We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

5,300
Open access books available

130,000
International authors and editors

155M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS BOOK CITATION INDEX INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Chapter

Harnessing the Genetic Diversity and Metabolic Potential of Extremophilic Microorganisms through the Integration of Metagenomics and Single-Cell Genomics

Deepika Goyal, Shiv Swaroop and Janmejay Pandey

Abstract

Microorganisms thriving under extreme environments have proven to be an invaluable resource for metabolic products and processes. While studies carried out on microbial characterization of extremophilic environments during golden era of microbiology adapted a 'reductionist approach' and focused on isolation, purification and characterization of individual microbial isolates; the recent studies have implemented a holistic approach using both culture-dependent and culture-independent approaches for characterization of total microbial diversity of the extreme environments. Findings from these studies have unmistakably indicated that microbial diversity within extreme environments is much higher than anticipated. Consequently, unraveling the taxonomic and metabolic characteristics of microbial diversity in extreme environments has emerged as an imposing challenge in the field of microbiology and microbial biotechnology. To a great extent, this challenge has been addressed with inception and advancement of next-generation sequencing and computing methods for NGS data analyses. However, further it has been realized that in order to maximize the exploitation of genetic and metabolic diversity of extremophilic microbial diversity, the metagenomic approaches must be combined synergistically with single-cell genomics. A synergistic approach is expected to provide comprehensions into the biology of extremophilic microorganism, including their metabolic potential, molecular mechanisms of adaptations, unique genomic features including codon reassignments etc.

Keywords: extremophilic environments, metabolic diversity, metagenomics, single cell genomics, small molecule secondary metabolites

1. Introduction

There are a number of extreme ecosystems present on Earth that harbor an array of microorganisms with unique genetic diversity and metabolic capabilities [1, 2].

These unique capabilities enable them not only survive but also thrive in extremes of physicochemical parameters [3–6]. The idea that microorganisms might survive in such extreme environments and the term 'extremophile' was first proposed in the 1970s by Robert MacElroy. Conventionally, extremophilic microorganisms have been defined by their ability to grow optimally under environments characterized by extreme temperature, pH, pressure, and salinity etc. [7, 8]. It is argued that survival and growth under extreme environments require stabilization of cellular components and enzymes so that their optimal functionality is maintained. Therefore, extremophilic microorganisms are proposed to be one of the greatest reservoirs of the wide spectrum of exclusive enzymes and metabolites with significant biotechnological applications [9–15]. In addition, the extremophilic microorganisms are now also being regarded to have the pivotal role in maintaining the balance of global biogeochemical cycles [16–18]. With this understanding, there has been a continued increase in the scientific interest in isolation and characterization of extremophilic microorganisms. The same is clearly reflected by the fact that many new extremophilic microorganisms have been isolated and cultured in laboratories all over the world during the past 2–3 decades [19, 20]. Still, much of the physiological and phylogenetic diversity of extremophilic microorganisms remains rather unexplored. Given the ability of extremophilic microorganisms to thrive in the extreme environments; their taxonomic, genetic and metabolic characterization is widely regarded as an indispensable step towards harnessing their true potential. The progress in this line of scientific endeavor has remained hampered due to the vast majority of microbial biodiversity within extremophilic environments comprising of the lineages that are recalcitrant to traditional culturing techniques based isolation and purification approaches [21]. In absence of purified cultures of extremophilic microorganisms, the access to their genetic and metabolic diversity has remained obscure as only until recently, cultivability was the single most important prerequisite for having access to the genetic complement of individual organisms.

This limitation has been circumvented to a great extent with the implementation of culture-independent approach (i.e. metagenomics). The 'state of the art metagenomics technologies,' allow not only to develop a theoretical and mechanistic understanding of the possible role of extremophilic microorganisms in biogeochemical cycles but also assess the genetic & metabolic potentials (e.g. discover novel enzymes and proteins for industrial applications) of the uncultured extremophilic microbial population [22–25]. Having mentioned that, it is also pertinent to remark that even with the implementation of improved cultivation methodologies and metagenomics characterization, the understanding of the 'black box of extremophilic microbial diversity' has improved only marginally over the period of last 2 decades. The optimal exploitation of their potential still remains elusive. This situation could be attributed to the following reasons: (i) despite the ever-improving cultivation methodologies, most of the extremophilic microorganisms are not yet amenable to laboratory culturing which use traditional reductionist culturing approaches; (ii) the microbial biomass densities within extremophilic environments are often too less to yield enough DNA for carrying out effective culture-independent analyses (e.g. metagenomics, metatranscriptomics, and recombinant cloning of a gene of interest); and (iii) inability to annotate novel genetic complements during post-sequencing analyses of metagenomic due to lack of reference sequences in the nucleotide databases [24].

This situation demands continued improvement of technical methodologies towards assessing and harnessing the genetic and metabolic diversity of extremophilic microorganisms from even the minute quantities of retrievable metagenomic DNA. Some of the developments in this aspect have focused on improving the recovery of metagenomic DNA from extremophilic environments [26]. Yet another

most important developments in this aspect has been the development of Single Cell Genome Analyses (SCGA) and its synergistic application with metagenomics [27]. The synergistic application of both of these approaches enables for assembly and annotation of draft genomes of even the uncultivated phyla. Therefore, these approaches could be effectively used to harness the genetic and metabolic potential of the extremophilic environments even without the need for extensive laboratory manipulation [28, 29]. Till date, such studies focusing on extreme environments have revealed substantial genomic information for several candidate extremophilic phyla, encompassing putative acidophiles, halophiles, thermophiles, and piezophiles. These data have also provided substantial insights (including catabolic and anabolic potential, molecular mechanism for adaptations to extreme environments, unique genomic features such as stop codon reassignments, and predictions about cell ultrastructure) into the biology of extremophilic microorganism. It is suggested that if metagenomics and SCGA methodologies are coupled with other "omics" technologies, such as transcriptomics, proteomics and metabolomics (i.e. study and quantification of mRNA transcript levels, proteins and cellular metabolites respectively), it could lead to further development of scientific capabilities for harnessing the genetic and metabolic potential of the extremophilic microbial diversity [30, 31].

2. Extremophilic environments and associated microbial diversity

The physicochemical characteristics of extremophilic environments observed on the planet Earth are quite diverse and they are often studied with regards to temperature, pH, salt concentration, nutrient availability etc. Some of the typical extremophilic environments widely studied include thermophilic environments, psychrophilic environments, halophilic environments, acidophilic environments, subterranean habitats, and hyper-arid environments [2, 8, 32]. Representative niches for each of these environments have been scanned with both cultivation-dependent and cultivation-independent approaches [19, 20, 24, 33]. A brief description of some of the representative extremophilic environments and the associated microbial diversity is presented below.

2.1 Thermophilic environments

Studies pertaining to thermophilic environments initiated in the 1970s and 1980s with the isolation of several novel hyperthermophiles. Subsequent studies led to the discovery of deep-sea hydrothermal vents and consequent addition of isolation of a wide range of hyper thermophilic microorganisms belonging to the 'archaeal' domain of the life [34]. During the 1990s, with the advent of culture-independent characterization of microbial diversity using 16S rRNA gene pool sequencing, the thermophilic environments e.g. hydrothermal vents were analyzed [35–37]. These studies could define the composition and diversity of the microbial communities present within the representative thermophilic environments and characterized the prokaryotic phylotypes amongst diverse thermophilic environments representing the temperature gradients from 60°C to 120°C [35–37]. However, the understanding of the functions associated with microbial diversity and the intra-species, inter-species interaction remained poorly defined.

