contribution in exploration of bioremediation process.

Hence, bioinformatics with its multidisciplinary approach has assisted in understanding the bioremediation by unveiling the pathways, chemistry of toxic chemicals that are undisclosed for making it a process for control of environmental contamination. The aim of this chapter is to provide a complete overview of the bioinformatic approaches and its applications present in relation to bioremediation.

27.2   Introduction

Bioremediation is the deliberate use of microorganisms, which act as biological catalysts for removing pollutants from the environment. In general, we can say that it is an environmental science approach where natural biological actions are used to remediate the polluted groundwater and contaminated soil. A variety of pollutants like xenobiotics, polycyclic aromatic hydrocarbons known as PAHs, and chlorinated and nitro-aromatic compounds are present, which can be cancer-causing and mutagenic to all the present life forms (Zhang and Bennett 2005; Samanta et al. 2002). With the use of these microbes for biodegradation, the natural environmental conditions can be maintained efficiently. So, the role of these microorganisms (bacteria, fungi, insects, worms, etc.) in bioremediation technique has proven to maintain our planet with its greenery.

The general microbial mode of action to perform bioremediation activity is done by metabolization of a compound to another metabolite, which is not harmful to the environment. The basic principle involved in biodegradation of pollutants is either biotic or abiotic conditions. It can be done by number of known processes, such as bioventing, biopiles, bioaugmentation, biostimulation, and bioattenuation. So, the bioremediation can be effective only where environmental conditions permit microbial growth and activity, and its application often involves the manipulation of environmental parameters to allow microbial growth and degradation to proceed at a faster rate (Kumar et al. 2011; Abatenh et al. 2017).

Being the natural process, it is cheap, harmless to the ecosystem, needs less labor requirement, eco-friendly, and sustainable (Dell Anno et al. 2012).

Thus, the use of bioremediation technique is an environmental-friendly approach for restoring and sustaining the contamination-free environment for future generations.

27.2.1 Introduction to Bioinformatics

Bioinformatics is the combination of biology and information technology. It involves the knowledge of both. The field of bioinformatics does the computer-based analysis of biological datasets followed by its interpretation. This is done by using statistical tools and algorithms.

In understanding the bioinformatics and its applications, it is important to know about the various approaches used for performing analysis. This includes genomics, proteomics, data mining, biological databases, phylogenetic analysis, and (trancriptomics, metabolomics) system biology. All of it together plays a significant role. Figure 27.1 below shows its various branches.

27.2.2 Integrating Bioinformatics with Bioremediation

The role of microbes in soil and water-based biodegradation and cleaning of the environment has shown us the way to maintain and sustain a greener earth. The use of bioinformatic domain for the study of bioremediation has shown in the past the suggested promising results. With the help of bioinformatic-based applications only, it has been made possible to perform the in silico studies and analyzation of data. For uplifting the technique of bioremediation and the study of specific microbes at the molecular level including the gene-to-gene interactions, pre-requirement of conditions needed to be used for the changes at genetic level can be done only with bioinformatic strategies. Also, the bioremediation process can be enhanced using databases for gene identification and microbial degradation pathways of compounds (Ellis et al. 2001).

Thus, bioinformatics along with its branches is revolutionizing and will continue to do so in its future prospects. The pictorial representation above in Fig. 27.2 is depicting the use of bioinformatic approach for the improvement of bioremediation process.
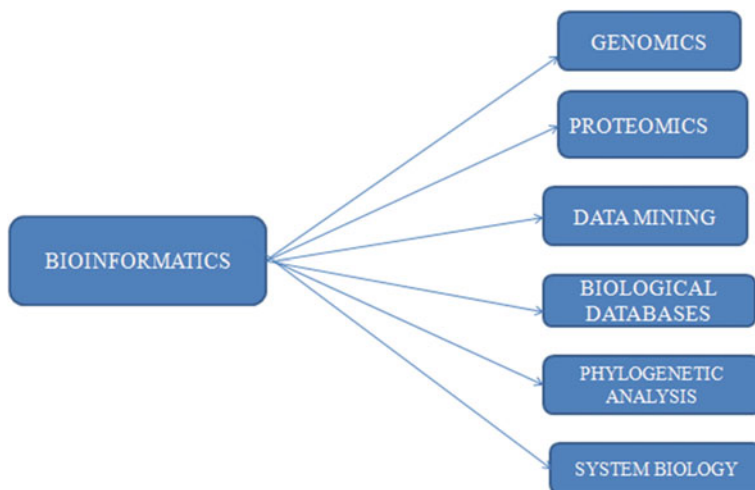
Fig. 27.1 Branches of bioinformatics

27.3   Bioinformatics in Improving Bioremediation

Although microbes are known for their potential to perform biodegradation, still the process has its own limitation. And this is because of scarcity of data for factors, which control the growth and metabolism of microbes with bioremediation potential (Dua et al. 2002). Therefore, bioinformatics aids in using microarray data by enhancing the structural characterization of microbial proteins with contamination degradable capabilities (Singh 2006).

Hence, by understanding the microbial process at the molecular level with use of bioinformatic analyses we can learn about the following below mentioned aspects of bioremediation in more depth.

1. Prediction of Degradation Pathways
2. Omic-Based Approaches
3. Prediction of Toxicity of Chemicals
4. Databases

27.3.1 Prediction of Degradation Pathways

For the bioremediation process, a microbe undergoes enzymatic reactions to change the pollutant into a metabolite, which is not harmful. For all this, the study of enzymatic kinetic aspect is important. This includes the physical and chemical characteristics of the degradation pathway (Okoh 2006).
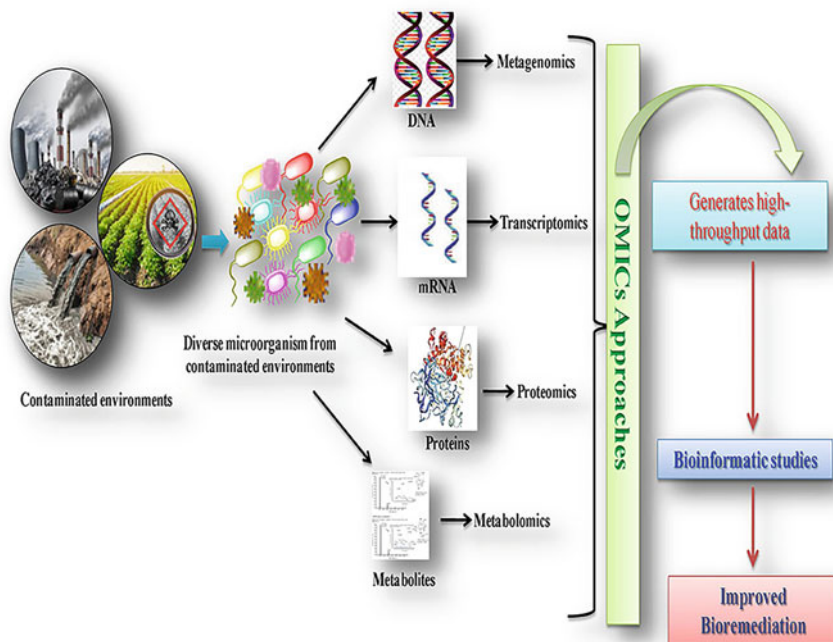
Fig. 27.2 Pictorial representation of integrated approach of advanced technologies in biodegradation of xenobiotic compounds. (Source: Mishra, S., Lin, Z., Pang, S., Zhang, W., Bhatt, P., & Chen, S. (2021). Recent Advanced Technologies for the Characterization of Xenobiotic-Degrading Microorganisms and Microbial Communities. Frontiers in bioengineering and biotechnology, 9, 632059. https://doi.org/10.3389/fbioe.2021.632059. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC-BY). The use, distribution, or reproduction in other forums is permitted, provided the original author(s) and the copyright owner (s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice)

But for the prediction of products and pathways associated with microbial degradation by in silico methods, classification approach is required (Wicker et al. 2010). This classification can be done as knowledge-based and machine learning-based approach. Both of which have some limitations and strengths. Taking into account the *machine learning approach*:

Firstly, this approach does prediction for a biotransformation when it has a quite a general class (Gomez et al. 2007) or whether it is the substrate of some broad reaction class, e.g., oxidoreductase catalyzed reactions (Mu et al. 2006).

Next is knowledge-*based approach*:

META: META is a knowledge-based expert system that simulates the biotransformation of xenobiotics. It operates with the help of a dictionary (knowledge base) to seek target fragments in a compound and transform them to products (Klopman et al. 1997).

METEOR: It is a knowledge-based expert system for prediction of metabolism (Marchant et al. 2008).

CATALOGIC: It is a platform for models targeting environmental fate of chemicals. It explicitly aims at probability estimates (Dimitrov et al. 2010).

UM-PPS: It stands for the University of Minnesota Pathway Prediction System (UM-PPS) and comes under the UM-BBD (University of Minnesota Biocatalysis/Biodegradation Database). It is available at http://umbbd.msi.umn.edu/predict/. Presently, it contains information on almost 1200 compounds, over 800 enzymes, almost 1300 reactions, and almost 500 microorganism entries (Gao et al. 2011).

The UM-PPS predicts plausible biodegradation pathways for organic compounds on the basis of sets of biotransformation rules derived from the UM-BBD database or from the scientific literature (Fenner et al. 2008). The user can predict both aerobic and anaerobic degradation pathways of chemicals and can select whether they will view all or only the more likely aerobic transformations. Users can also obtain the most accurate prediction for those compounds similar to compounds with biodegradation pathways that have been reported in the scientific literature (Gao et al. 2011; Arora and Bae 2014).

Usage

1. Prediction can be made both for aerobic and anaerobic degradation pathways of chemicals, and it can be selected that whether the user will view all or only the more likely aerobic transformations.
2. Also, we can obtain the most accurate prediction for those compounds similar to those biodegradation pathways that have been reported in the scientific literature.
3. For the prediction, users may enter a compound into the system by either drawing the structure and generating SMILES or entering SMILES directly.
4. For example, the degradation pathways of 4-nitrophenol have been thoroughly investigated, while those of 2-fluro-4-nitrophenol and 2-bromo-4-nitrophenol have not. However, the structures of 2-fluro-4-nitrophenol and 2-bromo-4-nitrophenol are similar to 4-nitophenol. Therefore, PPS can provide very accurate predictions for degradation of 2-flouro-4-nitrophenol and 2-bromo-4-nitrophenol (Arora and Bae 2014).

27.3.1.1 PathPred

It is a knowledge-based prediction system, which uses data derived from the Kyoto Encyclopedia of Genes and Genomes (KEGG) in the form of KEGG REACTION and KEGG RPAIR database. The KEGG RPAIR database has collection of biochemical structure transformation patterns, called RDM patterns, and chemical structure alignments of substrate–product pairs (reactant pairs) in all known enzyme-catalyzed reactions taken from the enzyme nomenclature and the KEGG PATHWAY database (Moriya et al. 2010).

It is a web-based server available at http://www.genome.jp/tools/pathpred/. It predicts plausible pathways of multi-step reactions starting from a query compound, based on the local RDM pattern match and the global chemical structure alignment

against the reactant pair library. The server provides transformed compounds and reference transformation patterns in each predicted reaction and displays all predicted multi-step reaction pathways in a tree-shaped graph (Moriya et al. 2010). It basically aims at predicting pathway for microbial biodegradation of environmental compounds and biosynthesis of plant secondary metabolites.

Usage

The PathPred server can be used for predicting microbial biodegradation pathways of xenobiotics in bacteria and biosynthesis pathways of secondary metabolites in plants. This can be done by

1. Selecting Reference Pathway—the user is requested to choose the reference pathway for either of biosynthesis and biodegradation, which determines the subset of RDM patterns to be utilized.
2. Query Format—the query can be inputted as a query compound in the MDL mol file format, in the SMILES representation, or by the KEGG compound/drug identifier (C/D number). This compound, termed initial compound, corresponds to the compound to be degraded or the compound to be synthesized.
3. Output—The output of the PathPred server shows the prediction results as tree-shaped graph. For example, the biodegradation prediction of glycolate (C00160) from 1,2,3,4-tetrachlorobenzene. The output tree graph predicts the other possible pathways including biodegradations through known compounds such as 3,4,6-trichlorocatechol (C12831), 6-chlorobenzene-1,2,4-triol(C06328), and 1,2,4-trichlorobenzene (C06594) (Fig. 27.3).

27.3.1.2 BNICE

It stands for Biochemical Network Integrated Computational Explorer, a computational approach developed to generate every possible biochemical reaction based on a set of enzyme reaction rules of the enzyme commission (EC) and starting compounds (Finley et al. 2009). In general, it predicts whether a particular compound is biodegradable and whether alternate routes can be engineered for compounds already known to be biodegradable.

The BNICE screens out all possible pathways for thermodynamic feasibility based on the Gibbs free energies of the reaction and selects feasible novel thermodynamic pathways (Soh and Hatzimanikatis 2010). Hence, it is used to (1) study the combinatorial nature of polyketide synthesis (Gonzalez-Lergier et al. 2005), (2) to provide systematic framework for linking of enzymatic chemistry and reactive sites of metabolic compounds (Hatzimanikatis et al. 2004), and (3) for the prediction of biodegradation pathways of compounds, which represent various classes of xenobiotics.

