Insights into novel antimicrobial compounds and antibiotic resistance genes from soil metagenomes

Alinne P. de Castro[1], Gabriel da R. Fernandes[2] and Octávio L. Franco[1,2]*

[1] Programa de Pós-Graduação em Biotecnologia, Universidade Católica Dom Bosco, Laboratórios Inova, Campo Grande, Brazil

[2] Programa de Pós-Graduação em Ciências Genômicas e Biotecnologia, Centro de Analises Proteomicas e Bioquimicas, Universidade Católica de Brasília, Brasilia, Brazil

In recent years a major worldwide problem has arisen with regard to infectious diseases caused by resistant bacteria. Resistant pathogens are related to high mortality and also to enormous healthcare costs. In this field, cultured microorganisms have been commonly focused in attempts to isolate antibiotic resistance genes or to identify antimicrobial compounds. Although this strategy has been successful in many cases, most of the microbial diversity and related antimicrobial molecules have been completely lost. As an alternative, metagenomics has been used as a reliable approach to reveal the prospective reservoir of antimicrobial compounds and antibiotic resistance genes in the uncultured microbial community that inhabits a number of environments. In this context, this review will focus on resistance genes as well as on novel antibiotics revealed by a metagenomics approach from the soil environment. Biotechnology prospects are also discussed, opening new frontiers for antibiotic development.

Keywords: antibiotics, metagenome, soil, resistance genes, drug development

INTRODUCTION

Nowadays one of the most intractable worldwide health problems involves treating infections that are resistant to antibiotics. Resistant pathogens are able to cause high mortality and consequently impose huge healthcare costs (Carlet et al., 2011). Until recently, antibiotics were used only for treating human infections. Now, however, antibiotics are being extensively used in agriculture, food industries, or veterinary practices, causing a high impact on natural environments and consequently on human health (Radhouani et al., 2014). This situation of *bacterial* gene resistance created by widespread and imprudent use of antibiotics has triggered a more energetic search for alternative compounds with deleterious activities against microbial infectious diseases, as well as the identification of pathways or genes related to resistance to traditional antibiotics.

In this regard, over the last decade, the developments of culture-independent approaches have allowed additional insights into the diversity of antimicrobial compounds and antibiotic resistance genes from different environments. Basically, culture-independent analyses are based on molecular methods, including the extraction, amplification, sequencing, and analysis of nucleic acids from environmental samples. Among these, the metagenomics approach has revolutionized knowledge about the vast majority of not-yet-culturable microbial communities. This idea, coined by Handelsman et al. (1998), briefly consists of direct or indirect DNA extraction from a microbial community in its natural habitat, bypassing microbial isolation, and traditional culturing methods.

In recent decades, cultured microorganisms were the exclusive source from which to isolate and clone antibiotic resistance genes or identify antimicrobial activity, since most of the microbial diversity was lost when researchers tried to grow them in standard laboratory culture medium (Hugenholtz et al., 1998). For this reason, metagenomics is a reliable alternative approach to reveal the potential reservoir of antimicrobial compounds and antibiotic resistance genes in the uncultured microbial community that inhabits the environments (**Figure 1**).

Metagenomics has also been considered a promising approach for the isolation of unusual antibiotics from environmental samples, as well as in identifying the mechanisms of *bacterial* resistance in the *in situ* microbial community. The combination of metagenomics with next-generation deep sequencing has brought great progress to the field of antimicrobial resistance and compounds, giving a more feasible representation of the origins and mechanisms of resistance genes (Forsberg et al., 2012; McGarvey et al., 2012). Although the metagenomic approach has many applications in biological sciences, this approach has several limitations, especially for data analysis through a homology-based approach (Prakash and Taylor, 2012). Usually, the functional annotations from metagenomic sequences are reached by homology-based approaches to a publicly available reference sequence (Prakash and Taylor, 2012). Although there are numerous pipelines dedicated to functional analysis of metagenomic data, these pipelines can only detect previously characterized genes that are similar to the newly identified genes, and therefore unknown novel bioactive molecules are missed. For this reason, activity-based functional screening of metagenomic libraries is sometimes still the most appropriate way to identify and characterize resistance genes and antimicrobial compounds (Su et al., 2014) rather than searching for genes
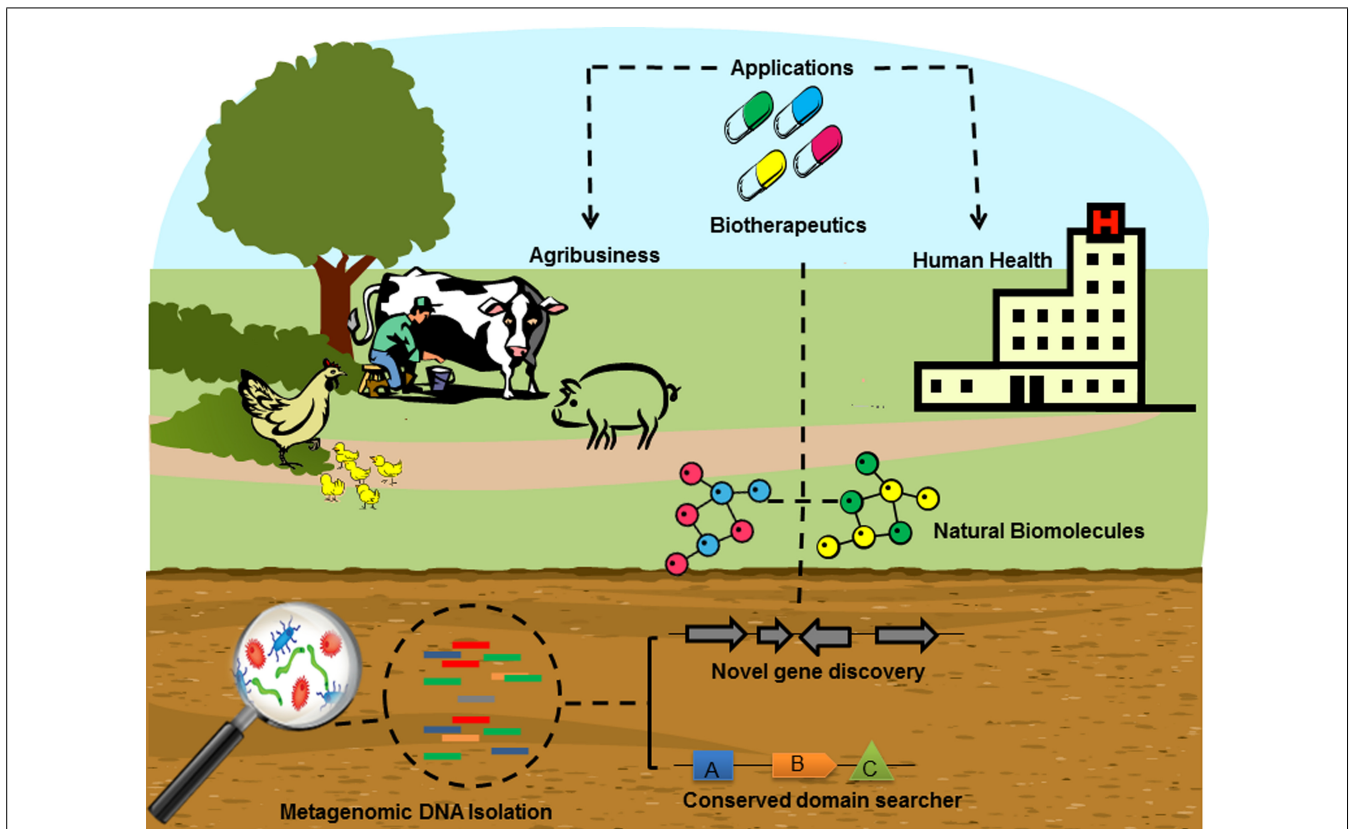
FIGURE 1 | Schematic depiction of the typical screening for novel antimicrobial compounds and antibiotic resistance genes from the soil environment through metagenomics. After collecting soil samples, the metagenomic DNA is extracted and sequenced from a microbial community in its natural habitat, bypassing microbial isolation and traditional cultivation methods, generating several million reads. Once coding sequences have been obtained, their corresponding antimicrobial compounds can be sought through conserved domain search or novel gene discovery in the reference functional databases by *in silico* analysis. Complementary methods reconstruct the identification of the biomolecules of interest. Large-scale production of the target molecule is then carried out for various biotechnological applications including agribusiness and human health.

in public sequence bases. Nevertheless, using this activity-based approach, the genes that have been discovered are evolutionarily distant from known resistance genes in the public databases (Su et al., 2014).