A few of the culture-independent studies on thermophilic environments, which analyzed the sequence of the entire metagenomic DNA pool rather than just the phylogenetic marker gene, identified dominance of sulfur- recycling genes amongst the dominant phylotypes within the sulfur-rich deep-sea vents [35]. Similarly, the

prevalence of hydrogen oxidation genes was observed in hydrogen-rich deep-sea hyperthermophilic vents [38–40]. A few other studies have identified the critical genetic signatures (e.g. genes for alternative mechanisms of nitrogen utilization) of the microbial communities surviving within the thermophilic environments. Some of the recent culture-independent studies on samples collected from thermophilic environments have indicated for the occurrence of the significantly higher diversity of CRISPR compared to the metagenomes of the mesophilic microbial diversity [41–43].

Even with increasing frequency of reports showing the identification of novel genetic and metabolic mechanisms prevalent in thermophilic environments; the comprehensive understanding about key genetic elements which determine the composition as well the function of the microbial diversity within the thermophilic environments is only poorly understood. It is not yet established how physicochemical factors contribute to shaping up the composition and structure of the microbial diversity of any thermophilic environment. The scenario is expected to improve only through the inclusion of physicochemical information along with full community metagenome data.

2.2 Psychrophilic environments

The psychrophilic environments are characterized by extremely low temperatures. Just like the thermophilic environments, they also represent one of the most thoroughly investigated extreme environments [21, 44]. It is noteworthy that unlike the thermophilic environments, the microbial diversity within psychrophilic environments consists of both eubacteria and archaea [45]. The biodiversity and adaptive strategies of psychrophilic microorganisms have been extensively studied. Results from some of the representative metagenomic studies on the psychrophilic environment have shown microbial community diversity and complexity to be significantly higher than other environments [45, 46]. The most note-worthy studies on psychrophilic environments have been carried out on samples from Antarctic continent, which harbors sub-glacial ice habitat. These studies have reported the dominance of 'chemoautotrophs' that are capable of tapping reduced iron and reduced sulfur compounds as the source of energy [47]. Other studies with psychrophilic environments have recognized the presence of 'chemolithotrophic' bacterial and archaeal communities [45, 47]. These share a close phylogenetic relationship with microorganisms able to use reduced nitrogen, and iron compounds as the source of energy. With regards to the psychrophilic environments, it is generally accepted that 'availability of organic metabolizable carbon' is the single most dominant factor determining the microbial activity, diversity, and dynamics.

2.3 Acidophilic environments

Acidophilic environments have emerged as 'extremophilic environments of choice' for studies on mechanisms and genetic elements determining the survival of life under extreme environments. A number of studies had reported attempts for isolation of microorganisms from acidophilic environments. Culture-independent studies with respect to acidophilic environments were first carreid out with a natural acidophilic biofilm sample [48]. Subsequent studies in this regard were carried out on samples collected from an Acid Mine Drainage located at different parts of the world [49–52]. The data obtained with these samples showed the microbial community structure to have a poor diversity with presence of only chemoautotrophic consortia largely comprising members of genera *Leptospirillum* and *Ferroplasm* [48]. The genetic signatures observed within the Acid Mine Drainage

metagenomes indicated for molecular mechanisms for acidophilic survival through implmentation of unique carbon metabolic pathways for Carbon metabolism, Nitrogen fixation and iron oxidation [53]. The community composition of Acid Mine Drainage samples were found to have significant contrast to the naturally occuring acidophilic biofilms that has *Acidithiobacillus*, *Acidimicrobium* and *Ferrimicrobium* as the dominant genera present within the community [49, 52, 54]. The other noticeably dominant microbial extremophilic taxa in acid mine drainages was Ferroplasm and Thermoplasmatales archaea [55, 56].

2.4 Halophilic environments

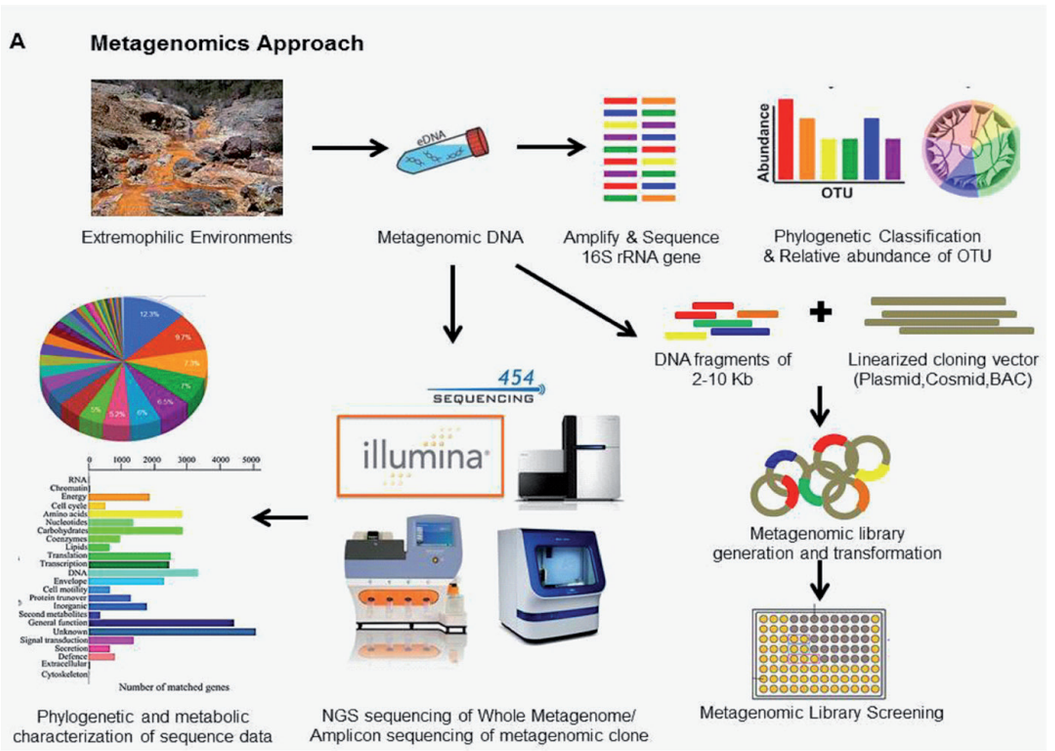
Like other extremophilic environments, the microbial community structure and diversity of the halophilic environments has also been subject of great scientific curiosity. Several culture dependent and culture independent studies have been carried in past 2–3 decades for the assessment of the microbial diversity thriving within the halophilic environments [57–59]. The research findings from some of the most important studies have been thoroughly reviewed. Studies pertaining to halophilic microorganisms have got greatly benefitted with the implementation of cultivation independent approaches for microbial diversity analyses. Metagenomic analyses of the samples collected from multiple hypersaline systems (e.g. Tyrell Lake, Crystallizer Ponds) have indicated presence of high phylotypic diversity with the dominance of halophilic archaeon in particular [60–63]. The whole DNA pool metagenome sequencing of halophilic samples followed by *de novo* assembly and annotation resulted in discovery of a dominant novel uncultivated archaeal class viz., Nanohaloarchaea [60]. This study also revealed occurrence of a unique combination of amino acids which increase the structural flexibility and osmo-resistance of the protein elements. Another characteristic feature of the genetic resources associated with microbial diversity within halophilic environment was discovered in an independent study and it was observed to be the prevalence of Halo-resistance mechanisms orchestrated through synthesis of solutes (such as glycine, betaine, ectoine and trehalose etc.) that are compatible with high salt concentrations [64].