Further, it has been also suggested by Soh and Hatzimanikatis et al. that the pathways generated by BNICE can be further evaluated using established pathway analysis approaches, such as thermodynamic-based flux balance analysis (FBA) Grow match allows investigation of the overall effects of these novel pathways on metabolic network performance in host organisms. FBA can help predict maximum yield, phenotypic changes, effects of gene knockouts, changes in bioenergetics of the
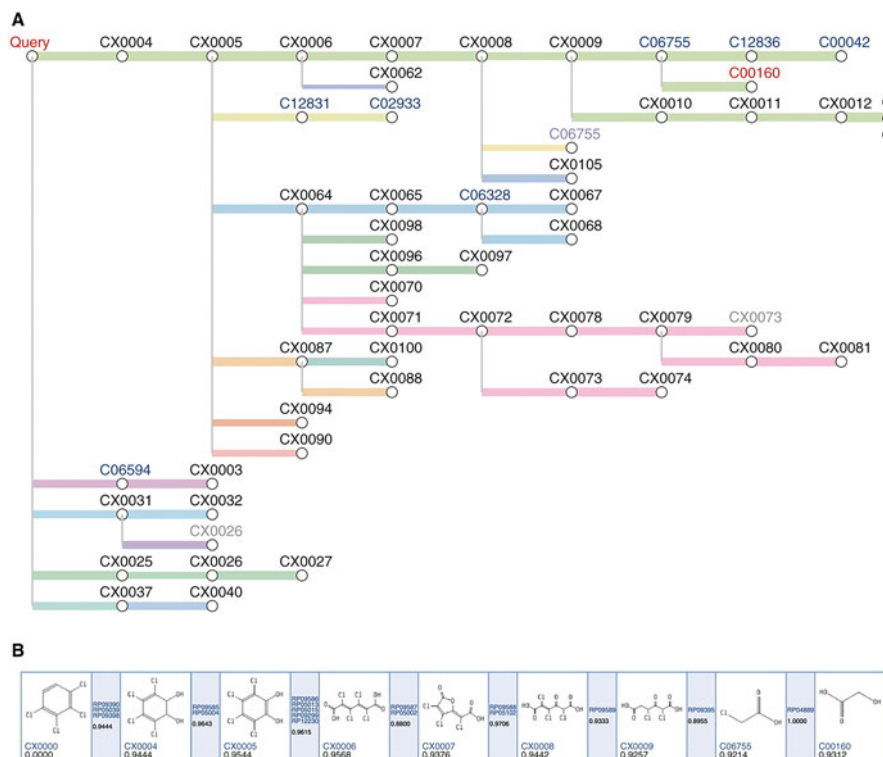
Fig. 27.3 Example of the predicted pathway tree of tetrachlorobenzene biodegradation (**a**) and the detail of the top green pathway from the query compound (query) to the final compound (C00160) (**b**). Structure images popup when the mouse is moved over nodes and edges in the tree if JavaScript is enabled in the web browser. (Source: Moriya Y, D. Shigemizu, M. Hattori, T. Tokimatsu, M. Kotera, S. Goto, Minoru Kanehisa, PathPred: an enzyme-catalyzed metabolic pathway prediction server, *Nucleic Acids Research*, Volume 38, Issue suppl_2, 1 July 2010, Pages W138–W143, https://doi.org/10.1093/nar/gkq318. Reused with Licence Number 5125780225626 dated August 11, 2021)

system for metabolic engineering, and synthetic biology (Soh and Hatzimanikatis 2010).

Usage

The BNICE framework searches for pathways by considering the starting compound and/or products, the requested length of the pathway, and the range of reactions to search over (Henry et al. 2010; Medema et al. 2012).

The user can also choose to search for a number of possibilities, either by searching for a pathway using enzyme reactions from known pathways, by combination of multiple pathways, or the whole metabolic network (Henry et al. 2010; Hatzimanikatis et al. 2005). A set of molecules is given as an input and every molecule is evaluated to determine whether it has the appropriate functionality to

undergo reactions corresponding to the specified reaction classes (Bashir Sajo and Mohd 2015).

While predicting the possible pathways the BNICE predicts more than 10,000 different pathways for the biosynthesis and degradation of the compound of interest, due to the fact that the system relies on few criteria. However, Henry et al. had pioneered a prioritization approach in this framework, in which generated pathways are ranked according to four criteria: pathway length, thermodynamic feasibility, maximum achievable yield, and maximum achievable activity.

Output: The output of the BNICE is a graph-theoretic matrix representation of biochemical compound, enzyme reaction rules, and molecules. It is represented using the bond-electron matrix (BEM) where each atom in a molecule is represented by a row and column. The BEM is characterized by diagonal elements, which denote the non-bonded valence electrons and non-diagonal elements, which give the connectivity via bonding between different atoms and the bond order between atoms (Hatzimanikatis et al. 2005).

27.3.1.3 DESHARKY

It is a Monte Carlo algorithm, which finds a metabolic pathway from a target compound by exploring database of enzymatic reactions. It predicts a possible route connecting the specified target metabolism with the host metabolism, instead of using pathway selection by enumeration of possible metabolic routes. It finds pathway within shortest possible time by computing its associated genetic burden. Also, it can be used also in distributed computing to sample most of the solution spaces (Rodrigo et al. 2008).

Usage

The algorithm is implemented in C/C+ +, and it is easily compiled and runs in UNIX environment (e.g., in Linux or in Windows using Cygwin). The algorithm calculates thermodynamic favorability and energy loss in transcription and translation.

The input of the algorithm is usually the target compound, while its output is the designed metabolic pathway together with quantification of the transcriptional, translational, and metabolic load (Rodrigo et al. 2008). This framework also provides the sequence of amino acids of the enzyme involved in the pathway.

Output: The output is the designed metabolic pathway together with the quantification of the transcription, translation, and metabolic load. It provides the sequence of amino acids of the enzymes involved in the pathway. These amino acid sequences provided are usually the closest phylogenetically to Escherichia coli according to KEGG classification of organisms (Rodrigo et al. 2008).

27.3.1.4 FMM

It stands for from metabolite to metabolite, a web server. It is available freely at http://FMM.mbc.nctu.edu.tw/. It can reconstruct metabolic pathways from one metabolite to another metabolite among different species, based mainly on the Kyoto Encyclopedia of Genes and Genomes (KEGG) database and other integrated biological databases (Chou et al. 2009). Even though KEGG maps utilized in many

metabolic tools, none of them can connect metabolites from different KEGG maps. FMM supports the connection of different KEGG maps.

FMM has many applications in synthetic biology and metabolic engineering. For example, the reconstruction of metabolic pathways to produce valuable metabolites or secondary metabolites in bacteria or yeast is a promising strategy for drug production. FMM provides a highly effective way to elucidate the genes from which species should be cloned into those microorganisms based on FMM pathway comparative analysis (Chou et al. 2009).

Usage

1. Data collection and Integration:
   Reaction definitions, species-specific reactions, reaction maps, and enzyme list can be obtained from KEGG/LIGAND and KEGG/PATHWAY databases recent releases. Information such as gene names, enzyme commission numbers, and species-specific enzymes can be retrieved from UniProtKB/Swiss-Prot and NCBI taxonomy databases. Additionally, the data in FMM are usually updated on a regular basis.

2. Construction of reaction matrix information on reactions and enzymes can be obtained from KEGG maps and the equation of each reaction can be determined. Therefore, reaction matrices can be constructed based on maps, reactions, and enzyme data.
   The workflow of FMM in above Fig. 27.4 shows the reaction matrix, which was developed to identify numerous reaction processes from one metabolite to another. Enzyme annotations from UniProtKB/Swiss-Prot (Boutet et al. 2007) were employed to identify enzymes from different species in comparative analysis.

3. Reconstruction of metabolic pathway from various KEGG pathway maps: After all possible reaction paths were identified, the number of pathway maps was calculated. Usually, found paths occurred not in only a single pathway map, but also in a complicated fashion in several maps. Pathway maps that contain the most paths are selected and the one pathway map that has only one reaction is avoided. A matrix of maps versus reactions was employed to reconstruct metabolic pathway from different KEGG maps.

4. Comparative Analysis: Comparative analysis provided in FMM is useful in synthetic biology. Comparative analysis provides an easy way to elucidate which genes from which species should be cloned into those microorganisms. First, the enzymes identified in the reconstructed pathway were processed to search for orthologous encoding genes from various species. Then, the presence or absence of the pathway in a particular species can be known.

27.3.1.5 RetroPath

It is a server, which applies a retrosynthetic approach, a concept originally proposed for synthetic chemistry, which uses reverse chemical transformations (reverse enzyme-catalyzed reactions in the metabolic space) starting from the desired target compound to identify the reactants (precursors) that are indigenous to the selected
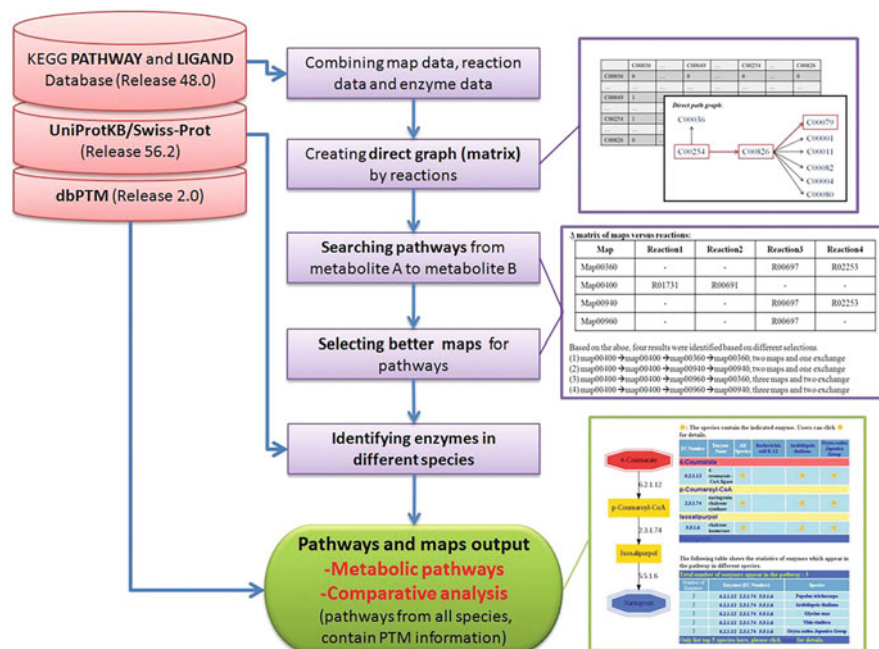
Fig. 27.4 Workflow of FMM. (Source: Chou, C. H., Chang, W. C., Chiu, C. M., Huang, C. C., & Huang, H. D. (2009). FMM: a web server for metabolic pathway reconstruction and comparative analysis. *Nucleic acids research*, *37*(Web Server issue), W129–W134. https://doi.org/10.1093/nar/gkp264. Reproduced with License no 5125960782113 dated August 11, 2021)

host (Carbonell et al. 2012). It is available at http://www.issb.genopole.fr/~faulon/retropath.php.

This method of metabolic pathway design is unique because it addresses the complexity problem by coding substrates, products, and reactions into molecular signatures. The approach used by RetroPath is characterized by metabolic maps, which are represented in hypergraphs. The complexity involved in the reactions is controlled by varying the specificity of the molecular signature. Each signature has different "heights," *h*, that correspond to levels of structural detail. The height can be varied, which reduces the number of reactions that can be generated (Carbonell et al. 2011).

The proliferation of several metabolic databases with rich information is considered to be a significant breakthrough. KEGG that is a database resource integrated with chemical and systematic functional information and genomics is linked to RetroPath, where information on the reactions predicted using this framework can be found in KEGG. BRENDA (Schomburg et al. 2013) is another database that contains one of the largest collections functional enzyme data. Incomplete knowledge or gaps still exist in many cases, especially when looking for novel ways to synthesize a target compound of interest (Carbonell et al. 2013).

To successfully achieve a heterologous pathway design, the process needs to be rationalized by following the principles of synthetic biology: modeling of the biological system of interest, modular design through standardization, goal-oriented optimization, and experimental validation (Carbonell et al. 2013).

Usage

In the research study done by Carbonell et al. (2013), they have suggested that for retrosynthetic design of heterologous pathways, the following steps will be required: (1) host chassis selection, (2) in silico model selection for the chassis from BiGG (Schellenberger et al. 2010) or biomodels (Le Novere et al. 2006), (3) definition of the metabolic space, (4) pathway enumeration, (5) gene selection, (6) estimation of yields by metabolic analysis software, e.g., COBRA, OptFlux (Rocha et al. 2010), and COPASI (Hoops et al. 2006; Schaber 2012), (7) toxicity prediction of pathway metabolite, (8) definition of an objective function to select the best pathway to engineer, and (9) pathway implementation and validation (Fehér et al. 2014).

27.3.1.6 Metabolic Tinker

It is a web tool used to design synthetic metabolic pathways between user-defined target and source compounds. The interface is available at http://osslab.ex.ac.uk/tinker.aspx. It uses a tailored heuristic search strategy to search for thermodynamically feasible paths in the entire known metabolic universe. The program contains a directed graph known as universal reaction network (URN), which represents the entire set of known reactions and compounds from the Rhea database (McClymont and Soyer 2013). Nodes and edges on this graph represent metabolites and reactions, respectively, and thus, the entire graph represents the current known metabolic universe. This tool searches possible biochemical paths between two compounds within this URN using standard search algorithms developed in computer science and graph theory (McClymont and Soyer 2013). To complete the search, the Rhea/CHEBI identification codes of both the source and target compounds are needed.

27.3.1.7 Carbon Search

It is an algorithm-based approach, which identifies pathways within existing metabolic networks by tracking the conservation of atoms moving through them. On the basis of this approach, two algorithms are developed that find metabolic pathways by using atom mapping data to track the movement of atoms through metabolic networks. One algorithm finds linear pathways, and the other algorithm finds branched pathways. They both take as input atom as mapping data, a start compound, a target compound, and a minimum number of atoms to conserve and a maximum number of pathways to return (Heath et al. 2010). In the output, a set of metabolic pathways, which conserve at least given number of atoms from the start compound to the target compound, are returned. They have also demonstrated that this carbon search tool based on the algorithms has efficiently identified both linear and branched metabolic pathways, in which a certain threshold of atoms is conserved. The resulting metabolic pathways are validated on known functional

pathways. The algorithms are having the potential to find novel or alternative pathways that may span multiple organisms (Heath et al. 2010).

Using this atom tracking approach, earlier Pitkänen et al. in 2009 have also enabled graph theoretical-based method for finding biologically meaningful linear and branched metabolic pathways in genome-scale metabolic networks.

