Proc Indian Natn Sci Acad **85** No. 4 December 2019 pp. 999-1023
© *Printed in India*. DOI: 10.16943/ptinsa/2019/49722

Review Article

Microbial Genomics and Metagenomics in India: Explorations and Perspectives

PRINCY HIRA[1], ABHAY BAJAJ[1,2], AKSHITA PURI[1], CHANDNI TALWAR[1], HIMANI KHURANA[1], RAM KRISHAN NEGI[1], YOGENDRA SINGH[1], RUP LAL[1,3,*] and MALLIKARJUN SHAKARAD[1*]
[1]*Department of Zoology, University of Delhi, Delhi 110 007, India*
Present Address: [2]*CSIR-National Environmental Engineering Research Institute (CSIR-NEERI), Nehru Marg, Nagpur 440 020, India*
[3]*The Energy and Resources Institute, Darbari Seth Block, IHC Complex, New Delhi 110 003, India*

(Received on 03 April 2019; Accepted on 15 October 2019)

The discipline of genomics and metagenomics has been rapidly growing as an emerging field of research since the last decade. This development is largely due to the rapid advances in sequencing technologies that generate enormous amount of high throughput data at low costs. While there is huge information on genomics and metagenomics from different parts of the world, there have been few studies from India due to the underlying limitation of data analyses. In this review, major genomic and metagenomic studies to predict the bacterial diversity and their ecological significance, and the functional roles played by microbes in different environments in India have been summarized. The integration of culture dependent and culture-independent approaches has expanded our knowledge on the composition and function of microbial communities in specific niches. However, there still exist challenges and gaps in our understanding about the microbial resource in Indian subcontinent as their importance is over-looked by the scientific community despite quite a few reports on metagenomics/ genomics from India.

Keywords: Genomics; Metagenomics; Microbial Diversity; Next-generation Sequencing

Introduction

Concurrent with the advances in sequencing technologies and computational tools, the discipline of microbial diversity, genomics and metagenomics has grown exponentially during the past two decades. First genome of a bacterium *Haemophilus influenza* was sequenced in 1995 (Fleischmann *et al*., 1995). Since then, the number of complete prokaryotic genomes sequenced globally till March 2018 are approximately 7600 (ncbi.nlm.nih.gov/genome).

Despite the challenges in genomics methods and data analyses the science of genomics has rapidly extended our knowledge on distribution and functional capabilities of microbes. The data on genomics and metagenomics is of particular interest for many biotechnological and industrial applications. One of the applications of genomics lies in the area of medicine where integration of genomics data with interdisciplinary approaches has drastically broadened our understanding of disease and drug response (Manolio *et al*., 2013). Besides, studies on microbial genomics of a particular niche by culture-independent approaches, called metagenomics, has also taken new dimensions in microbial studies.

While there is flood of information on genomics and metagenomics that has poured in during recent years from different parts of the world, there have been very few studies on these aspects from India. In this review, we summarize the work that has been carried out in India to unravel the microbial gene pool inhabiting different environments using both culture-dependent and culture-independent approaches.

Genomics

Microbial genomics studies from India are scanty at

**Authors for Correspondence: E-mail: ruplal@gmail.com; beelab.ms@gmail.com*

best due to lack of concerted effort in exploring the genomes of microbes from different environments. For the sake of suitability, we have grouped the meager studies under subsections pathogens, industrially important actinobacteria and cultivable from numerous stressed niches like contaminated soils, saline and hot-springs (Supplementary Table 1). Despite the challenges in genomic methods and data analyses, the science of genomics has rapidly extended our knowledge of distribution and functional capabilities of microbes from the western world and Europiean countries. The advancement in genome sequencing has made available genomic information of a large number of bacteria, which can now be used for taxonomic characterization of a strain. This new field of "Taxo-genomics" has upgraded the standards of taxonomical classification by complementing the results inferred from classical polyphasic approach of taxonomic classification with that of genomic information (Mahato *et al.,* 2017).

Pathogens

India represents major hub of human diseases-notably tuberculosis and leprosy, among many other infectious diseases. Various research institutions have initiated programs to decipher the genomic makeup of the causative organisms to gain better understanding of their evolutionary relationships, functional divergence and acquisition of virulence. Although the initial genome sequencing from India was done for a uropathogenic strain of *Escherichia coli* NA114 (Avasthi *et al.*, 2011), a resonably detailed information was provided for *Mycobacterium indicus pranii* (*MIP*) through genome wide comparisons (Saini *et al*., 2012). *MIP* genome was sequenced from India in 2012 by collaborative efforts at Centre for DNA Fingerprinting and Diagnostics (CCFD), University of Delhi and University of Hyderabad. This study involved genome wide comparisons along with molecular phylogenetic analyses by fluorescent amplified fragment length polymorphism (FAFLP), enterobacterial repetitive intergenic consensus (ERIC) based genotyping and candidate orthologues sequencing methods to understand the role of evolutionary dynamics in facilitating species divergence among pathogenic and saprophytic mycobacteria. Genome sequence and extensive characterization using several molecular tools and markers revealed that MIP has been the predecessor of *Mycobacterium avium intracellulare complex* (MAIC) *bacilli* and shared a common aquatic phase with early pathogenic forms of *Mycobacteria* thus, presenting a holistic picture of *Mycobacterium* evolution (Ahmed *et al.,* 2007). Further, extension of this work was carried out to understand the evolutionary and genomic mechanisms accountable for pathogenicity of MIP - a soil derived *Mycobacterium*. Mosaic architecture of *MIP* genome was reported for the first time with approximately 50.5% of genes to be laterally acquired. Comparative genomics suggested high antigenic potential of *MIP* as well as genome fluidity in habitat diversification and evolutionary divergence (Saini *et al.,* 2012).

Furthermore, many clinical isolates of pathogenic bacteria from diverse genera *Acinetobacter, Burkholderia, Corynebacterium, Helicobacter, Klebsiella, Neisseria, Pseudomonas,* and *Staphylococcus* have been sequenced and analysed in the last decade (Supplementary Table 1). For instance, *Acinetobacter baumannii* is clinically most important species of the genus *Acinetobacter* known for causing nosocomial infections. The genomic data of different strains of *A. baumannii* provided an insight into the genetic repertoire responsible for antibiotic resistance and pathogenicity. Variation in antibiotics resistance coding genes were identified in two strains of *A.baumannii* isolated from patients with different infections- a patient with bloodstream infection and and another with ventilator associated pneumonia (Balaji *et al.,* 2015). In another report (Vijaykumar *et al.,* 2015a, Vijaykumar *et al.,* 2015b), *A. baumannii* strain also isolated from bloodstream infection was identified to encode only $bla_{OXA\text{-}65}$ gene and no other antibiotic resistance gene. Genome analysis revealed drug resistance in a human pathogenic strain of *Brevundimonas diminutadiminuta* (Ghosh *et al.,* 2015). Through this study, it was identified that the drug resistance in the said strain of *B. diminuta* was due to change in amino acid residues at two positions of QRDR region in GyrA subunit. *Campylobacter fetus* is another pathogen reported for causing several diseases in humans and animals. The genome analysis of a *C. fetus* strain isolated from kidney disease patient with sepsis identified genes coding for virulence, antibiotic resistance and toxic compounds (Rohit *et al.,* 2015). Similarly, genome analysis of pathogenic strains of *Pseudomonas aeruginosa* (Malathi *et al.,* 2013) and

Staphylococcus aureus (Soni *et al.,* 2015) helped in identification of genes confering multi-drug resistance. Further, the first report on genome sequence analysis of diphtheria causing human pathogen, *Corynebacterium diphtheria* from India revealed the variation in the number of antibiotic resistance genes and virulence factors from different strains (Veeraraghavan *et al.,* 2016). Furthermore, the first report on the genome sequence of *Klebsiella pneumoniae* strains from India was reported in 2017 (Mathur *et al.,* 2017) unravelling the presence of carbapenem resistance gene cluster.

Although genome sequences of many pathogens reported from India have provied useful information, a detailed comparative genomic and phylogenetic analysis of these pathogenic strains is likely to provide valuable information about evolution and acquisition of pathogenecity and multi-drug resistance thus providing impetus to identify suitable drug targets as well.

Industrially Important Microbes (Actinobacteria)

Actinobacteria represent a very important group of bacteria which contribute approximately 70% of drugs in use today and form basis for the pharmaceutical industry. The secondary metabolites produced by this group of bacteria have been used as antibiotics, anticancer drugs, anti-depressants and immune-suppressants. Modern biotechnology techniques based on "omics approach" *viz.* genomics, metagenomics, proteomics and transcriptomics have further aided in exploring the metabolic pathways for synthesis of bioactive compounds from actinobacteria. In this area, a significant contribution has been made by Lal and team in deciphering the complete genomes of species belonging to genus *Amycolatopsis, namely, A. mediterranei* S699 (Verma *et al.,* 2011) and high-quality draft genomes of *A. mediterranei* DSM46095 (Saxena *et al.,* 2014), *A. mediterrani* DSM40773 (Mukherjee *et al.,* 2014) and *A. mediterranei* DSM46096 (Singh *et al.,* 2014). Strain *A. mediterranei* S699 produces an antibiotic rifamycin B, which is the mainstay of the first line anti-tuberculosis (anti-TB) drug regimen-Directly Observed Therapy Short-course (DOTS). The other three strains also produce rifamycin related compounds with significant antibacterial activity. The ongoing comparative genomic analysis of these strains including other sequenced genomes of the members of the genus *Amycolatopsis* will add to the knowledge about bioactive synthesis of pharmacologically important secendory metabolites. Another group from the Institute of Microbial Technology (CSIR-IMTech) have reported genomic studies of *Amycolatopsis* with draft genomes of *Amycolatopsis decaplanina* DSM 44594^T (Kaur *et al.,* 2013)*, A. azurea* $DSM43854^T$ (Khatri *et al.,* 2014) *and A. vancoresmycina* DSM 44592^T (Kaur *et al.,* 2014). Researchers have also reported the genomes of species belonging to genus *Streptomyces* (Jose *et al.,* 2013; Rajeswari *et al.,* 2015). Research group at University of North Bengal performed comparative genomic analysis of the actinobacterium *Prauserella sp.* Am3, isolated from root nodules of the actinorhizal plant *Alnusnepalensis* (Bose *et al.,* 2016). The study revealed the genomic similarities and dissimilarities of the newly sequenced *Prauserella* sp. Am3 with the type strain, *Prauserellarugosa* DSM 43194^T, and its relationship with *Amycolatopsis*, a closest neighbour of the genus *Prauserella*

Agriculturally Important Microbes (PGPRs)

Progress in molecular biology methods have contributed significantly to a better understanding of agricultural biology. Use of microbes in agriculture has proven to be a promising alternative to chemical fertilizers and pesticides as the latter causes severe damage to the environment as well as ecosystem. These agriculturally important microbes are generally referred to as plant growth promoting rhizobacteria (PGPRs) (Glick, 1995). Strains from diverse genera namely *Azotobacter, Azospirillum, Arthrobacter, Acinetobacter, Agrobacterium, Bacillus, Burkholderia, Bradyrhizobium, Enterobacter, Flavobacterium, Frankia, Pseudomonas, Rhizobium, Serratia* are classified as PGPR, with *Pseudomonas, Rhizobium* and *Bacillus* spp. predominantly playing an important role in plant growth promotion (Beneduzi *et al.,* 2012). There are innumerable studies from past decade which are focussed on identifying the molecular mechanisms of plant growth enhancement by the rhizobacteria and their diversity using sequencing of only housekeeping genes. However, only few studies are available from India which focus on genomics, comparative genomic and functional genomics aspects. For instance, a salt

tolerant bacterium *Bacillus safensis* strain VK from a saline desert area of Gujarat was sequenced and genes for salt tolerance and strain's plant growth-promoting potential were identified (Kothari *et al.,* 2013). Similarly, a rhizobial strain *Rhizobium lupini* HPC(L) sequenced from saline desert soil showed the presence of genes involved in oligotrophy and heterotrophy mode of life (Agarwal and Purohit, 2013). Another study isolated plant growth beneficial rhizobacteria from saline tolerant pokkali rice and has functionally evaluated their abilities to promote plant growth under saline conditions (Krishnan *et al.,* 2016). This strain consisted of the genes for 1-aminocyclopropane-1-carboxylic acid (ACC), production of indole acetic acid (IAA) and siderophore that help in plant growth. A study led by Gupta and co-workers (2014) focussed on genome sequencing of three rhizobacteria (CPCRI-1, CPCRI-2, CPCRI-3) from coconut, cocoa and arecanut and functional annotation predicted presence of genes for siderophores, acetoin, chitinases, phenazines, catalases and several other in all the three isolates (Gupta *et al.,* 2014).

In addition to the draft genomes, a population genomics study by Kumar *et al*. (2015) was done with an aim to characterize population structure of local *Rhizobium leguminosarum* associated with two different hosts by using core gene phylogeny. In this study, genome sequencing was carried out for 72 isolates. By using this information, the isolates were categorized into five genospecies on the basis of average nucleotide identity (ANI), core genes and mobile genes (Kumar *et al.,* 2015).

In another study, a complete genome of a PGPR strain isolated from Assam, *P. fluorescens* Pt14 was sequenced and core genome wide analysis revealed that sulphur metabolism is multifold in strain Pt14 in comparison to its closest phylogenetic neighbour *P. fluorescens* A506 (Rani *et al.,* 2017). Higher folds of sulphur metabolism in Pt14 can be attributed to its ability to survive in acidic environments (pH 4.65) as it is an isolate from an anaerobic rice soil where sulphur is present in reduced forms like sulphides and may thus be an adaptive feature.

Draft genomes of several strains belonging to different species and variable host plants have been sequenced and the information from genomic repertoire is used for sustainable agriculture in India (Supplementary Table 1). However, despite the recent progress in our understanding of colonization ability, mechanism of action and diversity of these rhizobacteria, we need extensive genomic studies for these agriculturally important microbes to reveal the presence or absence of important genetic traits involved in overall growth and pathogen suppression for their efficient use in ecological bio-agriculture.

