

Reeta Goel, Deep Chandra Suyal, Narayan, Biplab Dash,
and Ravindra Soni

Abstract

Microorganisms are considered as an unlimited, untapped, and intriguing source for development of novel genes, antibiotics, bioinoculants, biocatalysts, etc. Therefore, the identification and characterization of the microbial wealth becomes inevitable. Besides other features of bacteria that are a challenge to health, including increase in pathogenicity, antibiotic resistance, mounting host spectra, and possibility of usage for bioterrorism, it becomes a compulsion. Furthermore, a sustainable agriculture development program that frequently exploits the plant growth-promoting rhizobacteria (PGPR) also requires a correct identification and characterization of agriculturally important microorganisms. Metagenomics is defined as the study of all available genomes in an environment that contains considerably more genetic information than provided by the cultured subset. It also helps in bridging the gap between genetics and ecology, indicating that the genes of a single microorganism are connected to the genes of other members of the community.

Keywords

Metagenomics • Agriculture • Himalayan Soil

R. Goel • D.C. Suyal

Department of Microbiology, CBSH, G. B. Pant University of Agriculture and Technology,
Pantnagar 263 145, Uttaranchal, India

Narayan • B. Dash • R. Soni (✉)

Department of Agricultural Microbiology, College of Agriculture, Indira Gandhi Krishi
Vishwavidyalaya, Raipur, Chhattisgarh, India

e-mail: rs31693@gmail.com

13.1 Introduction

Soil biodiversity analysis is a very important aspect in the environmental sciences due to its significant interlinkages with other areas, like agriculture. As evident from available literature, soil biodiversity characterizes a huge underground world that contains a wide range of organisms, from prokaryotes to eukaryotes (e.g., archaea, bacteria, fungi, nematodes, insects, and earthworms). One gram of soil has been reported to contain up to 10^{10} prokaryotic cells and thousands of different species (Raynaud and Nunan 2014). This diverse microbial ecosystem plays a central role in the nutrient cycling, soil structure formation, decomposition of organic matter, soil health indicator, soilborne diseases, and plant growth promotion and thus is responsible for maintaining the biosphere integrity. Microbial diversity present in soil can be explored either through culture-based methods or recent indirect biotechnological approaches. Many indirect methods for overcoming the limitations of cultivation techniques are developed which are mainly based on nucleic acid isolation from soil and their characterization without the culture of microbes.

13.2 Metagenomics

Culture-dependent methods limit analysis of those microorganisms that can grow under laboratory conditions. It is widely accepted that only 0.1–1% (depending upon the environmental sample) of bacteria can be cultured by laboratory cultivation methods which leaves 99% or more microbial diversity unexplored. Furthermore, under environmental stress bacteria can enter a state called “viable but unculturable” which again limits the accessibility of these bacteria to traditional cultivation method. Thus, cultivation-dependent microbial identification can underestimate the microbial diversity. As per the available reports, *in vitro* culturability of the total microbes in freshwater and in sediments is 0.25% (Jones 1977), in sea water 0.001–0.1% (Surmann and Efferth 2014), and only less than 1% in soil (Pham and Kim 2012; Ferrari et al. 2005). Therefore, “metagenomics,” i.e., direct extraction of genetic material from environment, was conceptualized (Handelsman et al. 1998) for analyzing similar but not identical genomes, in the environment.

Metagenomics may target the structure of the metagenome by cloning and sequencing strategies (structural metagenomics) and/or characterize the functions of environmental DNA by direct cloning for heterologous expression in a surrogate host organism (functional metagenomics). As the functional metagenomics method relies on the ability of the cloned environmental DNA to confer a phenotypic function to the host, no sequence homology to previously characterized genes or other *a priori* sequence information is required. Functional metagenomics can therefore be considered as a true discovery tool for identifying and characterizing novel gene families (Nacke et al. 2011), metabolic traits (McGarvey et al. 2012), bioactive compounds (Craig et al. 2010), or pathways (Illegheems et al. 2015) from uncultured soil microbes. With suitably long genomic fragments, functional metagenomics

may also be used to define the genomic context of the functions of interest and enable their taxonomic assignment (Treusch et al. 2005).

13.3 Essential Steps of Metagenomics

13.3.1 Environmental Sampling

Sample collection is the first step of a soil metagenomic library construction for which details of physicochemical properties of the soils are required. Sometimes, enzyme activity assays or metagenomic sequencing can also be used to determine the functional diversity of the respective environment.