27.3.1.8 The Furusawa Platform

It is an in silico platform that uses a developed algorithm for finding feasible heterologous pathways by which non-native target metabolites are produced by microorganisms, using Escherichia coli, *Corynebacterium glutamicum,* and *Saccharomyces cerevisiae* as templates (Chatsurachai et al. 2012).

Usage

The usage of this platform for heterologous pathway design comprises of following four steps:

1. Construction of an in-house database of metabolic reactions—This is done by considering known metabolic reactions from KEGG ligand section database and BRENDA. These metabolic reactions are considered as candidate heterologous reactions that could be added to the host metabolic networks (Chatsurachai et al. 2012). All metabolic reaction information regarding genes, enzymes, pathways, and organism in the KEGG database can be collected into the database. The information collected the information in a constructed database using PostgresQL. The enzymatic information employed can be retrieved from BRENDA, and python script can be used to access the constructed in-house database (Chatsurachai et al. 2012).

2. Genome-scale metabolic models of host microorganisms—The microorganisms that are widely used in industry were adopted as chassis templates to demonstrate the viability of it on in silico platform. This includes *Escherichia coli*, *C. glutanicum,* and *S. cerevisiae*, which were selected based on a number of criteria such as having high growth activity under various conditions, ease of genetic manipulation, and hence are considered as ideal hosts for bioengineered products (Chatsurachai et al. 2012).

3. Heterologous pathway identification for target production—The developed platform can be used to screen all producible target metabolites listed in the database by adding heterologous reactions to host microorganisms. For all producible target metabolites, the user can estimate the production yields using FBA, assuming steady-state conditions and the maximum biomass production rate (Chatsurachai et al. 2012). The entire list of producible target metabolites in different hosts can be analyzed, and a set of rational heterologous pathways and hosts can be selected that will likely produce the desired targets.

4. Flux balance analysis (FBA)—FBA is based on a genome-scale metabolic model and optimization of a specific objective flux by linear programming. One can use FBA to estimate the metabolic flux profile of metabolic networks expanded with heterologous reactions. All FBA simulations in this framework can be performed under the MATLAB interface (Chatsurachai et al. 2012).

27.3.2 Omic-Based Approaches

27.3.2.1 Proteomics

According to Keller and Hettich (2009) and Aslam et al. (2017), proteomics has emerged as an interesting and fruitful technology for the study of protein expression (it includes post-translational modifications, protein turnover, proteolysis, and changes in the corresponding gene expression) of the microbial world. Proteomics has been used to identify microbial communities/microorganisms in various ecosystems including soil and sediment, activated sludge, marine and groundwater sediment, acid mine biofilms, and wastewater plants (Williams et al. 2013; Colatriano et al. 2015; Grob et al. 2015; Bastida et al. 2016; Jagadeesh et al. 2017). Thus, the inclusion of a proteomic approach helps to identify related enzymes and their metabolic pathways in the bioremediation of xenobiotics from various contaminated sites (Liu et al. 2017; Wei et al. 2017). Studies also revealed important and hidden information related to protein synthesis, gene expression stability, mRNA turnover, and protein–protein interaction networks in microbial communities in stress environments Aslam et al. (2017). Hence, the studies related to proteomic analysis plays important role for bioremediation process.

Protein Analysis: Generally, there are four primary steps that involve proteomic analysis of microbial communities:

1. preparation of a biological sample;
2. extraction and separation of proteins by using two-dimensional gel electrophoresis (2D-GE);
3. protein gel images are examined by means of image analysis software such as ImageMaster 2D or PDQuest; and
4. proteins are identified by using mass spectroscopy (MS)/MALDI-TOF/MS or LC-MS (Yates et al. 2009; Chakka et al. 2015; Velmurgan et al. 2017).

The workflow of proteomic analysis is shown below in Fig. 27.5.

Applications of Proteomics in Bioremediation

1. The bioremediation of compounds done by microorganisms has shown involvement of several proteins. This is demonstrated by the study done by Vandera et al. (2015). In their study, they have done comparative proteomic analysis of *Arthrobacter phenanivorans* Sphe3 on aromatic compounds phenanthrene and phthalates. The proteomic approach confirmed the involvement of several proteins in aromatic substrate degradation by identifying those mediating the initial ring hydroxylation and ring cleavage of phenanthrene to phthalate. This study also revealed the presence of both the ortho- and meta-cleavage pathways for the degradation of these aromatic compounds, and it also identified all proteins that take part in these pathways and are highly upregulated upon phthalate growth in comparison with phenanthrene growth.
2. The proteomic analysis of pyrene-degrading bacterium *Achromobacter xylosoxidans* PY4 done by Nzila et al. (2018) has identified a total of 1094
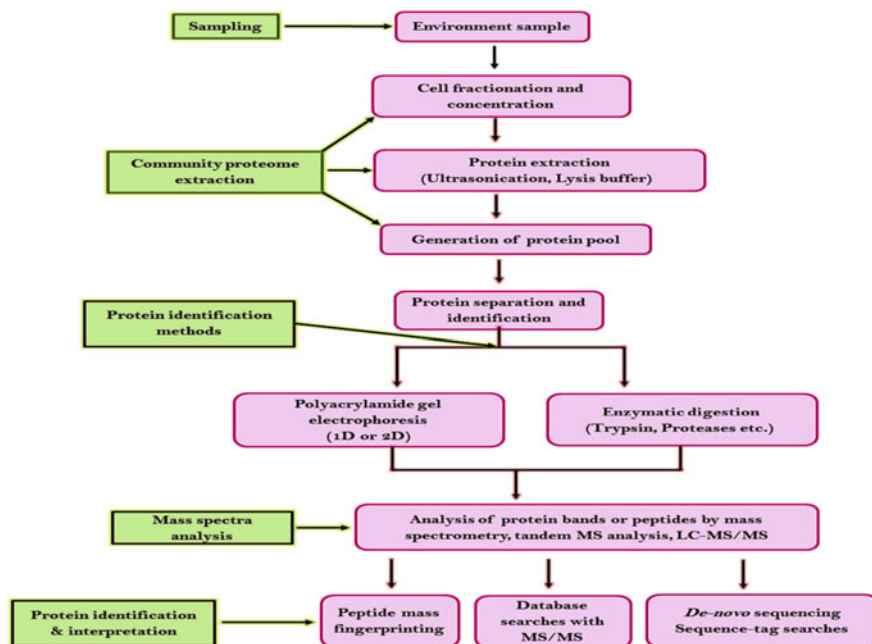
Fig. 27.5 Workflow of proteomic analysis. (Source: Chandran, H., Meena, M., & Sharma, K. (2020). Microbial Biodiversity and Bioremediation Assessment Through Omics Approaches. *Frontiers in Environmental Chemistry*, *1*, 9. https://doi.org/10.3389/fenvc.2020.570326. Frontiers is fully compliant with open-access mandates, by publishing its articles under the Creative Commons Attribution Licence (CC-BY))

proteins. Out of which, 95 proteins were detected in glucose supplementation, and 612 proteins were detected in the presence of pyrene. Furthermore, they have found 25 upregulated proteins to be involved in stress response and the progression of genetic information. Two upregulated proteins, 4-hydroxyphenylpyruvate dioxygenase and homogentisate 1,2-dioxygenase, are implicated in the lower degradation pathway of pyrene. Enzyme 4-hydroxyphenylpyruvate dioxygenase may catalyze the conversion of 2-hydroxybenzalpyruvic acid (metabolite of pyrene) to homogentisate. Homogentisate 1,2-dioxygenase is involved in the incorporation of 2 oxygen atoms to produce 4-maleyacetoacetate, which is an intermediate in several metabolic pathways (Nzila et al. 2018).

3. Lee et al. (2016) have performed proteomic analysis of PAH-degrading bacterial isolate *Sphingobium chungbukense* DJ77. This strain exhibited outstanding degradation capability for various aromatic compounds. With this study, it was demonstrated that the degradation of three xenobiotic compounds, i.e., phenanthrene, naphthalene, and biphenyls (PNB), and their associated proteins was analyzed by 2-DE and MALDI-TOF/MS analysis. During PNB biodegradation by bacterial cells, an alteration was observed in protein expression to cope with the stress condition.

4. In year 2019, Chen et al. have investigated a biodegradation mechanism of tetrabromobis-phenol (TBBPA) in *Phanerochaete chrysosporium* by using a proteomic approach. With aid of this approach, they have found that compared to control TBBPA, stress caused 148 differentially expressed proteins in *P. chrysosporium*, among which 90 proteins were upregulated and 58 proteins were downregulated. The upregulation of cytochrome p450 monooxygenase, glutathione-*S*-transferase, *O*-methyltransferase, and other oxidoreductases is responsible for the biotransformation of TBBPA via oxidative hydroxylation and reductive debromination.

5. Another bioremediation study with proteomic analysis was performed by Yu et al. (2019). It was of decabromodiphenyl ether (BDE-209). It was explored in *Microbacterium Y2* in a polluted water-sediment system. The results of study have shown that the overexpression of haloacid dehalogenases, glutathione *S*-transferase, and ATP-binding cassette (ABC) transporter might occupy important roles in BDE-209 biotransformation. Moreover, heat-shock proteins (HSPs), ribonuclease E, oligoribonuclease (Orn), and ribosomal proteins were activated to counter the BDE-209 toxicity. Thus, it is suggested that these proteins are implicated in microbial degradation, antioxidative stress, and glycolysis.

6. Another application of proteomics in bioremediation is researched by Gregson et al. (2020). It was reported that LC–MS/MS shotgun proteomics is used to determine variations in the proteome of hydrocarbon-degrading psychrophile *Oleispira antarctica RB-8* when grown on *n*-alkanes in cold temperatures.

27.3.2.2 Genomics and Metagenomics

Genomics is the powerful computer technology used to understand the structure and function of all genes in an organism based on knowing the organism's entire DNA sequence. This field includes intensive efforts to determine the entire DNA sequence of organisms and in-depth genetic mapping efforts (Fulekar and Sharma 2008).

Whereas metagenomic studies unblock the traditional ways of uncultured microorganisms and explore their genetic advantages in the process of bioremediation (Rahimi et al. 2018; Nascimento et al. 2020). It uses the pool of environmental genomes of microorganisms, which increases the probability to discover unique genes and diverse pathways with new enzymes containing highly specific catalytic properties (Scholz et al. 2012; Yergeau et al. 2017; Awasthi et al. 2020). This technology gives a new parade to microbiologists for understanding unculturable microbiota with a genetic variability of microbial communities (Devarapalli and Kumavnath 2015; Zhu et al. 2018; Awasthi et al. 2020). Hence, metagenomic information will enable researchers to integrate pure culture study with genomics (Hodkinson and Grice 2015). Current metagenomic practices allow for identifying the whole-genome structure of microorganisms and specifying particular genes that are attributed to encode degradative enzymes for the mineralization of xenobiotics (Zafra et al. 2016; Zhu et al. 2020). This clearly highlights the crucial role of novel genes in connecting the entire microbial population with functional diversity and structural identity. Based on it, the metagenomics involves the manufacturing of

metagenomic libraries. With the help of these, biological information can be retrieved from these metagenomic libraries by two types of analysis:

1. Sequence-Driven Analysis: This analysis is based on the sequencing of clones with phylogenetic anchors or conserved DNA sequences, which is the plausible origin of the DNA fragment (Wu et al. 2010; Felczykowska et al. 2015; Wong 2018). This type of analysis is increasingly used owing to the availability of several software packages for data analysis and the ease to assess metagenomic sequencing data. This approach is predominantly influenced by the precision of genome annotation, the integrity of the available data, algorithms, and facts in databases to ascertain the function of novel genes (Ferrer et al. 2009).
   The complete genome analysis or sequence analysis is progressed through three technical transformations:
   (a) *First-Generation Sequencing*—Frederick Sanger and Allen Maxam Walter Gilbert sequencing techniques were categorized as the first-generation DNA sequencing methods. Sanger sequencing uses denatured DNA template, radioactively labeled primer, DNA polymerase, and chemically modified nucleotides called di-deoxynucleotides to generate DNA fragments with various lengths.
   (b) *Next-Generation Sequencing*—It is also called high-throughput sequencing. Next-generation sequencing involves library preparation, sequencing, base calling, alignment to the established genome, and assorted annotation. Library preparation begins with the fragmentation of DNA into multiple fragments by sonication, enzymatic digestion, or transposase followed by ligation with adaptor sequences. The prepared library is then amplified using clonal amplification and PCR methods to generate DNA replicas. DNA replicas are then sequenced using different approaches (Samorodnitsky et al. 2015). The major platforms used for microbiome studies in next-generation sequencing are pyrosequencing (Roche/454 sequencing), Illumina, SOLiD, Ion Torrent, PacBio RS, etc.
   These are high-throughput sequencing techniques of ribosomal genes that quantify community structures and functions at a higher resolution, e.g., 16S rRNA in prokaryotes, and 5S or 18S rRNA genes, or the internal-transcribe-spacer (ITS) region in eukaryotes (Luo et al. 2012). The effectiveness of such NGS technologies in analyzing microbial communities from diverse environments was elucidated, validated, and documented in many studies (Brown et al. 2013; Shokralla et al. 2014; Zhou et al. 2015; Niu et al. 2016; Scholer et al. 2017).
   (c) *Third-Generation Sequencing*—It is also called single-molecule long-read sequencing. It offers lower sequencing charge and contented sample preparation without PCR amplification. The two most widely used sequencing platforms in third-generation sequencing are Pacific Biosciences, Oxford Nanopore Technology, and Helioscope technology.
   The competitive analysis of platforms used in second and third-generation sequencing is discussed below in Table 27.1.