Bacteria from Different Habitats

Contaminated Soils

Bacterial diversity studies generally provide the information about the microbial potential. Studies from a stressed environment like contaminated soils have been extensively performed to uncover the microbial enzymatic functions and adaptations in context of the extreme conditions. For almost a decade, Lal and co-workers have been analysing bacterial diversity present in Hexachlorocyclohexane dumpsite (HCH: 450 mg/g of soil) (Sangwan *et al.,* 2014) located at Ummari village, Lucknow, India. Synthesis and purification of one ton of insecticidal lindane (γ-HCH) from HCH mixture (α-,β-,γ-,δ-isomers) yields 8-12 tons of HCH muck. The unusual production of the insecticidal γ-isomer of HCH also called lindane and unregulated disposal of HCH muck has created various dumpsites all over the world including a dumpsite at Ummari village. High dose of HCH along with dumping of salts make the environment stressful and toxic. Various bacterial strains have been isolated, identified, characterized and sequenced from this HCH dumpsite (Supplementary Table 1).

Further, to understand the genome organization of HCH tolerating bacterial strains and the mode of acquisition of *lin* genes responsible for the degradation of HCH isomers, various bacterial species were sequenced and analysed. The analyses revealed the complete absence of *lin* genes from the genome of *Sphingobium lactosutens* DS20^T (Kumar *et al.,* 2013) and the presence of a single copy of *linA* (HCH dehydrochlorinase) in *Novospingobium lindaniclasticum* LE124T (Saxena *et al.,* 2013). The genome of *S. baderi* LL03^T (Kaur J *et al.,* 2013) was also found to be deficient in *linB* (Haloalkane dehalogenase). The absence of the key enzymes involved in the degradation of HCH revealed that these

bacteria have mechanisms to tolerate high levels of HCH but cannot degrade. The remaining bacterial species sequenced, possess the entire complement of *lin* genes which aid in HCH degradation (Supplementary Table 1).

In addition, there are different chemical enrichments in the industrial effluents which eventually contaminate the soil and water. A study for the treatment of dibutyl phosphate contaminated water was performed using aerobic microbial granules (Reddy GKK *et al.,* 2014). Dibutyl phosphite is an organophosphorous compound which has applications in different chemical industries and processes. Study showed that using aerobic granules in a sequencing batch reactor (SBR) with a 24 h cycle feeding with dibutyl phosphite as a co-substrate along with acetate resulted in complete biodegradation of 1.4, 2 and 3 mM of dibutyl phosphite in 4, 5 and 8 h, respectively, accompanied by stoichiometric release of phosphite (H3 PO3). Molecular analysis using t-RFLP showed the presence of 12 different bacterial types, of which two bacterial strains capable of growth on dibutyl phosphite as sole carbon source were isolated and characterized as *Acidovorax* sp. and *Sphingobium* sp. (Reddy GKK *et al.,* 2014).

Another study of a treatment plant receiving waste water from multiple drug manufacturers showed the presence of highly multi-drug resistant (MDR) integron bearing bacteria living under extreme antibiotic pressure (Marathe *et al.,* 2013). In one such study it has been reported that antibiotic containing water coming from multiple pharmaceutical companies near Hyderabad became the hotspot of bacterial antibiotic resistance in an Indian river (Lubick *et al.,* 2011).

Hot Springs

Hot springs represent the most unique naturally stressed aquatic niche formed by emergence of geothermally heated groundwater from Earth's crust. Hot springs usually harbour moderate to extreme thermophiles and hold interest as hot environment represent the conditions for origin of life. Additionally, these springs harbour rich bacterial diversity that are source of commercially important products specially enzymes, sugars, compatible solutes and antibiotics (Satyanarayana *et al.,* 2005). Thus, hot-springs can be regarded as an enormous treasure of yet unexplored information which can be mined for beneficial purposes.

India is home to around 340 thermal hot water springs (Ghelani *et al.,* 2015; Mangrola *et al.,* 2015a) as described by the Geological Survey of India (GSI) in 1991 in the "Geothermal Atlas of India". Among these, many are being studied in the hope of exploring their immense potential. Diversity analysis is now being supplemented by genomic and metagenomic studies. Additionally, the GSI has described more than 300 sites with geothermal potential with immense use in power generation and are estimated to have a potential of 10,600 MWe (Craig *et al.,* 2013).

Studies on Himalayan hot spring located at Manikaran, Himachal Pradesh, India revealed rich pool of functionally adapted microbial diversity from Microbial mats (Tripathi *et al.,* 2016, Sharma *et al.,* 2014) and sediment samples (Mahato *et al.,* 2014). Interestingly, a *Thermus* sp. RL was isolated from hot spring water (90°C-98°C) as bacteria belonging to this genus produce thermostable enzymes. The draft genome was sequenced in 2012 to understand bacterial life and adaptations (Dwivedi *et al*., 2012). However, the complete genome for strain RL was sequenced and compared with other *Thermus* spp. to decipher genetic variability, evolution and survival strategies at higher temperature (Tripathi *et al*., 2017). Further, an aerobic, non-motile and coccoid strain *Lampropedia cohaerens* $CT6^T$ was isolated from microbial mats and the whole genome sequencing of this strain using Illumina HiSeq 2000 technology has revealed genes involved in Entner-Duodoroff (ED) pathway and non-phosphorylated ED pathway along with arsenic, copper, cobalt, zinc, cadmium and magnesium tolerance genes. Also, it was found that diverse genetic potential for survival of this strain at arsenic rich hot spring was due to the presence of genes associated with biofilm formation, pyrroloquinoline-quinone production, isoquinoline degradation and mineral phosphate solubilization (Tripathi *et al.,* 2017). Similarly, a bacterial strain, *Cellulosimicrobium* sp. MM was isolated from microbial mats and annotated draft genome (3.85 Mb) having 3,718 CDS and 273 subsystems was announced (Sharma *et al.,* 2014). Pathogenicity islands (PAIs) having 49 potential marker genes with known association to human infections were identified and using synteny-based annotation, it was established

that gene transfer from non-pathogenic bacteria is a key factor in the evolution of PAIs (Sharma *et al.,* 2016). A moderately thermophilic bacterium *Deinococcus* sp. strain RL was isolated from sediment of Manikaran hot springs and its draft genome sequence showed presence of genes involved in base excision repair, nucleotide excision repair, mismatch repair and homologous recombination (Mahato *et al.,* 2014). In another study, four thermophilic strains, *Brevibacillus thermoruber* PS1, *Brevibacillus thermoruber* PS2, *Paenibacillus* sp., PS3 and *Bacillus licheniformis* PS4 were isolated using culture-based techniques followed by their morphological and biochemical characterization (Verma *et al.,* 2014). Using amplified ribosomal DNA restriction analysis (16S rRNA-ARDRA) technique, a total of 85 thermophilic isolates were screened and 42 phylogenetic clusters of Firmicutes, Actinobacteria and Proteobacteria were identified in the samples collected from Manikaran hot spring. Out of total 42 representatives, the study reported functionally diverse populations with 26 % of the isolates being amylase producers and 45 % being protease producers (Kumar *et al.,* 2014).

Hot spring (temperature, 43.5 °C) located in Pachmarhi, central India was also explored for microbial diversity and genomics studies. Draft genome of a proteobacterium *Gulbenkiania mobilis* strain MB1, a sulfur-metabolizing thermophile isolated from this hot spring (temperature, 43.5 °C); revealed genome size of 3.31 Mb with average %G+C content of 62 and 2905 protein coding genes. Genes for sulfur metabolism required for the assimilation of sulfate to sulfide, the *sox* gene for sulfur-oxidizing proteins and *soxD* encoding cytochrome c were also present indicating that MB1 can utilize sulfur in addition to oxygen as the terminal electron acceptor in thermophilic environment (Saxena *et al.,* 2015). Another draft genome (28,49,160bp) of a chemolithotrophic thermophile, *Tepidimonasta iwansis* strain MB2, isolated from a hot spring in central India, was reported. It was found to have 2,503 protein coding genes and showed properties of sulfur metabolism, nitrogen fixation, ammonia metabolism, assimilation of organic acids and a wide variety of proteases (Dhakan *et al.,* 2016). Culture-dependent survey of microbial diversity present in three alkaline and mesophilic hot springs of Odisha, India revealed genetic and functional variability among the isolates. A total of 48 isolates belonging to family *Bacillaceae, Paenibacillaceae, Planococcaceae, Pseudomonadaceae* and *Enterobacteriaceae* were analyzed. Functional diversity discovered that isolates produced a mixture of extracellular enzymes such as amylase, cellulase, lipase, phosphatase and protease whereas genus *Bacillus* dominated for extracellular enzymatic activity (Sen and Maiti, 2014).

In a study on Taptapani hot springs, Odisha, metagenomics was employed to gain insights into microbial communities and their enzyme activities (Sen *et al.,* 2015). Different dominating bacterial taxa were marked at different seasons of the year. The results reported that temperature and pH were two major environmental factors that were found to modulate prokaryotic community structures in hot springs (Sen *et al.,* 2015).

Halophiles

Another stressed environment is high salinity water from where bacterial isolates have been studied thoroughly using genome sequencing tools. In India, many studies have described the screening and characterization of the extracellular hydrolytic enzymes produced by halophilic microorganisms (Saju *et al*., 2011; Kumar *et al.,* 2012). Five such halophilic bacterial strains *Halobacillustrueperi* SS1, *Halobacillustrueperi* SS3, *Shewanella algae* SS2, *Halomonasvenusta* SS5 and *Marinomonas* sp. SS8 were isolated from the soil sediment of Lunsu, a salt water body of Himachal Pradesh, India (Gupta *et al*., 2016). The isolates SS1 and SS3 exhibited halophilic amylase activity; SS1, SS2, and SS3 exhibited protease activity; halotolerant lipase activity was exhibited by SS2 and glutaminase activity by all except SS1. Kovalam solar salt works in Kanyakumari of India have also been screened for diversity of halophiles and a biosurfactant producing *Kocuria marina* BS-15 was isolated which highlights the increased stability of bacteria at such extreme environments (Sarafin *et al*., 2014). Another extremely halophilic bacterium *Salinicoccus* sp. JAS4 that grows in the presence of high NaCl concentrations of upto 25% (w/v) was isolated from arable soil of west coast of Karnataka, India (Jayachandra *et al*., 2012). Further, two bacterial strains were isolated and sequenced from Lonar lake situated at Maharashtra, India: *Indibacteral kaliphilus* LW1^T (Singh *et al.,*

2013) and *Methylophaga lonarensis* MPL^T (Shetty *et al.,* 2013a). Draft genome sequence of LW1^T revealed presence of 48 genes involved in resistance against antibiotics and toxic compounds, 12 genes for intracellular resistance, and 81 genes for different stress responses, including 12 heat-shock, 3 cold-shock, and 12 periplasmic stress-response genes. Draft genome of strain MPL^T emphasized methanol as its sole carbon and energy source (Shetty *et al.,* 2013a). Similarly, genome of *Bacillus okhensis* strain Kh10-101T from a salt pan near port of Okha, India was sequenced and used as a potential model to study the molecular response of bacteria to salt as well as alkaline stress (Krishna *et al.,* 2015).

Psychrophiles and Aerophiles

Microorganisms growing well in cold temperature (15 to -20 C) are known as Psychrophiles often termed as cold loving organisms. A substantial amount of work has been done by group led by Shivaji *et al*., for bacterial diversity and determining genome sequences of pychrophiles from antarctica (Pinnaka *et al*., 2013a; Shivaji *et al*., 2013a) and arctic soil (Pinnaka *et al.,* 2013b; Shivaji *et al*., 2013b). One such psychrophile is a gram positive bacterium, *Cryobacterium roopkundensis* RuGl7 isolated from soil sample in the vicinity of Roopkund glacial lake, Himalayas, India. Draft genome analysis of RuGl7 showed presence of 117 stress response genes out of which 3belong to the *csp*A family of genes involved in cold stress (Reddy GS *et al.,* 2014). Also, efforts have been made to isolate and sequence bacteria from upper atmosphere with an objective of determining the microbial diversity and the underlying adaptations as an aerophile. For example, *Bacillus isronensis* strain B3W22 was isolated from air collected at an altitude ranging from 27 to 30 km above the city of Hyderabad, India (Shivaji *et al*., 2012). Genome Sequence of *Janibacter hoylei* MTCC8307, isolated from stratospheric air at an altitude of 41.4 km over Hyderabad, India was announced and genome analysis revealed the presence of putative genes responsible for glycolysis/gluconeogenesis, the tricarboxylic acid cycle, the pentose phosphate pathway, ABC transporters, DNA repair, osmotic stress, oxidative stress, cold shock, heat shock, and resistance to toxic compounds (Pawar *et al.,* 2012).

Linking Genomics and Metagenomics

The advancements in sequencing technologies over the past two decades have not only revolutionized the field of genomics but also opened the way to analyse the genetic material directly from the environment samples using culture-independent methodologies, popularly called as metagenomics. The major drawback in using culture dependent approaches is that these techniques cannot be used to decipher the complete diversity at a niche as 99% of the microorganisms cannot be cultured using established culture techniques. Metagenomic approaches on the other hand provide insights into the study of microbial communities in an environment. Thus, with application of metagenomics tools and techniques, the hidden (uncultivable) diversity can be explored with ease.With the newly available and cost-effective sequencing tools, the major bottleneck of sequencing of metagenome has been resolved. However, unprecedented challenges in metagenome sequence analysis and its interpretation continues. The workflow for metagenomics projects run parallel to genomics pipeline *i.e.* sequencing a metagenome with profound coverage, assembling of raw data into continuous stretches of contigs/scaffolds and finally annotation of these sequences to investigate the structure and function of microbial communities inhabiting a particular niche.

