African Journal of Biotechnology Vol. 11(92), pp. 15823-15834, 15 November, 2012
Available online at http://www.academicjournals.org/AJB
DOI: 10.5897/AJB12.2544
ISSN 1684–5315 ©2012 Academic Journals

Review

Tapping uncultured microorganisms through metagenomics for drug discovery

Abdelnasser Salah Shebl Ibrahim*, Ali Abdullah Al-Salamah, Ashraf A. Hatamleh, Mohammed S. El-Shiekh, and Shebl Salah S. Ibrahim

Department of Botany and Microbiology, College of Science, King Saud University, P.O. 2455 Riyadh 11351, Kingdom of Saudi Arabia.

Accepted 17 October, 2012

Natural products have been an important historical source of therapeutic agents. Microorganisms are major source of bioactive natural products, and several microbial products including antibiotics, anti-inflammatory, anti-tumour, immunosuppressants and others are currently used as therapeutic agents for human and domestic animals. Most of these products were obtained from cultured environmental microorganisms. However, it is widely accepted that a very large majority of the microorganisms present in natural environments are not readily cultured under laboratory conditions, and therefore are not accessible for drug discovery. Metagenomics is a recent culture-independent approach that has been developed to access the collective genomes of natural bacterial populations. It enables discovery of the diverse biosynthetic pathways encoded by diverse microbial assemblages that are known to be present in the environment but not-yet cultured. Recently, several new bioactive molecules and proteins have been discovered using a metagenomic approach. This review highlights the recent methodologies, limitations, and applications of metagenomics for the discovery of new drugs. Moreover, it shows how a multidisciplinary approach combining metagenomics with other technologies can expedite and revolutionize drug discovery from diverse environmental microorganisms.

Key words: Microbial diversity, metagenomics, natural products, drug discovery, microbial ecology.

INTRODUCTION

The advent of high throughput screening (HTS) and the availability of combinatorial compound libraries has led to a decline in the emphasis on natural products by the major pharmaceutical companies over the past decade (Rouhi, 2003; Gullo et al., 2005). Some companies terminated their natural products screening program despite the unproven success of the new technologies, but this was a premature decision. The perceived failure of combinatorial chemistry was primarily due to unrealistic expectations (Ortholand and Ganesan, 2004; Baltz, 2007). The promised avalanche of new drugs from high-throughput technologies such as combinatorial chemistry is yet to materialize and very few drug leads have reached the market using this new technology. For these reasons and others, the interest in natural products has been rekindled and there has been a growing recognition of natural products that were abandoned prematurely in drug discovery (Ortholand and Ganesan, 2004; Baltz, 2007).

Natural products have proven to be a productive source of lead structures for the development of new therapeutic agents for human, plants and animals therapies. More than 60% of the total drugs currently in the market, and 70% of antibiotics currently used are either natural products or derivatives of natural products. Furthermore, the drugs that were in clinical investigation in the period between 2001 and 2008 were mainly natural products or derivatives of natural products including 74% of all new chemical entities for cancer, 70% of anti-migraine drugs and 66% of anti-hypertensive agents (Koehn and Carter, 2005; Li and Vederas, 2009; Singh and Macdonald, 2010).

Cultured microorganisms have been traditionally used as a recognized source of active therapeutic molecules

*Corresponding author. E-mail: ashebl@ksu.edu.sa. Tel: (+966)-0597359415. Fax: (+966)-014675833.