3. Extremophilic microorganisms: invaluable source of novel metabolites

Microorganisms surviving in the extreme environments are being looked up to as they could help treat a wide spectrum of human illnesses, from ovarian cancer, migraine, high blood pressure, ovarian cancer and lung cancer to Alzheimer's disease. This doctrine has emerged out of the understanding that extremophilic environments present very hostile conditions that impose serious threat to survival of any organism exposed to them [8, 65]. However, extremophilic microorganisms which thrive under such hostile environment must be doing it by synthesizing unusual, but potentially very useful, secondary metabolites. Probably, the best studied molecules produced by extremophilic microorganisms are (i) biocatalytic proteins that are often referred as extremozymes; and (ii) secondary metabolites that are not directly required growth of the microorganism, yet they often perform many helpful functions, such as enabling defense mechanisms etc. [66–69],

It is suggested that extremophile enzymes would be more suitable and stable for use in industrial biotechnology applications than those obtained from mesophilic microbial species [9, 21, 70]. Also, the unusual secondary metabolites isolated from extremophilic microorganisms are steadily being characterized as drug molecules with unique potential and applications. One of the recently published studies

reported characterization of a secondary metabolite (viz., dihydrogranaticin) from a thermophilic fungus exhibits wide spectrum antibiotic functions. Similarly, secondary metabolites isolated from a psychrophilic bacterium from the Arctic glaciers have been reported to inhibit the growth of human colon cancer cells. Another secondary metabolite (psychrophilin D) isolated from a psychrophilic microorganism, exhibits inhibitory activity against mouse leukemia cell line [71].



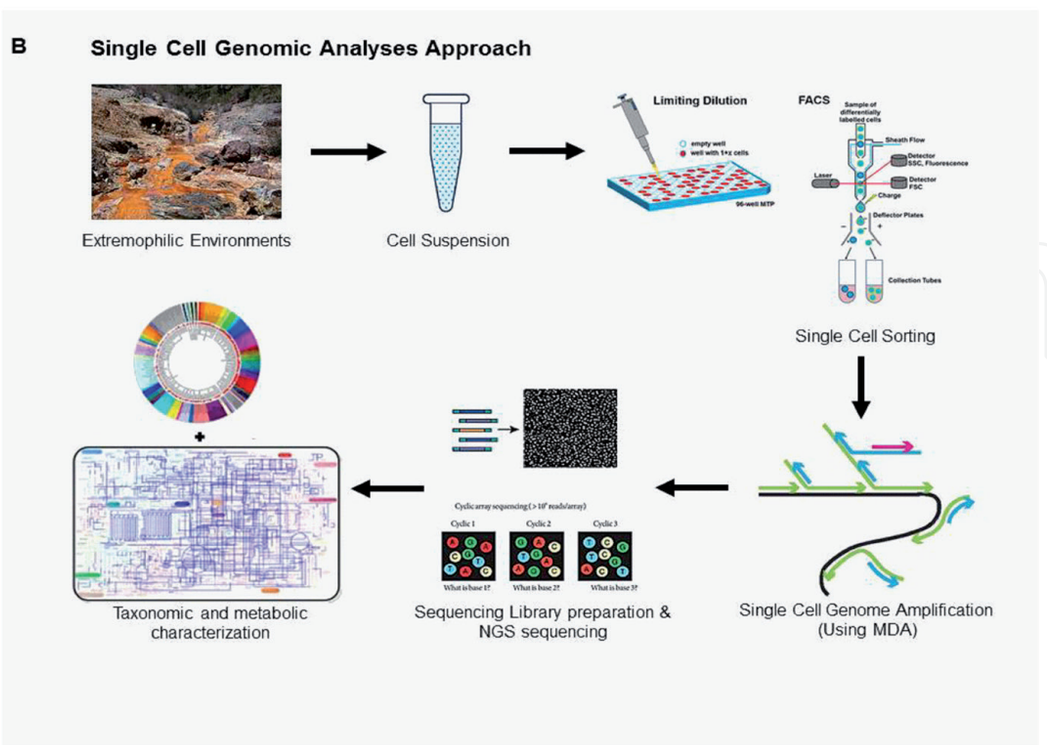


Figure 1.
Schematic representation of the workflow used for culture independent approaches for characterization of microbial diversity viz., metagenomics (A) and single cell genomics (B).

In accordance to the other extremophilic environments, the secondary metabolites produced by organisms that thrive at acidophilic environments, are a valuable source of novel metabolites. According to a recent survey, more than 20 previously unknown natural products have been isolated from acidophilic microbial diversity. Another valuable type of metabolites that is being proposed to have significant technological application is the 'natural inhibitors' of therapeutic target proteins. A representative example of such natural inhibitors was reported as berkeleyamide A, a secondary metabolite isolated from acidophilic strains *Penicillium rubrum* species [72]. This inhibits the proteases caspase-1 and matrix metalloproteinase 3 (MMP3), both of which are implicated in malignancy of some of the cancer types [72]. Another molecule (i.e. Berkelic acid,) isolated from an extremophilic microorganism, has a very unusual tetracyclic structure and it also inhibits both caspase-1 and MMP-3 [73]. Consequently, it exhibits selective inhibitory activity against an ovarian cancer cell line which has implication of abovementioned genes in cancer progress. Unfortunately, there is significantly less information available relevant to the secondary metabolites produced by extremophilic microorganism thriving at high pH and high salt concentrations. It has been often suggested that the enzymes from these microorganisms would be quite useful a biological detergents.

Considering the well-established potentials of the metabolites of the extremophilic origin, there is a need to develop fundamental understanding with respect to their physiological role in the growth and survival of extremophilic microorganism as well as their adaptation to the hostile environment. Many of the metabolites remain 'cryptic' during the cultivation of the extremophiles under the *in vitro* conditions since recreating the physicochemical conditions observed in the extreme environments within the laboratory is technically challenging, complicated and expensive [74]. Metagenomic Analyses and Single Cell Genomic Analyses., which enable the assessment of genetic and metabolic diversity without the need of cultivating the microorganisms, have helped to circumvent the limitations caused by the cryptic nature of secondary metabolic genes [75–77]. **Figure 1** presents a schematic representation of the workflow used for the metagenomics (**Figure 1A**) and single cell genomics (**Figure 1B**). As of now, a number of studies have already been carried out with metagenomes and single cell genomes from the extreme environments for studying extremozymes and cryptic metabolites.

4. Cultivation-independent approaches: tapping extremophilic metabolites

Cultivation-independent approaches are based on direct isolation of whole metagenomic DNA/environmental DNA. Subsequent downstream treatments of metagenomic DNA are broadly classified into 2 categories, i.e. (i) Metagenomic library generation and its functional screening; and (ii) Direct sequencing of the whole metagenomic DNA content [78–82]. The same technological framework is applicable to metagenomic analyses of samples collected from extremophilic environments. However, the complex nature of extremophilic matrix presents certain unique technical challenges with respect to isolation of metagenomic DNA. The methodologies successfully implemented to mesophilic sites for metagenomic DNA isolation often tend to be non-sufficient for isolation of high quality and high quantity metagenomic DNA from extremophilic samples. Even with non- optimal metagenomic DNA isolation procedures, the cultivation independent approach has enabled identification and characterization of several valuable extremozymes and extremophilic metabolites [20, 75].

4.1 Functional screening of extremophilic metagenomic libraries

Using rather simple and direct readout assays (e.g. appearance of either a halo or a color), the functional screening of the metagenomic libraries have been carried out for a number of extremophilic environments. For example, in a recent study, the Antarctic desert soil metagenomic library was screened for psychrophilic esterases using agar plates based screening approach [83, 84]. The positive clone with desired activity was selected on the basis of formation of a clear halo around the metagenomic clone. The halo formation indicated tributyrin hydrolysis; and resulted in identification and characterization of a novel cold-active psychrophilic esterase. Noticeably, it was found to be only distantly related to previously reported lipases.