Initially, the field of metagenomics started exploring the microbial community from different environments by using gene-centric approaches. Among all, 16S rRNA gene sequence has been predominantly used to identify the bacterial diversity of various ecological niches (Singh *et al.,* 2009). With the ease and availability of sequencing techniques and analyses tools as discussed previously, the study of microbial diversity using culture-based approaches has been superseded by study of microbial genomics and metagenomics at these niches using Next Generation Sequencing (NGS) methods. Metagenomics is an important tool to comprehensively analyse and decipher the microbial diversity and functions in an environment without resorting to culturing. Thus, moving from genomics of what is visible to eyes to analysing metagenomes, represents a major leap in technological innovations and knowledge advancement.

Metagenomics of Environmental Niches

Metagenomics has played a vital role in understanding the ecology and function of various ecosystems. For the past decade, with the availability of efficient and economic NGS platforms and improved analyses tools; many culture independent studies have been accomplished in India from different geographical locations and stressed sites which led to interesting findings on bacterial community dynamics. For instance, parallel studies were done to catalogue the unculturable diversity from Sunderbans mangrove (Basak *et al.,* 2015); Little Rann of Kutch, Gujarat (Patel *et al.,* 2015); Paradip Port, Odisha (Pramanik *et al.,* 2016); Farpuk caves, Mizoram (DeMandal *et al.,* 2015b); Khuangcherapuk caves, Mizoram (DeMandal *et al.,* 2015a) and Bat guano, Pnahkyndeng caves, Meghalaya (DeMandal *et al.,* 2015c). By screening metagenomic library for biocatalytic properties, novel proteins can be identified, for example a novel α-amylase was isolated and characterized from fosmid vector library from Western Ghats of Kerela (Vidya *et al.,* 2011). Another interesting finding from Banduhurang open cast Uranium mine and Jaduguda Uranium mine suggests that soils surrounding the ore deposits are characterized by distinct geochemical and microbiological features which can serve to scale the impact of U-mining and the native bacterial populations sustaining in these hazardous environments can be used as potential candidates for *in-situ* bioremediation (Islam *et al.,* 2010; Dhal *et al.,* 2014). Metagenomics also plays a significant role in identification of plethora of potential pathogens involved in disease and antibiotic resistance blowout. The first report using shotgun metagenomics revealed presence of pathogens like *Staphylococcus aureus, Corynebacterium glutamicum, Enterococcus faecalis* on paper currency notes which spread these infectious agents all over the country (Jalali *et al.,* 2015). Another study comes from the metagenomic sequence of Indian one-rupee coin as a source of spread of pathogenic bacteria like *Corynebacterium accolens*, *Propionibacterium granulosum*, *Staphylococcus aureus, Finegoldia magna* and *Listeria monocytogenes* as currency coins are extensively traded all around the country (Devi *et al.,* 2017). In the subsequent sections of this review, bioprospecting of micro-organisms that thrive in stressed environments and their functional attributes are discussed in detail.

Land Sources

Metagenomic studies to understand the unexplored diversity and functions of microbial communities in different types of soils are discussed in details in this section. A recent study investigated the genes involved in osmoadaptation to environmental soil by halophilic bacteria (Ahmed *et al.,* 2018). BCAA_ABCtp (branched chain amino acid ABC transporter gene), GSDH (glucose/sorbosone dehydrogenase protein), STKc_PknB (catalytic domain of bacterial Serine/ Threonine kinases), and duf3445 were identified as important genes for osmotolerance using salt stress resistant clones. The knowledge about the genetic repertoire of microbial communities thriving in salinity can be exploited in developing crop varieties capable of growing under saline conditions. Metagenomics approach had also been used to investigate the pathways of methanogenesis and methanotrophy in a 13-year old manure fed soil in lowland fields of rice paddy (Bhattacharyya *et al.,* 2017). Comparisons were made between control (absence of any manure), farm yard manure and green manure aided with *Sesbania aculeata*. The predominance of genera *Methanolobus* and *Methanotorris* in only control sample fields indicate the possibility of methanogenesis by these bacteria. However, the functional potential of the manure rich communities was higher than control soil in terms of C pools and methane production (Bhattacharyya *et al.,* 2017). Similar study was done by Bhattacharyya *et al.* (2016) on lowland rhizosphere of rice to understand the shifts in microbial dynamics under ambient CO_2 and elevated CO_2 and temperature conditions. As a consequence, methane producing genera and methanogenesis pathways were abundant in elevated carbon dioxide conditions.

The microbial communities in a given environment can exhibit the three Rs, resistance, resilience and redundancy. The study by Patel *et al.* (2016) illustrated the mechanism of response and resilience of microbial communities inhabiting the oil fields near industrial regions of Ahmedabad. The metabolic pathways related to fatty acid biosynthesis were predominant in such bacterial communities.

Enzymes like cellulases, proteases, amylases and lipases have industrial importance and isolation of such enzymes with ability to tolerate harsh

conditions of high salt, heat and acid attacks have been in great demand. Garg *et al.* (2016) isolated novel cellulase Cel5R from soil metagenome. Cel5R exhibited halo tolerance and thermal stability, and hence is a potential candidate for industrial application. The soil is a reservoir of novel compounds and rhodanese gene has been cloned from the soil metagenome of cold desert in North-West Himalayan regions (Bhat *et al.,* 2015).

The metagenomics approach indicated the dominance of Proteobacteria followed by Firmicutes, Chlorofexi, Bacteroidetes, Acidobacteria, Nitrospirae and Actinobacteria in Sundarbans mangrove (Basak *et al.,* 2015). There have been reports of application of metagenomics to unravel the microbial communities in the desert areas of Little Rann of Kutch, Gujarat (Patel *et al.,* 2015; Pandit *et al.,* 2015).

In another example of pesticide contaminated soil, Lal and co-workers investigated the microbial dynamics by using metagenomic approaches at the HCH dumpsite (Sangwan *et al.,* 2012). Comparative metagenomic survey of heavily contaminated HCH dumpsite (450 mg HCH g^{-1}), 1 km away (0.7 mg HCH g^{-1}) and 5 km away (0.04 mg HCH g^{-1}) soil samples revealed interesting results in terms of microbial diversity as well as community potential of HCH degradation. Bacterial diversity at the three sites revealed notable differences at the genera level, with *Pseudomonas*, *Sphingomonas*, *Novosphingobium*, *Sphingopyxis*, *Marinobacter*, *Chromohalobacter* among the most abundant at HCH dumpsite. This led to the conclusion that besides sphingomonads (avid HCH degraders), other indigenous bacterial species of the community should also be bio-stimulated for the development of an effective bioremediation technology. Furthermore, metagenome data at the HCH dumpsite was used along with genome sequence data of two HCH degrading *Sphingobium* species (*Sphingobium japonicum* UT26 and *Sphingobium indicum* B90A) to reconstruct last common ancestor (LCA) genotype (Sangwan *et al.,* 2014). Comparison of LCA genotype with these two subspecies in terms of genes repertoire revealed absence of two genes-*linA* and *linB* encoding for enzymes involved in the upper pathway of HCH degradation, which indicated that the descendants acquired these genes under HCH stress via transposon mediated lateral transfers.

Water Sources

Since long, water quality is often estimated by measuring the total coliform bacterial counts that serve as indicators of pollution. However, the diversity of microorganisms residing in water sources provide much deeper environmental cues corresponding to the types of contamination. Employing metagenomics approach to study the unculturable bacterial diversity and understanding their metabolic interactions holds much potential especially in the field of bio-remediation. Various Indian research groups have widely employed metagenomic sequencing technique and analyses tools to understand the microbiology of water resources and analyse various microbial genes, enzymes, functions and element cycling. From pristine water sources (Rajeev *et al.,* 2018) to sludge (Jadeja *et al.,* 2014), hot thermal springs (Ghelani *et al.,* 2015; Sangwan *et al*., 2015; Sharma *et al*., 2016) to contaminated waters (Yadav *et al.,* 2015), Indian researchers have performed metagenomic studies to identify the distribution and role of microbial communities with an aim to identify the unique functions of microorganisms at different habitats. The extreme environments such as hot springs, where water temperatures reach as high as 95 ºC, diversity analysis revealed novel genotypes which were reconstructed from the metagenome data (Sangwan *et al.,* 2015). Researchers also studied the functional dynamics that uncovered the necessary genes for survival at such extreme environments. The functional screening of metagenomic libraries constructed from activated sludge and ground water aquifer systems and further cloning of the selected sequences have resulted in the discovery of novel arsenic resistance genes (Chauhan *et al.,* 2009; Das *et al.,* 2017) and several other enzymes involved in environmental detoxification (More *et al.,* 2014; Sharma *et al.,* 2012; Jadeja *et al.,* 2014). It is for this reason that the metagenomic analysis of various water sources continues to interest researchers.

Hot-springs

In order to unlock the potential of the hot-springs, it is essential to decipher the microbial diversity and interactions existing at these sites. Microbial genomics and metagenomics of hot spring have been thoroughly investigated by different research groups across the country. The 16S rRNA gene amplicon metagenomic

sequencing to study the bacterial and archaeal diversity of a Himalayan hotspring located at Manikaran, revealed predominance of Firmicutes, Aquificae and Deinococcus-Thermus in this stressed niche. Crenanrchaeota being hyperthermophilic and metabolically versatile was the main archaeal phylum as identified in the study (Bhatia *et al.,* 2015). Using cultivation-independent comprehensive survey of bacterial diversity in Tulsi Shyam Hot Springs, India; Ghelani *et al*. (2015) reported the abundance, diversity, distribution and coexisting organisms in the hot spring. A total of 16 bacterial phyla demonstrating 97 families and 287 species were revealed in the hot spring metagenome. Most abundant phyla were Firmicutes (65.38%), Proteobacteria (21.21%) and unclassified bacteria (10.69%) whereas; *Peptostrepto coccaceae* (37.33%), *Clostridiaceae* (23.36%), and *Enterobacteriaceae* (16.37%) were highest reported families in the metagenome. Ubiquitous species were *Clostridium bifermentans* (17.47%), *Clostridium lituseburense* (13.93%) and uncultured bacterium (10.15%). Microbial biodiversity composition of the hot-spring sediment of Deulajhari hot-spring cluster (temperature 69 °C) located in the Angul district of Odisha, was described by employing Illumina sequencing based on amplicon metagenome sequencing of 16s rRNA targeting V3-V4 region. Over 28 phyla were detected of which Proteobacteria were predominantfollowed by Bacteriodetes, Firmicutes, Spirochetes and Chloroflexi (Singh and Subudhi, 2015). Two other alkaline Indian hot springs, Jakrem (Water temperature 46 ^{o}C, pH 9, Meghalaya) and Yumthang (Sikkim, water temperature 39 °C, pH 8) were studied through metagenomics approach by sequencing the amplified V4 region of the 16S rRNA gene from cDNA followed by their classification in different operational taxonomic units (OTUs) (Panda *et al.,* 2016). A total of 19 distinct phyla dominated by Proteobacteria, Bacteroidetes and Thermus were reported in the Yumthang 16S rRNA library whereas Jakrem 16S rRNA library was dominated by Firmicutes, Chloroflexi and Thermus (Panda *et al.,* 2016). Microbial diversity of Jakrem hot spring located in West Khasi hills, Meghalaya, India was analyzed by V3 hyper variable region of 16S rRNA gene sequencing. Sequence reads were clustered into 694 OTUs comprising of 14 bacterial phyla dominated by Firmicutes, Chloroflexi and Cyanobacteria (Panda *et al.,* 2015). Likewise, diversity at Bakreshwar hot-spring was studied and *Shewanella*-related thermophiles were cultured (Ghosh D *et al.,* 2003).

The use of metagenomics has not remained limited to only the microbial diversity analysis of the hot-springs and much recent studies have attempted to gain insights into the functional roles of these micro-organisms. The three major hot springs Badi Anhoni, Chhoti Anhoni, and Tattapanilocated at two geographically distinct regions (Anhoni and Tattapani) in central India were studied recently through metagenomic approach using both the amplicon and shotgun sequencing to uncover the resident microbial community as well as the functional dynamics at these thermal springs (Saxena *et al*., 2016). The samples of hot water with temperatures ranging from 43.5 to 98°C were assessed in the study. The Anhoni hot springs (43.5-55 °C), were dominated by *Pseudomonas stutzeri* and *Acidovorax* sp. with the genes involved in hydrocarbon degradation pathways, such as benzoate, xylene, toluene, and benzene being abundant. This suggested the presence of chemoorganotrophic thermophilic community with the ability to utilize complex hydrocarbons as a source of energy. At the Chhoti Anhoni hot-spring, methane metabolism pathway genes were abundant, with high abundance of *Methylococcus capsulatus* (Saxena *et al.,* 2016). This explains for the methane gas which is reported to constitute >80% of all the emitted gases at this hot spring. Due to high-temperature range (61.5-98 °C) of the Tattapani hot spring, a lower microbial diversity was dominated by a nitrate-reducing archaeal species *Pyrobaculum aerophilum*. A higher abundance of cell metabolism pathways was observed at Tattapani which is necessary for the survival of microorganisms in extreme conditions (Saxena *et al.,* 2016).

