Encyclopedia of Metagenomics
DOI 10.1007/978-1-4614-6418-1_623-3

Antarctic Soil Metagenome

Pablo Power[a]* and Renaud Berlemont[b]
[a]Department of Microbiology, University of Buenos Aires – CONICET, School of Pharmacy and Biochemistry, Buenos Aires, Argentina
[b]Department of Earth System Science and Department of Ecology and Evolutionary Biology, University of California, Irvine, Irvine, CA, USA

Introduction

If we performed a census to assess the number of microorganisms living in our planet, we would figure out that they represent no less than half the total biomass on Earth. Recent studies in the field of environmental microbiology have revealed the tremendous genetic and functional diversity of natural microbial consortia. Some evidences suggest that millions of bacterial lineages may inhabit the soil, the ocean, and the atmosphere. In soil, the amount of present microorganisms reaches the exorbitant number of billions (10^9–10^{10}) of cells by milliliter (Daniel 2005). Such a number deserves a little respect.

For a long time, scientists made efforts trying to develop novel culture methodologies for growing different types of microorganisms in pure cultures. Nevertheless, and according to what we could reasonably suppose, the percentage of microorganisms (talking about specific taxonomic groups) we are able to recover by in vitro culture techniques is almost negligible (no more than 0.1–1 % depending on the sample) (Amann et al. 1995). By epifluorescence microscopy and DAPI (4′,6-diamidino-2-phenylindole) stain, the number of cells in a soil sample has been estimated at 4×10^{10} per gram of soil; after trying to recover them in an enriched culture medium, no more than 10^6 CFU (colony-forming units) per gram were observed (0.01 %) (Handelsman 2004).

This discrepancy originated what has been called the "great plate count anomaly" (Staley and Konopka 1985), whose meaning relies in that concept: no matter the number of living microorganisms inhabiting a specific sample (soil, water, tissue, sediment, bone, etc.), we will lose more than 90 % of them in attempts to culture them. This makes that from approximately 300,000 taxonomic groups known today, less than half of them possesses a single member able to grow on culture media (Rappe and Giovannoni 2003).

Therefore, if the capability of recovering living cells by culture techniques is so scarce, how could we manage to face ecological, metabolic, biochemical, and genetic studies if we will not isolate them?

Back in 1998, Handelsman et al. used bacterial artificial chromosomes (BAC) to introduce genomic DNA fragments directly isolated from soil in *E. coli*. The so-called "metagenomic" DNA libraries have been screened for new enzymatic activities (Handelsman et al. 1998). Thus, metagenomics could be defined as a technique that requires common microbiology/molecular biology methodological approaches for accessing the genetic information (so-called the metagenome) from environmental microbial consortia, including uncultured microorganisms representing the vast majority of the total, without the need of previous isolation techniques (Handelsman 2004).

*Email: ppower@ffyb.uba.ar
*Email: pablopowermetal@gmail.com

DOI 10.1007/978-1-4614-6418-1_623-3

This relatively new technology provided access to previously hidden genetic information in genomes from uncultured organisms, to the isolation of novel genes, proteins, and to the analysis of genomes and metabolic pathways from uncultured soil microorganisms, paving the way to elucidate functions of microbial communities (Riesenfeld et al. 2004; Daniel 2005).

Since evolution and natural selection have been occurring in the Earth environment for billions of years, the metagenomic approach allows the isolation of enzymes that harbor tailor-made properties, fitting the physicochemical conditions of the habitats studied. Among them, extreme environments are particularly attractive because they potentially contain a vast majority of microbial taxa that are not easily cultivated but display many interesting fundamental and biotechnological features.

Cold-adapted microorganisms have successfully faced the negative effect of low temperatures on the reaction rate. These organisms display metabolic fluxes comparable to those of their mesophilic counterparts, at a temperature close to (or below) the freezing point. Indeed, bacterial growth has been detected at a temperature of −12 °C and bacterial metabolism at −20 °C. Overall, these microorganisms are important for the global functioning of many ecosystems (e.g., primary production and nutrient cycling) since up to two-thirds of the Earth is characterized by low temperatures (5 °C or less); polar-regions account for 14 % of the planet's surface and deep-sea water/sediment, together with alpine ecosystems, is also permanently cold (Margesin and Miteva 2011).

Compared to thermophilic organisms, cold-adapted microorganisms were less investigated. However, studies on cold-adapted proteins at the biochemical level (i.e., kinetic, folding, structure), together with considerations of their membrane permeability and resistance to freezing, have now been widely recognized.