While the abovementioned example for isolation and characterization of a novel psychrophilic esterase clearly highlights the value of 'functional screening' of the metagenomic libraries of the extremophilic origin, yet, it is also well acknowledged that many of the extremozymes and extremophilic metabolites are not expressed from the clones of the metagenomics library and therefore, they are not amenable to identification by library screening assays [85]. Several attempts have been made to evade the apparent limitations associated with library screening approach to metagenomics. Screening and development of alternative host for functional metagenomics screening [86] and development and application of 'Reporter Vectors' has been one of the most distinct attempts in this regard. One such Reporter Vector for metagenomic library screen was developed to have product-induced gene-expression of a reporter gene. It is done by coupling of the reporter gene to a product-sensitive transcription factor; Thus upon formation of a desired product, the transcription of the reporter gene is initiated, which could be subsequently monitored through standard reporter gene assay (e.g. fluorescence) [87]. Complementation assays have also been used as a strategy for functional screening and isolation of novel biocatalysts from metagenomic libraries [88].

4.2 Homology search based screening of the extremophilic metagenome

The alternative approach is based on direct sequencing and homology search based screening of the gene(s), protein(s) and secondary metabolites of interest. This approach generally involves DNA amplification (PCR) step as a necessary step of sequencing procedures. Even the next generation sequencing platforms (i.e. Pyrosequencing, Sequencing by Synthesis, and Ion Sequencing) involve the step for PCR amplification of the metagenomic Pool DNA [20, 53, 79]. An earlier approach for homology search based screening of metagenomic DNA used 'heterologous probe –hybridization'; however, that approach has given way to NGS approaches. With advancement in the field of genome informatics, and metagenome informatics, it is now easily feasible to detect conserved enzymatic sequence motifs in metagenomic DNA sequences including the metagenomes of extremophilic environments [24]. The most noticeable advantage of this approach over the functional screening of the metagenomic library is the inherent high throughput and flexibility to extend the scope of screening using *in silico* homology search and screening [82]. The sequence homology search based screening approach has been used with primary metagenomic sequence data as well as with the pre- existing metagenomic datasets. The homology search based screening of metagenomic sequences gets limited only in terms of the 'existing sequence databases'. In other words any novel sequence(s) with significant divergence from the previously characterized/reference sequences or not having homology gets identified as "sequence with unknown function".

5. Diversity of products screened from extremophilic metagenomes

In addition to the screening for extremozymes, the extremophilic metagenomes have also been subjected to screening for identification of small molecules and secondary metabolites which could have potential pharmaceutical applications as antibiotics, antifungal, anti-inflammatory, anti-tuberculosis, anti-cancer and immunosuppressive etc. Both functional screening of metagenome library and homology search based screening approaches have been successfully used for this purpose [22, 24, 36, 40, 62]. In comparison the extremozymes, there are relatively fewer high throughput assays available for detecting metagenomic clones that can produce small molecules and/or secondary metabolites. Thus functional screening has not been used very often for metagenomic libraries with the objective of identifying novel secondary metabolites. Therefore, there is a constant need for development of innovative functional screening methods for identification of small molecules and secondary metabolites of extremophilic origin. A few discreet studies have shown examples of novel screening approaches. In one such example a novel screening method was developed with use of indicator "Chrome Azurol-S" (CAS), which undergoes chromogenic change from orange to blue in the presence of iron. This screening method was subsequently used for identification of metagenomic clones (as well as cultivable isolates) encoding siderophores (the iron chelators). In these studies gene clusters encoding novel siderophores were identified from novel uncultivable strains.

In comparison to the functional screening, the homology search screening has been more frequently used for screening of metagenomes for the extremophilic metabolites. For the homology search screening, the metagenomic sequence data is probed to identify gene(s)/gene cluster(s) containing conserved domains or sequence that are predicted to be associated with biosynthesis of a secondary metabolites of interest. The most prominent secondary metabolites identified through homology search screening of extremophilic metagenome sequences has successfully led to the identified and characterization of: (i) glycopeptide antibiotics; (ii) cyanobactins cytotoxins; (iii) type –II polyketides antibiotics and anticancer molecules; and (iv) Trans-acyltransferse (trans-AT) polyketides [89–93]. While each of these classes of small molecules/secondary metabolites have been previously identified and characterized from the cultivable microbial diversity (more specifically actinobacterial diversity), however, with use of homology search screening of the extremophilic metagenomes, a number of novel representatives of the chemical scaffolds have been successfully identified and characterized.

5.1 Identification and characterization of glycopeptide

Glycopeptides are small molecule secondary metabolites produced by diverse organisms ranging from Proteobacteria to higher plants with Actinobacteria being the single most important source. These small molecules exhibit antibacterial activity against some of the most resistant Gram-positive pathogenic bacteria [94]. Consequently, glycopeptide are molecules of great scientific and industrial significance. The assortment of glycopeptides isolated and characterized from cultivable bacterial diversity is only very limited; therefore, several studies have been carried out with the objective of widening the catalogue of the glycopeptides through exploitation of culture- independent approaches. In one such study, soil metagenome was used as the DNA template and used for amplification a gene corresponding to OxyC, an oxidation coupling enzyme which is highly conserved and catalyzes a vital intermediate reaction during synthesis of many glycopeptides. This approach

resulted in identification of multiple predicted glycopeptide-encoding gene clusters from the soil metagenomic libraries. In the follow up studies, the novel glycopeptide synthesis related gene(s) and gene cluster(s) identified from the metagenomic DNA were transformed and heterologously expressed in a *Streptomyces* expression host [95, 96]. Such technical intervention resulted in several new derivative glycopeptide antibiotics (with methyl, sulfur and sugar substitution) were generated being synthesized.

5.2 Identification and characterization of cyanobactins

Cyanobactins are a family of small, cyclic peptides produced by cyanobacteria and consist of N-to-C macro-cylization of a 6–20 amino acid chain. They are generally assembled through the cleavage and modification of short precursor proteins. Many of these peptides show antimalarial or antitumor activity [97]. It is speculated that close to 30% of all cyanobacterial strains contain genes corresponding to synthesis of cyanobactins [98, 99]. It is also speculated that, bacterial diversity other than cyanobacteria may also have harbor the gene(s) and gene cluster(s) for synthesis of cyanobactins [98]. However, access to such cyanobactins gene cluster(s) is limited due to the non- cultivability of the vast microbial majority. A few metagenomic studies have reported cloning and heterologous expression of biosynthetic gene clusters for the cyanobactins. In one such example study, the gene cluster for 'patellamide' was cloned and heterologously expressed from metagenomic libraries of uncultured cyanobacterial symbionts associated with marine sponge [100, 101]. In other studies, the structural diversity of diversity was enriched with subtle changes in the gene encoding for precursor peptide and employed it in combination with multiple strategies e.g. (i) orthogonal loading of unnatural amino acids; (ii) mutagenesis of precursor peptide; (iii) generation of a library of hybrid cyanobactins [90].

5.3 Identification and characterization of Type II polyketides

Type II polyketides are a group of small molecules with aromatic rings and contain alternating carbonyl and methylene groups (-CO-CH2-).Many of the Type II polyketides (e.g. tetracycline and doxorubicin) are well documented for antimicrobial and ant cancerous activities [90]. Gene clusters involved in synthesis of these small molecules are rather divergent and exhibit low levels of DNA sequence homology, yet each of them contain at least a 'polyketide synthetase', encoded by three highly conserved genes, i.e. 2 genes for ketosynthases (KSs) and one gene for a acyl carrier protein. These 3 genes are referred as 'minimal PKS synthesis gene cluster'. Studies carried out with metagenomes in general and extremophilic metagenome in particular have shown a rich diversity of novel 'minimal PKS synthesis gene cluster' [102]. In subsequent studies, gene clusters with minimal PKS synthesis genes were identified in soil metagenomes [103]. The transformation and heterologous expression in different strains belonging to genus *Streptomyces* and lead to synthesis and identification of several new polyketide metabolites with previously unknown and rare carbon skeletons [93].