The four tropical hot-springs located at Atri, Athamallik, Tapttapani, and Tarabalo in Odisha were also studied through shotgun approach (Badhai *et al.,* 2015). The most abundant phyla reported at these sites were Chloroflexi and Proteobacteria. Other phyla such as Acetothermia, Nitrospirae, Acidobacteria, Firmicutes, Deinococcus-Thermus, Bacteroidetes, Thermotogae, Euryarchaeota, Verrucomicrobia, Ignavibacteriae, Cyanobacteria, Actinobacteria, Planctomycetes, Spirochaetes, Armatimonadetes, Crenarchaeota, and Aquificae were also reported (Badhai *et al.,* 2015). Further, the study showed that

environmental variables such as temperature, dissolved ions and solids, total hardness and conductivity shaped the overall microbial community composition. The functional profiles at the different springs had only little variations and the genes involved in the metabolism of carbohydrates and carbon fixation were the most abundant (Badhai *et al.,* 2015). Metagenomic approach further provided new insight into physiology of yet-unknown members of phylum Acetothermia (Badhai *et al.,* 2015). Shotgun metagenomics approach-based diversity analysis of microbial mat from a Himalayan hotspring located at Manikaran, revealed the reconstruction of two novel genomes of potential predator (*Bdellovibrio bacteriovorus*) and prey (*Enterobacter cloacae*) and established predator prey relationship existing in this environment. Genus *Bdellovibrio* has always raised interest of researchers as it feeds on other Gram-negative bacteria (especially pathogenic), thus validating great potential as a *'live'* antibiotic. These data were used to construct a theoretical model describing potential predator avoidance strategies, whereby the *E. cloacae* strains can move between anaerobic and aerobic niches by quorum sensing population size, which is modulated by a 'kill the winner' viral mechanism and predation by the obligate aerobe, *B. bacteriovorus* (Sangwan *et al.,* 2015). This study was the first comprehensive report on assessing population level dynamics and genotype reconstruction of prokaryotes across microbial mats of any Indian hot springs (Sangwan *et al.,* 2015).Complete metagenomic analyses of bacterial diversity and functional insights from basaltic hot spring of Unkeshwar, Maharashtra, revealed 41 phyla and 719 different species. In taxonomic analysis, the dominant phyla were found as, Actinobacteria (56%), Verrucomicrobia (24%), Bacteriodes (13%), Deinococcus-Thermus (3%) and Firmicutes (2%). Additionally, functional annotation using pathway information showed dynamic potential of hot spring community in terms of metabolism and environmental information processing (Mehetre *et al.,* 2016). Community and functional metagenome analyses of Lasundra hot springs (42-52 C), Kheda District, Gujarat, has also been done by shotgun metagenomic sequencing of community DNA isolated from water. Community profile was dominated by Firmicutes and Proteobacteria. Subsystem based functional annotations depicted 14.0% was carbohydrates, 7.0% was protein metabolism and 3.0% genes belonged to stress responses *viz.,* oxidative stress, periplasmic stress, osmotic stress, heat shock, cold shock, acid stress and detoxification (Mangrola *et al.,* 2015a). The same group of researchers also studied the Tuwa hot-spring, Gujarat (Mangrola *et al.,* 2015b). 22 bacterial phyla including 90 families and 201 species were identified in the metagenome. Among these, the dominant phyla reported were Firmicutes (97.0%), Proteobacteria (1.3%) and Actinobacteria (0.4%). Functional analysis of the metagenome revealed about one fourth of the total sequences were poorly characterized. This suggested that the Tuwa hot-spring is a potential source for novel microbial species and their products (Mangrola *et al.,* 2015b).

Such large data generated through metagenomics and thereafter its analysis has revealed enormous microbial diversity, population composition, community dynamics and factors influencing the dynamics at hot thermal springs across India; which otherwise could not be comprehended by using culture dependent methods.

Industrial Effluents and Sludge

Treatment of waste water has always been a challenge. Studying the bacterial community dynamics may provide clues for designing water treatment strategies. The metagenomic study of microbial diversity in samples from a wastewater treatment plant receiving antibiotic-containing waste revealed that the Proteobacteria was the most dominant phylum (Marathe *et al.,* 2016). The diversity in the aeration tanks was found to be considerably lower when compared with the corresponding samples from regular municipal waste water treatment plants. *Alcaligenaceae* and *Pseudomonadaceae*, reported to be highly multidrug resistant, and were the dominant families suggesting a strong selection pressure from antibiotics on the community structure (Marathe *et al.,* 2016).

Sidhu *et al.* (2017) explored the microbial community and their associated functions in the pre-treated and post-treated sludge. It revealed that ε-proteobacteria (~45.80%) dominate the pre-treated water while post treated sludge has the dominance of β-(30.23%) and δ-(13.38%) classes of Proteobacteria. Virulence factors and pathogenic

genes were shown to be abundant in the pre-treated than the post-treated sludge, therefore, highlighting the need for wastewater treatment. Key genes responsible for the degradation of polycyclic aromatic hydrocarbons were also identified in the metagenomes (Sidhu *et al.,* 2017). A similar study used comparative metagenomics to reveal different degradative capacity of activated biomass treating hydrocarbon contaminated waste water (Yadav *et al.,* 2015). The study examined the effect of different levels of total dissolved solids (TDS) and seasonal variations on the diversity and degradative capacity of activated biomass. It revealed that in low TDS metagenome, metabolic pathways related to degradation of aromatics via the central and peripheral pathways were dominant; while in high TDS sample, pathways corresponding to central carbohydrate metabolism, nitrogen and organic acids were predominant (Yadav *et al.,* 2015).

Mining of microbial diversity using metagenomics approaches from the industrial effluent waste sites have led to identification of novel enzyme which can be used for different treatment processes. Devi *et al.* (2016) analysed metagenomic library of tannery activated sludge and, identified and characterized novel alkaline serine protease designated as Prt1A. Investigation of catalytic site of this novel enzyme revealed that protein is related to S8A family subtilisin. Further, the enzyme activity was reported to be stable in the presence of anionic detergent, oxidizing agent and various organic solvents and displayed high affinity and catalytic efficiency for casein under standard assay conditions and the enzyme was also compatible with commercial detergents and could act as an efficient enzyme in various industrial applications (Devi *et al.,* 2016). A similar study constructed a metagenomic library from effluent treatment plant and screened two arsenic-resistant clones that showed 8 and 18 fold higher resistance to sodium arsenate (Chauhan *et al.,* 2009). Molecular analysis revealed these clones to be putative arsenate reductases and arsenite efflux pumps. A novel arsenate resistance gene (*ars*N) was further identified in this study from one of the clones (Chauhan *et al.,* 2009). These studies reflect a tremendous genetic potential of the yet uncultured microorganisms and the potential of functional metagenomics in environmental detoxification. The metagenomic functional data mining approach has been successfully applied from lab-scale experiments to pilot scale treatments at the waste water treatment plants that displayed improved efficiency of the treatments procedures (More *et al.,* 2014). In this study, metagenomics of a common effluent treatment plant in South India was performed. The most abundant phyla reported were Proteobacteria (65.68%), followed by Bacteroidetes/Chlorobi group (5.94%), Deinococcus-Thermus (2.98%), Chloroflexi (1.67%), Actinobacteria (1.26%) and Firmicutes (1.10%) along with representatives from two new phyla, Synergistetes and Elusimicrobia. The community functional profiles revealed the dominance of central meta-cleavage pathway. The study further investigated the effect of induction of the activated biomass with central aromatic intermediates-catechol, phenol, salicylate and resorcinol where highest oxygen uptake was shown with the use of salicylate as an inducer (More *et al.,* 2014). The sequence data mining also revealed enriched salicylate degradation pathway. A pilot experiment was thus carried out at the treatment plant using salicylate as an inducer that resulted in reduction of chemical oxygen demand by more than 50% (More *et al.,* 2014). In another study, metagenomic library screening led to identification of two novel oxygenases which showed similarity to flavin monooxygenases from *Mesorhizobium loti* and *Sphingomonas wittichi* and displayed potential as biocatalysts when expressed in *Escherichia coli* (Singh *et al.,* 2010). Several unique enzymes that can be implemented for bioremediation of waste material were also identified from the study of industrial effluent treatment plant, treating waste water generated at a pharmaceutical industry (Sharma *et al.,* 2012). Jadeja *et al.* (2014), explored the oxygenase coding gene sequences from the metagenome of activated biomass and identified several oxygenases and cellulases. The study focused on the catabolic capacity of the activated sludge for degradation of naphthalene, anthracene, phenol, biphenyl and o-toluidine among different metagenomic datasets. It was found that despite different geographical locations and source, many genes coding for oxygenases were common between treatment plants (Jadeja *et al.,* 2014). These studies highlighted the need for mining specific targets from the dataset and explore the potential of microbes.

Rivers and Lakes

Metagenomics analysis of lake water provides understanding about the microbial dynamics of water ecosystem. Chilika Lake, the largest lagoon of India was studied and analysed for the bacterial communities with an adaptability of tolerance to saline stress in brackish water using 454-pyrosequencing platform (Pramanik *et al.,* 2015). The study reported that the soil sediment of lake harbored 16,212 species belonging to 45 different phyla, with Proteobacteria, Chloroflexi, Firmicutes, Acidobacteria, Actinobacteria, Bacteroidetes and Planctomycetes as dominating phyla. Two different saline soil sediments analysed revealed significant differences in bacterial community composition and diversity value suggesting dynamic ecosystem in Chilika Lake (Pramanik *et al.,* 2015). Similar attempts were made to decipher the overall prokaryotic diversity (bacterial as well as archaeal) from Lonar Soda Lake sediments which identified genus *Caxiella*, *Fibrobacter* and *Candidatus* to be predominant (Dudhagara *et al.,* 2015). This approach was further extended to study health condition of Mahananda river at Siliguri by linking the abundance of specific species of bacteria associated with pollution in environmental waters (Mukherjee *et al.,* 2013). The microbial diversity of the Periyar river, the longest perennial river in the Western Ghats studied through amplicon metagenome sequencing revealed Proteobacteria (33.12%), Actinobacteria (14.58%), Acidobacteria (12.81%), and Bacteroidetes (9.89%) as the dominant phyla (Rajeev *et al.,* 2018).

In order to unravel the functional potential of the microbial communities at freshwater habitats, whole genome shotgun sequencing based metagenomic investigation of the Loktak lake was employed (Puranik *et al.,* 2016). It is the largest freshwater lake in Northeast India known for its floating islands or phumdis. Proteobacteria (51%) was found to be the most dominant bacterial phylum followed by Acidobacteria (10%), Actinobacteria (9%) and Bacteroidetes (7%). The study also compared the Loktak metagenome data with the other aquatic metagenomes from pristine to highly polluted environments. It reported selective domination of bacterial genera and prevalence of corresponding functions at the Loktak (Puranik *et al.,* 2016). In another study that employed bioprospecting through metagenomics of brackish water habitats to discover the novel enzymes and bioactivities, Pangong Lake located at Ladakh was studied (Rathour *et al.,* 2017). The study revealed abundance of bacteria in the phyla: Proteobacteria (54.36%), Bacteroidetes (24.01%), Firmicutes (1.14%), Actinobacteria (0.85%), Balneolaeota (0.79%), Cyanobacteria (0.59%), Verrucomicrobia (0.47%), Euryarchaeota (0.21%), Planctomycetes (0.19%) and Ascomycota (0.10%) with Methylophaga (10.19%) to be the most abundant genus. The study further revealed carbohydrate metabolism, energy metabolism, lipid metabolism and nucleotide metabolism to be the enriched functions at the habitat (Rathour *et al.,* 2017). Recently, the arsenic contaminated groundwater of the Ganges Brahmaputra Delta aquifer system was also studied using metagenomics approach (Das *et al.,* 2017). In the metagenome, Proteobacteria were dominant (62.6%) followed by Bacteroidetes (11.7%), Planctomycetes (7.7%), Verrucomicrobia (5.6%), Actinobacteria (3.7%) and Firmicutes (1.9%). The functional analysis revealed genes regulating the metabolic functions and cellular processes to be abundant (Das *et al.,* 2017). A considerable amount of sequences were identified as genes involved in arsenic resistance mechanism while half of them coded for the arsenate reductase enzyme which is the dominant enzyme of *ars*-operon. The study revealed the crucial role of microbial diversity in arsenic geocycle in contaminated ground water of Assam (Das *et al.,* 2017).

The rapid improvement in sequencing technology combined with reduction in cost lead to greater understanding of the microbial diversity of different habitats. Besides, the metagenomic analyses tools heleped in understanding not only the functional aspects of the micobial diversity but also facilitated identification of important enzymes, metabolites and such other molecules that can now be employed in bioremidiation as well as development of novel compounds for the benefit of human society.

Microbiome Projects

Human Microbiome

The importance of association between microbes and human as a host has been well established in the present state of health and disease concept. This fact is evident from the Human Microbiome Project (2007) as most alterations in the indigenous microbiome were

found linked with the diseased state. The details of dysbiosis in humans have not been well documented in India as compared to the Western countries. But few attempts have been made to decipher the association between humans and disease phenotypes with subjects in India.

The Human Microbiome Project that connects the microbial diversity in and out of human body has crossed the threshold in terms of research across the globe. Where, the Western world excels in the area of research in human microbiome; India suffers from lack of such studies. So far there are only few prominent reports which include a group from National Centre for Cell Sciences (NCCS)-Pune, India. Comparative study of faecal microbial diversity of healthy infants born through normal vaginal delivery and through caesarean section were found to be similar (Panday *et al.,* 2012a; Pandey *et al.,* 2012b). Generally, using the 16S rRNA marker gene sequencing technology, variations in gut microbiome associated with different age groups were evaluated (Marathe *et al.,* 2012). Another study reported the microbial community associated with obese individuals (Patil *et al.,* 2012). Nutrition has been a major challenge for healthy lifestyle. Researchers have highlighted the impact of varying nutritional status on gut microbiome and demonstrated that the microbial communities and their associated functions vary with nutritional status (Ghosh *et al.,* 2014; Shetty *et al.*, 2013b). Walujkar *et al.* (2014) studied the gut microbiota in patients suffering from different stages of Ulcerative colitis (UC) and reported a 10-fold increase in the total bacterial count in patients suffering from severe inflammatory stage when compared with patients with moderate and mild stages of inflammation. Another report reviewed the association of microbial community in relation to pediatric diseases like infantile colic, necrotizing enterocolitis, asthma, atopy, obesity, type-I diabetes, and autism (Arora *et al.,* 2015).