13.3.2 Metagenomic Library Construction

13.3.2.1 High-Quality DNA Isolation from Soil Samples

Metagenomic library construction requires a high-quality DNA from the environment. Soil heterogeneity, microbial diversity, and other soil properties make DNA extraction challenging. Moreover, soil DNA extraction often contains humic substances, which interfere and reduce the efficiency of downstream processes (Premalatha et al. 2009; Soni and Goel 2010).

To ensure efficient cloning, the isolated DNA should be purified of contaminants such as humic acids or phenolic compounds that can inhibit enzymatic reactions. The capture of full-length genes using small-insert libraries requires DNA fragments of at least 2 kb in length, whereas over 25 kb fragments are required for identification of operons using cosmid and fosmid libraries. While cell lysis directly in the soil matrix enables rapid recovery of greater amounts of DNA, indirect extraction methods typically recover DNA with larger fragment sizes, higher purity, and higher representation of many bacterial and archaeal taxa but lower representation of filamentous organisms such as fungi and *Actinobacteria* and microbes attached to soil matrix (Delmont et al. 2011).

13.3.2.2 Cloning Vector

Metagenomic libraries can be roughly categorized into small- and large-insert libraries as per the cloned DNA fragment size. Using plasmid vectors, construction of small-insert libraries is made which is able to contain up to 10 kb DNA fragments, thus making them suitable for identification of functional traits encoded by a single gene or small operon (Surmann and Effertth 2014). These plasmid-based libraries provide high transformation efficiency ($>10^5$ clones) and efficient expression systems using vectors that can be induced to high copy numbers and that contain promoters. Higher copy number is especially advantageous for characterization of genes, cloned without promoters or with low activity. Large-insert libraries use cosmid or fosmid vectors or bacterial artificial chromosomes (BAC) that can accommodate 25–35, 25–40, or 100–200 kb DNA fragments, respectively. Due to the

large size of the cloned DNA fragments, these libraries are well suited for identification of multi-domain traits or pathways and provide linkage information for the identification of functions encoded by multiple genes and potentially allow taxonomic linkages to be determined. The vectors used to construct large-insert libraries are generally present in cells in low copy number, which allows their stable replication in the screening host and reduces the risk of overexpression of toxic gene products.

13.3.2.3 Screening Host

In the majority of applications to date, the functional screening of metagenomic libraries relies on expression in *Escherichia coli*. This well-characterized laboratory model organism has stable replication of vectors and low rates of restriction and recombination, making an attractive host microorganism for cloning and expression of foreign DNA. A variety of expression systems for *E. coli* are available as are a large number of genetically modified *E. coli* strains for highly controlled and optimized cloning and expression (Sørensen and Mortensen 2005). Alternative screening hosts for functional mining of metagenomic libraries include *Streptomyces lividans* (Wang et al. 2000), *Bacillus subtilis* (Troeschel et al. 2010), *Sulfolobus solfataricus* (Albers et al. 2006), *Thermus thermophilus* (Angelov et al. 2009), *Saccharomyces cerevisiae* (Bailly et al. 2007), etc. When using these hosts, the library is generally constructed and maintained in *E. coli* and transformed to the alternative expression hosts for screening (Craig et al. 2010).

13.3.3 Library Screening

Metagenomics libraries can be screened by several techniques based either on functional activity or on nucleotide sequence. Soil-metagenomic libraries are screened by using target-specific probes. This approach is being used extensively to identify phylogenetic markers as well as other functional genes with highly conserved domains (Premalatha et al. 2009; Soni and Goel 2010, 2011). Moreover, microarray technology is also useful for soil metagenome analysis.

Enzymatic function of clones can be monitored by adding chemical dyes or chromophore-bearing enzyme substrate derivatives into culture medium. Thus, this type of sensitive nature of screening helps in the detection of rare clones.

13.4 Metagenomics of Himalayan Soils

The entire Himalayan mountain range is well known for its biodiversity due to its unique environment. Scattered habitations, inaccessibility, and uneconomic holdings keep these regions free from any anthropogenic contamination. Nevertheless, preference for traditional farming system over chemical-based farming is responsible for emergence of hilly agriculture lands as a gold mine for potential soil microorganisms.