5.4 Identification and characterization of trans-acyltransferse polyketides

This class of small molecule polyketides is biosynthesized through activity of a freestanding acyltransferases and constitutes one of the most important groups of pharmacologically interesting polyketides. Considering their pharmaceutical

implication and rather limited catalogue from the cultivated microorganism, the metagenomic route of discovery has been adapted. In this approach, the metagenomes from various environments including the extremophilic environments have been probed for presence of a conserved trans-Acyltransferase specific DNA sequences [104]. Using this approach, a single amplicons have been identified which would produce the novel Trans-acyltransferse polyketides. Unlike the Type II polyketides, studies with heterologous transformation and expression of the Trans-acyltransferse polyketides are relative obscure, yet, a few discreet studies have shown genesis of hybrid Trans-acyltransferse polyketides. In one such examples study, a gene encoding for O-methyltransferase from the pederin gene clusters was transformed in a mycalamide-A producing strains. Upon expression the O-methyltransferase catalyzed a site-specifically methylation which resulted in production of a hybrid compound 18-O-methylmycalamide which showed significantly improved antitumor activities [105].

6. Single Cell Genome Analyses of the extremophilic microbial diversity

A recent concept in the field of the culture- independent approaches for identification and characterization of microbial genetic and metabolic diversity is "Single Cell Genome Analyses (SCGA)" [106]. This approach accesses genomes from one cell at a time. Therefore, this approach allows the analyses of the microbial genetic and metabolic diversity at the level of the most fundamental biological unit. The central technical aspect of this approach involves separation of individual cells from a complex mixture of environmental matrix using a cell sorting methods such as fluorescence-activated cell sorting (FACS). Cell separation is followed by cell lysis and recovery of the femtogram levels of DNA from a Single cell. The recovered single cell DNA is amplified using multiple displacement amplification (MDA) and amplification of single cell genomic DNA, such that the quantities of DNA increases to 100s of nano grams – 10s of micro grams (a 10^3-10^6 fold increase) [107, 108]. The single amplified genomes (SAGs) are subsequently used for screening by PCR amplification and NGS sequencing. The taxonomic identity of the concerned extremophilic microbial cell is ascertained with 16S rRNA gene sequencing, whereas subsequent shotgun or NGS sequencing, assembly and annotation is carried out with single amplified genomes of interest identified through preliminary phylotype characterization [106–109].

Despite its tremendous scientific capabilities, the SCGA is yet to make outreaching impact on microbial genomics in general and extremophilic microbiology in particular. The technical procedure used for SCGA faces many challenges that are not yet completely addressed. The most critical challenges include: (i) technical limitation in precise and reproducible separation of single bacterial cells with available methodologies; (ii) low amounts of starting DNA recoverable from single bacterial cell; (iii) requirement of a high degree of amplification; (iv) possibility of cross contamination; (v) introduction of chimeric artifacts and biases in genomic coverage during single genome amplification; and (v) poor post-sequencing quality control, data analyses and sequence assembly [110]. Due to these limitations, the resulting composite assemblies from SCGA can often represent incomplete or inaccurately characterized genomes for a given strain or species [107, 111]. However, several technological updates are being made to circumvent these limitations of the SCGA, which would soon enable highly accurate data generation and its physiological interpretation based on the absence as well as presence of genes and pathways [108].

6.1 Combining single cell genomics and metagenomics

Despite the individual technical limitations of both the approaches, it is regarded that the combined synergistic application of single-cell genomics and metagenomics can offer great opportunities, since the advantages offered by each of these techniques are complementary in nature. To highlight, it is underlined that one hand metagenomics is not known to suffer from any problem associated with chimera generation during strand displacement and genome amplification or separation of individual microbial cells from a complex heterogeneous mixture. On the other hand single-cell genomics overcomes the limitation of metagenomics by leading to a direct and unambiguous association of phylogeny and metabolic functions. Information obtained from SCGA can be effectively used to assign taxonomy to individual metagenome contigs with high accuracy [107, 112–114]. SCGA may also be used for retrieving complete genomes of candidate taxon from the metagenomic data. Similarly, the metagenomic reads can be mapped back to scaffolds for closely related SAG and therefore significantly improve their annotation.

The synergistic application of metagenomics and single cell genomics is regarded to have a unified and far reaching implication in harnessing the biotechnological potential of the extremophilic microbial diversity. As a matter of fact, extremophilic environments have already featured prominently in studies implementing both metagenomics and single-cell genomics studies. The most note-worthy set of studies were performed on acidophilic biofilms of Richmond Mine, California, USA, wherein initial metagenomic studies led to the identification of dominant microbial communities, while subsequent single cell genomics studies could identify even novel, low-abundance archaeal lineages that were later named as archaeal richmond mine acidophilic nanoorganisms (ARMAN) [115, 116]. The nanoorganisms have since been the matter of investigation throughout the world. In the same vein, the synergistic application of metagenomics and single cell genomics has led to identification of three previously uncultivated and uncharacterized halophilic phylotypes that represent the candidate phylum Nanohaloarchaeota from studies carried out on samples collected from halophilic Pola salterns, Alicante, Spain. Apart from the taxonomic and phylogenetic characterization of novel extremophiles, the synergistic application of metagenomics and single cell genomics also led to identification of their critical metabolic functions e.g. presence of rhodopsin and genes for a photoheterotrophic lifestyle.

7. Conclusion

The advent of 'culture independent' approaches for characterization of microbial diversity and their dynamics has been the single most significant development in the field of microbiology in general and microbial ecology, microbial biotechnology in particular. It has also greatly accelerated the research pertaining to extremophilic microbial diversity. With use of present 'state of the art' technologies viz., metagenomics and single cell genomics, a number of vital discoveries have been made that would not have been possible without the use of these technologies. Thus, it could be proposed that although, considerable progress has been made, yet there is a lot of scope for better application of metagenomics and single-cell genomics approaches to not only access genomes for discovering novel taxonomic lineages of extremophilic microorganisms but also harness their genetic and metabolic potential towards discovery of novel high value metabolites.

As an eventual future objective, the application of metagenomics and single cell genomics would be expected to complement the traditional cultivation approaches

and follow suit with 'genomics guided –microbial culturing' towards' establishment of knowledge of the metabolic interactions circuits within mesophilic environments and more specifically within the extremophilic environments. Exploration of cultivation- independent approaches promise an exciting future for assessment and exploitation of extremophilic microbial diversity.

Conflict of interest

Authors declare 'no conflict of interest' with respect to publication of this book chapter.

Author details

Deepika Goyal[1†], Shiv Swaroop[2†] and Janmejay Pandey[1*]

1 Department of Biotechnology, School of Life Sciences, C.U. Rajasthan, India

2 Department of Biochemistry, School of Life Sciences, C.U. Rajasthan, India

*Address all correspondence to: janmejay@curaj.ac.in

† Both the authors contributed equally.