Since India is multi-ethnic and geographically and culturally diverse country, studies have been performed to identify the variations in the microbiome composition in light of these differences. *In-silico* analysis was conducted for Carbohydrates Active enzyme (CAZyme) profiles in the gut microbiota of 448 individuals belonging to different geographical locations.The study detected several geographical and age specific tendencies in gut CAZyme repertoire of the individuals which were also linked with BMI. The authors predicted that the abundance of CAZymes is lower in Indian infants/children as compared to those belonging to Japanese, Venezuela and Malawi population (Bhattacharya *et al.,* 2015). Another group studied the gut bacterial diversity of the tribes of India to understand the effect of ethnicity and geography on gut microbiota and carried out meta-analysis to compare them with the world-wide data host (Dehingia *et al.,* 2015). The prevalence of lifestyle related disorders like obesity, diabetes, IBD etc. have been known to be significantly lower compared to the non-tribal (urbanized) populations across the globe (Jain *et al.,* 2015). Studies have also focused on differences in gut microbial communities of healthy Indian subjects compared to the microbiota from other populations (Bhute *et al*., 2016). A study to understand the distinctive features of healthy Indian gut microbiome was also performed using 399 gut metagenomes from 8 different countries to describe several cross-geography trends (Yadav *et al.,* 2016). Other human microbiome studies focussed on linking microbiome to the diseased state of an individual, for instance using 16S rRNA gene-based sequencing, a group identified association of gut microbiota with type 2 diabetes mellitus (T2DM) on 30 subjects (Pushpanathan *et al.,* 2016; Suryavanshi *et al.,* 2016). A longitudinal study was carried out by Dinh *et al.* (2016) to study the alteration in gut microbiota in persistently stunted young children (from birth up to 2 years) in south India. This study demonstrated higher abundance of *Bacteroidetes* phylum in stunted children as compared to the control group (Dinh *et al.,* 2016). Most of the work mentioned above is a basic representation of microbial diversity analysis using amplicon sequencing. Recently a study by Maji *et al*. (2017) complemented the results of amplicon and shotgun-based metagenomics for deciphering the diversity as well as functional dysbiosis in gut of Indian individuals during tuberculosis infection. This study revealed that *Prevotella* is highly abundant in healthy individuals, whereas TB patients have high abundance of bacteria involved in short chain fatty acid production (Propionate and butyrate) which subsequently leads to reduced uptake of nutrients, lowered cholesterol and BMI levels (Maji *et al*., 2017). Future work can explore the dysbiosis in infected patients and study how therapy modulates the health state by affecting

the microbiome.

Fisheries and Aquaculture

The aquaculture industry primarily focuses on farming commercially important fish species for obtaining food and other important biomolecules such as protein and lipids from visceral organs. One important research area that has gained much popularity in recent years is the analyses of the diversity of microorganisms residing inside the gastro-intestinal (GI) tract of fish due to their vast functional roles that influence the host (Tarnecki *et al.,* 2017). The gut associated bacteria might be allochthonous that come from the ingested food or water and remain in the luminal contents or may adhere to the mucosal wall to become autochthonous (Ghanbari *et al.*, 2015). The gut microbial flora interacts with the host in a complex manner to influence host responses. These gut bacteria take part in the break-down of nutrients and provide the host with physiologically active materials, such as enzymes. Studying the gut microbiota structure and function has emerged as an important research tool for enhancing the fish health status in the last two decades (Ghosh *et al.,* 2002a). In order to develop our understanding of the gut microbial composition of the fish, culture-dependent approaches have been widely employed. These methods rely upon homogenizing sections of gut and plating the homogenate onto a range of selective media to obtain cultivable bacteria (Ghosh *et al.,* 2002a, Saha *et al.,* 2006, Ghosh *et al.,* 2010). However, electron microscopic examinations of intestinal wall using scanning electron microscopy and/or Transmission electron microscopy also serve as important tool for studying the mode of association of these microbial communities (Banerjee G *et al.,* 2015; Banerjee *et al.,* 2016). The bacterial populations isolated from GI tract are shown to aid in the nutrition of the fish (Ghosh *et al.,* 2002a, Kar and Ghosh 2008; Ganguly and Prasad, 2012) by producing important proteolytic, cellulolytic, lipolytic and amylolytic enzymes within the gut of freshwater fish species (Bairagi *et al.,* 2002; Nagvenkar *et al.,* 2006; Mondal *et al.,* 2008; Esakkiraj *et al.,* 2009; Mondal *et al.,* 2010; Ray *et al.,* 2010; Banerjee *et al.,* 2013; Dey *et al.,* 2016) and marine fishes (Das *et al.,* 2014). Roy *et al.* (2009) isolated phytase-producing bacterial flora from the foregut and hindgut regions of 10 freshwater teleost species with different feeding habits, namely rohu (*Labeo rohita*), catla (*Catla catla*), mrigal (*Cirrhinus mrigala*), bata (*Labeo bata*), kalbasu (*Labeo calbasu)*, Nile tilapia (*Oreochromis niloticus*), climbing perch (*Anabas testudineus*), common carp (*Cyprinus carpio*), silver carp (*Hypophthalmichthys molitrix*) and grass carp (*Ctenopharyn godonidella*). Dan and Ray (2014) identified three potent phytase-producing strains from four freshwater teleost fish species. These studies have been further extended to quantitative assessment of amylase, cellulase, lipase and protease activities by the selected enzyme producing bacterial strains. Likewise, chitinolytic gut bacteria have been isolated from carps (Banerjee *et al.,* 2015). Recently, the gut microbiota of 10 freshwater teleost species was investigated for the presence of tannase-producing bacterial strains (Mandal and Ghosh, 2013). Tannins are plant-derived antinutrients that readily bind with protein and other macromolecules to form indigestible complexes. The study identified one bacterial and three yeast strains with tannase producing ability. Since the gut associated bacteria serve diverse functions, several attempts have been made to screen the growth promoting, immune-stimulatory and pathogen inhibitory gut bacteria as potential probiotics for the fish (Ghosh *et al.,* 2002b; Ghosh K *et al.,* 2003; Kar *et al.,* 2008; Dutta and Ghosh, 2015; Ganguly *et al.,* 2010; Mukherjee and Ghosh, 2014). On similar grounds, gut microbiota of marine fishes has been explored for identifying potential probiotic candidates and isolating enzyme producing isolates (Sivasubramanian *et al.,* 2012; Velmurugan and Rajagopal, 2009). These studies revealed that the use of probiotics drastically increase the beneficial microbial load. The microorganisms identified as possible enzyme producers in fish gut include *Bacillus*, and *Acinetobacter*, *Aeromonas*, *Flavobacterium*, *Photobacterium*, *Pseudomonas*, *Vibrio*, *Microbacterium*, *Micrococcus*, *Staphylococcus*, Enterobacteriaceae, unidentified anaerobes and yeast (Ray *et al.,* 2012). Compared to the fish diversity in India, the studies involving metagenomic analyses of fish gut microbial communities to elucidate not only the structural composition but also the complete functional potential are far from few. Studying the complex nature of fish-microbial interactions will help in improving the nutritional value of fishes.

Plant Microbe Interaction

Association of microbes with plants whether pathogenic, symbiotic, or commensal may affect plant growth, directly or indirectly, and its susceptibility to diseases as well as the stress tolerance. The most widely studied ecological zone in the light of plant-microbe interactions is the rhizosphere, characterized by a narrow layer of soil in the close vicinity of roots. Rhizosphere serves as one of the most varied habitats for microbes in terms of community structure and species richness due to high nutrient and mineral content. This rich mineral source thus is critical for the survival and maintenance of the bacterial populations. In addition, the abundance of Plant Growth Promoting Rhizobacteria (PGPR) involved in suppression of plant diseases, seedling growth and development, production of phytohormones, increased efficiency of nutrient uptake, drought stress tolerance and nitrogen fixation in crop plants hold agronomic significance (Bora *et al.,* 2016). It is also important to study the diversity and function of these microbial communities inhabiting rhizosphere of various non-cultivated and wild species of plants as well for the proper understanding of their individual roles. However, less than 1% of microbes can be isolated using culture methods and majority of them cannot be cultured using the established approaches. Hence, the focus has shifted to examine these unculturable microorganisms using metagenomics by employing available advanced sequencing techniques and computational tools. Initially gene centric approach-based metagenome analysis was performed to identify the microbial diversity in the paddy fields in Kerala, using 16S rRNA metagenomic clone library preparation, followed by restriction fragment length polymorphism (RFLP). It revealed predominance of Proteobacteria and Firmicutes (Arjun and Harikrishnan, 2011). Another study on metagenome of north Indian soils has identified the presence of different gram negative and gram positive genera, mostly Proteobacteria and Firmicutes, (Nahid and Ali, 2016). Also, abundance of Proteobacteria in Saffron rhizosphere suggested that Proteobacteria is the predominant phylum in rhizospheric conditions. Other than 16S rRNA gene sequence, *nif*H gene also has been exploited to identify the divergence of nitrogen-fixing bacteria in rhizospheric soil (Suyal *et al.,* 2015; Soni *et al.,* 2016). In Western Indian Himalayas, the predominance of diazotrophs specially *Agrobacterium tumefaciens, Methylococcus capsulatus, Geobacter bemidjiensis*, *Dechloromonas aromatica*, *Burkholderia xenovorans*, *Xanthobacter autotrophicus* and *Sideroxydans lithotrophicus* was predicted based on *nif*H gene sequences. (Suyal *et al.,* 2015). Along with these proteobacterial genus, strains of Actinobacteria and Firmicutes, Cyanobacteria, Methanotrophs and Archaea were also detected (Soni *et al.,* 2016). However, due to biasness of these studies by using a nitrogen fixating gene, *nif*H only the diversity of diazotrophs were highlighted. Another study used both 16S rRNA gene sequences and ribulose-1,5-bisphosphate carboxylase/oxygenase (*cbb*L) gene to analyze the bacterial diversity from the rhizosphere of *Arachis hypogaea* from an agricultural field located at Bhavnagar, Gujarat. In this study, *cbb*L gene-based analysis showed the abundance of Proteobacteria such as *Rhizobium leguminosarum*, *Bradyrhizobium* sp., *Sinorhizobium meliloti* and *Ochrobactrum anthropi*, whereas 16S rRNA showed the abundance of bacteria in the order of Firmicutes (34.4%), Proteobacteria (18.3%), Actinobacteria (17.2%) and Bacteroidetes (16.1%) (Yousuf *et al.,* 2012). The product of gene *cbb*L is a crucial enzyme in Calvin-Benson-Bassham cycle for autotrophic carbon fixation in bacteria (Tourova *et al.,* 2010). Hence, the observed variation in species richness could be due to inherent biasness.

Bacterial communities are known to grow even in nutritionally poor soils in association with exotic weeds like *Prosopis juliflora* and *Parthenium hysterophorus* which can be studied for novel species and diverse functions for adaptability to nutritional stress. Principal genera inhabiting rhizosphere in this study included Acidobacteria, Gammaproteobacteria, and Bacteriodetes in Prosopis, and Acidobacteria, Betaproteobacteria, Nitrospiraein in *Parthenium* (Jothibasu *et al.,* 2012). Science of metagenomics has also been applied in evaluating the role of rhizobacteria in disease suppression. For example, a comparative analysis of rhizosphere soils from diseased and disease-free apple trees from orchards in Kinnaur district of Himachal Pradesh, India revealed increased production of chitinase and β-1,3 glucanase enzymes and enhanced activity in disease free samples; however, there was no significant difference in the composition or diversity of bacterial

communities (Shanmugam *et al.,* 2011).

Nowadays, focus has shifted towards the complete metagenomics profile of a site rather than on a single gene or group of genes. But in India, most of the studies have explored the ecological niches by using gene markers. One study from India was undertaken in Central Rice Research Institute, Cuttack, Odisha, India where whole genome metagenomics approach was used to demonstrate bacterial diversity and population dynamics under ambient CO_2 and elevated CO_2 + temperature. Pathways like nitrogen fixation, assimilatory and dissimilatory nitrate reduction and denitrification were studied for this fact (Bhattacharyya *et al.,* 2016).

Conclusions

Genomic and metagenomic sequencing and data analyses have attracted considerable attention owing to their potential to unravel microbial distribution, abundance and community dynamics from diverse environments. The science of "omics" has enabled identification of novel enzymes and proteins having innumerable industrial and agricultural applications which otherwise could not be identified using traditional microbiological approaches. Genomics has given accurate clues about "what is there" while metagenomics has been elemental in identifying "what 'else' is there". By using meta-omics approaches, besides identifying function of microbial communities in different environments, hetherto undiscovered hidden genes can now be validly established. Both, genomics and metagenomics has added to the knowledge of microbial ecology and provided clues for the evolution of metabolic process in the environment. However, there still exist unprecedented challenges and gaps in our understanding about the microbial resource in Indian subcontinent although there are quite a few reports on metagenomics/ genomics from India. However, with the availablity of efficient and low cost sequencing methods and advanced computational tools for high throughput data analyses, it would be possible in the future to develop better understanding of microbial resources and their complete metabolic potential existing in Indian subcontinent.

Acknowledgements

RL is thankful to The National Academy of Sciences, India (NASI) for providing the NASI Senior Scientist Platinum Jubilee Fellowship.