One of the keys for success to adaptation to cold is to achieve a suitable enzymatic fitness to face low temperatures, liquid-water depletion, and the associated increase of viscosity (Marx et al. 2007). Proteins derived from organisms living in cold environments display increased activities at low temperatures generally associated with a high thermal instability (Gerday et al. 2000). Additional key points are the ability to modulate membrane fluidity (e.g., modulating the fatty acid saturation), the ability to produce "cryoprotectants," and the capacity to prevent oxidative stress (Margesin and Miteva 2011), among others.

Many microorganisms have successfully overcome the low temperature challenge. In this regard, Antarctica represents a very attractive location for the application of metagenomic approaches aimed at the search for novel cold-adapted enzymes. The uniqueness of Antarctica relies in its already demonstrated capability of sheltering viable microorganisms in spite of the dominating low temperatures, associated with a human-associated activity that is kept to a minimum while cell densities in cold environment may be elevated (10^5–10^6 cells/ml in Antarctic sea waters) (Marx et al. 2007).

Soil Metagenomics in Practice

Methodologically, and as depicted in Fig. 1, metagenomics is based on the extraction of total DNA from the soil sample, digestion and cloning of resulting fragments, and transformation of ligation mixture into cultivable hosts like *Escherichia coli* so that "metagenomic libraries" can be achieved (Handelsman et al. 2002; Daniel 2004).

Each of the steps for succeeding on the construction of a metagenomic library includes a couple of tips that deserve attention.

The first of them is of course the proper selection of the sample which, besides the logical reasons of choice (interest in studying the microbial diversity from a particular location, the screening of a given set of genes or proteins matching specific features, etc.), is subjected to luck due to the often

DOI 10.1007/978-1-4614-6418-1_623-3

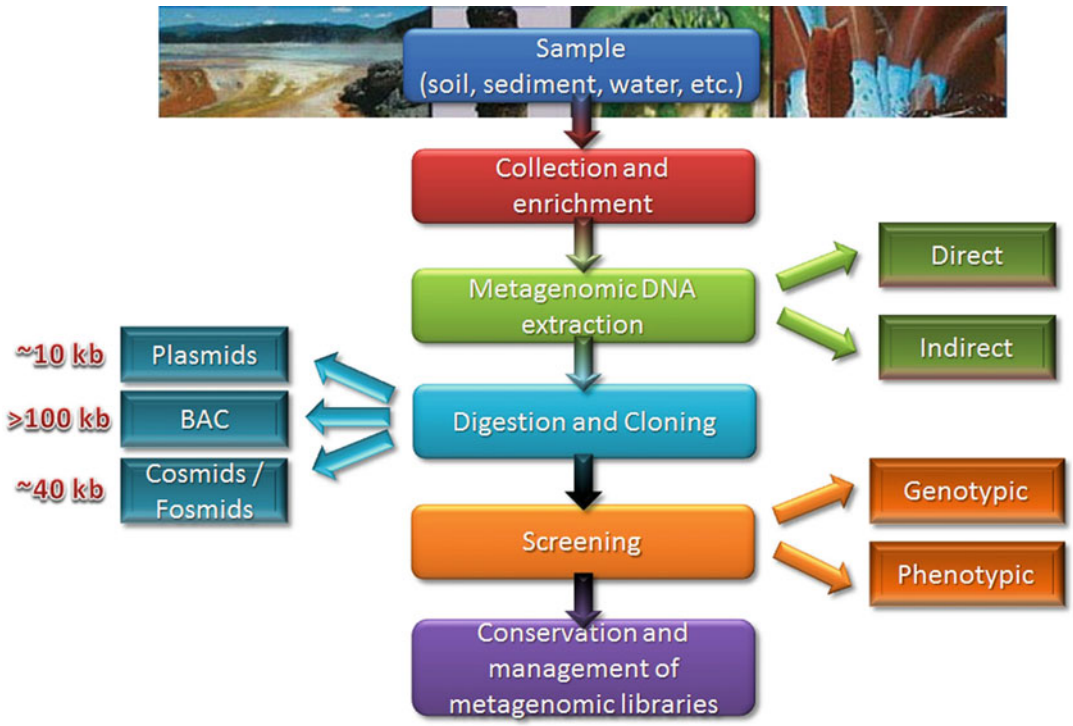

Fig. 1 Methodology chart depicting the main steps for achieving a metagenomic library from environmental samples

subjective and random selection of the specific point from which we will collect it: 10 cm to the right, left, or depth can sometimes change the whole nature and composition of our sample. Another key point during the sample collection, after we decided from where to take it, relies in the need for keeping the integrity of the sample and avoiding exogenous contaminations during manipulation, for which we might use the best aseptic technique we could (considering the obvious difficulties we face when working "in the wild").