In Uttarakhand Himalaya perspective, a triphasic approach, viz., real-time PCR (qPCR), denaturation gradient gel electrophoresis (DGGE), and temporal gradient gel electrophoresis (TGGE), has been used for evaluation of bacterial population in different rhizospheric soil systems (Soni and Goel 2010). Moreover, several *nifH* homologs have been identified from Himalayan rhizospheric soil metagenome (Soni and Goel 2011). Recently, Goel and co-workers made two 16S rDNA clone libraries, i.e., SB1 and SB2, using rhizospheric soil samples from two different locations of Western Indian Himalaya, namely, Chhiplakot (30.70°N/80.30°E) and Munsyari (30.60°N/80.20°E), selected on the basis of qPCR analysis to characterize the total bacterial population and their community structure (Suyal et al. 2015a). The phylum *Proteobacteria* was the dominant phylum in the Himalayan soils along with *Bacteroidetes*, *Nitrospira*, *Acidobacteria*, *Chloroflexi*, *Firmicutes*, *Cyanobacteria*, *Gemmatimonadetes*, *Planctomycetes*, *BRC1*, *Actinobacteria*, and *Chlorobi*. Comparative study on the bacterial diversity observed in this region with that of other Himalayan cold habitats like Tibetan plateau glacier (Liu et al. 2009), Drass, cold desert of the Western Himalaya (Shivaji et al. 2011), Puruogangri ice (Zhang et al. 2008), and Roopkund glacier (Pradhan et al. 2010) indicated that the bacterial diversity in both soils was comparable with each other.

13.5 Metagenomics as a Tool for Sustainable Agriculture

An integral constituent of integrated nutrient management (INM) and soil biodiversity system (SBS) is the soil-inhabiting microflora and microfauna, thereby playing an important role in plant growth and all-round development. Hazardous effects of chemical fertilizers and pesticides on soil and plant health having deleterious environmental impact are being frequently observed in recent years. Beneficial microbial wealth of agricultural importance can serve as a crucial alternative for achieving sustainable agriculture production. Metagenomics help in the prediction of microbial community structure and, therefore, can tackle and address fundamental scientific questions related to agriculturally important microorganisms. This approach has been successfully explored for the assessment of the diazotrophs belonging to the rhizosphere of native red kidney beans (RKB) of the Western Indian Himalaya by targeting *nifH* (Suyal et al. 2015b). This metagenomic effort has examined the community structure and diversity of N₂-fixing microorganisms in a Himalayan RKB rhizosphere, which can be explored to provide the backbone for further studies. Moreover, previous metagenomic efforts indicated the pervasiveness of *csj* and *nif* from the Himalayan soils of (Premalatha et al. 2009; Soni and Goel 2010).

The emergence of metagenome information for a rhizosphere is beginning to expose detailed information about associated community structure, dynamics, and functional activities, thereby allowing an improved perceptive of community development, interspecies coordination and competition for essential nutrients, and distribution of metabolic activities across the community members. Moreover, functional metagenomics can also be explored for reshaping the composition of

rhizospheric microbial population and to readdress microbial activity, which can be referred to as “rhizosphere engineering.”

13.6 Current Scenario of Metagenomics

Metagenomics has significant potential as a discovery and annotation tool for linking genes with functions and processes and providing valuable phylogenetic context that may enable the role and ecological niches of soil microorganisms to be determined. However, in order to keep up with advances in sequencing technologies and to enable discovery from the broad diversity of soil microorganisms, there are challenges to be faced. An important development would be to increase the taxonomic coverage of genes that can be expressed and screened. Most of the currently used expression vectors replicate only in *E. coli*. Given the limited ability of *E. coli* to express genes from distant taxonomic groups of organisms, shuttle vectors with extended host range are needed (Aakvik et al. 2009; Craig et al. 2010). Maintenance of these vectors is conveniently done inside *E. coli*, which can then be transferred by conjugation into another expression host for repeated screening. It will also be necessary to modify the structure and expression mechanisms of the vectors in order to accommodate larger DNA fragments, to enable expression from different orientations depending on the orientation of cloned DNA insert, and to control the gene expression level via copy number adjustment and induction (Lämmle et al. 2007). Genetic modification of existing expression hosts by ribosome engineering, co-expression of molecular chaperones, or engineering of transcription and translation factors as well as secretion systems represent a means to increase the rate of expression of foreign genes in *E. coli* (Bernstein et al. 2007). However, screening in taxonomically distant expression hosts or the use of multiple different hosts has been shown to significantly increase detection frequency for novel functional traits. Screening in a physiologically suitable host is especially useful when mining extreme environments for novel biochemical properties because proteins that function at extreme temperatures or under high salinity often require additional modifications to ensure protein stability (Angelov et al. 2009). Ideally, cell-free expression systems for universal expression of DNA originating from taxonomically different hosts could be used.