This review will focus on soil samples, which are characterized as a complex and dynamic environmental system, comprising higher microbial diversity of *bacteria,* archaea, fungi, viruses, and protozoa (Young and Crawford, 2004), when compared to other natural environments such as freshwater or extreme habitats (Sleator et al., 2008).

The ecology and activity of soil microbial communities depend on biotic or/and abiotic factors such as soil pH, nutrient availability, water availability, and vegetation cover aboveground (Fierer et al., 2007).

In this context, the present review will attempt a broad literature investigation into diversity and abundance of resistance genes as well as novel antimicrobial compounds, as revealed by the metagenomics approach in the soil microbial community.

SOIL ANTIBIOTIC RESISTANCE GENES

Although the origin of the genes associated with resistance is considered a mystery, there is a link between antibiotic resistance genes in human pathogens and those found in commensal microorganisms, with several common *bacteria* resistance taxa such as *Staphylococcus aureus, Pseudomonas aeruginosa,* and *Klebsiella pneumoniae* coming from the natural environment (Wright, 2010). Generally, *bacterial* resistance to antibiotics can be acquired by horizontal gene transfer (HGT) or by spontaneous mutation in target gene (Hassan et al., 2012). In fact, antibiotic resistance genes could be associated with a transposable element. The mobility of antibiotic resistance genes involves the transference of genetic material to other *bacteria* of the same or different species (Thomas and Nielsen, 2005).

It has previously been reported that antibiotic resistance is everywhere, and consequently efforts are being devoted to understanding the origin of resistance genes, particularly among the vast majority of not-yet-culturable environmental *bacteria.* For instance, the close association between people, animals, and the environment can be responsible for the evolution and spread of antibiotic resistance. For more details, see references (Gautam and Morten, 2014) for an extensive and intense review.

In this context, the environment is constantly exposed to bioactive chemicals produced by a high genetic diversity of plants, protists, *bacteria,* or fungi. For this reason, it is not unexpected

that their microbial community possesses specialized mechanisms to respond to and metabolize small molecules, including antibiotics (Knapp et al., 2010), indicating that the environment is a massive reservoir in which to search for novel resistant organisms (Allen et al., 2010). Indeed, the concept of the resistome on a global scale consists of antibiotic resistance genes detected either free living or as commensals in the environment (D'Costa et al., 2006).

Recently, there have been many investigations into resistant microorganisms in natural environments. These environments include deep sea sediments (Song et al., 2005), pristine cave ecosystems (Bhullar et al., 2012), permafrost soil at depths from 2 to 40 m (Mindlin et al., 2008), urban sewage and wastewater (Graham et al., 2011; Novo et al., 2013), and mainly in the human habitat such as the oral cavity (Xie et al., 2010) or gut microbiota (Schjorring and Krogfelt, 2011).

Although a range of environments have been focused, in this review, special attention will be paid to the soil environment because it has been discussed with more detail in the literature. Some recent studies have demonstrated that soil microbial communities serve as a reservoir for understanding the diversity, abundance, and origins of resistance genes (Knapp et al., 2010, 2011; Ma et al., 2014; Su et al., 2014) and that the *bacterial* community is characterized as opportunistic pathogens, including members of the genera *Escherichia*, *Klebsiella*, *Pseudomonas*, and *Streptomyces*, which are known to be multi-drug resistant and ubiquitous in soil samples worldwide (D'Costa et al., 2007).

In this context, the pioneering study carried out by Riesenfeld et al. (2004) was one of the first to use metagenomics to reveal the potential of soil microorganisms in relation to antibiotic resistance genes. This analysis was based on the construction of a metagenomic library from Wisconsin remnant oak savannah soil. Among the gigabase of cloned DNA, nine clones were confirmed to have resistance against aminoglycoside and tetracycline antibiotics (Riesenfeld et al., 2004). Importantly, all DNA sequences from antibiotic resistance genes were significantly different from those previously reported in the database sequences, demonstrating that there was a higher and unexplored *bacterial* genetic diversity in the soil environment.

The same methodology was used in undisturbed Alaskan soil to search for genes that mainly mediate resistance to β-lactam antibiotics (Allen et al., 2009). This analysis was important for future conclusions because the soil that was collected had no previous exposure to antibiotic compounds, never having undergone anthropic pollution. Their results showed that β-lactamases characterized from soil in Alaska were divergent from those isolated in a clinical environment, indicating that they are more related to β-lactamases ancestral homologs (Allen et al., 2009). This finding was attributed to the uncontaminated nature of the sampling site. Although phylogenetic analysis indicated that the β-lactamases sequences from Alaskan soil metagenomic libraries are not related to β-lactamases sequences from the clinical environment, they were still capable of conferring resistance on *Escherichia coli* strains (Allen et al., 2009).

Another example of the characterization of resistance genes in the soil environment was found by Torres-Cortes et al. (2011). Their results showed a high number of clones that conferred resistance to various types of antibiotics. Nevertheless, only one clone screened codified a dihydrofolate reductase denominated *Tm8-3* with approximately 26.8 kDa, which conferred resistance to a synthetic antibiotic such as trimethoprim. In this regard, it was established that the *Tm8-3* enzyme had similar dihydrofolate reductase activity, but with low amino acid identity (41%) against *dhfr* genes, demonstrating an unassigned reductase targeting trimethoprim resistance. The authors emphasize that a functional metagenomics approach enabled them to discover unassigned enzymes with unexpected activities without prior knowledge about their resistance gene sequences (Torres-Cortes et al., 2011).

Although there is no clear evidence if the resistance genes flow from the soil environment to the clinical, or vice-versa, the recent analysis provided by Forsberg et al. (2012) demonstrated that seven resistance gene cassettes from non-pathogenic cultured Proteo *bacteria* from soil conferred tolerance to five antibiotic classes. Additionally these resistance genes had 100% nucleotide identity in relation to those resistance genes detected in pathogens from clinical isolates, suggesting that soil *bacteria* may share antibiotic resistance genes with human pathogens (Forsberg et al., 2012).

It is now becoming accepted that resistance genes are extremely abundant, diverse, and widely distributed in the soil environment, and that they also show clear similarities to those found in common pathogens. All of the studies mentioned above have recovered multidrug resistance cassettes with formerly unknown functions, based on the entries in the GenBank database. Therefore, the vast majority of them were previously identified as antibiotic gene resistance targeting multiple classes of antibiotics, including bacteriocins, β-lactams, aminoglycosides, vancomycin, chloramphenicol, or tetracyclines.