Table 27.1 Comparative analysis of different platforms used for second- and third-generation Sequencing. Source: Chandran, H., Meena, M., & Sharma, K. (2020). Microbial Biodiversity and Bioremediation Assessment Through Omics Approaches. *Frontiers in Environmental Chemistry, 1*, 9. https://doi.org/10.3389/fenvc.2020.570326. Frontiers is fully compliant with open-access mandates, by publishing its articles under the Creative Commons Attribution Licence (CC-BY)

S No:	Omics technologies	Principle	Advantages	Disadvantages	Applications
I	Second generation sequencing platforms				
1.	Pyrosequencing	Sequencing by synthesis established on the identification of pyrophosphate discharged after nucleotide amalgamation in the newly synthesized DNA strand	Fast and firm method with real-time read out appropriate for sequencing short fragments of DNA, low cost, nucleotide dispensation easily programmable, alterations in the pyrogram pattern reveal mutations, deletions and insertions	Generation of long homopolymers generating sequencing errors, difficulty in deciding the quantity of integrated nucleotides in homopolymeric areas	Identification of microbes, whole genome sequencing
2.	Illumina	Sequence-by-synthesis method	Cells not required, highest throughput, produce relatively short reads with length up to 300 bp, lowest cost per base output compatible with most applications	Slow, short reads, high cost of reagents, aberrant incorporation of incorrect dNTPs by polymerases	Gene expression studies to identify isoforms, novel transcripts, gene fusions, exocrine sequencing, whole genome sequencing
3.	Ion Torrent	Sequencing established through the recognition of hydrogen ions released during polymerization of DNA	Produces reads up to 400 bp length, less run time, reliable and cost-effective tool, high accuracy and short run time	High error rates for specific regions	To study microbial diversity in complex ecosystems, targeted, exome, transcriptome, *de novo*, small RNA sequencing
4.	ABI SOLiD system	Sequencing technology based on ligation of DNA fragments	High accuracy since each base is read twice	relatively short reads and log run times	Whole genome sequencing, targeted sequencing, transcriptome research, epigenome analysis

II	Third generation sequencing platforms				
1.	Pacific biosciences SMRT sequencing	Sequencing through synthesis method and real time detection of integrated fluorescently labelled nucleotides	Long reads, high accuracy, uniform coverage, single-molecule resolution	Longer reads make create scaffolds in repeat regions	Whole genome sequencing targeted sequencing, RNA sequencing, epigenetic studies, study complex populations
2.	Oxford nanopore sequencing	Measurement of physical changes when DNA sequence translocates through nanometer size pores under influence of electric domain	Cheap, fast and accurate DNA sequencing, longer reads, better resolution, small size	Based errors, high cost per read	DNA, RNA or protein analyses
3	HeliScope	Single-molecule sequencing platform using a highly sensitive fluorescence detection system	Large number of single molecules read, to reduce high error rates	Short reading sequence, the sequencing process deteriorate from numerous biases due to inadequate clonel amplification and DNA extension devastation	Re sequencing, transcript counting

Shotgun Sequencing

It is also called shotgun metagenomic sequencing. It is a powerful technique in microbial ecology because it provides a vigorous and reliable evaluation of microbial diversity (Hillmann et al. 2018). It does not depend on PCR amplification and is used to examine the functional potential and microbial composition of the community.

Importance of shot-gun sequencing in bioremediation

(a) It is the only way to study the microbial community with no markers like viruses (Quince et al. 2017; Vermote et al. 2018).

(b) It allows strain-level remodeling in the taxonomic analysis and pathway predictions for the functional annotation of the microbiome under study (Han et al. 2020).

(c) It is an emerging molecular method to bridge the gap amid community structure and functional competence.

(d) It also helps in understanding the strategies adopted by microorganisms to thrive in adverse conditions (Sharpton 2014; Peabody et al. 2015; Ranjan et al. 2016).

(e) This techniques workflow for taxonomy analysis consists of quality pruning and evaluation of a reference database involving whole genomes or specifically designed marker genes to create a taxonomy profile. Since it contains all genetic information in a sample, the information can be used for supplementary analyses like metagenomic assembly and binning, metabolic function profiling, and antibiotic-resistant gene profiling (Chandran et al. 2020).

(f) Shotgun metagenomic analysis of microbial communities from deep seabed petroleum seeps in the Eastern Gulf of Mexico revealed the presence of diverse communities of chemoheterotrophs and chemolithotrophs (Dong et al. 2019).

(g) Whole-genome shotgun sequencing was engaged to identify the taxonomic diversity and gene repertoire of bacteria isolated from tannery effluents and petrol-polluted soil samples for degradation of persistent organic pollutants like naphthalene, toluene, petrol, and xylene (Muccee and Ejaz 2020).

2. Function-Driven Analysis: The function-driven analysis is based on the identification of clones that express their functional activity. If the sequence analogy does not complement to a functional association or the original gene has less analogy to some genes whose products have been investigated biochemically or a specific gene is capable to accomplish diverse tasks in the cell (Hallin et al. 2008), then in such cases, function-driven screening is preferred to discover genes with novel functions or to explore the sequence variation of protein families (Singh et al. 2009; Meena et al. 2016). The workflow below is showing the general methodology used for metagenomic research in Fig. 27.6.

Applications of Metagenomics

1. With metagenomic analysis, research area has increased to analyze microbial communities, their genetic diversity, and metabolic pathways. It has provided opportunities to discover microbial consortia and genes involved in the
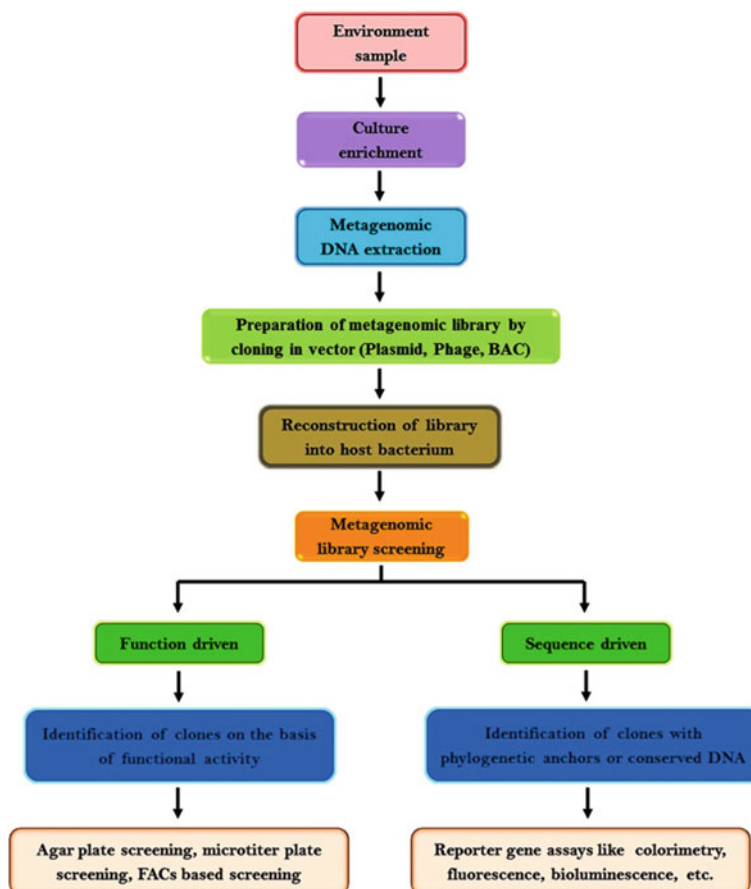
Fig. 27.6 Workflow of Metagenomic Research. (Source: Chandran, H., Meena, M., & Sharma, K. (2020). Microbial Biodiversity and Bioremediation Assessment Through Omics Approaches. *Frontiers in Environmental Chemistry*, *1*, 9. https://doi.org/10.3389/fenvc.2020.570326. Frontiers is fully compliant with open-access mandates, by publishing its articles under the Creative Commons Attribution Licence (CC-BY))

bioremediation of xenobiotic compounds. For example, phenol-degrading pathways of uncultivated bacteria in activated sludge were studied using metagenomics (Sueoka et al. 2009).

2. The metagenomic approach was used by Silva et al. (2013) to characterize genes and metabolic pathways associated with the degradation of phenol and other aromatic compounds in sludge from a petroleum refinery wastewater treatment system.

3. Also, Jeffries et al. in 2018 have employed metagenomic analysis to outline the functional potential and taxonomic community composition, and to predict the

breakdown of chemical compounds of soils with organophosphorus pesticide exposure.

4. A combined physical and chemical analysis along with metagenomics was done by Gaytán et al. (2020) to explicate probable metabolic pathways associated with polyurethane-degrading to alleviate plastics and xenobiotic pollution.

5. Studies are done by Aubé et al. (2020), using metagenome and enriched mRNA metatranscriptome sequencing on the persistent impact of petroleum pollutants on the taxonomic and metabolic structure of microbial mats.

6. Auti et al. (2019) have demonstrated that 16S rRNA gene sequencing analysis is a highly recommended cost-effective technique for the phylogenetic resolution and taxonomic profiling of microbial communities. As 16S rRNA gene sequence similarity between two strains provides a simple yet robust criterion for the identification of newly isolated strains, whereas phylogenetic analyses can be used to elucidate the overall evolutionary relationship between related taxa (Johnson et al. 2019).

7. Using metagenomic approach, Zhu et al. (2020) have explored microbial assemblage and functional genes potentially involved in upstream and downstream phthalate degradation in soil. Results of which indicate that bacterial taxon Actinobacteria (*Pimelobacter*, *Nocardioides*, *Gordonia*, *Nocardia*, *Rhodococcus*, and *Mycobacterium*) was a major degrader under aerobic conditions, and bacterial taxa Proteobacteria (*Ramlibacter* and *Burkholderia*), Acidobacteria, and Bacteroidetes were involved under anaerobic conditions.

8. By metagenomic analysis, Hidalgo et al. (2020) in their research have exposed that the members of *Geobacteraceae* and *Peptococcaceae* microbiota present in the jet-fuel-contaminated site could be exploited for their remarkable metabolic potential for the mitigation of toluene and benzene.

27.3.2.3 Transcriptomics

Transcriptomics is the study of an organism's transcriptome, i.e., the sum of all of its RNA transcripts. The information content of an organism is recorded in the DNA of its genome and expressed through transcription. Here, mRNA serves as a transient intermediary molecule in the information network, while noncoding RNAs perform additional diverse functions. A transcriptome captures a snapshot in time of the total transcripts present in a cell (Lowe et al. 2017).

It is also called gene expression profiling because it provides the understanding of up- or downregulation of genes under various environments in microbial communities. mRNA analysis provides a direct vision of cell and tissue-specific gene manifestation like (1) the existence, nonexistence, and assessment of transcript, (2) assessment of alternative splicing to foresee protein isoforms, and (3) quantitative evaluation of genotype impact on gene expression via expression assessable trait loci analysis or allele-specific expression (Chandran et al. 2020). Thus, transcriptomic analysis provides a large amount of gene information about the potential function of microbial communities in adaptation and survival in extreme environments (Singh et al. 2018).

There are a number of techniques in transcriptomics that supports in reviewing and evaluating mRNA expression of an organism. This includes the following:

1. Microarrays: DNA microarray is a powerful technique in transcriptomics that supports in reviewing and evaluating mRNA expression of every single gene existing in an organism. The technique has been employed to evaluate variance in metabolic and catabolic gene expressions, to analyze the microbial community physiology from diverse environments, identify new bacterial species, etc. (Dennis et al. 2003; Greene and Voordouw 2003).

2. RNA Sequencing: RNA sequencing uses next-generation sequencing to determine the amount of RNA in a sample. It is very extensive as it facilitates different types of RNA at a much-advanced coverage and broad discovery studies (Shendure 2008; Nagalakshmi et al. 2010).
   The generation of raw transcriptome data involves purification of fine RNA of interest followed by transformation of RNA to complementary DNA (cDNA), fragmenting cDNA to build a library using sequence by synthesis (RNA sequencing), running the microarray or sequence through superior software platform and carrying out ad hoc QC (Chandran et al. 2020). Thus, it a better approach to understand the basic nature and mechanism of differently expressed genes in the host and symbiotic microbes at a time (Kaur and Kaur 2016).

3. GeoChip: It is a high-throughput tool, which analyzes microbial community composition, structure, and functional activity. It uses key enzymes or genes to spot various microbe-mediated mechanisms for biogeochemical cycles, resistance mechanism for heavy metals, and degradation pathways of xenobiotics (He et al. 2010; Xiong et al. 2010; Xie et al. 2011).

4. DNA and RNA-SIP: These are both stable isotope probing technologies. They are used for probing hydrocarbon degraders. They are also valuable to uncover the microbial taxa and catabolic genes that are important for the bioremediation of polluted environments (Lueders 2015).

5. microRNAs: The regulation of gene expression can be studied also by the collective analysis of mRNA and microRNA levels. MicroRNAs (mRNAs) are short, noncoding RNA molecules that control transcription of mRNA. The precise binding of mRNAs to a target mRNA (by sequence homology) either impedes mRNA binding to the ribosome or targets it for degradation. mRNA profiling along with miRNA expression can be used to explore variations in the transcriptome profile, particularly to identify the miRNA transcripts that are subjected to regulation, emphasizing the probable molecular pathways supporting a particular trait or condition (Chandran et al. 2020).

Applications of Transcriptomics

1. Comparative transcriptomics have been used to reveal highly upregulated degradation pathways and putative transporters for phenol to improve phenol tolerance and utilization by lipid-accumulating *Rhodococcus opacus PD630* (Yoneda et al. 2016).

2. Hong et al. in year 2016 have studied hydrocarbon-degrading bacterium *Achromobacter* sp. using transcriptomics. The species was isolated from seawater and indicated that the upregulation of enzymes such as dehydrogenases, monooxygenases, and novel genes associated with fatty acid metabolism is responsible for its enormous capability for hydrocarbon degradation and survival.

3. The investigation done by Lima-Morales et al. in year 2016 using transcriptomic approach on the microbial organization and catabolic gene diversity. They have worked on three types of contaminated soil under continuous long-term pollutant stress with benzene and benzene/toluene/ethylene/xylene (BTEX). The results obtained have shown shifts in community structure and the prevalence of key genes for catabolic pathways. Moreover, de novo transcriptome synthesis gives new insights into and reveals basic information about nonmodel species without a genome reference.