Eukaryotic genes are still in minority among the discovered genes derived from environmental DNA, although construction of functional metagenomic libraries from soil transcriptomes, i.e., total RNA isolated from a soil sample, may be used to overcome limitations in mining eukaryotic genes containing nonbacterial genetic elements and introns. Messenger RNA is captured with polyadenylated primers and reverse-transcribed to double-stranded cDNA. This cDNA is then cloned to an expression vector for expression and functional screening in either bacterial or eukaryotic hosts. Although cDNA libraries have been successfully used for identification of fungal genes from soil metatranscriptomes, RNA instability and challenges in RNA isolation often result in low recovery of full-length transcripts (Bailey et al. 2007). The separate cloning of single transcripts

also limits the recovery of entire biosynthetic pathways. Regardless of the type of methodology used for library preparation or expression screening, it is obvious that the scale of a screening effort that is required to capture less abundant microbial groups and functions, and to maintain a rate of functional annotation consistent with that of metagenome sequence data acquisition, greatly exceeds the capacity of current functional metagenomic implementations. This capacity could be increased by development of more sensitive screening substrates and by increasing the throughput of screening assays. Increased assay sensitivity and throughput could be achieved through the combination of novel substrates and multiplexed assays for analysis using FACS, high-throughput liquid chromatography, or mass spectrometry, with miniaturized systems using microfluidic devices to enable nanoliter reaction volumes to improve throughput and speed and substantially reduce cost.

Conclusions

The importance of the microorganisms in the growth and development of plant ecosystems has been well known; however, the major portion of rhizosphere population is still uncharacterized and unexplored. Coupling traditional with advanced metagenomic methodologies to evaluate community structure and function will bring new insights to explore microbial life in the soil. Further, identification of the plant signals, exudates, and key factors in the rhizosphere microbial ecosystem will provide chemical and microbial markers to explain how plants recruit and stimulate beneficial microorganisms. Moreover, soil metagenomics also holds prospective to improve crop production and to uncover several yet unexploited soil microorganisms, their functions, and genes for diverse applications.

Acknowledgment The work mentioned in this chapter from author group was supported by the National Bureau of Agriculturally Important Microorganisms, India (NBAIM/ICAR), grant to R.G.

References

- Aakvik T, Degnes KF, Dahlsrud R, Schmidt F, Dam R, Yu L, Völker U, Ellingsen TE, Valla S (2009) A plasmid RK2-based broad-host-range cloning vector useful for transfer of metagenomic libraries to a variety of bacterial species. *FEMS Microbiol Lett* 296:149–158. doi:[10.1111/j.1574-6968.2009.01639.x](https://doi.org/10.1111/j.1574-6968.2009.01639.x)
- Albers SV, Jonuscheit M, Dinkelaker S, Urlich T, Kletzin A, Tampe R, Driessen AJM, Schleper C (2006) Production of recombinant and tagged proteins in the hyperthermophilic archaeon *Sulfolobus solfataricus*. *Appl Environ Microbiol* 72:102–111. doi:[10.1128/AEM.72.1.102-111.2006](https://doi.org/10.1128/AEM.72.1.102-111.2006)
- Angelov A, Mientus M, Liebl S, Liebl W (2009) A two-host fosmid system for functional screening of (meta) genomic libraries from extreme thermophiles. *Syst Appl Microbiol* 3:177–185. doi:[10.1016/j.syapm.2008.01.003](https://doi.org/10.1016/j.syapm.2008.01.003)
- Bailey J, Fraissinet-Tachet L, Verner MC, Debaud JC, Lemaire M, Wésolowski-Louvel M, Marmeisse R (2007) Soil eukaryotic functional diversity, a metatranscriptomic approach. *ISME J* 1:632–642. doi:[10.1038/ismej.2007.68](https://doi.org/10.1038/ismej.2007.68)