References

Agarwal L and Purohit HJ (2013) Genome sequence of *Rhizobium lupini* HPC (L) isolated from saline desert soil, Kutch (Gujarat) *Genome Announc* **1** e00071-12

Ahmed N, Saini V, Raghuvanshi S, Khurana JP, Tyagi AK *et al.* (2007) Molecular Analysis of a Leprosy Immunotherapeutic *Bacillus* Provides Insights into *Mycobacterium* Evolution *PLoS One* **2** e968

Ahmed V, Verma MK, Gupta S, Mandhan V and Chauhan NS (2018) Metagenomic Profiling of Soil Microbes to Mine Salt Stress Tolerance Genes *Front Microbiol* **9** 159

Arjun JK and Harikrishanan K (2011) Metagenomic analysis of bacterial diversity in the rice rhizosphere soil microbiome *Biotechnol Bioinf Bioeng* **1** 361-367

Arora SK, Dewan P and Gupta P (2015) Microbiome: Paediatricians' perspective *Indian J Med Res* **142** 515

Avasthi TS, Kumar N, Baddam R, Hussain A, Nandanwar N, Jadhav S, and Ahmed N (2011) Genome of multidrug-resistant uropathogenic *Escherichia coli* strain NA114 from India *J Bacteriol* **193** 4272-4273

Badhai J, Ghosh TS and Das SK (2015) Taxonomic and functional characteristics of microbial communities and their correlation with physicochemical properties of four geothermal springs in Odisha, India *Front Microbiol* **6** 1166

Bairagi A, Ghosh KS, Sen SK and Ray AK (2002) Enzyme producing bacterial flora isolated from fish digestive tracts *Aquaculture International* **10** 109-121

Balaji V, Rajenderan S, Anandan S and Biswas I (2015) Genome sequences of two multidrug-resistant *Acinetobacter baumannii* clinical strains isolated from southern India *Genome Announc* **3** e01010-15

Banerjee G, Dan S K, Nandi A, Ghosh P and Ray AK (2015) Autochthonous gut bacteria in two Indian air-breathing fish, climbing perch (*Anabas testudineus*) and walking catfish (*Clariasbatrachus*) mode of association, identification, and enzyme producing ability *Pol J Microbiol* **64** 361-368

Banerjee G, Nandi A, Dan S K, Ghosh P and Ray AK (2016) Mode of association, enzyme producing ability and

identification of autochthonous bacteria in the gastrointestinal tract of two Indian air-breathing fish murrel (*Channa punctatus*) and stinging catfish (*Heteropneustesfossilis*) *Proceedings of the Zoological Society* **70** 132-140

Banerjee G, Ray A K, Askarian F and Ringo E (2013) Characterisation and identification of enzyme producing autochthonous bacteria from the gastrointestinal tracts of two Indian air-breathing fish *Benef Microbes* **4** 277-284

Banerjee S, Mukherjee A, Dutta D and Ghosh K (2015) Evaluation of chitinolytic gut bacteria in some carps and optimization of culture conditions for chitinase production by the selected bacteria *J Microbiol Biotechnol Food Sci* **5** 12-19

Basak P, Pramanik A, Roy R, Chattopadhyay D and Bhattacharyya M (2015) Cataloguing the bacterial diversity of the Sundarbans mangrove, India in the light of metagenomics *Genom data* **4** 90-92

Beneduzi A, Ambrosini A and Passaglia LM (2012) Plant growth-promoting rhizobacteria (PGPR): their potential as antagonists and biocontrol agents *Genet Mol Biol* **35** 1044-1051

Bhat A, Riyaz-Ul-Hassan S, Srivastava N and Johri S (2015) Molecular cloning of rhodanese gene from soil metagenome of cold desert of North-West Himalayas: sequence and structural features of the rhodanese enzyme *3 Biotech* **5** 513-521

Bhatia S, Batra N, Pathak A, Green SJ, Joshi A and Chauhan A (2015) Metagenomic evaluation of bacterial and archeal diversity in the geothermal hot springs of Manikaran, India *Genome Announc* **3** e01544-14

Bhattacharya T, Ghosh T S and Mande SS (2015) Global profiling of carbohydrate active enzymes in human gut microbiome *PLoS One* **10** 1-20

Bhattacharyya P, Roy KS, Das M, Ray S, Balachandar D, Karthikeyan S, Nayak AK and Mohapatra T (2016) Elucidation of rice rhizosphere metagenome in relation to methane and nitrogen metabolism under elevated carbon dioxide and temperature using whole genome metagenomic approach *Sci Total Environ* **542** 886-898

Bhattacharyya P, Roy KS, Nayak AK, Shahid M, Lal B, Gautam P and Mohapatra T (2017) Metagenomic assessment of methane production-oxidation and nitrogen metabolism of long-term manured systems in lowland rice paddy *Sci Total Environ* **586** 1245-1253.

Bhute S, Pande P, Shetty SA, Shelar R, Mane S, Kumbhare S, Gawali A, Makhani H, Navandar M, Dhotre D *et al*. (2016) Molecular characterization and meta-analysis of gut microbial communities illustrate enrichment of *Prevotella* and *Megasphaera* in Indian Subjects *Front Microbiol* **7** 1-14

Bora SS, Keot J, Das S, Sarma K, and Barooah M (2016) Metagenomics analysis of microbial communities associated with a traditional rice wine starter culture (*Xaj-pitha*) of Assam, India *Biotech* **6** 153-166

Bose D, Sarkar I, Labar R, Oshone R, Ghazal S, Morris K, Abebe-Akele F, Thomas WK, Tisa LS and Sen A (2016) Comparative genomics of *Prauserella* sp. Am3, an actinobacterium isolated from root nodules of *Alnusnepalensis* in India *Symbiosis* **70** 49-58

Chauhan NS, Ranjan R, Purohit HJ, Kalia VC and Sharma R (2009) Identification of genes conferring arsenic resistance to *Escherichia coli* from an effluent treatment plant sludge metagenomic library *FEMS Microbiol Ecol* **67** 130-139

Craig J, Absar A, Bhat G, Cadel G, Hafiz M, Hakhoo N, Kashkari R, Moore J, Ricchiuto TE, Thurow J, and Thusu B (2013) Hot springs and the geothermal energy potential of Jammu and Kashmir State, NW Himalaya, India *Earth Sci Rev* **126** 156-177

Dan S K and Ray AK (2014) Characterization and identification of phytase-producing bacteria isolated from the gastrointestinal tract of four freshwater teleosts *Ann Microbiol* **64** 297-306

Das P, Mandal S, Khan A, Manna SK and Ghosh K (2014) Distribution of extracellular enzyme-producing bacteria in the digestive tracts of 4 brackish water fish species *Turkish J Zool* **38** 79-88

Das S, Bora SS, Yadav RNS and Barooaha M (2017) A metagenomic approach to decipher the indigenous microbial communities of arsenic contaminated groundwater of Assam *Genom Data* **12** 89-96

De Mandal S, Panda AK, Bisht SS and Kumar NS (2015c) First report of bacterial community from a bat guano using Illumina next-generation sequencing *Genom data* **4** 99-101

De Mandal S, Panda AK, Lalnunmawii E, Bisht SS and Kumar NS (2015a) Illumina-based analysis of bacterial community in Khuangcherapuk cave of Mizoram, Northeast India *Genom data* **5** 13-14

De Mandal S, Sanga Z and Kumar NS (2015b) Metagenome sequencing reveals *Rhodococcus*dominance in Farpuk Cave, Mizoram, India, an Eastern Himalayan biodiversity hot spot region *Genome Announc* **3** e00610-15

Dehingia M, Thangjamdevi K, Talukdar NC, Talukdar R, Reddy N, Mande SS, Deka M and Khan MR (2015) Gut bacterial diversity of the tribes of India and comparison with the worldwide data *Sci Rep* **5** 18563

Devi SG, Fathima AA, Sanitha M, Iyappan S, Curtis WR and

Ramya M (2016) Expression and characterization of alkaline protease from the metagenomic library of tannery activated sludge *J Biosci Bioeng* **122** 694-700

Devi SN, Mahalakshmi G, Rethika S and Asraf SS (2017) Metagenome of Indian one rupee coin reveals plethora of microbiota *IIOABJ* **8** 69-75

Dey A, Ghosh K and Hazra N (2016) Evaluation of extracellular enzyme-producing autochthonous gut bacteria in walking catfish, *Clariasbatrachus* (L.) *J Fisheries* **4** 345-352

Dhakan DB, Saxena R, Chaudhary N and Sharma VK (2016) Draft Genome Sequence of *Tepidimonasta iwanensis* Strain MB2, a Chemolithotrophic Thermophile Isolated from a Hot Spring in Central India *Genome Announc* **4** e01723-15

Dhal PK and Sar P (2014) Microbial communities in uranium mine tailings and mine water sediment from Jaduguda U mine, India A culture independent analysis *J Environ Sci Health* **49** 694-709

Dinh DM, Ramadass B, Kattula D and Sarkar R (2016) Longitudinal Analysis of the Intestinal Microbiota in Persistently Stunted Young Children in South India Longitudinal Analysis of the Intestinal Microbiota in Persistently Stunted Young Children in South India *PloS One* **11** e0155405

Dudhagara P, Ghelani A, Patel R, Chaudhari R and Bhatt S (2015) Bacterial tag encoded FLX titanium amplicon pyrosequencing (bTEFAP) based assessment of prokaryotic diversity in metagenome of Lonar soda lake, India *Genom data* **4** 8-11

Dutta D and Ghosh K (2015) Screening of extracellular enzyme-producing and pathogen inhibitory gut bacteria as putative probiotics in mrigal, *Cirrhinusmrigala* (Hamilton, 1822) *Int J Fisheries Aquat Studies* **2** 310-318

Dwivedi V, Sangwan N, Nigam A, Garg N, Niharika N, Khurana P, Khurana JP and Lal R (2012) Draft genome sequence of *Thermus* sp. RL isolated from hot water spring located atop the Himalayan Ranges at Manikaran, India *J Bacteriol* **194** 3534-3535

Esakkiraj P, Immanuel G, Sowmya SM, Iyapparaj P and Palavesam A (2009) Evaluation of protease-producing abilityof fish gut isolate *Bacillus cereus* for aqua feed *Food Bioproc Technol* **2** 383-390

Fleischmann RD, Adams MD, White O, Clayton RA, Kirkness EF, Kerlavage AR, Bult CJ, Tomb JF, Dougherty BA and Merrick JM (1995) Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd *Science* **269** 496-512

Ganguly S and Prasad A (2012) Microflora in fish digestive tract plays significant role in digestion and metabolism *Rev Fish Biol Fisheries* **22** 11-16

Ganguly S, Paul I and Mukhopadhayay SK (2010) Immunostimulant, probiotic and probiotic-their applications and effectiveness in aquaculture *Israeli J Aquaculture (Bamidgeh)* **62** 130-138

Garg R, Srivastava R, Brahma V, Verma L, Karthikeyan S and Sahni G (2016) Biochemical and structural characterization of a novel halotolerant cellulase from soil metagenome *Sci Rep* **6** 39634

Ghanbari M, Kneifel W and Domig KJ (2015) A new view of the fish gut microbiome advances from next-generation sequencing *Aquaculture* **448** 464-475

Ghelani A, Patel R, Mangrola A and Dudhagara P (2015) Cultivation-independent comprehensive survey of bacterial diversity in TulsiShyam hot springs, India *Genom Data* **4** 54-56

Ghosh A, Chandratre K, Chaudhary A, Chaudhary S, Badani N, Chaudhary PS, Dhawan D, Vudathala S and Chikara SK (2015) Whole-genome sequencing of *Brevundimonas diminuta* XGC1, isolated from a tuberculosis patient in Gujarat, India *Genome Announc* **3** e00686-15

Ghosh D, Bal B, Kashyap VK and Pal S (2003) Molecular phylogenetic exploration of bacterial diversity in a Bakreshwar (India) hot spring and culture of *Shewanella*-related thermophiles *Appl Environ Microbiol* **69** 4332-4336

Ghosh K, Roy M, Kar N and Ringo E (2010) Gastrointestinal bacteria in rohu, *Labeo rohita* (Actinopterygii Cypriniformes Cyprinidae) scanning electron microscopy and bacteriological study *Acta Ichthyologica et Piscatoria* **40** 129-135

Ghosh K, Sen SK and Ray AK (2002a) Characterization of *bacilli* isolated from the gut of rohu, *Labeorohita*, fingerlings and its significant in digestion *J Appl Aquaculture* **12** 33-42

Ghosh K, Sen SK and Ray AK (2002b) Growth and survival of rohu, *Labeorohita* (Hamilton) spawn fed diets supplemented with fish intestinal microflora *Acta Ichthyologica et Piscatoria* **32** 83-92

Ghosh K, Sen SK and Ray AK (2003) Supplementation of an isolated fish gut bacterium, *Bacillus circulans*, in formulated diets for rohu, *Labeorohita*, fingerlings *Israeli J Aquaculture* (*Bamidgeh*) **55** 13-21

Ghosh TS, Gupta SS, Bhattacharya T, Yadav D, Barik A, Chowdhury A, Das B, Mande SS and Nair GB (2014) Gut microbiomes of Indian children of varying nutritional status *PLoS One* **9** 1-13

Glick BR (1995) The enhancement of plant growth by free-living

bacteria *Can J Microbiol* **41** 109-117

Gupta A, Gopal M, Thomas GV, Manikandan V, Gajewski J, Thomas G, *et al*. (2014) Whole Genome Sequencing and Analysis of Plant Growth Promoting Bacteria Isolated from the Rhizosphere of Plantation Crops Coconut, Cocoa and Arecanut *PLoS One* **9** e104259

Gupta S, Sharma P, Dev K and Sourirajan A (2016) Halophilic Bacteria of Lunsu Produce an Array of Industrially Important Enzymes with Salt Tolerant Activity *Biochem Res Int* **2016** 9237418

Islam E, Dhal PK, Kazy SK and Sar P (2011) Molecular analysis of bacterial communities in uranium ores and surrounding soils from Banduhurang open cast uranium mine, India A comparative study *J Environ Sci Health* **46** 271-280

Jadeja NB, More RP, Purohit HJ and Kapley A (2014) Metagenomic analysis of oxygenases from activated sludge *Bioresour Technol* **165** 250-256

Jain Y, Kataria R, Patil S, Kadam S, Kataria A, Jain R, Kurbude R and Shinde S (2015) Burden & pattern of illnesses among the tribal communities in central India A report from a community health programme *Indian J Med Res* **141** 663-672

Jalali S, Kohli S, Latka C, Bhatia S, Vellarikal SK, Sivasubbu S, Scaria V and Ramachandran S (2015) Screening currency notes for microbial pathogens and antibiotic resistance genes using a shotgun metagenomic approach *PLoS One* **10** e0128711