4. Metatranscriptomic analysis of the wheat rhizosphere identified dominant bacterial communities of diverse taxonomic phyla, including *Acidobacteria*, *Cyanobacteria*, *Bacteroidetes*, *Streptophyta*, *Ascomycota*, and *Firmicutes*, having functional roles in the degradation of various xenobiotic pollutants (Singh et al. 2018).

5. Transcriptomic along with genomic approaches was used by Sengupta et al. in year 2019 for studying mechanistic insights of 4-nitrophenol (4-NP)-degrading bacterium *Rhodococcus* sp. strain BUPNP1. The study identified a catabolic 43 gene cluster named *nph* that harbors not only mandatory genes for the breakdown of 4-NP into acetyl co-A and succinate by nitrocatechol, but also for other diverse aromatic compounds.

6. Transcriptome analysis of activated sludge microbiomes decoded the role of the nitrifying organisms in heavy oil degradation (Sato et al. 2019).

7. Also, studies done by Das et al. in year 2020 using transcriptome analyses of crude oil-degrading Pseudomonas aeruginosa strains revealed the significance of differentially expressed genes implicated in crude oil degradation.

27.3.2.4 Metabolomics

A metabolome is the total metabolites in an organism, and the study of the metabolite profile of a cell within a given condition is called metabolomics (Beale et al. 2017). Metabolomics explores the relationships between organisms and the environment, such as organismal responses to abiotic stressors, including both natural factors such as temperature, and anthropogenic factors such as pollution, to investigate biotic–biotic interactions such as infections, and metabolic responses (Lindon et al. 2006; Griffiths 2007; Mallick et al. 2019).

Metabolomics analyzes the metabolites produced by the cell in response to changing environmental conditions, which in turn provide information about the regulatory events in a cell (Krumsiek et al. 2015). A metabolomic analysis workflow starts with sample acquisition and preparation followed by separation and detection of analytes. Detection and quantification of metabolites are normally accomplished through an amalgamation of chromatography techniques (liquid chromatography

and gas chromatography) and detection systems like mass spectrometry and nuclear magnetic resonance (Aldridge and Rhee 2014).

Applications of Metabolomics

1. Seo et al. in 2013 have investigated the degradation mechanism of carbaryl and other *N*-methyl carbamates pesticides in *Burkholderia* sp. strain C3 by using metabolomic approach. The result of this study has shown that the metabolic adaptation of *Burkholderia* sp. C3 to carbaryl in comparison with glucose and nutrient broth. The metabolic changes were notably associated with the biosynthesis and metabolism of amino acids, sugars, PAH lipids, and cofactors. Thus, this metabolomic study could provide detailed insights into bacterial adaptation to different metabolic networks and the metabolism of toxic pesticides and chemicals.

2. Wang et al. in 2019 have applied comparative metabolic approach for studying the microbial degradation of cyfluthrin by *Photobacterium ganghwense*. This approach has explored the biotransformation pathway of cyfluthrin with the identification of 156 metabolites during the biodegradation process.

3. In 2018, Li et al. on the basis of interactions of indigenous soil microorganisms to PAH-contaminated soil have that the majority of microbial metabolic functions were adversely affected to cope with PAH pollution. This study includes the combined study of enzyme activity and sequencing analysis with metabolomics, which further exposed the specific inhibition of soil metabolic pathways associated with carbohydrates, amino acids, and fatty acids due to microbial community shifting under PAH stress.

4. High-throughput sequencing and soil metabolomics were used by Song et al. in 2020 for investigating the differential structures and functions of soil bacterial communities in the pepper rhizosphere and bulk soil under plastic greenhouse vegetable cultivation (PGVC). In the study, a total of 245 metabolites were identified, among which 11 differential metabolites were detected between rhizosphere and bulk soil, including organic acids and sugars that were positively and negatively correlated with the relative abundances of the differential bacteria. A starch and sucrose metabolic pathway was the most differentially expressed pathway in rhizospheric soil. The main functional genes participating in this pathway were predicted to be downregulated in rhizosphere soil.

5. Wright et al. in 2020 also evaluated the metabolomic characterization of two potent marine bacterial isolates, *Mycobacterium* sp. *DBP42* and *Halomonas* sp. ATBC28, capable of the degradation of phthalate and plasticizers such as ATBC, DBP, and DEHP. This research study presented the molecular analysis of metabolites generated during biodegradation. It also confirmed that DBP and ATBC were degraded through the sequential removal of ester side chains and generated monobutyl phthalate and phthalate in the case of DBP degradation and citrate in the case of ATBC degradation in Mycobacterium species.

6. Metabolite pathway databases and repositories are there, which can be used to supervise and investigate the information about metabolites and their pathways. They provide a databank on metabolic information and help in the unification of

complex data into metabolic pathways. These databases and repositories also help in modeling metabolic pathways that can be investigated and prompted using mathematical modeling techniques (Chandran et al. 2020).

27.3.3 Prediction of Chemical Toxicity

Determination of chemical toxicity level, which is lethal for the survival of the degrading microbes, is very important. Several tools and computational models are present, which can predict the toxicity of chemicals involved.

QSAR-Based Models: It stands for quantitative structure regulatory activity relationship. This calculates toxicity based on the physical characteristics of the structure of chemicals such as the molecular weight or the number of benzene rings (molecular descriptors) using mathematical algorithms (Eriksson et al. 2003). There is number of tools based on QSAR:

1. VirtualToxLab—It is for prediction of the toxic potential of drugs, chemicals, and natural products. This includes endocrine and metabolic disruption, and some aspects of carcinogenicity and cardiotoxicity (Vedani et al. 2009).
2. Toxicity Estimation Software Tool (TEST)—This tool is for prediction of the acute toxicity of organic chemicals based on their molecular structures. It allows a user to estimate toxicity without requiring any external programs. Users input a chemical to evaluate by drawing it in an included chemical sketcher window, entering a structure text file, or importing it from an included database of structures. Once entered, the toxicity is estimated using one of several advanced QSAR methodologies (http://www.epa.gov/nrmrl/std/qsar/qsar.html).
3. Sarah Nexus—It is a statistical-based model used for prediction of the mutagenicity of chemicals (Barber et al. 2016).
4. TOPKAT—It is for prediction of the ecotoxicity, mutagenicity, and reproductive or developmental toxicity of chemicals (Prival 2001).
5. Ecological Structure–Activity Relationships (ECOSAR)—The Ecological Structure–Activity Relationships (ECOSAR) Class Program is a computerized predictive system that estimates aquatic toxicity. The program estimates a chemical's acute (short term) toxicity and chronic (long term or delayed) toxicity to aquatic organisms, such as fish, aquatic invertebrates, and aquatic plants, by using computerized structure–activity relationships (SAR) (http://www.epa.gov/oppt/newchems/tools/21ecosar.htm). This software is available for free.
6. Estimation Programs Interface (EPI)—The Estimation Programs Interface (EPI) Suite is a Windows-based suite of physical/chemical property and environmental fate estimation programs. It is a screening-level tool. It uses a single input to run the following estimation programs: KOWWIN, AOPWIN, HENRYWIN, MPBPWIN, BIOWIN, BioHCwin, KOCWIN, WSKOWWIN, WATERNT, BCFBAF, HYDROWIN, KOAWIN, and AEROWIN, and the fate models WVOLWIN, STPWIN, LEV3EPI, and ECOSAR (http://www.epa.gov/opptintr/exposure/pubs/episuite.htm).

7. CAESAR—The CAESAR QSAR model is developed for assessment of chemical toxicity under the REACH (Cassano et al. 2010).
8. ToxinPred—It is a web server available for prediction of aqueous toxicity of small chemical molecules in *Tetrahymena pyriformis*. It is available at http://crdd. osdd.net/raghava/toxipred. It is used for environmental risk assessment of small chemical compounds based on quantitative structure–toxicity relationship (QSTR) model (Mishra et al. 2014).
9. ACD/TOx suite—It is a tool for potential bacterial system to be employed in textile dye decolorization and degradation studies (Srinivasan et al. 2017).

27.3.4 Databases

In relation to bioremediation, the number of databases has been developed to provide information regarding chemicals and their biodegradation. Given below is the list of chemical databases:

1. TOXNET—It is developed by the National Library of Medicine (NLM), is a Web-based system of databases providing information on toxicology, hazardous chemicals, and the environment. Databases fall under the general headings of Toxicology Data, Toxicology Literature, Toxic Releases, and Chemical Identification/Nomenclature (Wexler 2001). There are various databases under it, and this includes:
   (a) CCRIS—It stands for Chemical Carcinogenesis Research Information System. The database contains chemical records with carcinogenicity, mutagenicity, tumor inhibition test results. It was developed by the National Cancer Institute (NCI). Data are derived from studies cited in primary journals, current awareness tools, NCI reports, and other sources. Test results have been reviewed by experts in carcinogenesis and mutagenesis (http://toxnet. nlm.nih.gov/cgibin/sis/htmlgen?CCRIS).
   (b) Developmental and Reproductive Toxicology Database (DART)—It provides references related to developmental and reproductive toxicology literature (http://toxnet.nlm.nih.gov/cgi-bin/sis/htmlgen?DARTETIC).
   (c) Genetic Toxicology Data Bank (GENE-TOX)—It provides genetic toxicology (mutagenicity) test data from expert peer review of open scientific literature for more than 3000 chemicals from the United States Environmental Protection Agency (EPA). It was established to select assay systems for evaluation, review data in the scientific literature, and recommend proper testing protocols and evaluation procedures for these systems (http://toxnet.nlm.nih.gov/cgi-bin/sis/htmlgen?GENETOX).
   (d) Integrated Risk Information System (IRIS)—This program supports the mission by identifying and characterizing the health hazards of chemicals found in the environment. Each IRIS assessment can cover a chemical, a group of related chemicals, or a complex mixture. IRIS assessments are an important source of toxicity information used by EPA, state and local health

agencies, other federal agencies, and international health organizations (http://toxnet.nlm.nih.gov/cgi-bin/sis/htmlgen?IRIS).

2. Biodegradative Strain Database (BSD)—It is a Web-based database that provides detailed information about biodegradative bacteria and the hazardous chemicals that they degrade (Urbance et al. 2003). It is available at http://www.bsd.cme.msu.edu/.

3. MetaRouter—It maintains varied information regarding biodegradation networks, predicting biodegradative pathways for chemical compounds (Pazos et al. 2005). It is available at http://pdg.cnb.uam.es/MetaRouter.

4. ECHA Classification & Labeling Inventory—It gives the information about the classification and labeling of substances reported and registered by manufacturers and importers (Schöning 2011).

5. N-CLASS—It stands for the Nordic N-Class Database on Environmental Hazard Classification. It provides information describing chemicals that have been or are currently being considered by the European Commission on classification and labeling for environmental effects (http://apps.kemi.se/nclass/default.asp).

6. International Toxicity Estimates for Risk (ITER)—It provides risk information for 600 chemicals from authoritative groups worldwide (Wullenweber et al. 2008).

7. ProteoWizard—It is used for rapid proteomic analysis (Kessner et al. 2008). It is available at http://proteowizard.sourceforge.net/

8. SuperToxic—It is a Web database having collection of about 60,000 toxic compounds and their structures. With the aid of implemented similarity searches, it can provide information about possible biological interactions. Also, connections to the Protein Data Bank, UniProt, and the KEGG database are available, to allow the identification of targets and the pathways, the searched compounds that are involved in Ref. Schmidt et al. (2009). This database is available online at http://bioinformatics.charite.de/supertoxic.

9. Acutoxbase—It aims to optimize and prevalidate an in vitro testing strategy for predicting acute human toxicity. The database consists of two principal parts for archiving in vitro and in vivo data, respectively. The in vitro part, designed following the principles of Good Cell Culture Practice (GCCP), provides a standard format for collection of in vitro data, together with detailed descriptions of methodologies (Standard Operating Procedures, SOPs), generated by research laboratories participating in the project (Kinsner-Ovaskainen et al. 2009).

10. Biodegradation Network-Molecular Biology (Bionemo)—The Bionemo database is available at http://bionemo.bioinfo.cnio.es. It was developed by the structural Computational Biology Group at the Spanish National Cancer Research Center. Bionemo is a manually curated database that provides information regarding proteins and genes involved in biodegradation metabolism. The protein information involves sequences, domains, and structures for proteins, whereas the genomic information involves sequences, regulatory elements, and transcription units for genes (Carbajosa et al. 2009). It complements UM-BBD, which focuses on the biochemical aspects of

biodegradation. Bionemo has been developed by manually associating sequence database entries to biodegradation reactions based on the information extracted from published articles.

11. OxDBase—It is an enzymatic database that contains all literature-cited information related to oxygenases (Arora et al. 2009). It is available at www.imtech.res.in/raghava/oxdbase/.

12. PAHbase—The PAH database contains significant information on PAH-degrading bacteria, their occurrence phylogeny, metabolic pathways, and the genetic basis of their biodegradation capability (Surani et al. 2011). It is available at http://www.pahbase.in.

13. BioRadBase—It is a comprehensive knowledge database that provides detailed information about the bioremediation of radioactive waste through microorganisms (Reena et al. 2012). It is available at http://biorad.igib.res.in.

14. BiofOmics—It is a novel, systematic, and large-scale database for the management and analysis of biofilm data from high-throughput experiment studies of microorganisms (Lourenco et al. 2012). It is available at www.biofomics.org.

15. Kyoto Encyclopedia of Genes and Genomes (KEGG)—It provides information regarding genetic, metabolic, enzymatic, and cellular progressions of microorganisms (Kanehisa et al. 2017). It is available at http://genome.ad.jp/kegg/

16. Proteomics Identifications (PRIDE)—It is a world's largest database for analysis of mass spectrometry-based proteomic data. It includes generic standard-based format that can be annotated to capture data generated using any proteomic pipeline (Vizcaino et al. 2016). It is available at http://www.ebi.ac.uk/pride/.

17. MetaboLights—It is a database for metabolomic studies that provide primary research data and metadata for cross-platform and species metabolomic studies (Kale et al. 2016). It is available at http://www.ebi.ac.uk.