- Bernstein JR, Bulter T, Shen CR, Liao JC (2007) Directed evolution of ribosomal protein S1 for enhanced translational efficiency of high GC *Rhodospseudomonas palustris* DNA in *Escherichia coli*. *J Biol Chem* 282:18929–18936. doi:[10.1074/jbc.M701395200](https://doi.org/10.1074/jbc.M701395200)
- Craig JW, Chang FY, Kim JH, Obiajulu SC, Brady SF (2010) Expanding small-molecule functional metagenomics through parallel screening of broad-host-range cosmid environmental DNA libraries in diverse proteobacteria. *App Environ Microbiol* 76:1633–1641. doi:[10.1128/AEM.02169-09](https://doi.org/10.1128/AEM.02169-09)
- Delmont TO, Robe P, Clark I, Simonet P, Vogel TM (2011) Metagenomic comparison of direct and indirect soil DNA extraction approaches. *J Microbiol Methods* 86:397–400. doi:[10.1016/j.mimet.2011.06.013](https://doi.org/10.1016/j.mimet.2011.06.013)
- Ferrari BC, Binnerup SJ, Gillings M (2005) Microcolony cultivation on a soil substrate membrane system selects for previously uncultured soil bacteria. *Appl Environ Microbiol* 71(12):8714–8720. doi:[10.1128/AEM.71.12.8714-8720.2005](https://doi.org/10.1128/AEM.71.12.8714-8720.2005)
- Handelsman J, Rondon MR, Brady SF, Clardy J, Goodman RM (1998) Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. *Chem Biol* 5:245–249. doi:[10.1016/S1074-5521\(98\)90108-9](https://doi.org/10.1016/S1074-5521(98)90108-9)
- Illegheems K, Weckx S, Vuyst LD (2015) Applying meta-pathway analyses through metagenomics to identify the functional properties of the major bacterial communities of a single spontaneous cocoa bean fermentation process sample. *Int J Food Microbiol* 50:54–63. doi:[10.1016/j.fm.2015.03.005](https://doi.org/10.1016/j.fm.2015.03.005)
- Jones JG (1977) The effect of environmental factors on estimated viable and total populations of planktonic bacteria in lakes and experimental enclosures. *Freshwater Biol* 7:67–91. doi:[10.1111/j.1365-2427.1977.tb01659.x](https://doi.org/10.1111/j.1365-2427.1977.tb01659.x)
- Lämmle K, Zipper H, Breuer M, Hauer B, Buta C, Brunner H, Rupp S (2007) Identification of novel enzymes with different hydrolytic activities by metagenome expression cloning. *J Biotechnol* 127:575–592. doi:[10.1016/j.jbiotec.2006.07.036](https://doi.org/10.1016/j.jbiotec.2006.07.036)
- Liu Y, Yao T, Jiao N, Kang S, Xu B, Zeng Y, Huang S, Liu X (2009) Bacterial diversity in the snow over Tibetan Plateau Glaciers. *Extremophiles* 13:411–423. doi:[10.1007/s00792-009-0227-5](https://doi.org/10.1007/s00792-009-0227-5)
- McGarvey KM, Queitsch K, Fields S (2012) Wide variation in antibiotic resistance proteins identified by functional metagenomic screening of a soil DNA library. *Appl Environ Microbiol* 78:1708–1714. doi:[10.1128/AEM.06759-11](https://doi.org/10.1128/AEM.06759-11)
- Nacke H, Will C, Herzog S, Nowka B, Engelhaupt M, Daniel R (2011) Identification of novel lipolytic genes and gene families by screening of Metagenomic libraries derived from soil samples of the German Biodiversity Exploratories. *FEMS Microbiol Ecol* 78:188–201. doi:[10.1111/j.1574-6941.2011.01088.x](https://doi.org/10.1111/j.1574-6941.2011.01088.x)
- Pham VH, Kim J (2012) Cultivation of unculturable soil bacteria. *Trends Biotechnol* 30(9):475–484. doi:[10.1016/j.tibtech.2012.05.007](https://doi.org/10.1016/j.tibtech.2012.05.007)
- Pradhan S, Srinivas TNR, Pindi PK, Kishore KH, Begum Z, Singh PK, Singh AK, Pratibha MS, Yasala AK, Reddy GS, Shivaji S (2010) Bacterial biodiversity from Roopkund glacier, Himalayan mountain ranges, India. *Extremophiles* 14:377–395. doi:[10.1007/s00792-010-0318-3](https://doi.org/10.1007/s00792-010-0318-3)
- Premalatha K, Soni R, Khan M, Marla SS, Goel R (2009) Exploration of csp gene(s) from temperate and glacier soils of Indian Himalaya and *in silico* analysis of encoding proteins. *Curr Microbiol* 58:343–348. doi:[10.1007/s00284-008-9344-0](https://doi.org/10.1007/s00284-008-9344-0)
- Raynaud X, Nunan N (2014) Spatial ecology of bacteria at the microscale in soil. *PLoS One* 9(1):e87217. doi:[10.1371/journal.pone.0087217](https://doi.org/10.1371/journal.pone.0087217)
- Shivaji S, Pratibha MS, Sailaja B, Kishore KH, Singh AK, Begum Z, Anarasi U, Prabakaran SR, Reddy GSN, Srinivas TNR (2011) Bacterial diversity of soil in the vicinity of Pindari glacier, Himalayan mountain ranges, India, using culturable bacteria and soil 16S rRNA gene clones. *Extremophiles* 15:1–22. doi:[10.1007/s00792-010-0333-4](https://doi.org/10.1007/s00792-010-0333-4)
- Soni R, Goel R (2010) Triphasic approach for assessment of bacterial population in different soil systems. *Ekologija* 56:94–98. doi:[10.2478/v10055-010-0014-8](https://doi.org/10.2478/v10055-010-0014-8)
- Soni R, Goel R (2011) *nifH* homologous from soil metagenome. *Ekologija* 57:87–95. doi:[10.6001/ekologija.v57i3.1914](https://doi.org/10.6001/ekologija.v57i3.1914)