References

[1] Hu X. Ciliates in extreme environments. The Journal of eukaryotic microbiology. 2014;**61**:410-418

[2] Harrison JP, Gheeraert N, Tsigelnitskiy D, Cockell CS. The limits for life under multiple extremes. Trends in microbiology. 2013;**21**:204-212

[3] Burtscher M, Gatterer H, Burtscher J, Mairbaurl H. Extreme Terrestrial Environments: Life in Thermal Stress and Hypoxia. A Narrative Review. Frontiers in physiology. 2018;**9**:572

[4] Bang C, Dagan T, Deines P, Dubilier N, Duschl WJ, Fraune S, et al. Metaorganisms in extreme environments: do microbes play a role in organismal adaptation? Zoology (Jena, Germany). 2018;**127**:1-19

[5] Chevin LM, Hoffmann AA. Evolution of phenotypic plasticity in extreme environments. Philosophical transactions of the Royal Society of London. Series B, Biological sciences. 2017;**372**

[6] Morozkina EV, Slutskaia ES, Fedorova TV, Tugai TI, Golubeva LI, Koroleva OV. Extremophilic microorganisms: biochemical adaptation and biotechnological application (review). Prikladnaia biokhimiia i mikrobiologiia. 2010;**46**:5-20

[7] Stetter KO. Extremophiles and their adaptation to hot environments. FEBS letters. 1999;**452**:22-25

[8] Rothschild LJ, Mancinelli RL. Life in extreme environments. Nature. 2001;**409**:1092-1101

[9] Littlechild JA. Enzymes from Extreme Environments and Their Industrial Applications. Frontiers in bioengineering and biotechnology. 2015;**3**:161

[10] Antranikian G, Vorgias CE, Bertoldo C. Extreme environments as a resource for microorganisms and novel biocatalysts. Advances in biochemical engineering/biotechnology. 2005;**96**:219-262

[11] van den Burg B. Extremophiles as a source for novel enzymes. Current opinion in microbiology. 2003;**6**:213-218

[12] Elleuche S, Schroder C, Sahm K, Antranikian G. Extremozymes--biocatalysts with unique properties from extremophilic microorganisms. Current opinion in biotechnology. 2014;**29**:116-123

[13] Cavicchioli R, Siddiqui KS, Andrews D, Sowers KR. Low-temperature extremophiles and their applications. Current opinion in biotechnology. 2002;**13**:253-261

[14] Ferrer M, Golyshina O, Beloqui A, Golyshin PN. Mining enzymes from extreme environments. Current opinion in microbiology. 2007;**10**:207-214

[15] Schiraldi C, De Rosa M. The production of biocatalysts and biomolecules from extremophiles. Trends in biotechnology. 2002;**20**:515-521

[16] Sorokin DY, Berben T, Melton ED, Overmars L, Vavourakis CD, Muyzer G. Microbial diversity and biogeochemical cycling in soda lakes. Extremophiles: life under extreme conditions. 2014;**18**:791-809

[17] Offre P, Spang A, Schleper C. Archaea in biogeochemical cycles. Annual review of microbiology. 2013;**67**:437-457

[18] Bao P, Li GX, Sun GX, Xu YY, Meharg AA, Zhu YG. The role of sulfate-reducing prokaryotes in the coupling of element biogeochemical cycling. The Science of the total environment. 2018;**613-614**:398-408

[19] Rinker KD, Han CJ, Kelly RM. Continuous culture as a tool for investigating the growth physiology of heterotrophic hyperthermophiles and extreme thermoacidophiles. Journal of applied microbiology. 1998;**85**(Suppl 1): 118s-127s

[20] Vester JK, Glaring MA, Stougaard P. Improved cultivation and metagenomics as new tools for bioprospecting in cold environments. Extremophiles: life under extreme conditions. 2015;**19**:17-29

[21] Feller G. Life at low temperatures: is disorder the driving force? Extremophiles: life under extreme conditions. 2007;**11**:211-216

[22] Saxena R, Dhakan DB, Mittal P, Waiker P, Chowdhury A, Ghatak A, et al. Metagenomic Analysis of Hot Springs in Central India Reveals Hydrocarbon Degrading Thermophiles and Pathways Essential for Survival in Extreme Environments. Frontiers in microbiology. 2016;**7**:2123

[23] Vavourakis CD, Ghai R, Rodriguez-Valera F, Sorokin DY, Tringe SG, Hugenholtz P, et al. Metagenomic Insights into the Uncultured Diversity and Physiology of Microbes in Four Hypersaline Soda Lake Brines. Frontiers in microbiology. 2016;**7**:211

[24] Cowan DA, Ramond JB, Makhalanyane TP, De Maayer P. Metagenomics of extreme environments. Current opinion in microbiology. 2015;**25**:97-102

[25] Poli A, Finore I, Romano I, Gioiello A, Lama L, Nicolaus B. Microbial Diversity in Extreme Marine Habitats and Their Biomolecules. Microbiome. 2017;**5**

[26] Verma D, Satyanarayana T. An improved protocol for DNA extraction from alkaline soil and sediment samples for constructing metagenomic libraries. Applied biochemistry and biotechnology. 2011;**165**:454-464

[27] Chen Z, Chen L, Zhang W. Tools for Genomic and Transcriptomic Analysis of Microbes at Single-Cell Level. Frontiers in microbiology. 2017;**8**:1831

[28] Lutz S, Anesio AM, Field K, Benning LG. Integrated 'Omics', Targeted Metabolite and Single-cell Analyses of Arctic Snow Algae Functionality and Adaptability. Frontiers in microbiology. 2015;**6**:1323

[29] Zhang W, Li F, Nie L. Integrating multiple 'omics' analysis for microbial biology: application and methodologies. Microbiology (Reading, England). 2010;**156**:287-301

[30] Y. Seeleuthner, S. Mondy, Single-cell genomics of multiple uncultured stramenopiles reveals underestimated functional diversity across oceans, 9 (2018) 310.

[31] Kodzius R, Gojobori T. Single-cell technologies in environmental omics. Gene. 2016;**576**:701-707

[32] A. Poli, I. Finore, I. Romano, A. Gioiello, L. Lama, B. Nicolaus, Microbial Diversity in Extreme Marine Habitats and Their Biomolecules, Microorganisms, 5 (2017).

[33] C.D. Vavourakis, A.S. Andrei, M. Mehrshad, R. Ghai, D.Y. Sorokin, G. Muyzer, A metagenomics roadmap to the uncultured genome diversity in hypersaline soda lake sediments, 6 (2018) 168.

[34] Stetter KO. Hyperthermophilic procaryotes. FEMS microbiology reviews. 1996;**18**:149-158

[35] Xie W, Wang F, Guo L, Chen Z, Sievert SM, Meng J, et al. Comparative metagenomics of microbial communities inhabiting deep-sea hydrothermal vent chimneys with contrasting chemistries. The ISME journal. 2011;**5**:414-426

[36] Anantharaman K, Breier JA, Dick GJ. Metagenomic resolution

of microbial functions in deep-sea hydrothermal plumes across the Eastern Lau Spreading Center. The ISME journal. 2016;**10**:225-239

[37] Urich T, Lanzen A, Stokke R, Pedersen RB, Bayer C, Thorseth IH, et al. Microbial community structure and functioning in marine sediments associated with diffuse hydrothermal venting assessed by integrated meta-omics. Environmental microbiology. 2014;**16**:2699-2710

[38] Petersen JM, Zielinski FU, Pape T, Seifert R, Moraru C, Amann R, et al. Hydrogen is an energy source for hydrothermal vent symbioses. Nature. 2011;**476**:176-180

[39] Anantharaman K, Breier JA, Sheik CS, Dick GJ. Evidence for hydrogen oxidation and metabolic plasticity in widespread deep-sea sulfur-oxidizing bacteria. Proceedings of the National Academy of Sciences of the United States of America. 2013;**110**:330-335

[40] Brazelton WJ, Nelson B, Schrenk MO. Metagenomic evidence for h(2) oxidation and h(2) production by serpentinite-hosted subsurface microbial communities. Frontiers in microbiology. 2012;**2**:268

[41] Romano C, D'Imperio S, Woyke T, Mavromatis K, Lasken R, Shock EL, et al. Comparative genomic analysis of phylogenetically closely related Hydrogenobaculum sp. isolates from Yellowstone National Park. Applied and environmental microbiology. 2013;**79**:2932-2943

[42] Sorokin VA, Gelfand MS, Artamonova II. Evolutionary dynamics of clustered irregularly interspaced short palindromic repeat systems in the ocean metagenome. Applied and environmental microbiology. 2010;**76**:2136-2144