18. MetaCyc—It is a database of metabolic pathways derived from the scientific experimental literature that comprises more than 2097 experimentally determined metabolic pathways from more than 2460 different organisms. This is the largest curated database of metabolic pathways of all domains of life. This database provides information regarding the metabolic pathways involved in primary and secondary metabolism with associated compounds, enzymes, and genes (Capsi et al. 2016). This database is freely available at http://metacyc.org/. It provides multiple scientific applications:
    (a) provide reference data for computational prediction of the metabolic pathways of organisms from their sequenced genomes,
    (b) support metabolic engineering,
    (c) facilitate comparison of biochemical networks, and
    (d) serve as an encyclopedia of metabolism.

19. BioCyc—This database was developed and curated by the BioCyc group at SRI international. It is available at BioCyc (http://biocyc.org/). It is a collection of more than 2988 organism-specific Pathway/Genome Databases (PGDBs). Each PGDB contains the full genome and predicted metabolic pathway of a single

organism. The pathway tool software predicts pathways using MetaCyc as a reference database.

The BioCyc PGDBs contain information about predicted operons, transport systems, and pathway hole fillers. BioCyc pathway tool-based websites offer multiple tools for querying and analysis of PGDBs, including analysis of gene expression, metabolomics, and other large-scale datasets (Capsi et al. 2016).

20. Molecular Evolutionary Genetic Analysis (MEGA 7.0)—It is used for sequence alignment, hierarchical classification, and constructing phylogenetic trees (Kumar et al. 2016). It is available at www.megasoftware.net.

27.4   Conclusion and Future Prospective

With the advent of bioinformatics, the application area of bioremediation has increased. The progressive increase in research from last few decades to present has changed the scenario a lot. The applications of genomics, proteomics, transcriptomics, and metabolomics have given in-depth knowledge of genes, proteins, and enzymes with which the ability to understand the cellular mechanism of microbes has widened. Hence, it can be concluded that this interdisciplinary approach would be supporting the bioremediation by providing distinctive and comprehensive knowledge to build new biodegradative pathways at the molecular level, new hypotheses, postulations, and paradigm for the bioremediation of contaminated living habitat. But in view of future prospective, still research is required for recognition of specific genes and protein sequences of microbes for efficaciously eliminating contamination. Also, studies related to homogeneity shared by genes and proteins involved in bioremediation practice.

References

Abatenh E, Gizaw B, Tsegaye Z, Wassie M (2017) The role of microorganisms in bioremediation - a review. Open J Environ Biol 2(1):038–046

Aldridge BB, Rhee KY (2014) Microbial metabolomics: innovation, application, insight. Curr Opin Microbiol 19:90–96. https://doi.org/10.1016/j.mib.2014.06.009

Arora PK, Bae H (2014) Integration of bioinformatics to biodegradation. Biol Proc Online 16:8. https://doi.org/10.1186/1480-9222-16-8. PMID: 24808763, PMCID: PMC4012781

Arora PK, Kumar M, Chauhan A, Raghava GP, Jain RK (2009) OxDBase: a database of oxygenases involved in biodegradation. BMC Res Notes 2:67. https://doi.org/10.1186/1756-0500-2-67

Aslam B, Basit M, Nisar MA, Khurshid M, Rasool MH (2017) Proteomics: technologies and their applications. J Chromatogr Sci 55:182–196. https://doi.org/10.1093/chromsci/bmw167

Aubé J, Senin P, Bonin P, Pringault O, Jeziorski C, Bouchez O et al (2020) Meta-omics provides insights into the impact of hydrocarbon contamination on microbial mat functioning. Microb Ecol 80:286–295. https://doi.org/10.1007/s00248-020-01493-x

Auti AM, Narwade NP, Deshpande NM, Dhotre DP (2019) Microbiome and imputed metagenome study of crude and refined petroleum-oil contaminated soils: potential for hydrocarbon degradation and plant- growth promotion. J Biosci 44:114. https://doi.org/10.1007/s12038-019-9936-9

Awasthi MK, Ravindran B, Sarsaiya S, Chen H, Wainaina S, Singh E et al (2020) Metagenomics for taxonomy profiling: tools and approaches. Bioengineered 11:356–374. https://doi.org/10.1080/21655979.2020.1736238

Barber C, Cayley A, Hanser T, Harding A, Heghes C, Vessey JD, Werner S, Weiner SK, Wichard J, Giddings A, Glowienke S, Parenty A, Brigo A, Spirkl HP, Amberg A, Kemper R, Greene N (2016) Evaluation of a statistics-based Ames mutagenicity QSAR model and interpretation of the results obtained. Regul Toxicol Pharmacol 76:7–20. https://doi.org/10.1016/j.yrtph.12.006. PMID: 26708083

Bashir Sajo M, Mohd SS (2015) An overview of pathway prediction tools for synthetic design of microbial chemical factories. AIMS Bioeng 2(1):1–14. https://doi.org/10.3934/bioeng.2015.1.1

Bastida F, Jehmlich N, Lima K, Moris BE, Richnow HH, Hernandez T et al (2016) The ecological and physiological responses of the microbial community from a semiarid soil to hydrocarbon contamination and its bioremediation using compost amendment. J Proteomic 135:162–169. https://doi.org/10.1016/j.jprot.2015.07.023

Beale DJ, Karpe AV, Ahmed W, Cook S, Morrison PD, Staley C et al (2017) A community multi-omics approach towards the assessment of surface water quality in an urban river system. Int J Environ Res Public Health 14:E303. https://doi.org/10.3390/ijerph14030303

Boutet E, Lieberherr D, Tognolli M, Schneider M, Bairoch A (2007) UniProtKB/Swiss-Prot. Methods Mol Biol 406:89–112

Brown SP, Callaham MA, Oliver AK, Jumpponen A (2013) Deep ion torrent sequencing identifies soil fungal community shifts after frequent prescribed fires in a southeastern US forest ecosystem. FEMS Microbiol Ecol 86:557–566. https://doi.org/10.1111/1574-6941.12181

Capsi R, Billington R, Ferrer L (2016) The MetaCyc database of metabolic pathways and enzymes and the Bio Cyc collection of pathways/genome databases. Nucleic Acids Res 44:D471–D480. https://doi.org/10.1093/nar/gkv1164

Carbajosa G, Trigo A, Valencia A (2009) Cases I: Bionemo: molecular information on biodegradation metabolism. Nucleic Acids Res 37(Database Issue):D598–D602

Carbonell P, Planson AG, Fichera D et al (2011) A retrosynthetic biology approach to metabolic pathway design for therapeutic production. BMC Syst Biol 5:122

Carbonell P, Planson AG, Paillard E et al (2012) Compound toxicity screening and structure-activity relationship modeling in *Escherichia coli*. Biotechnol Bioeng 109:846–850

Carbonell P, Planson AG, Faulon JL (2013) Retrosynthetic design of heterologous pathways. In: Methods in molecular biology. Springer Science Business Media, LLC, New York, NY, pp 149–173

Cassano A, Manganaro A, Martin T, Young D, Piclin N, Pintore M, Bigoni D, Benfenati E (2010) CAESAR models for developmental toxicity. Chem Cent J 4(Suppl 1):S4

Chakka D, Gudla R, Madikonda AK, Pandeeti EVP, Partasarathy S, Nandavaram A et al (2015) The organophosphate degradation (opd) island-born esterase-induced metabolic diversion in Escherichia coli and its influence on p-nitrophenol degradation. J Biol Chem 290:29920–29930. https://doi.org/10.1074/jbc.M115.661249

Chandran H, Meena M, Sharma K (2020) Microbial biodiversity and bioremediation assessment through omics approaches. Front Environ Chem 1:570326. https://doi.org/10.3389/fenvc.2020.570326

Chatsurachai S, Furusawa C, Shimizu H (2012) An in silico platform for the design of heterologous pathways in nonnative metabolite production. BMC Bioinformatics 13:93

Chen Z, Yin H, Peng H, Lu G, Liu Z, Dang Z (2019) Identification of novel pathways for biotransformation of tetrabromobisphenol A by Phanerochaete chrysosporium combined with mechanism analysis at proteome level. Sci Total Environ 659:1352–1362. https://doi.org/10.1016/j.scitotenv.2018.12.446

Chou CH, Chang WC, Chiu CM et al (2009) FMM: a web server for metabolic pathway reconstruction and comparative analysis. Nucleic Acids Res 37:W129–W134

Colatriano D, Ramachandran A, Yergeau E, Maranger R, Gelinas Y, Walsh DA (2015) Metaproteomics of aquatic microbial communities in a deep and stratified estuary. Proteomics 15:3566–3579. https://doi.org/10.1002/pmic.201500079

Das D, Mawlong GT, Sarki YN, Singh AK, Chikkaputtaiah C, Boruah HPD (2020) Transcriptome analysis of crude oil degrading Pseudomonas aeruginosa strains for identification of potential genes involved in crude oil degradation. Gene 755:144909. https://doi.org/10.1016/j.gene.2020.144909

Dell Anno A, Beolchini F, Rocchetti L, Luna GM, Danovaro R (2012) High bacterial biodiversity increases degradation performance of hydrocarbons during bioremediation of contaminated harbor marine sediments. Environ Pollut 167:85–92. Link: https://goo.gl/RHnDWP

Dennis P, Edwards EA, Liss SN, Fulthorpe R (2003) Monitoring gene expression in mixed microbial communities by using DNA microarrays. Appl Environ Microbiol 69:769–778. https://doi.org/10.1128/AEM.69.2.769-778.2003

Devarapalli P, Kumavath RN (2015) Metagenomics – a technological drift in bioremediation. In: Advances in bioremediation of wastewater and polluted soil. IntechOpen. https://doi.org/10.5772/60749

Dimitrov S, Nedelcheva D, Dimitrova N, Mekenyan O (2010) Development of a biodegradation model for the prediction of metabolites in soil. Sci Total Environ 408:3811–3816

Dong X, Greening C, Rattray JE, Chakraborty A, Chuvochina M, Mayumi D et al (2019) Metabolic potential of uncultured bacteria and archaea associated with petroleum seepage in deep-sea sediments. Nat Commun 10:1816. https://doi.org/10.1038/s41467-019-09747-0

Dua M, Singh A, Sethunathan N, Johri AK (2002) Biotechnology and bioremediation: successes and limitations. Appl Microbiol Biotechnol 59:143–152

Ellis LB, Hershberger CD, Bryan MB, Wackett LP (2001) The University of Minnesota Biocatalysis/Biodegradation database: microorganisms, genomics and prediction. Nucleic Acids Res 29(1):340–343

Eriksson L, Jaworska J, Worth A, Cronin M, McDowell RM, Gramatica P (2003) Methods for reliability, uncertainty assessment, and applicability evaluations of regression based and classification QSARs. Environ Health Perspect 111:1361–1375

Fehér T, Planson AG, Carbonell P et al (2014) Validation of RetroPath, a computer-aided design tool for metabolic pathway engineering. Biotechnol J 9:1446–1457

Felczykowska A, Krajewska A, Zielínska S, Łós JM, Bloch SK, Nejman-Faléncvyk B (2015) The most widespread problems in the function-based microbial metagenomics. Acta Biochim Pol 62:161–166. https://doi.org/10.18388/abp.2014_917

Fenner K, Gao J, Kramer S, Ellis L, Wackett L (2008) Data-driven extraction of relative reasoning rules to limit combinatorial explosion in biodegradation pathway prediction. Bioinformatics 24:2079–2085. https://doi.org/10.1093/bioinformatics/btn378

Ferrer M, Beloqui A, Vieites JM, Guazzaroni ME, Berger I, Aharoni A (2009) Interplay of metagenomics and in vitro compartmentalization. Microb Biotechnol 2:31–39. https://doi.org/10.1111/j.1751-7915.2008.00057.x

Finley SD, Broadbelt LJ, Hatzimanikatis V (2009) Computational framework for predictive biodegradation. Biotechnol Bioeng 104:1086–1097

Fulekar MH, Sharma J (2008) Bioinformatics Applied in Bioremediation. Innov Roman Foor Biotechnol 2(2):28–36

Gao J, Ellis LB, Wackett LP (2011) The University of Minnesota pathway prediction system: multi-level prediction and visualization. Nucleic Acids Res 39(2):W406–W411

Gaytán I, Sánchez-Reyes A, Burelo M, Vargas-Suárez M, Liachko I, Press M et al (2020) Degradation of recalcitrant polyurethane and xenobiotic additives by a selected landfill microbial community and its biodegradative potential revealed by proximity ligation-based metagenomic analysis. Front Microbiol 10:2986. https://doi.org/10.3389/fmicb.2019.02986

Gomez MJ et al (2007) The environmental fate of organic pollutants through the global microbial metabolism. Mol Syst Biol 3:114

Gonzalez-Lergier J, Broadbelt LJ, Hatzimanikatis V (2005) Theoretical considerations and computational analysis of the complexity in poly-ketide synthesis pathways. J Am Chem Soc 127: 9930–9938

Greene EA, Voordouw G (2003) Analysis of environmental microbial communities by reverse sample genome probing. J Microbiol Methods 53:211–219. https://doi.org/10.1016/S0167-7012(03)00024-1

Gregson BH, Metodieva G, Metodiev MV, Golyshin PN, McKew BA (2020) Protein expression in the obligate hydrocarbon-degrading psychrophile Oleispira antarctica RB-8 during alkane degradation and cold tolerance. Environ Microbiol 22:1870–1883. https://doi.org/10.1111/1462-2920.14956

Griffiths W (2007) Metabolomics, metabonomics and metabolite profiling. Royal Society of Chemistry, Cambridge. https://doi.org/10.1039/9781847558107

Grob C, Taubert M, Howat AM, Burns OJ, Dixon JL, Richnow HH et al (2015) Combining metagenomics with metaproteomics and stable isotope probing reveals metabolic pathways used by a naturally occurring marine methylotroph. Environ Microbiol 17:4007–4018. https://doi.org/10.1111/1462-2920.12935

Hallin PF, Binnewies TT, Ussery DW (2008) The genome BLAST atlas-a Gene Wiz extension for visualization of whole-genome homology. Mol BioSyst 4:363–371. https://doi.org/10.1039/b717118h