The strict molecular biology techniques, namely, restriction digestion, cloning, and transformation in a suitable host (in most cases gentle *Escherichia coli* strains), should not represent major obstacles for the average researcher.

However, the "bottle neck" in the built of a soil metagenomic library is probably the problem posed by the presence of humic acids (and derivatives) which are present in different concentrations in soil (the darker brown the DNA solution is, the higher the humics content). These compounds are extremely deleterious even in trace concentrations (micrograms per ml), especially towards DNA polymerases and restriction endonucleases, the enzymes needed for processing the metagenomic DNA after being isolated (Tebbe and Vahjen 1993). Luckily, several methodologies aimed at removing these inhibitors were developed (ion exchange resins, precipitation with polyethylene glycol, etc.), including commercial kits.

Once the library is finally successfully built, the last obstacle to solve is how we will maintain our clones in a way that the whole genetic information contained in them will not be lost, especially when the library is composed by several thousand bacterial colonies. The most economical yet not so "fancy" way to achieve this is to conserve the metagenomic clones as "pools," each containing hundreds or even thousands of clones. The ideal way to conserve them is, of course, as individual clones (e.g., in microplates), although the handling will be undoubtedly harder (at least our budget is sufficient for buying a modern yet expensive "picking" robot).

Encyclopedia of Metagenomics
DOI 10.1007/978-1-4614-6418-1_623-3

Searching the "Needle in the Haystack" in the Antarctic Soil Metagenome

Once the soil-derived metagenomic library is finally achieved, the most exciting quest is probably to try to recover individual clones expressing a phenotype of interest. This approach aims at the direct or indirect detection of a phenotypic task or a biochemical activity reflecting the expression of a given gene, known as "phenotype-based," "activity-driven," or "functional metagenomics," being the proper expression of the encoding gene mandatory (Rondon et al. 1999; Handelsman 2004; Daniel 2005). The availability of a detection method (usually based on an enzymatic reaction) determines a priori if this approach is doable, which is not always easy to apply. Among the drawbacks, the most relevant are probably the need of having the complete gene cloned and a proper heterologous expression in the host used.

Alternatively (yet not exclusively), the identification of the gene of interest could be achieved by using molecular biology methodologies (PCR screening, random and large-scale shotgun sequencing, hybridization, etc.) and was therefore called "genotypic" or "sequence-driven metagenomics"; although there is no need of any activity-based detection system, a minimal information of the sequences of interest is often helpful, especially for the design of primers or probes, although it is a good approach when we do not have the possibility of performing a phenotype-based screening (Handelsman 2004).

A few studies conducted either phenotypic or genotypic screening of proteins/genes for different applications (Table 1): lipases/esterases, cellulases, proteases, and amylases.

Lipases (EC 3.1.1.3) and esterases (EC 3.1.1.1) are the most frequent Antarctic enzymes studied by metagenomic-based approaches. They catalyze the hydrolysis of esters and the clones that express these enzymes are easily detected on tributyrin-containing plates (Berlemont et al. 2011; Berlemont et al. 2013).

The MHlip enzyme is a 262 amino acid lipase detected in a short-insert Antarctic soil library and displayed high amino acid identity to a putative cytosolic α-/β-hydrolase derived from *Acidovorax delafieldii* and to several mesophilic homologs. The most remarkable feature of MHlip is its adaptation to cold temperatures probably due to discrete structural modifications leading to this

Table 1 Enzymatic activities screened in Antarctic soil metagenomic libraries by phenotypic-based approaches

Target	Location	Screening	Vector	Total clones[a]	Coverage (Mb)[b]	Positive (1/× Mb)[c]	References
Lipase/ esterase	Adélie Land	Phenotypic	BAC	32,000	163	14 (11)	(Berlemont et al. 2011; Berlemont et al. 2013)
Amylase	Adélie Land	Phenotypic	BAC	27,500	140	14 (10)	(Berlemont et al. 2011)
Cellulase	Adélie Land	Phenotypic	BAC	8,800	45	11 (4)	(Berlemont et al. 2009; Berlemont et al. 2011)
Protease	Adélie Land	Phenotypic	BAC	16,000	82	3 (27)	(Berlemont et al. 2011)
MTA phosphorylase	Adélie Land	Phenotypic	Plasmid	85,000	nr	1 (nr)	(Cieslinski et al. 2009)
Alkaliphilic esterase	Miers Dry Valley	Phenotypic	Fosmid	10,000	300	nr	(Heath et al. 2009)
Biodiversity	Mullins Valley/ Beacon Valley	Genotypic	BAC	nr	nr	–	(Bidle et al. 2007)