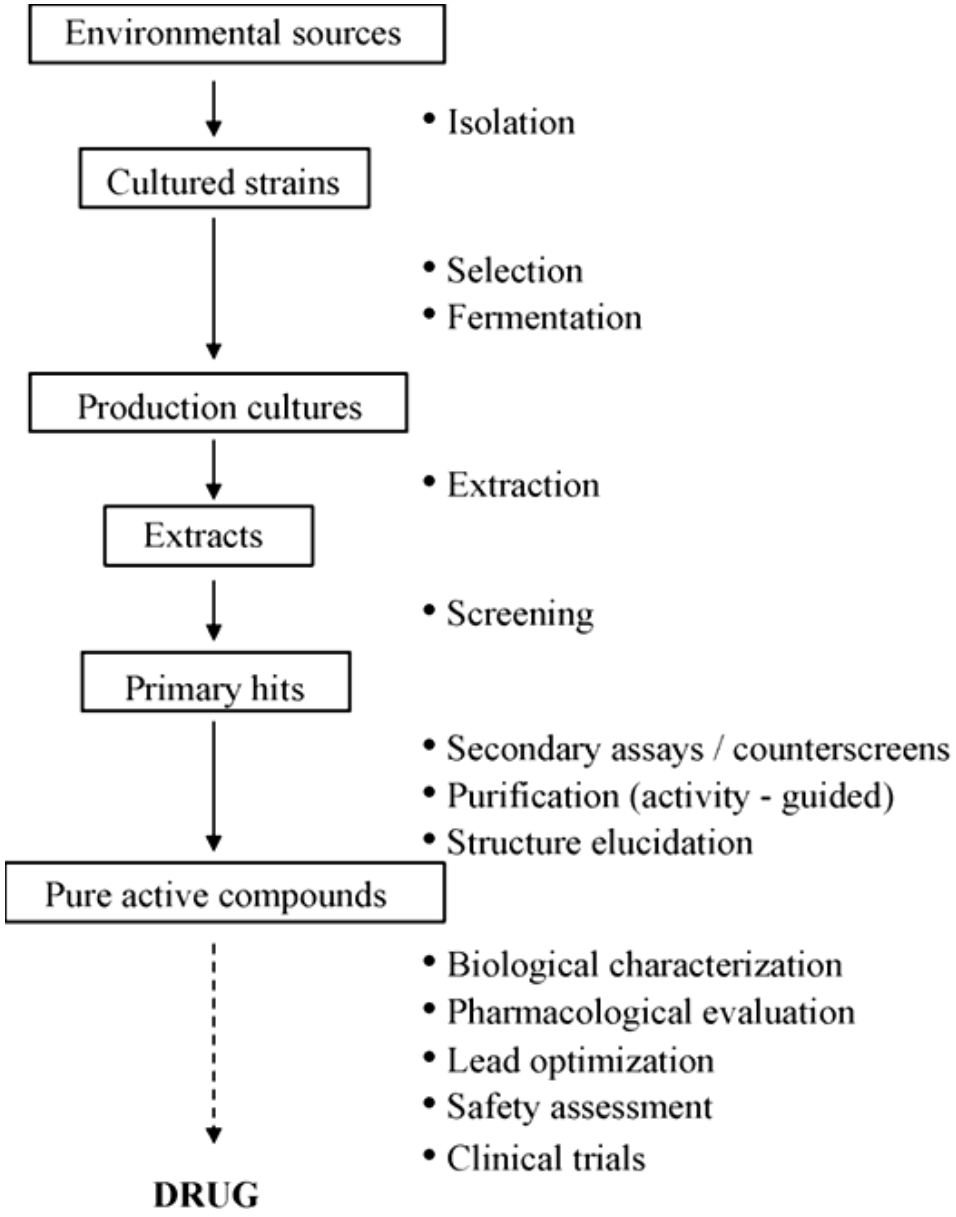

Figure 1. Cultivation-dependent approaches for drug discovery from microbial natural products (Pelaez, 2006).

with original chemical scaffolds. However, the originality of this biological resource has progressively dried up over the past decades, while the increasing request for novel active molecules arose from health care specialists to meet the unmet medical needs (Lefevre et al., 2008; Monier, et al., 2011). Metagenomics is a recent culture-independent approach that has been developed to access the collective genomes of natural bacterial populations (Hirsch et al., 2010). Metagenomic approaches may be integrated into a suitable discovery platform, including high throughput assays targeting pathogenic bacteria and cancer cell lines (Van Lanen and Shen, 2006; Singh and Macdonald, 2010).