In fact, the main targets of antibiotic molecules are DNA replication, protein synthesis and cell-wall biosynthesis (Walsh, 2000). The knowledge of antibiotic mechanisms of action is extremely valuable for antibiotic resistance prediction. In this context, functional metagenomics investigations have provided a direction for researchers to prospect not only genes that are functionally active against multiple classes of antibiotics in the non-clinical environments, but also to look for genes with unrelated functions that in some circumstances may confer resistance functions (Pehrsson et al., 2013). For example, in the first instance, a metagenomic library derived from South Korean wetland was screened for esterase activity using tributyrin agar plates, which are usually used for detection of lipolytic microorganisms (Jeon et al., 2011). Using functional screening, one clone of the 6912 cosmid clones had previously been screened with esterase activity revealing the presence of an open reading frame (ORF; *estU1*). Although the activity-based functional screening of metagenomic libraries is one quick approach to identify clones with a target of interest, it is important to note that a main limitation of this approach is that it requires expression of the function of interest in the host cell (Schloss and Handelsman, 2003), and for this reason, the frequency of active clones could be quite low.

Despite the low frequency of positive clones, the only positive clone identified in the metagenomic library derived from South

Korean wetland was overexpressed in *E. coli,* and the amino acid sequence analysis of EstU1 demonstrated the conserved S-X-X-K motif that is typical of family VIII carboxylesterases.

Furthermore, it had previously been reported that family VIII carboxylesterases and class C β-lactamases are phylogenetically related and both are involved in cleaving the lactamic acid ring, conferring *bacterial* resistance. (Wagner et al., 2002). For this reason, more detailed analysis of purified EstU1 protein showed that this single enzyme has both esterase and β-lactamase activities (Jeon et al., 2011) indicating that EstU1 was the first case of a single enzyme that exhibited notable catalytic features, cleaving the β-lactam ring of antibiotics as well as ester substrates.

In particular, functional metagenomic studies offer alternative opportunities to identify new genes and pathways that produce bioactive molecules. However, there are several technical difficulties in the construction of metagenomic libraries, including bias in the DNA extraction, size of inserts and choice of vectors (Delmont et al., 2011a). In fact, direct or indirect techniques used to extract DNA from soils are not completely efficient, due to the fact that different DNA extraction protocols can yield substantially contrast results with respect to purity, quantity, and high-quality DNA recovery (Delmont et al., 2011b). In relation to the size of the inserts, the metagenomic library can be classified into large-insert (>40 kb) or small-insert (<15 kb) depending on higher molecular weight of DNA. The advantages and disadvantages for each method were discussed in (Rajendhran and Gunasekaran, 2008).

Although the use of *bacterial* host has enabled great progress in biological activities from natural environments, the DNA fragments from other microorganisms such as fungi can be challenging to express in a *bacterial* host (Rondon et al., 1999). This problem is caused by incompatibilities between fungal and host molecular biology, such as the presence of introns in fungal genes and the need for protein glycosylation, which is important for visualizing the target activity. However, in order to overcome these obstacles, extensive efforts have been made to use *Streptomyces lividans* and *P. putida* as alternative hosts (Martinez, 2004).

NGS AND AMP/ANTIBIOTIC RESISTANCE GENES PROSPECTION

Since the price of sequencing has fallen with the advances in sequencing technologies, researchers have sought to discover alternatives for analyzing such a large amount of available data. Sequencing of the whole metagenome provides information about the functional potential in the environment and intraspecific variations (Schloissnig et al., 2013). Venter suggested sequences assembly as an alternative to identify new genomes in environmental samples (Rusch et al., 2007). Metagenome assembly is a critical step, since researchers are dealing with an unknown number of different genomes, and the possibility of assembling a chimeric sequence is real. It is well known that NGS platforms produce shorter reads than traditional dideoxynucletide sequencing, and short reads are more difficult to assemble, especially for metagenomics (Raes et al., 2007; Kunin et al., 2008). In order to minimize the effect of this sequence mosaic, bioinformaticians have been dedicated to discovering new assembly algorithms and pipelines, which will now be discussed.

MetaVelvet (Namiki et al., 2012) is an extension of the Velvet de novo assembler (Zerbino and Birney, 2008), which uses de Bruijn graphs to connect short reads in a high coverage to construct the contigs. The extension decomposes the graph built by Velvet into individual subgraphs and then assembles each of them separately. Using this approach, the tools can build longer scaffolds and improve the gene prediction. At a species level, MetaVelvet can cover 94.56% the total metagenome size. This is an improvement when compared to the regular Velvet coverage (60.29%) and SOAPdenovo (84.62%). The assembly procedure also leads to better gene prediction: 81842 genes predicted from MetaVelvet scaffolds, 38445 and 65176 predicted genes from Velvet, and SOAPdenovo scaffolds, respectively, (Namiki et al., 2012). Recently, new powerful pipelines to perform the full analysis have been released. Among them we highlight MOCAT (Kultima et al., 2012) and MetAMOS (Treangen et al., 2013). MOCAT is a modular pipeline developed for the processing, assembly, and gene prediction of metagenomics NGS reads. The first step – quality filtering – is essential to avoid assembly errors related to low quality reads. The scaffolds are built using SOAPdenovo (Li et al., 2010), and in the next step the reads are mapped to the assembled sequences. This remapping procedure resolves the chimeric regions and improves gene prediction. MetAMOS has a similar approach, allowing the user to choose from 20 different analysis tools. This tool also provides a functional annotation step that uses BLAST (Altschul et al., 1997) to annotate the predicted ORFs.

Gene prediction is an important step for metagenomics analysis. Identifying ORFs in mixed environment sequences can be a challenge. New prediction tools have been released to help bioinformaticians to identify genes accurately: MetaGUN (Liu et al., 2013) based on Support Vector Machine; MetaGeneMark (Zhu et al., 2010) uses Hidden Markov Models; Glimmer-MG (Kelley et al., 2012) with Interpolated Markov Models. The use of different approaches to reach consisent results increases the accuracy of gene prediction. Gene prediction will provide a set of candidates that can be tested to identify compounds of biotechnological interest. This procedure was used to describe the human gut microbial gene catalog with 3.3 billion different genes (Qin et al., 2010). Among all these predicted genes we can scan for a specific activity.

The characterization of the predicted genes is the most important step toward identifying new antimicrobial peptides and antibiotic resistance genes. To achieve this goal it is essential to use a good reference database for the annotation process. Gene sequences can be retrieved from databases such as NCBI RefSeq (Tatusova et al., 2014) or from secondary databases such as Antibiotic Resistance Genes Database (ARDB; Liu and Pop, 2009). ARDB is a manually curated database that collects information about antibiotic resistance from most of the public databases. The database organizes the information about 23137 genes into 380 different types (e.g., Beta Lactamases, Multi Drug Transporters). Each gene is also related to at least one of 249 antibiotics described in the ARDB and linked to 1737 species. ARDB comprises data about mechanism of action, ontology, orthology groups, and conserved domains. The gene sequence is provided as well, and this information can be used as a reference for sequence similarity

searches to infer homology to genes predicted in the metagenomic sample. Besides the characterization of the resistance genes, it is also important to identify nucleotides or amino acid variations and link them to the resistance profile.

Among the metagenomics sequences the identification of genes codifying antimicrobial peptides could also be performed. This identification follows two steps: a sequence similarity search using a reference database, and discovery of novel peptides. A secondary database containing curated peptides with antimicrobial activity can be used for a search using BLAST (Altschul et al., 1997). A profile search can also be performed using HMMR (Eddy, 2011) and the reference database. An example of a database that can be used is APD2 (Wang et al., 2009), which organizes known antimicrobial peptides according to families, sources, and targets. The database contains more than 2400 identified compounds: 1986 have antibacterial, 890 antifungal, and 148 antiviral activities. These data can be used as a reference to assign antimicrobial activity to a predicted gene.