[43] Anderson RE, Brazelton WJ, Baross JA. Using CRISPRs as a metagenomic tool to identify microbial hosts of a diffuse flow hydrothermal vent viral assemblage. FEMS microbiology ecology. 2011;**77**:120-133

[44] Price PB. A habitat for psychrophiles in deep Antarctic ice. Proceedings of the National Academy of Sciences of the United States of America. 2000;**97**:1247-1251

[45] Cowan DA, Russell NJ, Mamais A, Sheppard DM. Antarctic Dry Valley mineral soils contain unexpectedly high levels of microbial biomass. Extremophiles: life under extreme conditions. 2002;**6**:431-436

[46] Simon C, Wiezer A, Strittmatter AW, Daniel R. Phylogenetic diversity and metabolic potential revealed in a glacier ice metagenome. Applied and environmental microbiology. 2009;**75**:7519-7526

[47] Mikucki JA, Priscu JC. Bacterial diversity associated with Blood Falls, a subglacial outflow from the Taylor Glacier, Antarctica. Applied and environmental microbiology. 2007;**73**:4029-4039

[48] Bond PL, Druschel GK, Banfield JF. Comparison of acid mine drainage microbial communities in physically and geochemically distinct ecosystems. Applied and environmental microbiology. 2000;**66**:4962-4971

[49] Baker BJ, Banfield JF. Microbial communities in acid mine drainage. FEMS microbiology ecology. 2003;**44**: 139-152

[50] Bruneel O, Duran R, Casiot C, Elbaz-Poulichet F, Personne JC. Diversity of microorganisms in Fe-As-rich acid mine drainage waters of Carnoules, France. Applied and environmental microbiology. 2006;**72**:551-556

[51] He Z, Xiao S, Xie X, Zhong H, Hu Y, Li Q, et al. Molecular diversity

of microbial community in acid mine drainages of Yunfu sulfide mine. Extremophiles: life under extreme conditions. 2007;**11**:305-314

[52] S. Xiao, X. Xie, J. Liu, Microbial communities in acid water environments of two mines, China, Environmental pollution (Barking, Essex: 1987), 157 (2009) 1045-1050.

[53] Dick GJ, Andersson AF, Baker BJ, Simmons SL, Thomas BC, Yelton AP, et al. Community-wide analysis of microbial genome sequence signatures. Genome biology. 2009;**10**:R85

[54] Huang LN, Kuang JL, Shu WS. Microbial Ecology and Evolution in the Acid Mine Drainage Model System. Trends in microbiology. 2016;**24**:581-593

[55] Bruneel O, Pascault N, Egal M, Bancon-Montigny C, Goni-Urriza MS, Elbaz-Poulichet F, et al. Archaeal diversity in a Fe-As rich acid mine drainage at Carnoules (France). Extremophiles: life under extreme conditions. 2008;**12**:563-571

[56] Golyshina OV, Timmis KN. Ferroplasma and relatives, recently discovered cell wall-lacking archaea making a living in extremely acid, heavy metal-rich environments. Environmental microbiology. 2005;**7**:1277-1288

[57] Kamekura M. Diversity of extremely halophilic bacteria. Extremophiles: life under extreme conditions. 1998;**2**:289-295

[58] N.M. Mesbah, J. Wiegel, Life at extreme limits: the anaerobic halophilic alkalithermophiles, Annals of the New York Academy of Sciences, 1125 (2008) 44-57.

[59] Mesbah NM, Wiegel J. Life under multiple extreme conditions: diversity and physiology of the halophilic alkalithermophiles. Applied and environmental microbiology. 2012;**78**:4074-4082

[60] Narasingarao P, Podell S, Ugalde JA, Brochier-Armanet C, Emerson JB, Brocks JJ, et al. De novo metagenomic assembly reveals abundant novel major lineage of Archaea in hypersaline microbial communities. The ISME journal. 2012;**6**:81-93

[61] Anton J, Rossello-Mora R, Rodriguez-Valera F, Amann R. Extremely halophilic bacteria in crystallizer ponds from solar salterns. Applied and environmental microbiology. 2000;**66**:3052-3057

[62] Plominsky AM, Delherbe N, Ugalde JA, Allen EE, Blanchet M, Ikeda P, et al. Metagenome sequencing of the microbial community of a solar saltern crystallizer pond at cahuil lagoon, chile. Genome announcements. 2014;**2**

[63] M. Kambourova, I. Tomova, I. Boyadzhieva, N. Radchenkova, E. Vasileva-Tonkova, Unusually High Archaeal Diversity in a Crystallizer Pond, Pomorie Salterns, Bulgaria, Revealed by Phylogenetic Analysis, Archaea (Vancouver, B.C.), 2016 (2016) 7459679.

[64] Fernandez AB, Ghai R, Martin-Cuadrado AB, Sanchez-Porro C, Rodriguez-Valera F, Ventosa A. Prokaryotic taxonomic and metabolic diversity of an intermediate salinity hypersaline habitat assessed by metagenomics. FEMS microbiology ecology. 2014;**88**:623-635

[65] Pikuta EV, Hoover RB, Tang J. Microbial extremophiles at the limits of life. Critical reviews in microbiology. 2007;**33**:183-209

[66] Dumorne K, Cordova DC, Astorga-Elo M, Renganathan P. Extremozymes: A Potential Source for Industrial Applications. Journal of microbiology and biotechnology. 2017;**27**:649-659

[67] Demirjian DC, Moris-Varas F, Cassidy CS. Enzymes from extremophiles. Current opinion in chemical biology. 2001;**5**:144-151

[68] Raddadi N, Cherif A, Daffonchio D, Neifar M, Fava F. Biotechnological applications of extremophiles, extremozymes and extremolytes. Applied microbiology and biotechnology. 2015;**99**:7907-7913

[69] Lentzen G, Schwarz T. Extremolytes: Natural compounds from extremophiles for versatile applications. Applied microbiology and biotechnology. 2006;**72**:623-634

[70] Arakawa T, Yamaguchi R, Tokunaga H, Tokunaga M. Unique Features of Halophilic Proteins. Current protein & peptide science. 2017;**18**:65-71

[71] Dalsgaard PW, Larsen TO, Christophersen C. Bioactive Cyclic Peptides from the Psychrotolerant Fungus Penicillium algidum. The Journal of Antibiotics. 2005;**58**:141

[72] Stierle AA, Stierle DB, Patacini B. The Berkeleyamides, Amides from the Acid Lake Fungus Penicillum rubrum. Journal of Natural Products. 2008;**71**:856-860

[73] Stierle AA, Stierle DB, Kelly K. Berkelic Acid, A Novel Spiroketal with Selective Anticancer Activity from an Acid Mine Waste Fungal Extremophile. The Journal of Organic Chemistry. 2006;**71**:5357-5360

[74] Jančič S, Frisvad JC, Kocev D, Gostinčar C, Džeroski S, Gunde-Cimerman N. Production of Secondary Metabolites in Extreme Environments: Food- and Airborne Wallemia spp. Produce Toxic Metabolites at Hypersaline Conditions. PLOS ONE. 2016;**11**:e0169116

[75] Mende DR, Aylward FO, Eppley JM, Nielsen TN, DeLong EF. Improved Environmental Genomes via Integration of Metagenomic and Single-Cell Assemblies. Frontiers in microbiology. 2016;**7**

[76] Gomariz M, Martínez-García M, Santos F, Rodriguez F, Capella-Gutiérrez S, Gabaldón T, et al. From community approaches to single-cell genomics: the discovery of ubiquitous hyperhalophilic Bacteroidetes generalists. The ISME journal. 2014;**9**:16

[77] Rashid M, Stingl U. Contemporary molecular tools in microbial ecology and their application to advancing biotechnology. Biotechnology advances. 2015;**33**:1755-1773

[78] Raes J, Foerstner KU, Bork P. Get the most out of your metagenome: computational analysis of environmental sequence data. Current opinion in microbiology. 2007;**10**:490-498