Nr not reported
[a]Approximate total of clones screened for the given activity
[b]Average nucleotidic coverage in Mb
[c]Positive clones and (in parenthesis) equivalent frequency of occurrence for each positive clone

DOI 10.1007/978-1-4614-6418-1_623-3

behavior, since it lacks evident signature residues associated with cold adaptation (Berlemont et al. 2013). Another interesting enzyme belonging to the esterase, named CHA3, is showed to be active over a wide range of temperatures (7–50 °C) and has an alkaliphilic behavior (Heath et al. 2009).

While screening another metagenomic library for lipolytic activities, a serendipitous discovery occurred. A pinkish fluorescent clone expressing a methylthioadenosine (MTA) phosphorylase related to a homolog protein from *Psychrobacter arcticus* came out and could be interesting for being used as a new reporter gene for molecular biology applications (Cieslinski et al. 2009).

Among the other enzymes detected by functional metagenomics from Antarctic soil metagenomes (Berlemont et al. 2011), the cellulase RBcel1 represents the most exciting example. This enzyme, from the glycosyl hydrolases family five (GH5), presents dual activity: it displays a strong endocellulolytic activity on amorphous cellulose to yield cellobiose and cellotriose and synthesize non-reticulated cellulose using cellobiose as substrate (Berlemont et al. 2009).

Finally, genotypic screening on Antarctic soil metagenomic libraries was performed for microbial diversity studies as regards of the microbial activity. Metagenomic analysis of community DNA suggests that the Antarctic soil contains many orthologs belonging to extant metabolic genes (Bidle et al. 2007).

Sequencing the Antarctic Soil Metagenome

Pyrosequencing endorsed the scientists to a very valuable tool for accessing a huge amount of nucleotidic information in a short time (ca. 100-fold increase in throughput over Sanger sequencing technology), which in combination with powerful bioinformatics platforms allowed the assessment of the microbial status in almost every soil sample (Margulies et al. 2005).

This high-throughput (HT) sequencing technology has also been used on Antarctic soil samples for the analysis of microbial biodiversity and functional potential of the indigenous community. By this means, it was possible to assess which are the "active" genes and what groups are predominantly present and also to monitor changes in the relative abundance of genetic features (including taxonomic and functional traits) (Table 2).

A study performed on soil samples from Mars Oasis demonstrated that when HT-sequencing technologies are applied to cold environments, there seems to be a much greater heterogeneity at the sequence level in opposition to the apparent limitation at the generic level (with *Proteobacteria* and *Actinobacteria* being common) for Antarctic terrestrial environments than previously thought (Pearce et al. 2012).

A comparative study on the metagenomes of cyanobacterial mats from both Arctic and Antarctic ice shelves demonstrated similar protein-encoding gene distribution in both poles. *Proteobacteria*, *Actinobacteria*, and *Cyanobacteria* were found to be dominant whereas their relative abundances are different: cyanobacterial genes seem to be more prevalent in Antarctic metagenomes whereas *Actinobacteria* and *Alphaproteobacteria* genes are more abundant in the Arctic. Also, dependent on

Table 2 Sequenced Antarctic soil metagenomes

Sample	Library construction	Sequencing method	Reference
Sediment soil	Fosmid/metagenomic digested DNA	End sequencing/454 pyrosequencing	(Pearce et al. 2012)
Pond soil	Not performed	454 pyrosequencing	(Varin et al. 2012)
Frozen soil	Plasmid/amplified 16S rRNA fragments	DNA sequencing	(Stomeo et al. 2012)

DOI 10.1007/978-1-4614-6418-1_623-3

the type of stress or pollutants to which both ecosystems are subjected, different homeostasis genes are prevalently observed in both poles, evidencing that diverse mechanisms of adaptation to cold and stress are selected and are distributed in the major bacterial groups (Varin et al. 2012).

Finally, a study performed on the most extensive ice-free region in Antarctica showed abiotic variations in the bacterial communities and their metabolic activity depending on the location and depth from where they live (Stomeo et al. 2012).

Concluding Remarks

In opposition to what was believed, Antarctic soil is quite a rich area considering the bacterial diversity. The development of novel DNA-sequencing technologies in combination with efficient screening techniques has unlocked the access to the characterization of microbial communities from the white continent.

Metagenomics endorsed scientists with a set of extremely powerful approaches to reach what during decades remained hidden: the unknown microbial majority. From this milestone, an entirely new field of research opened before our eyes.