Han D, Gao P, Li R, Tan P, Xie J, Zhang R et al (2020) Multicenter assessment of microbial community profiling using 16S rRNA gene sequencing and shotgun metagenomic sequencing. J Adv Res 26:111. https://doi.org/10.1016/j.jare.2020.07.010

Hatzimanikatis V, Li C, Ionita JA, Broadbelt LJ (2004) Metabolic networks: enzyme function and metabolite structure. Curr Opin Struct Biol 14:300–306. PubMed 15193309

Hatzimanikatis V, Li C, Ionita JA et al (2005) Exploring the diversity of complex metabolic networks. Bioinformatics 21:1603–1609

He Z, Deng Y, van Nostrand JD, Xu M, Hemme LH, Tu Q et al (2010) GeoChip 3.0 as a high-throughput tool for analyzing microbial community composition, structure and functional activity. ISME J 4:1167–1179. https://doi.org/10.1038/ismej.2010.46

Heath AP, Bennett GN, Kavraki LE (2010) Finding metabolic pathways using atom tracking. Bioinformatics 26:1548–1555

Henry CS, Broadbelt LJ, Hatzimanikatis V (2010) Discovery and analysis of novel metabolic pathways for the biosynthesis of industrial chemicals: 3-hydroxypropanoate. Biotechnol Bioeng 106:462–473

Hidalgo KJ, Teramoto EH, Soriano AU, Valoni E, Baessa MP, Richnow HH et al (2020) Taxonomic and functional diversity of the microbiome in a jet fuel contaminated site as revealed by combined application of in situ microcosms with metagenomic analysis. Sci Total Environ 708:135152. https://doi.org/10.1016/j.scitotenv.2019.135152

Hillmann B, Al-Ghalith GA, Shields-Cutler RR, Zhu Q, Gohl DM, Beckman KB et al (2018) Evaluating the information content of shallow shotgun metagenomics. mSystems 3:e00069–e00018. https://doi.org/10.1128/mSystems.00069-18

Hodkinson BP, Grice EA (2015) Next-generation sequencing: a review of technologies and tools for wound microbiome research. Adv Wound Care 4:50–58. https://doi.org/10.1089/wound.2014.0542

Hong YH, Deng MC, Xu XM, Wu CF, Xiao X, Zhu Q (2016) Characterization of the transcriptome of Achromobacter sp. HZ01 with the outstanding hydrocarbon-degrading ability. Gene 584: 185–194. https://doi.org/10.1016/j.gene.2016.02.032

Hoops S, Sahle S, Gauges R et al (2006) COPASI--a COmplex PAthway SImulator. Bioinformatics 22:3067–3074

Jagadeesh DS, Kannegundla U, Reddy RK (2017) Application of proteomic tools in food quality and safety. Adv Anim Vet Sci 5:213–225. https://doi.org/10.17582/journal.aavs/2017/5.5.213.225

Jeffries TC, Rayu S, Nielsen UN, Lai K, Ijaz A, Nazaries L et al (2018) Metagenomic functional potential predicts degradation rates of a model organophosphorus xenobiotic in pesticide contaminated soils. Front Microbiol 9:147. https://doi.org/10.3389/fmicb.2018.00147

Johnson SJ, Spakowicz DJ, Hong B-Y, Petersen L, Demkowicz P, Chen L et al (2019) Evaluation of 16S rRNA gene sequencing for species and strain-level microbiome analysis. Nat Commun 10:5029. https://doi.org/10.1038/s41467-019-13036-1

Kale NS, Haug K, Conesa P, Jayseelam K, Moreno P, Rocca-Serra P et al (2016) Metabo Lights: an open-access database repository for metabolomics data. Curr Protoc Bioinformatics 53:14. https://doi.org/10.1002/0471250953.bi1413s53

Kanehisa M, Furumichi M, Tanabe M (2017) KEGG: new perspectives on genome, pathways, diseases and drugs. Nucleic Acids Res 45:D353–D361. https://doi.org/10.1093/nar/gkw1092

Kaur H, Kaur G (2016) Application of ligninolytic potentials of a white-rot fungus Ganoderma lucidum for degradation of lindane. Environ Monit Assess 188:588. https://doi.org/10.1007/s10661-016-5606-7

Keller M, Hettich R (2009) Environmental proteomics: a paradigm shift in characterizing microbial activities at the molecular level. Microbiol Mol Biol Rev 73:62–70. https://doi.org/10.1128/MMBR.00028-08

Kessner D, Chambers M, Burke R (2008) ProteoWizard: open source software for rapid proteomics tools development. Bioinformatics 24:2534–2536. https://doi.org/10.1093/bioinformatics/btn323

Kinsner-Ovaskainen A, Rzepka R, Rudowski R, Coecke S, Cole T, Prieto P (2009) Acutoxbase, an innovative database for in vitro acute toxicity studies. Toxicol in Vitro 23:476–485

Klopman G et al (1997) Meta 3 a genetic algorithm for metabolic transform priorities optimization. J Chem Inf Comput Sci 37:329–334

Krumsiek J, Mittelstrass K, Do KT, Stückler F, Ried J, Adamski J et al (2015) Gender-specific pathway differences in the human serum metabolome. Metabolomics 11:1815–1833. https://doi.org/10.1007/s11306-015-0829-0

Kumar A, Bisht BS, Joshi VD, Dhewa T (2011) Review on bioremediation of polluted environment: a management tool. Int J Environ Sci 1:1079–1093. https://goo.gl/P6Xeqc

Kumar SS, Shantkriti S, Muruganandham T, Murugesh E, Rane N, Govindwar SP (2016) Bioinformatics aided microbial approach for bioremediation of wastewater containing textile dyes. Ecol Info 31:112–121. https://doi.org/10.1016/j.ecoinf.2015.12.001

Le Novere N, Bornstein B, Broicher A et al (2006) BioModels Database: a free, centralized database of curated, published, quantitative kinetic models of biochemical and cellular systems. Nucleic Acids Res 34:D689–D691

Lee SY, Sekhon SS, Ban YH, Ahn JY, Ko JH, Lee L et al (2016) Proteomic analysis of polycyclic aromatic hydrocarbons (PAHs) degradation and detoxification in Sphingobiumchungbukense DJ77. J Microbiol Biotechnol 26:1943–1950. https://doi.org/10.4014/jmb.1606.06005

Li C, Ma Y, Mi Z, Huo R, Zhou T, Hai H et al (2018) Screening for Lactobacillus plantarum strains that possess organophosphorus pesticide-degrading activity and metabolomic analysis of phorate degradation. Front Microbiol 9:2048. https://doi.org/10.3389/fmicb.2018.02048

Lima-Morales D, Jauregui R, Camarinha-Silva A, Geffers R, Pieper DH, Vilchez-Vergas R (2016) Linking microbial community and catabolic gene structures during the adaptation of three contaminated soils under continuous long-term polluted stress. Appl Environ Microbiol 82:2227–2237. https://doi.org/10.1128/AEM.03482-15

Lindon JC, Nicholson JK, Holmes E (2006) The handbook of metabonomics and metabolomics. Elsevier Science, London

Liu S, Gu C, Dang Z, Liang X (2017) Comparative proteomics reveal the mechanism of Tween 80 enhanced phenanthrene biodegradation by Sphingomonas sp. GY2B. Ecotoxicol Environ Saf 137:256–264. https://doi.org/10.1016/j.ecoenv.2016.12.015

Lourenco A, Ferreira A, Veiga N, Machado I, Pereira MO, Azevedo NF (2012) Biof Omics: a web platform for the systematic and standardized collection of high-through put biofilm data. PLoS One 7:e39960. https://doi.org/10.1371/journal.pone.0039960

Lowe R, Shirley N, Bleackley M, Dolan S, Shafee T (2017) Transcriptomics technologies. PLoS Comput Biol 13(5):e1005457. https://doi.org/10.1371/journal.pcbi.1005457. PMID: 28545146, PMCID: PMC5436640

Lueders T (2015) DNA-and RNA based stable isotope probing of hydrocarbon degraders. In: McGenity TJ, Timmis KN, Nogales B (eds) Hydrocarbon and lipid microbiology protocols. Humana Press, New York, NY, pp 181–197. https://doi.org/10.1007/8623_2015_74

Luo C, Tsementzi D, Kyrpides N (2012) Direct comparison of Illumina vs. Roche 454 sequencing technologies on the same microbial community DNA sample. PLoS One 7:e30087. https://doi.org/10.1371/journal.pone.0030087

Mallick H, Franzosa EA, Mclver LJ, Banerjee S, Sirota-Madi A, Kostic AD et al (2019) Predictive metabolomic profiling of microbial communities using amplicon or metagenomic sequences. Nat Commun 10:3136. https://doi.org/10.1038/s41467-019-10927-1

Marchant CA, Briggs KA, Long A (2008) In silico tools for sharing data and knowledge on toxicity and metabolism: derek for windows, meteor, and vitic. Toxicol Mech Methods 18:177–187

McClymont K, Soyer OS (2013) Metabolic tinker: an online tool for guiding the design of synthetic metabolic pathways. Nucleic Acids Res 41(11):e113. https://doi.org/10.1093/nar/gkt234

Medema MH, van Raaphorst R, Takano E et al (2012) Computational tools for the synthetic design of biochemical pathways. Nat Rev Microbiol 10:191–202

Meena M, Zehra A, Dubey MK, Aamir M, Gupta VK, Upadhyay RS (2016) Comparative evaluation of biochemical changes in tomato (*Lycopersicon esculentum* Mill.) infected by Alternaria alternata and its toxic metabolites (TeA, AOH, and AME). Front Plant Sci 7:1408. https://doi.org/10.3389/fpls.2016.01408

Mishra NK, Singla D, Agarwal S (2014) Consortium OSDD, Raghava GPS: ToxiPred: a server for prediction of aqueous toxicity of small chemical molecules in T. Pyriformis. J Transl Toxicol 1: 21–27

Moriya Y, Shigemizu D, Hattori M, Tokimatsu T, Kotera M, Goto S, Kanehisa M (2010) Path Pred: an enzyme-catalyzed metabolic pathway prediction server. Nucleic Acids Res 38(Web Server Issue):W138–W143. https://doi.org/10.1093/nar/gkq318. PMCID: PMC2896155

Mu F et al (2006) Prediction of oxidoreductase-catalyzed reactions based on atomic properties of metabolites. Bioinformatics 22:3082–3088

Muccee F, Ejaz S (2020) Whole genome shotgun sequencing of POPs degrading bacterial community dwelling tannery effluents and petrol contaminated soil. Microbiol Res 238:126504. https://doi.org/10.1016/j.micres.2020.126504

Nagalakshmi U, Waern K, Snyder M (2010) RNA-Seq: a method for comprehensive transcriptome analysis. Curr Protoc Mol Biol 4:1–13. https://doi.org/10.1002/0471142727.mb0411s89

Nascimento FX, Hernandez G, Glick BR, Rossi MJ (2020) Plant growth-promoting activities and genomic analysis of the stress resistant Bacillus megaterium STB1, a bacterium of agriculture and biotechnological interest. Biotechnol Rep 25:e00406. https://doi.org/10.1016/j.btre.2019.e00406

Niu J, Rang Z, Zhang C (2016) The succession pattern of soil microbial communities and its relationship with tobacco bacterial wilt. BMC Microbiol 16:233. https://doi.org/10.1186/s12866-016-0845-x

Nzila A, Ramirez CO, Musa MM, Sankara S, Basheer C, Li QX (2018) Pyrene biodegradation and proteomic analysis in Achromobacter xylosoxidans, PY4 strain. Int Biodeterior Biodegrad 175: 1294–1305. https://doi.org/10.1016/j.ibiod.2018.03.014

Okoh A (2006) Biodegradation alternative in the Cleanup of petroleum hydrocarbon pollutants. Microbiol Mol Biol Rev 1:38–50

Pazos F, Guijas D, Valencia A, De Lorenzo V (2005) Meta Router: bioinformatics for bioremediation. Nucleic Acids Res 45:D588–D592. https://doi.org/10.1093/nar/gki068

Peabody MA, van Rossum T, Lo R, Brinkman FSL (2015) Evaluation of shotgun metagenomics sequence classification methods using in silico and in vitro simulated communities. BMC Bioinformatics 16:363. https://doi.org/10.1186/s12859-015-0788-5

Pitkänen E et al (2009) Inferring branching pathways in genome-scale metabolic networks. BMC Syst Biol 3:103

Prival MJ (2001) Evaluation of the TOPKAT system for predicting the carcinogenicity of chemicals. Environ Mol Mutagen 37(1):55–69

Quince C, Walker AW, Simpson JT, Loman NJ, Segata N (2017) Shotgun metagenomics, from sampling to analysis. Nat Biotechnol 35:833–844. https://doi.org/10.1038/nbt.3935

Rahimi T, Niazi A, Deihimi T, Taghavi SM, Avatollahi S, Ebrahimie E (2018) Genome annotation and comparative genomic analysis of Bacillus subtilis MJ01 a new biodegradation strain isolated from oil contaminated soil. Funct Integr Genomics 18:533–543. https://doi.org/10.1007/s10142-018-0604-1

Ranjan R, Rani A, Metwally A, McGee HS, Perkins DL (2016) Analysis of the microbiome: advantages of whole genome shotgun versus 16S amplicon sequencing. Biochem Biophys Res Commun 469:967–977. https://doi.org/10.1016/j.bbrc.2015.12.083

Reena R, Majhi MC, Arya AK, Kumar R, Kumar A (2012) BioRadBase: a database for bioremediation of radioactive waste. Afr J Biotechnol 11:8718–8721. https://doi.org/10.5897/AJB12.020

Rocha I, Maia P, Evangelista P et al (2010) OptFlux: an open-source software platform for in silico metabolic engineering. BMC Syst Biol 4:45

Rodrigo G, Carrera J, Prather KJ et al (2008) DESHARKY: automatic design of metabolic pathways for optimal cell growth. Bioinformatics 24:2554–2556