In the soil environment there is much information to be discovered. Due to the lack of reference genomes and curated sequences in databases it will not be possible to classify all the predicted genes using similarity searches. Models generated by support vector machines can be used to classify antimicrobial peptides that are not present in reference databases. AMPPred (Porto et al., 2012) is a software that uses five features to predict novel antimicrobial activity from the translated gene sequences. Using the combination of described tools it is possible to extract the maximum information from one's data, and characterize antibiotic resistance mechanisms and antimicrobial peptides.

ANTIMICROBIAL COMPOUNDS FROM SOIL

In addition to the efforts of metagenomics to identify resistance genes that are functionally active against multiple classes of antibiotic, another important attribute of this methodology is the ability to identify several small molecules with antimicrobial activities (MacNeil et al., 2001; Gillespie et al., 2002; Lim et al., 2005). The discovery and screening of novel antibiotics has been an enormous and essential task in recent years. Nowadays, antibiotic resistance poses a global threat to public health (Civljak et al., 2014). On the one hand, bacteria have become more resistant to traditional antibiotic compounds, while on the other, various research groups have selected antimicrobials from different sources including microorganisms, plants, and animals (Roy et al., 2013; de Souza Candido et al., 2014). Among them, the microbial community seems to be the most effective, being the source of many antimicrobial molecules that enable them to live in a very competitive microenvironment. Nevertheless, so far only some microorganisms can be easily detected and cultivated. Metagenomics is a remarkable approach to discover novel entities that could act as a repository of unusual antibiotics in complete environments, such as soil, as focused here.
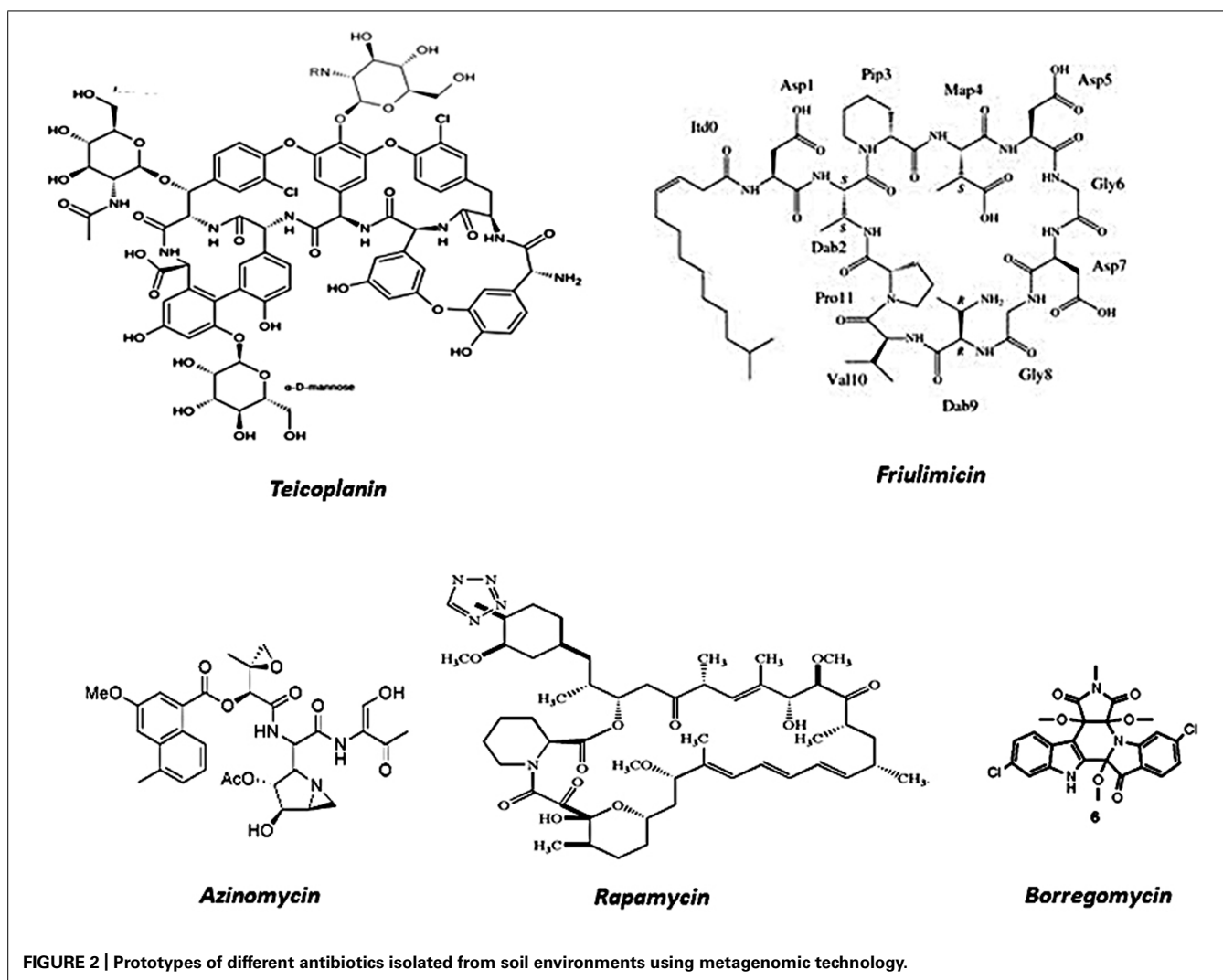
In most habitats, only a small fraction of all existing prokaryotes are acquiescent to cultivation and chemical study. Moreover, there is a strong body of evidence that uncultivated soil diversity characterizes a massive resource of novel biomolecules with biotechnological potential (Wilson and Piel, 2013). In this context, soil *bacteria* seem to be an important source of bioactive

natural products for anti-infective discovery. To select such compounds, several cutting-edge technologies have been applied, including metagenomic library construction, heterologous expression, sequencing of soil microbial communities and single-cell methods (Piel, 2011). Indeed, this workflow starts with DNA isolation, which is typically used for DNA library construction in a suitable host such as *E. coli* (Brady et al., 2009). Moreover, screening efforts rise as insert size decreases, since a high number of clones are needed to cover the target genome. However, the construction of large-insert libraries such as BAC (*bacterial artificial chromosome*) vectors is frequently challenging, due to the amount of DNA obtained and the insufficiency of clone numbers (Piel, 2011). After library construction, the next stage consists of clone identification by harboring genes of interest. The most straightforward screens involve detection of clones exhibiting modified phenotypes, which could be color or appearance based, such as inhibition zones near clones growing on fungal (Chung et al., 2008) or *bacterial* (Rondon et al., 2000) development plates.

An alternative method to phenotypic screens could be based on DNA sequence detection. In spite of higher sequencing costs, one significant advantage over the previous strategy is that natural product pathways can be recognized even in the absence of expression in the library host. This strategy was successfully used for identifying giant multimodular polyketide synthase (PKS) and nonribosomal peptide synthetase (NRPS) clusters (Cane and Walsh, 1999). After phenotype identification a large scale sequencing of enriched bacteria and metagenomes for eDNA gene discovery was applied. Additionally, single-cell analysis has emerged as a powerful strategy to investigate environmental *bacteria* (Wang and Bodovitz, 2010). Isolation of individual cells may be attained by micromanipulation (Frolich and Konig, 1999), flow cytometry (Muller and Nebe-von-Caron, 2010) or microfluidic devices (Weibel et al., 2007). One distinct single cell approach benefit is that genes can be associated with taxonomic information, which is habitually challenging to achieve with metagenomics. This knowledge is of essential importance, since it might lead to the detection of productive taxa and can be valuable for selecting a suitable expression host for cloned genes.