Biotechnology, health sciences, pharmacology, and many other disciplines gained the possibility to develop profitable novel enzymes discovered by functional metagenomics, making scientists to turn their attention at the extreme environments as very attractive locations to study.

References

Amann RI, Ludwig W, et al. Phylogenetic identification and in situ detection of individual microbial cells without cultivation. Microbiol Rev. 1995;59(1):143–69.

Berlemont R, Delsaute M, et al. Insights into bacterial cellulose biosynthesis by functional metagenomics on Antarctic soil samples. ISME J. 2009;3(9):1070–81.

Berlemont R, Pipers D, et al. Exploring the Antarctic soil metagenome as a source of novel cold-adapted enzymes and genetic mobile elements. Rev Argent Microbiol. 2011;43(2):94–103.

Berlemont R, Jacquin O, et al. Novel cold-adapted esterase MHlip from an Antarctic soil metagenome. Biology. 2013;2:177–88.

Bidle KD, Lee S, et al. Fossil genes and microbes in the oldest ice on earth. Proc Natl Acad Sci U S A. 2007;104(33):13455–60.

Cieslinski H, Dlugolecka A, et al. An MTA phosphorylase gene discovered in the metagenomic library derived from Antarctic top soil during screening for lipolytic active clones confers strong pink fluorescence in the presence of rhodamine B. FEMS Microbiol Lett. 2009;299(2):232–40.

Daniel R. The soil metagenome–a rich resource for the discovery of novel natural products. Curr Opin Biotechnol. 2004;15(3):199–204.

Daniel R. The metagenomics of soil. Nat Rev Microbiol. 2005;3(6):470–8.

Gerday C, Aittaleb M, et al. Cold-adapted enzymes: from fundamentals to biotechnology. Trends Biotechnol. 2000;18(3):103–7.

Handelsman J. Metagenomics: application of genomics to uncultured microorganisms. Microbiol Mol Biol Rev. 2004;68(4):669–85.

Handelsman J, Rondon MR, et al. Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. Chem Biol. 1998;5(10):R245–9.

DOI 10.1007/978-1-4614-6418-1_623-3

Handelsman J, Liles M, et al. Cloning the metagenome: culture-independent access to the diversity and functions of the uncultivated microbial world. In: Brendan W, Nick D (eds.) Methods in microbiology – functional microbial genomics. Academic Press, New York; 2002. p. 241–255.

Heath C, Hu XP, et al. Identification of a novel alkaliphilic esterase active at low temperatures by screening a metagenomic library from antarctic desert soil. Appl Environ Microbiol. 2009;75(13):4657–9.

Margesin R, Miteva V. Diversity and ecology of psychrophilic microorganisms. Res Microbiol. 2011;162(3):346–61.

Margulies M, Egholm M, et al. Genome sequencing in microfabricated high-density picolitre reactors. Nature. 2005;437(7057):376–80.

Marx JC, Collins T, et al. Cold-adapted enzymes from marine Antarctic microorganisms. Mar Biotechnol (NY). 2007;9(3):293–304.

Pearce DA, Newsham KK, et al. Metagenomic analysis of a southern maritime antarctic soil. Front Microbiol. 2012;3:403.

Rappe MS, Giovannoni SJ. The uncultured microbial majority. Annu Rev Microbiol. 2003;57:369–94.

Riesenfeld CS, Schloss PD, et al. Metagenomics: genomic analysis of microbial communities. Annu Rev Genet. 2004;38:525–52.

Rondon MR, Raffel SJ, et al. Toward functional genomics in bacteria: analysis of gene expression in *Escherichia coli* from a bacterial artificial chromosome library of *Bacillus cereus*. Proc Natl Acad Sci U S A. 1999;96(11):6451–5.

Staley JT, Konopka A. Measurement of in situ activities of nonphotosynthetic microorganisms in aquatic and terrestrial habitats. Annu Rev Microbiol. 1985;39:321–46.

Stomeo F, Makhalanyane TP, et al. Abiotic factors influence microbial diversity in permanently cold soil horizons of a maritime-associated Antarctic Dry Valley. FEMS Microbiol Ecol. 2012;82(2):326–40.

Tebbe CC, Vahjen W. Interference of humic acids and DNA extracted directly from soil in detection and transformation of recombinant DNA from bacteria and a yeast. Appl Environ Microbiol. 1993;59(8):2657–65.

Varin T, Lovejoy C, et al. Metagenomic analysis of stress genes in microbial mat communities from Antarctica and the High Arctic. Appl Environ Microbiol. 2012;78(2):549–59.