CULTIVATION-DEPENDENT APPROACHES FOR DRUG DISCOVERY

Competition for survival among microbes and environmental stress have driven the evolution of defense, attack and signaling diversity, resulting in chemical and biological diversity and potential new drugs. Therefore, microorganisms isolated from the natural environment are an amazing resource of diverse drugs (Singh and Macdonald, 2010). The diversity of environmental microorgani-sms has been exploited for many years based on the cultivation and isolation of microbial species. Figure 1 demonstrates the process of drug discovery from micro-

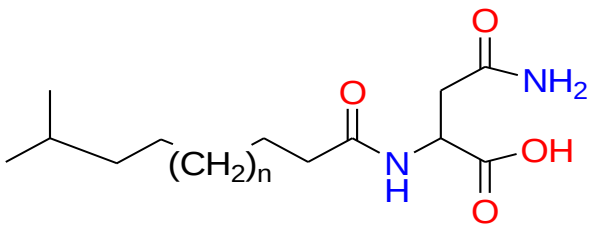

Figure 2. Chemical structures of lipoamides A–C (n = 3: lipoamide A; n = 4: lipoamide B and n = 5: lipoamide C).

bial natural products using a cultivation-dependent approach (Pelaez, 2006). Several bioactive compounds and drugs were discovered using such approach. Wang et al. (2006) reported the discovery of plastencimycin, a Gram-positive broad-spectrum antibiotic found after isolation of 83,000 cultivable strains. The corresponding author of this review and others has recently reported the discovery of new antimicrobial small compounds from marine bacteria using a cultivation-dependent approach (Berrue et al., 2009). Marine bacteria isolated were screened from sediment samples collected in the Bahamas for antimicrobial activity and identified a new *Bacillus pumilus* SP21 strain. A bioassay-guided fractionation of the culture supernatant led to isolation of three new active small compounds named lipoamides A–C (Figure 2). The antimicrobial activity of each compound was evaluated against a panel of pathogens, and showed a promising activity against several pathogens including methicillin-resistant *Staphylococcus aureus* (MRSA), (Berrue et al., 2009). The new compounds were purified from *B. pumilus* SP21 isolated from sediment samples collected in the Bahamas. Lipoamides A–C possesses an activity against several pathogens including MRSA (Berrue, et al., 2009).

However, the number of strains that have been isolated till date is negligible compared to the estimated number of existing bacterial taxa, which are estimated at between 10^7 and 10^9 (Schloss and Handelsman, 2004; Mocali and Benedetti; 2010). This situation has been explained through several molecular ecology studies, which have demonstrated that more than 99% of environmental bacteria are not readily cultivated with classically used culture methods. It has been estimated that less than 0.3% of soil bacteria can be cultured in the laboratory using a single medium type, and this fraction reduces to 0.00001% for water associated microorganisms (Amann et al. 1995; Oren, 2004; Abulencia et al., 2006; Schloss and Handelsman, 2004; Mocali and Benedetti; 2010). This huge reservoir of previously uncultured microorganisms had remained unknown and untapped until the recent development and combination of molecular ecology and molecular biology, which gave rise to a new approach called ‘‘metagenomics’’ (Gillespie et al., 2002; Rouhi, 2003, Schloss and Handelsman, 2004; Gullo et al., 2005; Hirsch et al., 2010; Monier et al., 2011).

METAGENOMICS VERSUS DRUG DISCOVERY

Metagenomic was initially developed as a fundamental research technology for investigation of bacterial population ecology. However, recent improvements and automation of the key steps in a metagenomic approach, including DNA extraction to functional and molecular screens have turned this innovative approach into a valuable discovery tool for the pharmaceutical industries (Banik and Brady, 2010; Fernandez-Arrojo et al., 2010). Drug discovery using a metagenomic approach begins with extraction of genomic DNA from environmental samples without the need for culturing. The DNA is then partially digested or fragmented, ligated into a suitable vector (e.g., fosmid or bacterial artificial chromosome (BAC)) and then transformed into a bacterial host (such as *Escherichia coli* or other) to generate a metagenomic library (Hirsch et al., 2010). These environmental DNA-based methods do not rely on isolation and cultivation of single microorganisms and thus reduce the cultivation bias (Daniel, 2004). A metagenomic library can then be used to study the structural and functional diversity of a microbial assemblage by DNA sequencing or to search for new products based on sequence mining or functional expression (Figure 3), (Daniel, 2004; Singh and Macdonald 2010; Mocali and Benedetti, 2010). Environmental DNA is extracted by direct or indirect lysis of cells from soil particles. Then, recovered DNA is purified, partially digested and ligated into the cloning vector which is introduced into a suitable host cell. The cloning library can be screened for specific functions or sequences (Mocali and Benedetti, 2010).