Samanta S, Singh O, Jain RK (2002) Polycyclic aromatic hydrocarbons: environmental pollution and bioremediation. Trends Biotechnol 20:243–248. https://doi.org/10.1016/S0167-7799(02)01943-1

Samorodnitsky E, Jewell BM, Hagopian R, Miya J, Wing MR, Lyon E et al (2015) Evaluation of hybridization capture versus amplicon-based methods for whole-exome sequencing. Hum Mutat 36:903–914. https://doi.org/10.1002/humu.22825

Sato Y, Hori T, Koike H, Navarro RR, Ogata A, Habe H (2019) Transcriptome analysis of activated sludge microbiomes reveals an unexpected role of minority nitrifiers in carbon metabolism. Commun Biol 2:179. https://doi.org/10.1038/s42003-019-0418-2

Schaber J (2012) Easy parameter identifiability analysis with COPASI. Biosystems 110:183–185

Schellenberger J, Park JO, Conrad TM et al (2010) BiGG: a Biochemical Genetic and Genomic knowledgebase of large scale metabolic reconstructions. BMC Bioinformatics 11:213

Schmidt U, Struck S, Gruening B, Hossbach J, Jaeger IS, Parol R, Lindequist U, Teuscher E, Preissner R (2009) SuperToxic: a comprehensive database of toxic compounds. Nucleic Acids Res 37(Database Issue):D295–D299

Scholer A, Jacquiod S, Vestergaard G (2017) Analysis of soil microbial communities based on amplicons sequencing of marker genes. Biol Fertil Soils 53:485–489. https://doi.org/10.1007/s00374-017-1205-1

Scholz MB, Lo CC, Chain PS (2012) Next generation sequencing and bioinformatic bottlenecks: the current state of metagenomic data analysis. Curr Opin Biotechnol 23:9–15. https://doi.org/10.1016/j.copbio.2011.11.013

Schomburg I, Chang A, Placzek S et al (2013) BRENDA in 2013: integrated reactions, kinetic data, enzyme function data, improved disease classification: new options and contents in BRENDA. Nucleic Acids Res 41:D764–D772

Schöning G (2011) Classification & labelling inventory: role of ECHA and notification requirements. Ann Ist Super Sanita 47(2):140–145

Sengupta K, Swain MT, Livingstone PG, Whiteworth DE, Saha P (2019) Genome sequencing and comparative transcriptomics provide holistic view of 4-nitrophenol degradation and concurrent fatty acid catabolism by Rhodococcus sp. strain BUPNP1. Front Microbiol 9:3209. https://doi.org/10.3389/fmicb.2018.03209

Seo J, Keum YS, Li QX (2013) Metabolomic and proteomic insights into carbaryl catabolism by Burkholderia sp. C3 and degradation of ten N-methylcarbamates. Biodegradation 24:795–811. https://doi.org/10.1007/s10532-013-9629-2

Sharpton TJ (2014) An introduction to the analysis of shotgun metagenomic data. Front Plant Sci 5: 209. https://doi.org/10.3389/fpls.2014.00209

Shendure J (2008) The beginning of the end for microarrays? Nat Methods 5:585–587. https://doi.org/10.1038/nmeth0708-585

Shokralla S, Gibson JF, Niknakht H (2014) Nest-generation DNA barcoding: using next generation sequencing to enhance and accelerate DNA barcode capture from single specimen. Mol Ecol Resour 14:892–901. https://doi.org/10.1111/1755-0998.12236

Silva CC, Hayden H, Sawbridge T, Mele P, De Paula SO, Silva LCF et al (2013) Identification of genes and pathways related to phenol degradation in metagenomic libraries from petroleum refinery wastewater. PLoS One 8:e61811. https://doi.org/10.1371/journal.pone.0061811

Singh OV (2006) Proteomics and metabolomics: the molecular make-up of toxic aromatic pollutant bioremediation. Proteomics 6:5481–5492

Singh J, Behal A, Singla N, Joshi A, Birbian N, Singh S et al (2009) Metagenomics: concept, methodology, ecological inference and recent advances. Biotechnol J 4:480–494. https://doi.org/10.1002/biot.200800201

Singh V, Gohil N, Ramírez García R, Braddick D, Fofié CK (2018) Recent advances in CRISPR-Cas9 genome editing technology for biological and biomedical investigations. J Cell Biochem 119:81–94. https://doi.org/10.1002/jcb.26165

Soh KC, Hatzimanikatis V (2010) Dreams of metabolism. Trends Biotechnol 28(10):501–508. https://doi.org/10.1016/j.tibtech.2010.07.002. PMID: 20727603

Song Y, Li X, Yao S, Yang X, Jiang X (2020) Correlations between soil metabolomics and bacterial community structures in the pepper rhizosphere under plastic greenhouse cultivation. Sci Total Environ 728:138439. https://doi.org/10.1016/j.scitotenv.138439

Srinivasan S, Shanmugam G, Surwase SV, Jadhav JP, Sadasivam SK (2017) In silico analysis of bacterial systems for textile azo dye decolorization and affirmation with wetlab studies. CLEAN Soil Air Water 45:1600734

Sueoka K, Satoh H, Onuki M, Mino T (2009) Microorganisms involved in anaerobic phenol degradation in the treatment of synthetic coke-oven wastewater detected by RNA stable-isotope probing. FEMS Microbiol Lett 291:169–174. https://doi.org/10.1111/j.1574-6968.2008.01448.x

Surani JJ, Akbari VG, Purohit MK, Singh SP (2011) Pahbase, a freely available functional database of polycyclic aromatic hydrocarbons (Pahs) degrading bacteria. J Bioremed Biodegrad 2:116–135. https://doi.org/10.4172/2155-6199.1000116

Urbance JW, Cole J, Saxman P (2003) BSD: the biodegradative strain database. Nucleic Acids Res 31:152–155. https://doi.org/10.1093/nar/gkg032

Vandera E, Samiotaki A, Parapouli M, Panayotou G, Koukkou AI (2015) Comparative proteomic analysis of Arthrobacter phenanthrenivorans Sphe3 on phenanthrene, Phthalate and glucose. J Proteomic 115:73–89. https://doi.org/10.1016/j.jprot.2014.08.018

Vedani A, Smiesko M, Spreafico M, Peristera O, Dobler M (2009) Virtual ToxLab–in silico prediction of the toxic (endocrine-disrupting) potential of drugs, chemicals and natural products: two years and 2,000 compounds of experience: a progress report. ALTEX 26(3):167–176

Velmurgan N, Lee H, Cha HJ, Lee YS (2017) Proteomic analysis of the marine-derived fungus Paecilomyces sp. strain SF-8 in response to polycyclic aromatic hydrocarbons. Bot Mar 60:101. https://doi.org/10.1515/bot-2016-0101

Vermote L, Verce M, de Vuyst L, Weckx S (2018) Amplicon and shotgun metagenomic sequencing indicates that microbial ecosystems present in cheese brines reflect environmental inoculation during the cheese production process. Int Dairy J 87:44–53. https://doi.org/10.1016/j.idairyj.2018.07.010

Vizcaino JA, Csordas A, del-Toro N, Dianes JA, Griss J, Lavidas I et al (2016) Update of the PRIDE database and its related tools. Nucleic Acids Res 44:D447–D456. https://doi.org/10.1093/nar/gkw880

Wang T, Hu C, Zhang R, Sun A, Li D, Shi X (2019) Mechanism study of cyfluthrin biodegradation by Photobacterium ganghwense with comparative metabolomics. Appl Microbiol Biotechnol 103:473–488. https://doi.org/10.1007/s00253-018-9458-7

Wei K, Yin H, Peng H, Liu Z, Lu G, Dang Z (2017) Characteristics and proteomic analysis of pyrene degradation by Brevibacillus brevis in liquid medium. Chemosphere 178:80–87. https://doi.org/10.1016/j.chemosphere.03.049

Wexler P (2001) TOXNET: an evolving web resource for toxicology and environmental health information. Toxicology 157:3–10

Wicker J, Fenner K, Ellis L, Wackett L, Kramer S (2010) Predicting biodegradation products and pathways: a hybrid knowledge-and machine learning-based approach. Bioinformatics 26:814–821

Williams TJ, Wilkins D, Long E, Evans F, DeMaere MZ, Raftery MJ et al (2013) The role of planktonic Flavobacteria in processing algal organic matter in coastal East Antartica revealed using metagenomics and metaproteomics. Environ Microbiol 15:1302–1317. https://doi.org/10.1111/1462-2920.12017

Wong DWS (2018) Gene targeting and genome editing. The ABCs of gene cloning. Springer, Cham, pp 187–197. https://doi.org/10.1007/978-3-319-77982-9_20

Wright R, Bosch R, Gibson MI, Christie-Oleza J (2020) Plasticizer degradation by a marine bacterial isolates: a proteogenomic and metabolomic characterization. Environ Sci Technol 54:2244–2256. https://doi.org/10.1021/acs.est.9b05228

Wu YR, Luo ZH, Kwok-Kei Chow R, Vrijmoed LLP (2010) Purification and characterization of an extracellular laccase from the anthracene-degrading fungus Fusarium solani MAS2. Bioresour Technol 101:9772–9777. https://doi.org/10.1016/j.biortech.2010.07.091

Wullenweber A, Kroner O, Kohrman M, Maier A, Dourson M, Rak A, Wexler P, Tomljanovic C (2008) Resources for global risk assessment: the International Toxicity Estimates for Risk (ITER) and Risk Information Exchange (RiskIE) databases. Toxicol Appl Pharmacol 233:45–53

Xie J, He Z, Liu X, Van Nostrand JD, Deng Y (2011) GeoChip based analysis of functional gene diversity and metabolic potential of microbial communities in acid mine drainage. Appl Environ Microbiol 77:991–999. https://doi.org/10.1128/AEM.01798-10

Xiong JB, Wu LY, Tu SX, Van Nostrand JD, He ZH, Zhou JZ et al (2010) Microbial communities and functional genes associated with soil arsenic contamination and the rhizosphere of arsenic-hyperaccumulating plant Pteris vittata L. Appl Environ Microbiol 76:7277–7284. https://doi.org/10.1128/AEM.00500-10

Yates JR, Ruse CL, Nakorchevsky A (2009) Proteomics by mass spectrometry: approaches, advances and applications. Annu Rev Biomed Eng 11:49–79. https://doi.org/10.1146/annurev-bioeng-061008-124934

Yergeau E, Michel C, Tremblay J, Niemi A, King TL, Wyglinski J et al (2017) Metagenomic survey of the taxonomic and functional microbial communities of seawater and sea ice from the Canadian Arctic. Sci Rep 7:42242. https://doi.org/10.1038/srep42242

Yoneda A, Henson WR, Goldner NK, Park KJ, Forsberg KJ, Kim SJ et al (2016) Comparative transcriptomics elucidates adaptive phenol tolerance and utilization in lipid-accumulating Rhodococcus opacus PD630. Nucleic Acids Res 44:2240–2254. https://doi.org/10.1093/nar/gkw055

Yu Y, Yin H, Peng H, Lu G, Dang Z (2019) Proteomic mechanism of decabromodiphenyl ether (BDE-209) biodegradation by Microbacterium Y2 and its potential in remediation of BDE-209 contaminated water-sediment system. J Hazard Mater 387:121708. https://doi.org/10.1016/j.jhazmat.2019.121708

Zafra G, Taylor TD, Absalon AE, Cortes-Espinosa DV (2016) Comparative metagenomic analysis of PAH degradation in soil by a mixed microbial consortium. J Hazard Mater 318:702–710. https://doi.org/10.1016/j.jhazmat.2016.07.060

Zhang C, Bennett GN (2005) Biodegradation of xenobiotics by anaerobic bacteria. Appl Microbiol Biotechnol 67:600–618. https://doi.org/10.1007/s00253-004-1864-3

Zhou J, He Z, Yang Y (2015) High-throughput metagenomic technologies for complex microbial community analysis: open and closed formats. MBio 6:e02288–e02214. https://doi.org/10.1128/mBio.02288-14

Zhu Y, Klompe SE, Vlot M, van der Oost J, Staals RH (2018) Shooting the messenger: RNA-targeting CRISPR-Cas systems. Biosci Rep 38:BSR20170788. https://doi.org/10.1042/BSR20170788

Zhu F, Doyle E, Zhu C, Zhou D, Gu C, Gao J (2020) Metagenomic analysis exploring microbial assemblages and functional genes potentially involved in di (2-ethylhexyl) phthalate degradation in soil. Sci Total Environ 715:137037. https://doi.org/10.1016/j.scitotenv.2020.137037

Surbhi Khanna is a Master's of Philosophy in Bioinformatics from University of Kerala. She has expertise in various bioinformatic computational tools, techniques, and databases along with computer languages like Python, HTML, and Perl. Also, she is skilled in analyzing large molecular datasets such as raw microarray data, next-generation sequencing (NGS), and omic data for clinical or basic research purposes. She has served impetusly for the research assignment from DBT (Department of Biotechnology)-funded Bioinformatics Infrastructure Facility, Gargi College, Delhi University. She has also worked as Patent Research Analyst in leading law firm in New Delhi. With knowledge and experience in the field of biotechnology and bioinformatics, she aims to elucidate its various applications in the field of biological sciences.

Ashwani Kumar Alexander von Humboldt Fellow (Germany), MSc (Botany) (Gold Medalist) (1967)), PhD (1971), post-doc AvH Fellow (Germany) and JSPS fellow (Japan) got selected in the Indian Administrative Services (IAS: IPS) (1972), but he opted career in botany. 37 years of teaching and research experience. He guided 39 PhD. research students and published 320 research papers and 26 books. He elected as fellow FBS, FPSI, FIMPS, and FIFS. He served as Editor in Chief in IJLS. H warded V. Puri Medal (2008) and CEE Award (2015). He served as President of CHEC (India), Indian Botanical Society, Society for Promotion of Plant Sciences. Professor Emeritus, Department of University of Rajasthan, Jaipur. Presently, he is Chair Professor at JVWU, Jaipur, and Adj. Professor Mody University, Sikar.