Jayachandra SY, Anil Kumar S, Merley DP and Sulochana MB (2012) Isolation and characterization of extreme halophilic bacterium *Salinicoccus* sp. JAS4 producing extracellular hydrolytic enzymes *Recent Res Sci Tech* **4** 46-49

Jose PA, Sivakala KK, and Jebakumar SR (2013) Formulation and statistical optimization of culture medium for improved production of antimicrobial compound by *Streptomyces* sp. JAJ06 *Int J Microbiol* **2013** 526260

Jothibasu K, Chinnadurai C, Sundaram SP, Kumar K and Balachandar D (2012) Molecular Profiling of Rhizosphere Bacterial Communities Associated with *Prosopis juliflora* and *Parthenium hysterophorus J Microbiol Biotechnol* **22** 301-310

Kar N and Ghosh K (2008) Enzyme producing bacteria in the gastrointestinal tracts of *Labeorohita* (Hamilton) and *Channa punctatus* (Bloch) *Turkish J Fisheries Aquat Sci* **8** 115-120

Kar N, Roy RN, Sen SK and Ghosh K (2008) Isolation and characterization of extracellular enzyme producing bacilli in the digestive tracts of rohu, *Labeorohita* (Hamilton) and Murrel, *Channa punctatus* (Bloch) *Asian Fisheries Sci* **21** 421-434

Kaur J, Verma H, Tripathi C, Khurana JP and Lal R (2013) Draft genome sequence of a hexachlorocyclohexane-degrading bacterium, *Sphingobium baderi* strain LL03^T *Genome Announc* **1** e00751-13

Kaur N, Kumar S and Mayilraj S (2014) Genome sequencing and annotation of *Amycolatopsis vancoresmycina* strain DSM 44592^T *Genom data* **2** 16-17

Kaur N, Kumar S, Bala M, Raghava GPS and Mayilraj S (2013) Draft genome sequence of *Amycolatopsis decaplanina* strain DSM 44594^T *Genome Announc* **1** e00138-13

Khatri I, Subramanian S and Mayilraj S (2014) Genome sequencing and annotation of *Amycolatopsis azurea* DSM 43854^T *Genom Data* **2** 44-45

Kothari VV, Kothari RK, Kothari CR, Bhatt VD, Nathani NM, Koringa PG, Joshi CG and Vyas BRM (2013) Genome sequence of salt-tolerant *Bacillus safensis* strain VK, isolated from saline desert area of Gujarat, India *Genome Announc* **1** e00671-13

Krishna PS, Sreenivas A, Singh DK, Shivaji S and Prakash JS (2015) Draft genome sequence of *Bacillus okhensis* Kh10-101T, a halo-alkali tolerant bacterium from Indian saltpan *Genom Data* **6** 283-284

Krishnan R, Menon RR, Tanaka N, Hans-Jürgen Busse HJ, Krishnamurthi S and Rameshkumar N (2016) *Arthrobacter pokkalii* sp. nov, a Novel Plant Associated Actinobacterium with Plant Beneficial Properties, Isolated from Saline Tolerant Pokkali Rice, Kerala, India *PLoS One* **11** e0150322

Kumar M, Yadav AN, Tiwari R, Prasanna R and Saxena AK (2014) Deciphering the diversity of culturable thermotolerant bacteria from Manikaran hot springs *Ann Microbiol* **64** 741-751

Kumar N, Lad G, Giuntini E, Kaye ME, Udomwong P, Shamsani NJ, Young JP and Bailly X (2015) Bacterial genospecies that are not ecologically coherent: population genomics of *Rhizobium leguminosarum Open Biol* **5** 140133

Kumar R, Dwivedi V, Negi V, Khurana JP and Lal R (2013) Draft genome sequence of *Sphingobium lactosutens* strain DS20 isolated from an hexachlorocyclohexane (HCH) dumpsite *Genome Announc* **1** 00753-13

Kumar S, Karan R, Kapoor S, Singh SP and Khare SK (2012) Screening and isolation of halophilic bacteria producing industrially important enzymes *Braz J Microbiol* **43** 1595-603

Lubick N (2011) Dumped drugs lead to resistant microbes *Nature* doi101038/news201146

Mahato NK, Gupta V, Singh P, Kumari R, Verma H, Lal R *et al.*

(2017) Microbial Taxonomy in the era of OMICS: application of DNA sequences, computational tools and techniques. *Antonie van Leeuwenhoek* **110** 1357-1371

Mahato NK, Tripathi C, Verma H, Singh N and Lal R (2014) Draft genome sequence of *Deinococcus* sp. strain RL isolated from sediments of a hot water spring *Genome Announc* **2** e00703-14

Maji A, Misra R, Dhakan D, Gupta V, Mahato NK, Saxena R, Mittal P, Thukral N, Sharma E, Singh A, Virmani R, Gaur M, Singh H, Hasija Y, Arora G, Agrawal A, Chaudhry A, Khurana JP, Sharma VK, Lal R, Singh Y (2017) Gut microbiome contributes to impairment of immunity in pulmonary tuberculosis patients by alteration of butyrate and propionate producers *Environ Microbiol* **20** 402-419

Malathi J, Murugan N, Umashankar V, Bagyalakshmi R and Madhavan HNR (2013) Draft Genome Sequence of Multidrug-Resistant *Pseudomonas aeruginosa* Strain VRFPA02, Isolated from a Septicemic Patient in India *Genome Announc* **1** e00425-13

Mandal S and Ghosh K (2013) Isolation of tannase-producing microbiota from the gastrointestinal tracts of some freshwater fish *J Appl Ichthyol* **29** 145-153

Mangrola AV, Dudhagara P, Koringa P, Joshi CG, Parmar M and Patel R (2015b) Deciphering the microbiota of Tuwa hot spring, India using shotgun metagenomic sequencing approach *Genom Data* **4** 153-155

Mangrola AV, Dudhagara P, Koringa P, Joshi CG and Patel RK (2015a) Shotgun metagenomic sequencing based microbial diversity assessment of Lasundra hot spring, India *Genom Data* **4** 73-75

Manolio TA, Chisholm RL, Ozenberger B, Roden DM, Williams MS, Wilson R, Bick D, Bottinger EP, Brilliant MH, Eng C and Frazer KA (2013) Implementing genomic medicine in the clinic: the future is here *Genet Med* **15** 258-267

Marathe NP, Shetty SA, Shouche YS and Larsson DG (2016) Limited bacterial diversity within a treatment plant receiving antibiotic-containing waste from bulk drug production *PLoS One* **11** e0165914

Marathe NP, Regina VR, Walujkar SA, Charan SS, Moore ERB, Larsson DGJ *et al*. (2013) A Treatment Plant Receiving Waste Water from Multiple Bulk Drug Manufacturers Is a Reservoir for Highly Multi-Drug Resistant Integron-Bearing Bacteria *PLoS One* **8** e77310

Marathe NP, Shetty SA, Lanjekar V, Shouche YS and Ranade D (2012) Changes in human gut flora with age an Indian familial study *BMC Microbiol* **12** 222

Mathur P, Veeraraghavan B, DevangaRagupathi NK, Inbanathan FY, Khurana S, Bhardwaj N, Kumar S, Sagar S and Gupta A (2017) First report on a cluster of colistin-resistant *Klebsiella pneumoniae* strains isolated from a tertiary care centre in India Whole-genome shotgun sequencing *Genome Announc* **5** e01466-16

Mehetre GT, Paranjpe AS, Dastager SG and Dharne MS (2016) Complete metagenome sequencing based bacterial diversity and functional insights from basaltic hot spring of Unkeshwar, Maharashtra, India *Genom Data* **7** 140-143

Mondal S, Roy T and Ray AK (2010) Characterization and identification of enzyme-producing bacteria isolated from the digestive tract of bata, *Labeo bata J World Aquaculture Society* **41** 369-377

Mondal S, Roy T, Sen SK and Ray AK (2008) Distribution of enzyme-producing bacteria in the digestive tracts of some freshwater fish *Acta Ichthyologica et Piscatoria* **38** 1-8

More RP, Mitra S, Raju SC, Kapley A and Purohit HJ (2014) Mining and assessment of catabolic pathways in the metagenome of a common effluent treatment plant to induce the degradative capacity of biomass *Bioresour Technol* **153** 137-146

Mukherjee A and Ghosh K (2014) Antagonism against fish pathogens by cellular components and verification of probiotic properties in autochthonous bacteria isolated from the gut of an Indian major carp, *Catlacatla* (Hamilton) *Aquaculture Res* **45** 1-13

Mukherjee S, Kumar D, Nanda AK and Chakraborty R (2013) 16S rRNA gene sequence analyses of the metagenome derived from waters of river Mahananda at Siliguri An approach to understand bacterial diversity *Indian J Biotechnol* **12** 80-87

Mukherjee U, Saxena A, Kumari R, Singh P and Lal R (2014) Draft genome sequence of *Amycolatopsis mediterranei* DSM 40773, a tangible antibiotic producer *Genome Announc* **2** e00752-14

Nagvenkar GS, Nagvenkar SS, Rivonker CU and Sangodkar UMX (2006) Microbial diversity and enzyme production in mullet *Mugil cephalus* L (Pisces) along Goa, west coast of India *Indian J Marine Sci* **35** 36-42

Nahid S and Ali A (2016) Genetic Diversity Analysis of Wheat Rhizobacterial Isolates from North Indian Soils Showed the Unique Occurrence of *Sporosarcina* Related Species *Int J Sci Technol* **4** 29

Panda AK, Bisht SS, Kumar NS and Mandal SD (2015) Investigations on microbial diversity of Jakrem hot spring, Meghalaya, India using cultivation-independent approach *Genom Data* **4** 156-157

Panda AK, Bisht SS, Mandal SD and Kumar NS (2016) Bacterial and archaeal community composition in hot springs from

Indo-Burma region, North-east India *AMB Expr* **6** 111

Pandey PK, Siddharth J, Verma P, Bavdekar A, Patole M S and Shouche YS (2012b) Molecular typing of fecal eukaryotic microbiota of human infants and their respective mothers *J Biosci* **37** 221-226

Pandey PK, Verma P, Kumar H, Bavdekar A, Patole MS and Shouche YS (2012a) Comparative analysis of fecal microflora of healthy full-term Indian infants born with different methods of delivery (vaginal vs cesarean) *Acinetobacter* sp. prevalence in vaginally born infants *J Biosci* **37** 989-998

Pandit AS, Joshi MN, Bhargava P, Shaikh I, Ayachit GN, Raj SR, Saxena AK and Bagatharia SB (2015) A snapshot of microbial communities from the Kutch: one of the largest salt deserts in the World *Extremophiles* **1** 973-987

Patel R, Mevada V, Prajapati D, Dudhagara P, Koringa P and Joshi CG (2015) Metagenomic sequence of saline desert microbiota from wild ass sanctuary, Little Rann of Kutch, Gujarat, India *Genom Data* **3** 137-139

Patel V, Sharma A, Lal R, Al-Dhabi NA and Madamwar D (2016) Response and resilience of soil microbial communities inhabiting in edible oil stress/contamination from industrial estates *BMC Microbiol* **16** 50

Patil D, Dhotre DP, Chavan S, Sultan A, Jain DS, Lanjekar VB, Gangawani J, Shah P, Stodkar J, Shah S *et al*. (2012) Molecular analysis of gut microbiota in obesity among Indian individuals *J Biosci* **37** 647-657

Pawar SP, Dhotre DP, Shetty SA, Chowdhury SP, Chaudhari BL and Shouchea YS (2012) Genome Sequence of *Janibacter hoylei MTCC8307*, Isolated from the Stratospheric Air *J Bacteriol* **194** 6629-6630

Pinnaka AK, Ara S, Singh A and Shivaji S (2013b) Draft genome sequence of *Winogradsky ellapsychrotolerans* RS-3T, isolated from the marine transect of Kongsfjorden, Ny-Ålesund, Svalbard, Arctic Ocean *Genome Announc* **1** e00630-13

Pinnaka AK, Singh A, Ara S, Begum Z, Reddy GS and Shivaji S (2013a) Draft Genome Sequence of *Leifsonia rubra* Strain CMS 76R^T, Isolated from a Cyanobacterial Mat Sample from a Pond in Wright Valley, McMurdo, Antarctica *Genome Announc* **1** e00633-13

Pramanik A, Basak P, Banerjee S, Sengupta S, Chattopadhyay D and Bhattacharyya M (2016) Metagenomic exploration of the bacterial community structure at Paradip Port, Odisha, India *Genom Data* **7** 94-96

Pramanik A, Basak P, Banerjee S, Sengupta S, Chattopadhyay D and Bhattacharyya M (2015) Pyrosequencing based profiling of the bacterial community in the Chilika Lake, the largest lagoon of India *Genom Data* **4** 112-114

Puranik S, Pal RR, More RP, Purohit HJ (2016) Metagenomic approach to characterize soil microbial diversity of Phumdi at Loktak Lake *Water Sci Technol* **74** 2075-2086

Pushpanathan P, Srikanth P, Seshadri KG, Selvarajan S, Pitani RS, Kumar TD and Janarthanan R (2016) Gut Microbiota in Type 2 Diabetes Individuals and Correlation with Monocyte Chemoattractant Protein1 and Interferon Gamma from Patients Attending a Tertiary Care Centre in Chennai, India *Indian J Endocrinol Metab* **20** 523-530

Rajeev A , Sahu N, Deori M, Arun Dev S, Pal Yadav V and Ghosh I (2018) Metagenomic exploration of microbial signatures on Periyar river sediments from the Periyar tiger reserve in the Western Ghats *Genome Announc* **6** e00154-18

Rajeswari P, Jose PA, Amiya R and Jebakumar SR (2015) Characterization of saltern based *Streptomyces* sp. and statistical media optimization for its improved antibacterial activity *Front Microbiol* **5** 753

Rani P, Mahato NK, Sharma A, Rao DLN, Kamra K and Lal R (2017) Genome Mining and Predictive Functional Profiling of Acidophilic Rhizobacterium *Pseudomonas fluorescens* Pt14 *Indian J Microbiol* **57** 155-161