Among the multiple antimicrobial compounds that may be found in nature, a key subject of antibiotic screening from soil samples is polyketides, which are natural products containing multiple β-hydroxyketone or β-hydroxyaldehyde $[-H_2C(=O)CH_2CH(OH)CH_2C(=O)-]$ functional groups. These natural metabolites include the basic chemical structure of multiple compounds with numerous functions such as anticholesteremic and anticancer agents, parasiticides, immunomodulators, and antibiotics (Weissman, 2009).

Since antimicrobial polyketides were discovered, pharmaceutical companies have invested millions of dollars in the search for potent and selective molecules, which have already included geldanamycin, doxycycline, azithromycin, and erythromycin (**Figure 2**; Robinson, 1991). Polyketides are regularly biosynthesized through the decarboxylative condensation of malonyl-CoA derivative extender units in a comparable fatty acid synthesis process. The polyketide chains synthesized by a PKS are frequently

FIGURE 2 | Prototypes of different antibiotics isolated from soil environments using metagenomic technology.

modified into the final version of bioactive natural compounds (Gomes et al., 2013).

The potential of polyketides was immediately associated with metagenomic techniques in order to explore uncultivable soil microorganisms as a remarkable polyketide reservoir (Feng et al., 2011). In order to find novel compounds with biotechnological compounds, (Owen et al., 2013) focused their efforts on a soil sample from the Chihuahuan Desert in southwestern New Mexico. With this aim, conserved regions of PKS and also NRPS biosynthetic machinery were selected to be the target sequences in an experimental framework. By using a combination of PCR amplifications and multiple bioinformatics tools, including Anti-SMASH (Antibiotics and Secondary Metabolite Analysis Shell), the authors were able to identify the gene clusters of biosynthetic pathways encoding novel derivatives of therapeutically applicable *bacterial* natural products (Owen et al., 2013). Amongst those with antitumoral action are thiocoraline-like and tallysomycin-like compounds, as well as several antibiotics including two forms of teicoplanin-like, azinomycin-like, two forms of friumilicin-like and a rapamycin-like product. Moreover, in the same report, one

glycopeptide yielded from analyses was expressed in order to complete the process of drug production, clearly showing that soil samples could be evaluated by an efficient data-generation pipeline associated with software tool analysis.

Similar analyses were performed in soils sampled from 96 locations throughout the southwestern and northeastern USA, considered the most biologically diverse regions in that country (Charlop-Powers et al., 2014). Once more, PKS were focused and a wide range of antibiotic gene clusters was also found. In this case, not only were the compounds focused but also a chemical biogeographic distribution map of medically valuable groups of natural biocompounds, demonstrating once more the potential of metagenome technology (Charlop-Powers et al., 2014).

Additionally, soil DNA isolated from the Anza-Borrego Desert in California was used to select an indolotryptoline-base for a biosynthetic gene cluster, encoding borregomycins (**Figure 2**; Chang and Brady, 2013). A similar approach was also applied to an antitumor substance BE-54017, also from an indolotryptoline gene cluster (Chang and Brady, 2011). Borregomycins may be yielded from a branched tryptophan dimer biosynthetic

metabolic pathway, in which one branch generates borregomycin A, an indolotryptoline antiproliferative agent with the capability to inhibit kinase, and a second branch leading to borregomycins B–D (2–4), which are dihydroxyindolocarbazole agents with anticancer and bactericidal activities. Moreover, BorR clusters were cloned in *Stryptomyces albus* and borregomycins were produced and functionally characterized, clearly showing the complete viability of the metagenome for antibiotic screenings.

Finally and no less importantly, efforts have been made to isolate useful variants from plantaricins. In a study realized by (Pal and Srivastava, 2014), gene-specific primers for plantaricin was used, followed by PCR amplification to select plnE, -F, -J, and -K structural gene amplicons from the soil DNA metagenome. Plantaricin is a cationic, heat-stable peptide belonging to the bacteriocin group (Moll et al., 1999), which is considered a possible alternative to traditional antibiotics against pathogen targets (Cotter et al., 2013). Thus, based on the heterologous production of recombinant plantaricin peptides, the work of (Pal and Srivastava, 2014) demonstrated that plantaricin mature peptides have wide antimicrobial activity spectra, being active against different pathogenic *bacterial* species (Pal and Srivastava, 2014). Plantaricins are short AMPs with the ability to permeate membranes, being extremely dependent on membrane lipid composition. Plantaricins are capable of controlling Gram-positive and -negative bacteria, as well fungi and tumor cells, binding to zwitterionic monolayers and liposomes without significantly penetrating their membranes. Although cholesterol attenuates peptide activity, the association of plantaricin and negatively charged lipid membrane is evident. Furthermore, plantaricins were also able to lead to clear liposome aggregation, suggesting a clear lipid binding mechanism (Zhao et al., 2006). It is important to note that, in accordance with this methodology, it is also possible to achieve a significantly enhanced yield of plantaricin peptides when compared with the purified plantaricin peptides obtained from culturable *Lb. plantarum* strain growth (Pal and Srivastava, 2014), indicating that this methodology can offer an alternative strategy for large-scale production of useful molecules from the natural environment with potential applications in clinical biotechnology research.

BIOTECHNOLOGICAL PROSPECTS AND CONCLUSIVE REMARKS

The problem of bacterial resistance and the absence of novel antibiotics are being felt on all the continents. Our age has been described as a "post-antibiotic era," where people die from relatively simple infections that have been curable for decades. Such resistant pathogens will have devastating implications unless significant action is urgently taken. This situation calls for new approaches to discovering novel antibiotics, and metagenome technology seems to be a useful approach to shed some light on bacterial resistance as well to help mine novel and unusual antimicrobial compounds. Nevertheless, metagenomics is a fledgling area of research, and at the moment only a few reports have reached the point of *in vitro* evaluation (Chang and Brady, 2013; Owen et al., 2013). None have been *in vivo* evaluated in animal models so far, but this is a question of time.

Another important question for the pharmaceutical industry concerns production. Since most antibiotics in the pipeline today have been produced by chemical synthesis, it is desirable to produce microorganisms with the ability to secrete such compounds on a large scale and at low cost. However, several pitfalls have been observed, such as folding posttranslational modifications in the case of antimicrobial proteins and peptides, and the dosage amount for antibiotics that at higher concentrations could kill the *bacterial*-host (Parachin et al., 2012).

Indeed, efforts for the development of novel production systems and fermentation process designs are essential for the establishment of a cost-effective methodology for large-scale production of antibiotics associated with metagenome pipelines. Finally, metagenome screening in soil and other environments must also be performed in order to find unusual compounds and not only molecules from known classes. Of course, this is a vast challenge, but a metagenomic approach could allow us to control the most dangerous resistant and infective bacteria. The contribution of the soil metagenome to this field is clear and, along with other under-explored environments, it might be the source of ground-breaking therapeutic agents in the near future.