[79] Simon C, Daniel R. Metagenomic analyses: past and future trends. Applied and environmental microbiology. 2011;**77**:1153-1161

[80] Ferrer M, Beloqui A, Timmis KN, Golyshin PN. Metagenomics for mining new genetic resources of microbial communities. Journal of molecular microbiology and biotechnology. 2009;**16**:109-123

[81] Moore-Connors JM, Dunn KA, Bielawski JP, Van Limbergen J. Novel Strategies for Applied Metagenomics. Inflammatory bowel diseases. 2016;**22**:709-718

[82] Sudarikov K, Tyakht A, Alexeev D. Methods for The Metagenomic Data Visualization and Analysis. Current issues in molecular biology. 2017;**24**:37-58

[83] Mirete S, Morgante V, Gonzalez-Pastor JE. Functional metagenomics of extreme environments. Current opinion in biotechnology. 2016;**38**:143-149

[84] De Santi C, Altermark B, Pierechod MM, Ambrosino L, de Pascale D, Willassen NP. Characterization of a cold-active and salt tolerant esterase identified by functional screening of Arctic metagenomic libraries. BMC biochemistry. 2016;**17**:1

[85] Troeschel SC, Drepper T, Leggewie C, Streit WR, Jaeger KE. Novel tools for the functional expression of metagenomic DNA. Methods in molecular biology (Clifton, N.J.). 2010;**668**:117-139

[86] Liebl W, Angelov A, Juergensen J, Chow J, Loeschcke A, Drepper T, et al. Alternative hosts for functional (meta) genome analysis. Applied microbiology and biotechnology. 2014;**98**:8099-8109

[87] Park SY, Kim GJ. Screening of functional promoter from metagenomic DNA for practical use in expression systems. Methods in molecular biology (Clifton, N.J.). 2010;**668**:141-152

[88] Bitok JK, Lemetre C, Ternei MA, Brady SF. Identification of biosynthetic gene clusters from metagenomic libraries using PPTase complementation in a Streptomyces host. FEMS microbiology letters. 2017;**364**

[89] Schmidt EW, Donia MS. Chapter 23. Cyanobactin ribosomally synthesized peptides--a case of deep metagenome mining. Methods in enzymology. 2009;**458**:575-596

[90] Iqbal HA, Feng Z, Brady SF. Biocatalysts and small molecule products from metagenomic studies. Current opinion in chemical biology. 2012;**16**:109-116

[91] Nguyen T, Ishida K, Jenke-Kodama H, Dittmann E, Gurgui C, Hochmuth T, et al. Exploiting the mosaic structure of trans-acyltransferase polyketide synthases for natural product discovery and pathway dissection. Nature biotechnology. 2008;**26**:225-233

[92] Donia MS, Ravel J, Schmidt EW. A global assembly line for cyanobactins. Nature chemical biology. 2008;**4**:341-343

[93] Feng Z, Kallifidas D, Brady SF. Functional analysis of environmental DNA-derived type II polyketide synthases reveals structurally diverse secondary metabolites. Proceedings of the National Academy of Sciences of the United States of America. 2011;**108**:12629-12634

[94] Reynolds PE. Structure, biochemistry and mechanism of action of glycopeptide antibiotics, European journal of clinical microbiology & infectious diseases: official publication of the European Society of Clinical. Microbiology. 1989;**8**:943-950

[95] Donadio S, Sosio M. Biosynthesis of glycopeptides: prospects for improved antibacterials. Current topics in medicinal chemistry. 2008;**8**:654-666

[96] Banik JJ, Brady SF. Cloning and characterization of new glycopeptide gene clusters found in an environmental DNA megalibrary. Proceedings of the National Academy of Sciences of the United States of America. 2008;**105**:17273-17277

[97] Jaspars M. The origins of cyanobactin chemistry and biology. Chemical communications (Cambridge, England). 2014;**50**:10174-10176

[98] Leikoski N, Liu L, Jokela J, Wahlsten M, Gugger M, Calteau A, et al. Genome mining expands the chemical diversity of the cyanobactin family to include highly modified linear peptides. Chemistry & biology. 2013;**20**:1033-1043

[99] Martins J, Leao PN, Ramos V, Vasconcelos V. N-terminal protease gene phylogeny reveals the potential for novel cyanobactin diversity in cyanobacteria. Marine drugs. 2013;**11**:4902-4916

[100] Long PF, Dunlap WC, Battershill CN, Jaspars M. Shotgun cloning and heterologous expression of the patellamide gene cluster as a strategy to achieving sustained metabolite production, Chembiochem: a European. journal of chemical biology. 2005;**6**:1760-1765

[101] Donia MS, Ruffner DE, Cao S, Schmidt EW. Accessing the hidden majority of marine natural products through metagenomics, Chembiochem: a European. journal of chemical biology. 2011;**12**:1230-1236

[102] Hertweck C, Luzhetskyy A, Rebets Y, Bechthold A. Type II polyketide synthases: gaining a deeper insight into enzymatic teamwork. Natural product reports. 2007;**24**:162-190

[103] Wawrik B, Kerkhof L, Zylstra GJ, Kukor JJ. Identification of unique type II polyketide synthase genes in soil. Applied and environmental microbiology. 2005;**71**:2232-2238

[104] Piel J. Biosynthesis of polyketides by trans-AT polyketide synthases. Natural product reports. 2010;**27**:996-1047

[105] Zimmermann K, Engeser M, Blunt JW, Munro MH, Piel J. Pederin-type pathways of uncultivated bacterial symbionts: analysis of o-methyltransferases and generation of a biosynthetic hybrid. Journal of the American Chemical Society. 2009;**131**:2780-2781

[106] Stepanauskas R. Single cell genomics: an individual look at microbes. Current opinion in microbiology. 2012;**15**:613-620

[107] T. Woyke, D.F.R. Doud, The trajectory of microbial single-cell sequencing, 14 (2017) 1045-1054.

[108] Blainey PC. The future is now: single-cell genomics of bacteria and archaea. FEMS microbiology reviews. 2013;**37**:407-427

[109] Rinke C, Lee J, Nath N, Goudeau D, Thompson B, Poulton N, et al. Obtaining genomes from uncultivated environmental microorganisms using FACS-based single-cell genomics. Nature protocols. 2014;**9**:1038-1048

[110] Xu Y, Zhao F. Single-cell metagenomics: challenges and applications. Protein & cell. 2018;**9**:501-510

[111] Nurk S, Bankevich A, Antipov D, Gurevich AA, Korobeynikov A, Lapidus A, et al. Assembling single-cell genomes and mini-metagenomes from chimeric MDA products. Journal of computational biology: a journal of computational molecular cell biology. 2013;**20**:714-737

[112] Kodzius R, Gojobori T. Single-cell technologies in environmental omics. Nature methods. 2016;**576**:701-707

[113] Yilmaz S, Singh AK. Single cell genome sequencing. Current opinion in biotechnology. 2012;**23**:437-443

[114] Hedlund BP, Dodsworth JA, Murugapiran SK, Rinke C, Woyke T. Impact of single-cell genomics and metagenomics on the emerging view of extremophile "microbial dark matter". Extremophiles: life under extreme conditions. 2014;**18**:865-875

[115] Baker BJ, Tyson GW, Webb RI, Flanagan J, Hugenholtz P, Allen EE, et al. Lineages of acidophilic archaea revealed by community genomic analysis. Science (New York, N.Y.). 2006;**314**:1933-1935

[116] Baker BJ, Comolli LR, Dick GJ, Hauser LJ, Hyatt D, Dill BD, et al. Enigmatic, ultrasmall, uncultivated Archaea. Proceedings of the National Academy of Sciences of the United States of America. 2010;**107**:8806-8811