Rathour R, Gupta J, Kumar M, Hiloidhari M, Mehrotra AK and Thakur IS (2017) Metagenomic sequencing of microbial communities from brackish water of Pangong Lake of the northwest Indian Himalayas *Genome Announc* **5** e01029-17

Ray AK, Ghosh K and Ringo E (2012) Enzyme-producing bacteria isolated from fish gut A review *Aquaculture Nutr* **18** 465-492

Ray AK, Roy T, Mondal S and Ringø E (2010) Identification of gut-associated amylase, cellulase and protease-producing bacteria in three species of Indian major carps *Aquaculture Res* **41** 1462-1469

Reddy GKK, Nancharaiah YV and Venugopalan VP (2014) Aerobic granular sludge mediated biodegradation of an organophosphorous ester, dibutyl phosphite *FEMS Microbiol Lett* **359** 110-115

Reddy GS, Sreenivas A and Shivaji S (2014) Draft Genome Sequence of *Cryobacterium roopkundensis* Strain RuGl7, Isolated from a Soil Sample in the Vicinity of Roopkund Lake, Himalayas, India *Genome Announc* **2** e01206-14

Rohit A, Kumar BK, Deekshit VK, *et al*. (2015) Draft Genome Sequence of *Campylobacter fetus* MMM01, Isolated from a Chronic Kidney Disease Patient with Sepsis *Genome Annouc* **3** e01055-15

Roy T, Mondal S and Ray AK (2009) Phytase-producing bacteria in the digestive tracts of some freshwater fish *Aquaculture*

Res **40** 344-353

Saha S, Roy RN, Sen SK and Ray AK (2006) Characterization of cellulase-producing bacteria from the digestive tract of tilapia, *Oreochromis mossambica* (Peters) and grass carp, *Ctenopharyngodonidella* (Valenciennes) *Aquaculture Res* **37** 380-388

Saini V, Raghuvanshi S, Khurana JP, Ahmed N, Hasnain SE, Tyagi AK and Tyagi AK (2012) Massive gene acquisitions in *Mycobacterium indicus pranii* provide a perspective on mycobacterial evolution *Nucleic Acids Res* **40** 10832-10850

Saju KA, Michael Babu M, Murugan M and ThiraviaRaj S (2011) Survey on Halophilic microbial diversity of Kovalam Saltpans in Kanyakumari District and its industrial applications *J Appl Pharm Sci* **1** 160-163

Sangwan N, Lambert C, Sharma A, Gupta A, Khurana P, Khurana JP, Sockett RE, Gilbert JA and Lal R (2015) Arsenic rich Himalayan hot spring metagenomics reveal genetically novel predator-prey genotypes *Environ Microbiol Rep* **7** 812-823

Sangwan N, Lata P, Dwivedi V, Singh A, Niharika N, Kaur J, Anand S, Malhotra J, Jindal S, Nigam A, Lal R (2012) Comparative metagenomic analysis of soil microbial communities across three hexachlorocyclohexane contamination levels *PLoS One* **7** e46219

Sangwan N, Verma H, Kumar R, Negi V, Lax S, Khurana P, Khurana JP, Gilbert JA and Lal R (2014) Reconstructing an ancestral genotype of two hexachlorocyclohexane-degrading *Sphingobium* species using metagenomic sequence data *ISME J* **8** 398-408

Sangwan N, Verma H, Kumar R, Negi V, Lax S, Khurana P, Khurana JP, Gilbert JA and Lal R (2014) Reconstructing an ancestral genotype of two hexachlorocyclohexane-degrading *Sphingobium* species using metagenomic sequence data *ISME J* **8** 398-408

Sarafin Y, Donio MBS, Velmurugan S, Michaelbabu M and Citarasu T (2014) *Kocuria marina* BS-15 a biosurfactant producing halophilic bacteria isolated from solar salt works in India *Saudi J Biol Sci* **21** 511-519

Satyanarayana T, Raghukumar C and Shivaji S (2005) Extremophilic microbes' diversity and perspectives *Curr Sci* **89** 78-90

Saxena A, Kumari R, Mukherjee U, Singh P and Lal R (2014) Draft Genome Sequence of Rifamycin producer *Amycolatopsis rifamycinica* 46095 *Genome Announc* **2** e00662-14

Saxena A, Nayyar N, Sangwan N, Kumari R, Khurana JP and Lal R (2013) Genome Sequence of *Novosphingobium lindaniclasticum* LE124T, Isolated from a Hexachlorocyclohexane Dumpsite *Genome Announc* **1** e00715-13

Saxena R, Chaudhary N, Dhakan DB and Sharma VK (2015) Draft genome sequence of *Gulbenkiania mobilis* strain MB1, a sulfur-metabolizing thermophile isolated from a hot spring in central India *Genome Announc* **3** e01295-15

Saxena R, Dhakan D B, Mittal P, Waiker P, Chowdhury A, Ghatak A and Sharma VK (2016) Metagenomic analysis of hot springs in central India reveals hydrocarbon degrading thermophiles and pathways essential for survival in extreme environments *Front Microbiol* **7** 2123

Sen R and Maiti NK (2014) Genomic and Functional Diversity of Bacteria isolated from hot springs in Odisha, India *Geomicrobiol J* **31** 541-550

Sen SK, Raut S, Satpathy S, Bandyopadhyay B, Mohapatra PKD and Raut S (2015) Exploration of Microbial Diversity of Taptapani (India) Hot Spring Through Molecular Phylogenetic Analysis *Arab J Sci Eng* **40** 51-61

Shanmugam V, Verma R, Rajkumar S and Naruka DS (2011) Bacterial diversity and soil enzyme activity in diseased and disease-free apple rhizosphere soils *Ann Microbiol* **61** 765-772

Sharma A and Lal R (2016) Survey of (Meta)genomic approaches for understanding microbial community dynamics *Indian J Microbiol* **57** 23-38

Sharma A, Gilbert JA and Lal R (2016) (Meta) genomic insights into the pathogenome of *Cellulosimicrobium cellulans Sci Rep* **6** 25527

Sharma A, Hira H, Shakarad M and Lal R (2014) Draft genomeSequence of *Cellulosimicrobium* sp. Strain MM,isolated from arsenic-rich microbial mats of a Himalayan hot spring *Genome Announc* **2** e01020-14

Sharma N, Tanksale H, Kapley A and Purohit HJ (2012) Mining the metagenome of activated biomass of an industrial waste water treatment plant by novel method *Indian J Microbiol* **52** 538-543

Shetty SA, Marathe NP and Shouche YS (2013b) Opportunities and challenges for gut microbiome studies in the Indian population *Microbiome* **1** 24

Shetty SA, Marathe NP, Munot H, *et al*. (2013a) Draft Genome Sequence of *Methylophaga lonarensis* MPL^T, a Haloalkaliphilic (Non-Methane-Utilizing) Methylotroph *Genome Announc* **1** e00202-13

Shivaji S, Ara S, Bandi S, Singh A and Pinnaka AK (2013a) Draft genome sequence of *Arthrobacter gangotriensis* strain Lz1yT, isolated from a penguin rookery soil sample collected in Antarctica, near the Indian station Dakshin Gangotri *Genome Announc* **1** e00347-13

Shivaji S, Ara S, Prasad S, Manasa BP, Begum Z, Singh A and Kumar AP (2013b) Draft Genome Sequence of *Arcticibacter svalbardensis* Strain MN12-7T, a Member of the Family Sphingobacteriaceae Isolated from an Arctic Soil Sample *Genome Announc* **1** e00484-13

Shivaji S, Ara S, Singh SK, Bandi S, Singh A and Pinnaka AK (2012) Draft genome sequence of *Bacillus isronensis* Strain B3W22, isolated from the upper atmosphere *J bacterial* **194** 6624-6625

Sidhu C, Vikram S and Pinnaka AK (2017) Unravelling the microbial interactions and metabolic potentials in pre-and post-treated sludge from a wastewater treatment plant using metagenomic studies *Front Microbiol* **8** 1382

Singh A and Subudhi E (2015) Structural insights of microbial community of Deulajhari (India) hot spring using 16S-rRNA based metagenome sequencing *Genom Data* **7** 101-102

Singh A, Jangir PK, Sharma R, Singh A, Pinnaka AK and Shivaji S (2013) Draft Genome Sequence of *Indibacteral kaliphilus* Strain LW1T, Isolated from Lonar Lake, a Haloalkaline Lake in the Buldana District of Maharashtra, India *Genome Announc* **1** e00513-13

Singh A, Singh Chauhan N, Thulasiram H V, Taneja V and Sharma R (2010) Identification of two flavin monooxygenases from an effluent treatment plant sludge metagenomic library *Bioresour Technol* **101** 8481-8484

Singh J, Behal A, Singla N, Joshi A, Birbian N, Singh S, Bali V and Batra N (2009) Metagenomics Concept, methodology, ecological inference and recent advances *Biotechnol J* **4** 480-494

Singh P, Kumari R, Mukherjee U, Saxena A, Sood U and Lal R (2014) Draft Genome Sequence of Rifamycin derivatives producing *Amycolatopsis mediterranei* DSM 46096/ S955 *Genome Announc* **2** e00837-14

Sivasubramanian K, Ravichandran S and Kavitha R (2012) Isolation and Characterization of Gut Micro Biota from Some Estuarine Fishes *Marine Sci* **2** 1-6

Soni I, Chakrapani H and Chopra S (2015) Draft genome sequence of methicillin-sensitive *Staphylococcus aureus* ATCC 29213 *Genome Announc* **3** e01095-15

Soni R, Suyal DC, Sai S and Goel R (2016) Exploration of nifH gene through soil metagenomes of the western Indian Himalayas *Biotech* **6** 25

Suryavanshi MV, Bhute SS, Jadhav SD, Bhatia MS, Gune RP and Shouche YS (2016) Hyperoxaluria leads to dysbiosis and drives selective enrichment of oxalate metabolizing bacterial species in recurrent kidney stone endures *Sci Rep* **6** 34712

Suyal DC, Yadav A, Shouche Y and Goel R (2015) Diversified diazotrophs associated with the rhizosphere of Western Indian Himalayan native red kidney beans (*Phaseolus vulgaris* L.) *Biotech* **5** 433-441

Tarnecki AM, Burgos FA, Ray CL and Arias CR (2017) Fish Intestinal Microbiome Diversity and Symbiosis Unraveled by Metagenomics *J Appl Microbiol* **1** 2-17

Tourova TP, Kovaleva OL, Sorokin DY, and Muyzer G (2010) Ribulose-1,5-bisphosphate carboxylase/oxygenase genes as a functional marker for chemolithoautotrophic halophilic sulfur-oxidizing bacteria in hypersaline habitats *Microbiol* **156** 2016-2025

Tripathi C, Mahato NK, Rani P, Singh Y, Kamra K and Lal R **(**2016) Draft genome sequence of *Lampropedia cohaerens* strain CT6^{T} isolated from arsenic rich microbial mats of a Himalayan hot water spring *Stand Genomic Sci* **11** 64

Tripathi C, Mishra H, Khurana H, Dwivedi V, Kamra K, Negi RK and Lal R (2017) Complete genome analysis of *Thermus parvatiensis* and comparative genomics of *Thermus* spp. provide insights into genetic variability and evolution of natural competence as strategic survival attributes *Frontiers Microbiol* **8** 1410

Veerarghavan B, Anandan S, RajamaniSekar SK, Gopi R, DevangaRagupathi NK, Ramesh S, Verghese VP, Korulla S, Mathai S, Sangal L and Joshi S (2016) First report on the draft genome sequences of *Corynebacterium diphtheriae* isolates from India *Genome Annouc* **4** e01316-16

Velmurugan S and Rajagopal S (2009) Beneficial uses of probiotics in mass scale production of marine ornamental fish *Afr J Micro Res* **3** 185-190

Verma A, Gupta M and Shirkot P (2014) Isolation and characterization of thermophilic bacteria in natural hot water springs of Himachal Pradesh (India) *Bioscan* **9** 947-952

Verma M, Kaur J, Kumar M, Kumari K, Saxena A, Anand S, Nigam A, Ravi V, Raghuvanshi S, Khurana P, Tyagi AK, Khurana JP and Lal R (2011) Whole genome sequence of the rifamycin B-producing strain *Amycolatopsis mediterranei* S699 *J Bacteriol* **193** 5562-5563

Vidya J, Swaroop S, Singh S, Alex D, Sukumaran R and Pandey A (2011) Isolation and characterization of a novel á-amylase from a metagenomic library of Western Ghats of Kerala, India *Biologia* **66** 939-944

Vijaykumar S, Balaji V and Biswas I (2015a) Complete genome sequence of *Acinetobacter baumannii* strain B8300, which displays high twitching motility *Genome Announc* **3**

e00956-15

Vijaykumar S, Balaji V and Biswas I (2015b) Complete Genome Sequence of *Acinetobacter baumannii* strain B8342, a motility-positive clinical isolate *Genome Announc* **3** e00925-15

Walujkar SA, Dhotre DP, Marathe NP, Lawate PS, Bharadwaj RS and Shouche YS (2014) Characterization of bacterial community shift in human Ulcerative Colitis patients revealed by Illumina based 16S rRNA gene amplicon sequencing *Gut Pathog* **6** 22

Yadav D, Ghosh TS and Mande SS (2016) Global investigation of composition and interaction networks in gut microbiomes of individuals belonging to diverse geographies and age-groups *Gut Pathog* **8** 17

Yadav TC, Pal RR, Shastri S, Jadeja NB and Kapley A (2015) Comparative metagenomics demonstrating different degradative capacity of activated biomass treating hydrocarbon contaminated wastewater *Bioresour Technol* **188** 24-32

Yousuf B, Keshri J, Mishra A and Jha B (2012) Application of targeted metagenomics to explore abundance and diversity of CO 2-fixing bacterial community using *cbb*L gene from the rhizosphere of *Arachis hypogaea Gene* **506** 18-24.