REFERENCES

Allen, H. K., Donato, J., Wang, H. H., Cloud-Hansen, K. A., Davies, J., and Handelsman, J. (2010). Call of the wild: antibiotic resistance genes in natural environments. *Nat. Rev. Microbiol.* 8, 251–259. doi: 10.1038/nrmicro2312

Allen, H. K., Moe, L. A., Rodbumrer, J., Gaarder, A., and Handelsman, J. (2009). Functional metagenomics reveals diverse beta-lactamases in a remote Alaskan soil. *ISME J.* 3, 243–251. doi: 10.1038/ismej.2008.86

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., et al. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389–3402. doi: 10.1093/nar/25.17.3389

Bhullar, K., Waglechner, N., Pawlowski, A., Koteva, K., Banks, E. D., Johnston, M. D., et al. (2012). Antibiotic resistance is prevalent in an isolated cave microbiome. *PLoS ONE* 7:e34953. doi: 10.1371/journal.pone.0034953

Brady, S. F., Simmons, L., Kim, J. H., and Schmidt, E. W. (2009). Metagenomic approaches to natural products from free-living and symbiotic organisms. *Nat. Prod. Rep.* 26, 1488–1503. doi: 10.1039/b817078a

Cane, D. E., and Walsh, C. T. (1999). The parallel and convergent universes of polyketide synthases and nonribosomal peptide synthetases. *Chem. Biol.* 6, 319–325. doi: 10.1016/S1074-5521(00)80001-0

Carlet, J., Collignon, P., Goldmann, D., Goossens, H., Gyssens, I. C., Harbarth, S., et al. (2011). Society's failure to protect a precious resource: antibiotics. *Lancet* 378, 369–371. doi: 10.1016/S0140-6736(11)60401–60407

Chang, F. Y., and Brady, S. F. (2011). Cloning and characterization of an environmental DNA-derived gene cluster that encodes the biosynthesis of the antitumor substance BE-54017. *J. Am. Chem. Soc.* 133, 9996–9999. doi: 10.1021/ja2022653

Chang, F. Y., and Brady, S. F. (2013). Discovery of indolotryptoline antiproliferative agents by homology-guided metagenomic screening. *Proc. Natl. Acad. Sci. U.S.A.* 110, 2478–2483. doi: 10.1073/pnas.1218073110

Charlop-Powers, Z., Owen, J. G., Reddy, B. V., Ternei, M. A., and Brady, S. F. (2014). Chemical-biogeographic survey of secondary metabolism in soil. *Proc. Natl. Acad. Sci. U.S.A.* 111, 3757–3762. doi: 10.1073/pnas.1318021111

Chung, E. J., Lim, H. K., Kim, J. C., Choi, G. J., Park, E. J., Lee, M. H., et al. (2008). Forest soil metagenome gene cluster involved in antifungal activity expression in *Escherichia coli*. *Appl. Environ. Microbiol.* 74, 723–730. doi: 10.1128/AEM.01911–1917

Civljak, R., Giannella, M., Di Bella, S., and Petrosillo, N. (2014). Could chloramphenicol be used against ESKAPE pathogens? A review of in vitro data in the literature from the 21st century. *Expert Rev. Anti Infect. Ther.* 12, 249–264. doi: 10.1586/14787210.2014.878647

Cotter, P. D., Ross, R. P., and Hill, C. (2013). Bacteriocins – a viable alternative to antibiotics? *Nat. Rev. Microbiol.* 11, 95–105. doi: 10.1038/nrmicro2937

D'Costa, V. M., Griffiths, E., and Wright, G. D. (2007). Expanding the soil antibiotic resistome: exploring environmental diversity. *Curr. Opin. Microbiol.* 10, 481–489. doi: 10.1016/j.mib.2007.08.009

D'Costa, V. M., Mcgrann, K. M., Hughes, D. W., and Wright, G. D. (2006). Sampling the antibiotic resistome. *Science* 311, 374–377. doi: 10.1126/science.1120800

de Souza Candido, E., e Silva Cardoso, M. H., Sousa, D. A., Viana, J. C., De Oliveira-Junior, N. G., Miranda, V., et al. (2014). The use of versatile plant antimicrobial peptides in agribusiness and human health. *Peptides* 55C, 65–78. doi: 10.1016/j.peptides.2014.02.003

Delmont, T. O., Robe, P., Cecillon, S., Clark, I. M., Constancias, F., Simonet, P., et al. (2011a). Accessing the soil metagenome for studies of microbial diversity. *Appl. Environ. Microbiol.* 77, 1315–1324. doi: 10.1128/AEM.01526–1510

Delmont, T. O., Robe, P., Clark, I., Simonet, P., and Vogel, T. M. (2011b). Metagenomic comparison of direct and indirect soil DNA extraction approaches. *J. Microbiol. Methods* 86, 397–400. doi: 10.1016/j.mimet.2011.06.013

Eddy, S. R. (2011). Accelerated profile HMM searches. *PLoS Comput. Biol.* 7:e1002195. doi: 10.1371/journal.pcbi.1002195

Feng, Z., Kallifidas, D., and Brady, S. F. (2011). Functional analysis of environmental DNA-derived type II polyketide synthases reveals structurally diverse secondary metabolites. *Proc. Natl. Acad. Sci. U.S.A.* 108, 12629–12634. doi: 10.1073/pnas.1103921108

Fierer, N., Breitbart, M., Nulton, J., Salamon, P., Lozupone, C., Jones, R., et al. (2007). Metagenomic and small-subunit rRNA analyses reveal the genetic diversity of *bacteria*, archaea, fungi, and viruses in soil. *Appl. Environ. Microbiol.* 73, 7059–7066. doi: 10.1128/AEM.00358–357

Forsberg, K. J., Reyes, A., Wang, B., Selleck, E. M., Sommer, M. O., and Dantas, G. (2012). The shared antibiotic resistome of soil *bacteria* and human pathogens. *Science* 337, 1107–1111. doi: 10.1126/science.1220761

Frolich, J., and Konig, H. (1999). Rapid isolation of single microbial cell from mixed natural and laboratory populations with the aid of a micromanipulator. *Syst. Appl. Microbiol.* 22, 249–257. doi: 10.1016/S0723-2020(99)80072-1

Gautam, D., and Morten, S. (2014). How to fight back against antibiotic resistance. *Am. Sci.* 102, 42–51. doi: 10.1511/2014.106.42

Gillespie, D. E., Brady, S. F., Bettermann, A. D., Cianciotto, N. P., Liles, M. R., Rondon, M. R., et al. (2002). Isolation of antibiotics turbomycin A and B from a metagenomic library of soil microbial DNA. *Appl. Environ. Microbiol.* 68, 4301–4306. doi: 10.1128/AEM.68.9.4301-4306.2002

Gomes, E. S., Schuch, V., and De Macedo Lemos, E. G. (2013). Biotechnology of polyketides: new breath of life for the novel antibiotic genetic pathways discovery through metagenomics. *Braz. J. Microbiol.* 44, 1007–1034. doi: 10.1590/S1517-83822013000400002

Graham, D. W., Olivares-Rieumont, S., Knapp, C. W., Lima, L., Werner, D., and Bowen, E. (2011). Antibiotic resistance gene abundances associated with waste discharges to the Almendares River near Havana, Cuba. *Environ. Sci. Technol.* 45, 418–424. doi: 10.1021/es102473z

Handelsman, J., Rondon, M. R., Brady, S. F., Clardy, J., and Goodman, R. M. (1998). Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. *Chem. Biol.* 5, R245–R249. doi: 10.1016/S1074-5521(98)90108-9

Hassan, M., Kjos, M., Nes, I. F., Diep, D. B., and Lotfipour, F. (2012). Natural antimicrobial peptides from *bacteria*: characteristics and potential applications to fight against antibiotic resistance. *J. Appl. Microbiol.* 113, 723–736. doi: 10.1111/j.1365-2672.2012.05338.x

Hugenholtz, P., Goebel, B. M., and Pace, N. R. (1998). Impact of culture-independent studies on the emerging phylogenetic view of *bacterial* diversity. *J. Bacteriol.* 180, 4765–4774.