- Sørensen HP, Mortensen KK (2005) Advanced genetic strategies for recombinant protein expression in *Escherichia coli*. *J Biotechnol* 115:113–128. doi:[10.1016/S0167-7799\(97\)01155-4](https://doi.org/10.1016/S0167-7799(97)01155-4)
- Surmann E-M, Efferth T (2014) Biodiversity and metagenomics. In: Kuete V, Efferth T (eds) Biodiversity natural, products and cancer treatment. World Scientific Publishing Co., Singapore, pp 35–69. doi:[10.1142/9789814583510_0002](https://doi.org/10.1142/9789814583510_0002)
- Suyal DC, Yadav A, Shouche Y, Goel R (2015a) Bacterial diversity and community structure of Western Indian Himalayan red kidney bean (*Phaseolus vulgaris* L.) rhizosphere as revealed by 16S rRNA gene sequences. *Biologia* 70:305–313. doi:[10.1515/biolog-2015-0048](https://doi.org/10.1515/biolog-2015-0048)
- Suyal DC, Yadav A, Shouche Y, Goel R (2015b) Diversified diazotrophs associated with the rhizosphere of Western Indian Himalayan native red kidney beans (*Phaseolus vulgaris* L.) *3Biotech* 5:433–441. doi:[10.1007/s13205-014-0238-5](https://doi.org/10.1007/s13205-014-0238-5)
- Treusch AH, Leininger S, Kletzin A, Schuster SC, Klenk HP, Schleper C (2005) Novel genes for nitrite reductase and Amo-related proteins indicate a role of uncultivated mesophilic crenarchaeota in nitrogen cycling. *Environ Microbiol* 7:1985–1995. doi:[10.1111/j.1462-2920.2005.00906.x](https://doi.org/10.1111/j.1462-2920.2005.00906.x)
- Troeschel SC, Drepper T, Leggewie C, Streit WR, Jaeger KE (2010) Novel tools for the functional expression of metagenomic DNA. *Methods Mol Biol* 668:117–139. doi:[10.1007/978-1-60761-823-2_8](https://doi.org/10.1007/978-1-60761-823-2_8)
- Wang GY, Graziani E, Waters B, Pan W, Li X, McDermott J, Meurer G, Saxena G, Andersen RJ, Davies J (2000) Novel natural products from soil DNA libraries in a streptomycete host. *Org Lett* 10:2401–2404. doi:[10.1021/ol005860z](https://doi.org/10.1021/ol005860z)
- Zhang X, Yao T, Tian L, Xu S, An L (2008) Phylogenetic and physiological diversity of bacteria isolated from Puruogangri ice core. *Microbiol Ecol* 55:476–488. doi:[10.1007/s00248-007-9293-3](https://doi.org/10.1007/s00248-007-9293-3)