Jeon, J. H., Kim, S. J., Lee, H. S., Cha, S. S., Lee, J. H., Yoon, S. H., et al. (2011). Novel metagenome-derived carboxylesterase that hydrolyzes beta-lactam antibiotics. *Appl. Environ. Microbiol.* 77, 7830–7836. doi: 10.1128/AEM.05363–5311

Kelley, D. R., Liu, B., Delcher, A. L., Pop, M., and Salzberg, S. L. (2012). Gene prediction with Glimmer for metagenomic sequences augmented by classification and clustering. *Nucleic Acids Res.* 40:e9. doi: 10.1093/nar/gkr1067

Knapp, C. W., Dolfing, J., Ehlert, P. A., and Graham, D. W. (2010). Evidence of increasing antibiotic resistance gene abundances in archived soils since 1940. *Environ. Sci. Technol.* 44, 580–587. doi: 10.1021/es901221x

Knapp, C. W., Mccluskey, S. M., Singh, B. K., Campbell, C. D., Hudson, G., and Graham, D. W. (2011). Antibiotic resistance gene abundances correlate with metal and geochemical conditions in archived Scottish soils. *PLoS ONE* 6:e27300. doi: 10.1371/journal.pone.0027300

Kultima, J. R., Sunagawa, S., Li, J., Chen, W., Chen, H., Mende, D. R., et al. (2012). MOCAT: a metagenomics assembly and gene prediction toolkit. *PLoS ONE* 7:e47656. doi: 10.1371/journal.pone.0047656

Kunin, V., Copeland, A., Lapidus, A., Mavromatis, K., and Hugenholtz, P. (2008). A bioinformatician's guide to metagenomics. *Microbiol. Mol. Biol. Rev.* 72, 557–578. doi: 10.1128/MMBR.00009–8

Li, R., Zhu, H., Ruan, J., Qian, W., Fang, X., Shi, Z., et al. (2010). De novo assembly of human genomes with massively parallel short read sequencing. *Genome Res.* 20, 265–272. doi: 10.1101/gr.097261.109

Lim, H. K., Chung, E. J., Kim, J. C., Choi, G. J., Jang, K. S., Chung, Y. R., et al. (2005). Characterization of a forest soil metagenome clone that confers indirubin and indigo production on *Escherichia coli*. *Appl. Environ. Microbiol.* 71, 7768–7777. doi: 10.1128/AEM.71.12.7768-7777.2005

Liu, B., and Pop, M. (2009). ARDB – antibiotic resistance genes database. *Nucleic Acids Res.* 37, D443–D447. doi: 10.1093/nar/gkn656

Liu, Y., Guo, J., Hu, G., and Zhu, H. (2013). Gene prediction in metagenomic fragments based on the SVM algorithm. *BMC Bioinformatics* 14(Suppl. 5):S12. doi: 10.1186/1471-2105-14-S5-S12

Ma, L., Li, B., and Zhang, T. (2014). Abundant rifampin resistance genes and significant correlations of antibiotic resistance genes and plasmids in various environments revealed by metagenomic analysis. *Appl. Microbiol. Biotechnol.* 98, 5195–5204. doi: 10.1007/s00253-014-5511–5513

MacNeil, I. A., Tiong, C. L., Minor, C., August, P. R., Grossman, T. H., Loiacono, K. A., et al. (2001). Expression and isolation of antimicrobial small molecules from soil DNA libraries. *J. Mol. Microbiol. Biotechnol.* 3, 301–308.

Martinez, A. (2004). Genetically modified *bacterial* strains and novel *bacterial* artificial chromosome shuttle vectors for constructing environmental libraries and detecting heterologous natural products in multiple expression hosts. *Appl. Environ. Microbiol.* 70, 2452–2463 doi: 10.1128/AEM.70.4.2452-2463.2004

McGarvey, K. M., Queitsch, K., and 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.06759–6711

Mindlin, S. Z., Soina, V. S., Ptrova, M. A., and Gorlenko Zh, M. (2008). Isolation of antibiotic resistance *bacterial* strains from East Siberia permafrost sediments. *Genetika* 44, 36–44.

Moll, G. N., Van Den Akker, E., Hauge, H. H., Nissen-Meyer, J., Nes, I. F., Konings, W. N., et al. (1999). Complementary and overlapping selectivity of the two-peptide bacteriocins plantaricin EF and JK. *J. Bacteriol.* 181, 4848–4852.

Muller, S., and Nebe-von-Caron, G. (2010). Functional single-cell analyses: flow cytometry and cell sorting of microbial populations and communities. *FEMS Microbiol. Rev.* 34, 554–587. doi: 10.1111/j.1574-6976.2010.00214.x

Namiki, T., Hachiya, T., Tanaka, H., and Sakakibara, Y. (2012). MetaVelvet: an extension of Velvet assembler to de novo metagenome assembly from short sequence reads. *Nucleic Acids Res.* 40:e155. doi: 10.1093/nar/gks678

Novo, A., Andre, S., Viana, P., Nunes, O. C., and Manaia, C. M. (2013). Antibiotic resistance, antimicrobial residues and *bacterial* community composition in urban wastewater. *Water Res.* 47, 1875–1887. doi: 10.1016/j.watres.2013.01.010

Owen, J. G., Reddy, B. V., Ternei, M. A., Charlop-Powers, Z., Calle, P. Y., Kim, J. H., et al. (2013). Mapping gene clusters within arrayed metagenomic libraries to expand the structural diversity of biomedically relevant natural products. *Proc. Natl. Acad. Sci. U.S.A.* 110, 11797–11802. doi: 10.1073/pnas.1222159110

Pal, G., and Srivastava, S. (2014). Cloning and heterologous expression of plnE, -F, -J and -K genes derived from soil metagenome and purification of active plantaricin peptides. *Appl. Microbiol. Biotechnol.* 98, 1441–1447. doi: 10.1007/s00253-013-5097–5091

Parachin, N. S., Mulder, K. C., Viana, A. A., Dias, S. C., and Franco, O. L. (2012). Expression systems for heterologous production of antimicrobial peptides. *Peptides* 38, 446–456. doi: 10.1016/j.peptides.2012.09.020

Pehrsson, E. C., Forsberg, K. J., Gibson, M. K., Ahmadi, S., and Dantas, G. (2013). Novel resistance functions uncovered using functional

metagenomic investigations of resistance reservoirs. *Front. Microbiol.* 4:145. doi: 10.3389/fmicb.2013.00145

Piel, J. (2011). Approaches to capturing and designing biologically active small molecules produced by uncultured microbes. *Annu. Rev. Microbiol.* 65, 431–453. doi: 10.1146/annurev-micro-090110–102805

Porto, W. F., Pires, A. S., and Franco, O. L. (2012). CS-AMPPred: an updated SVM model for antimicrobial activity prediction in cysteine-stabilized peptides. *PLoS ONE* 7:e51444. doi: 10.1371/journal.pone.0051444

Prakash, T., and Taylor, T. D. (2012). Functional assignment of metagenomic data: challenges and applications. *Brief. Bioinform.* 13, 711–727. doi: 10.1093/bib/bbs033

Qin, J., Li, R., Raes, J., Arumugam, M., Burgdorf, K. S., Manichanh, C., et al. (2010). A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* 464, 59–65. doi: 10.1038/nature08821

Radhouani, H., Silva, N., Poeta, P., Torres, C., Correia, S., and Igrejas, G. (2014). Potential impact of antimicrobial resistance in wildlife, environment and human health. *Front. Microbiol.* 5:23. doi: 10.3389/fmicb.2014.00023

Raes, J., Foerstner, K. U., and Bork, P. (2007). Get the most out of your metagenome: computational analysis of environmental sequence data. *Curr. Opin. Microbiol.* 10, 490–498. doi: 10.1016/j.mib.2007.09.001

Rajendhran, J., and Gunasekaran, P. (2008). Strategies for accessing soil metagenome for desired applications. *Biotechnol. Adv.* 26, 576–590. doi: 10.1016/j.biotechadv.2008.08.002

Riesenfeld, C. S., Goodman, R. M., and Handelsman, J. (2004). Uncultured soil *bacteria* are a reservoir of new antibiotic resistance genes. *Environ. Microbiol.* 6, 981–989. doi: 10.1111/j.1462-2920.2004.00664.x

Robinson, J. A. (1991). Polyketide synthase complexes: their structure and function in antibiotic biosynthesis. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 332, 107–114. doi: 10.1098/rstb.1991.0038

Rondon, M. R., August, P. R., Bettermann, A. D., Brady, S. F., Grossman, T. H., Liles, M. R., et al. (2000). Cloning the soil metagenome: a strategy for accessing the genetic and functional diversity of uncultured microorganisms. *Appl. Environ. Microbiol.* 66, 2541–2547. doi: 10.1128/AEM.66.6.2541-25 47.2000

Rondon, M. R., Raffel, S. J., Goodman, R. M., and Handelsman, J. (1999). 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.* 96, 6451–6455. doi: 10.1073/pnas.96. 11.6451

Roy, A., Mahata, D., Paul, D., Korpole, S., Franco, O. L., and Mandal, S. M. (2013). Purification, biochemical characterization and self-assembled structure of a fengycin-like antifungal peptide from *Bacillus thuringiensis* strain SM1. *Front. Microbiol.* 4:332. doi: 10.3389/fmicb.2013.00332

Rusch, D. B., Halpern, A. L., Sutton, G., Heidelberg, K. B., Williamson, S., Yooseph, S., et al. (2007). The Sorcerer II Global Ocean Sampling expedition: northwest Atlantic through eastern tropical pacific. *PLoS Biol.* 5:e77. doi: 10.1371/journal.pbio.0050077

Schjorring, S., and Krogfelt, K. A. (2011). Assessment of *bacterial* antibiotic resistance transfer in the gut. *Int. J. Microbiol.* 2011:312956. doi: 10.1155/2011/312956

Schloissnig, S., Arumugam, M., Sunagawa, S., Mitreva, M., Tap, J., Zhu, A., et al. (2013). Genomic variation landscape of the human gut microbiome. *Nature* 493, 45–50. doi: 10.1038/nature11711

Schloss, P. D., and Handelsman, J. (2003). Biotechnological prospects from metagenomics. *Curr. Opin. Biotechnol.* 14, 303–310. doi: 10.1016/S0958-1669(03)00067-3

Sleator, R. D., Shortall, C., and Hill, C. (2008). Metagenomics. *Lett. Appl. Microbiol.* 47, 361–366. doi: 10.1111/j.1472-765X.2008.02444.x

Song, J. S., Jeon, J. H., Lee, J. H., Jeong, S. H., Jeong, B. C., Kim, S. J., et al. (2005). Molecular characterization of TEM-type beta-lactamases identified in cold-seep sediments of Edison Seamount (south of Lihir Island, Papua New Guinea). *J. Microbiol.* 43, 172–178.

Su, J. Q., Wei, B., Xu, C. Y., Qiao, M., and Zhu, Y. G. (2014). Functional metagenomic characterization of antibiotic resistance genes in agricultural soils from China. *Environ. Int.* 65, 9–15. doi: 10.1016/j.envint.2013.12.010

Tatusova, T., Ciufo, S., Fedorov, B., O'neill, K., and Tolstoy, I. (2014). RefSeq microbial genomes database: new representation and annotation strategy. *Nucleic Acids Res.* 42, D553–D559. doi: 10.1093/nar/gkt1274

Thomas, C. M., and Nielsen, K. M. (2005). Mechanisms of, and barriers to, horizontal gene transfer between *bacteria*. *Nat. Rev. Microbiol.* 3, 711–721. doi: 10.1038/nrmicro1234

Torres-Cortes, G., Millan, V., Ramirez-Saad, H. C., Nisa-Martinez, R., Toro, N., and Martinez-Abarca, F. (2011). Characterization of novel antibiotic resistance genes identified by functional metagenomics on soil samples. *Environ. Microbiol.* 13, 1101–1114. doi: 10.1111/j.1462-2920.2010.02422.x

Treangen, T. J., Koren, S., Sommer, D. D., Liu, B., Astrovskaya, I., Ondov, B., et al. (2013). MetAMOS: a modular and open source metagenomic assembly and analysis pipeline. *Genome Biol.* 14:R2. doi: 10.1186/gb-2013-14-1-r2

Wagner, U. G., Petersen, E. I., Schwab, H., and Kratky, C. (2002). EstB from *Burkholderia gladioli*: a novel esterase with a beta-lactamase fold reveals steric factors to discriminate between esterolytic and beta-lactam cleaving activity. *Protein Sci.* 11, 467–478. doi: 10.1110/ps.33002

Walsh, C. (2000). Molecular mechanisms that confer antibacterial drug resistance. *Nature* 406, 775–781. doi: 10.1038/35021219

Wang, D., and Bodovitz, S. (2010). Single cell analysis: the new frontier in 'omics'. *Trends Biotechnol.* 28, 281–290. doi: 10.1016/j.tibtech.2010.03.002

Wang, G., Li, X., and Wang, Z. (2009). APD2: the updated antimicrobial peptide database and its application in peptide design. *Nucleic Acids Res.* 37, D933–D937. doi: 10.1093/nar/gkn823

Weibel, D. B., Diluzio, W. R., and Whitesides, G. M. (2007). Microfabrication meets microbiology. *Nat. Rev. Microbiol.* 5, 209–218. doi: 10.1038/nrmicro1616

Weissman, K. J. (2009). Introduction to polyketide biosynthesis. *Methods Enzymol.* 459, 3–16. doi: 10.1016/S0076-6879(09)04601–4601

Wilson, M. C., and Piel, J. (2013). Metagenomic approaches for exploiting uncultivated *bacteria* as a resource for novel biosynthetic enzymology. *Chem. Biol.* 20, 636–647. doi: 10.1016/j.chembiol.2013.04.011

Wright, G. D. (2010). Antibiotic resistance in the environment: a link to the clinic? *Curr. Opin. Microbiol.* 13, 589–594. doi: 10.1016/j.mib.2010.08.005

Xie, G., Chain, P. S., Lo, C. C., Liu, K. L., Gans, J., Merritt, J., et al. (2010). Community and gene composition of a human dental plaque microbiota obtained by metagenomic sequencing. *Mol. Oral Microbiol.* 25, 391–405. doi: 10.1111/j.2041-1014.2010.00587.x

Young, I. M., and Crawford, J. W. (2004). Interactions and self-organization in the soil-microbe complex. *Science* 304, 1634–1637. doi: 10.1126/science.1097394

Zerbino, D. R., and Birney, E. (2008). Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res.* 18, 821–829. doi: 10.1101/gr.074492.107

Zhao, H., Sood, R., Jutila, A., Bose, S., Fimland, G., Meyer, J. N., et al. (2006). Biochimica et Biophysica acta. *Biomembranes* 1758, 1461–1474. doi: 10.1016/j.bbamem.2006.03.037

Zhu, W., Lomsadze, A., and Borodovsky, M. (2010). Ab initio gene identification in metagenomic sequences. *Nucleic Acids Res.* 38:e132. doi: 10.1093/nar/gkq275