Extraction of environmental DNA

The initial extraction of genomic DNA from environmental samples is of fundamental importance if representative clone libraries and sequence information are to be obtained (Bertrand et al., 2005). High-quality DNA that has not been fragmented is needed since high-molecular weight DNA is desirable because biosynthetic gene clusters encoding antimicrobial and other compounds of industrial interest typically range from 30 to 100 kb (Singh and Macdonald, 2010). Two strategies are routinely used for metagenomic DNA extraction. The first involves direct *in situ* lysis of microbes from the environmental samples, using freezing and thawing or other lysis methods (Zhou et al., 1996; Lefevre et al., 2008). Although this method is rapid and efficient, it does not discriminate between genetic materials of bacterial, archaeal or eukaryotic origin (Roose-Amsaleg et al., 2001). This method, therefore, usually leads to metagenomic DNA libraries that contain a high proportion of unwanted sequences, and hence reduces the screening efficiency of these libraries. Moreover, humic acid, which is usually co-extracted in this direct lysis method, can cause inhibition of further DNA manipulation (Tsai and Olson, 1992; Robe et al.,

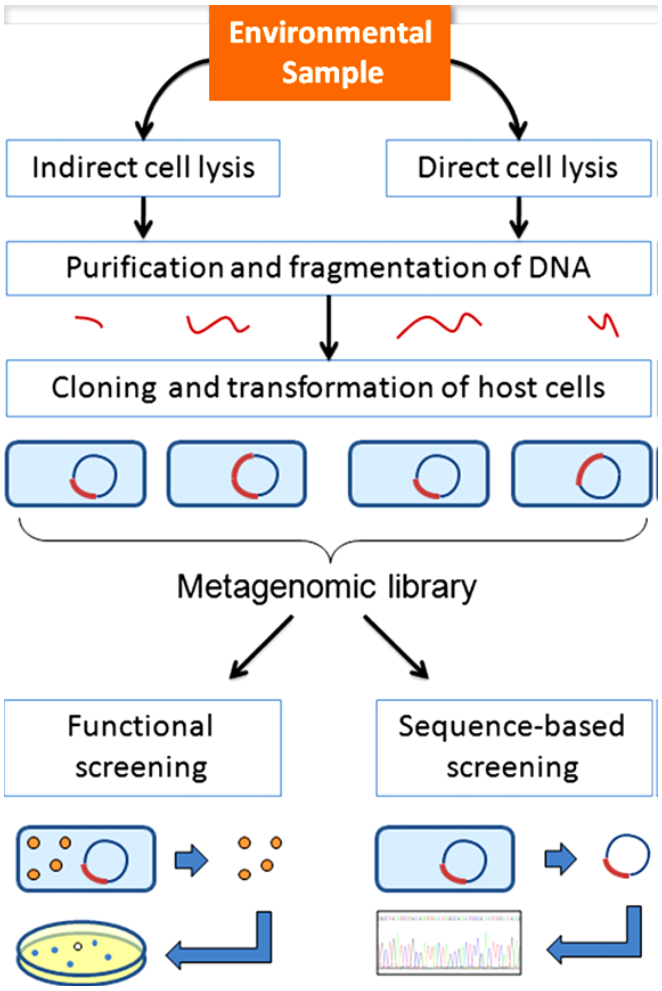

Figure 3. General scheme of the metagenomic approach.

2003; Cowan et al., 2005; Lefevre et al., 2008). Alternatively, an indirect extraction method can be used in which bacterial cells are first separated from the environmental samples before lysis. This approach can overcome problems associated with the direct lysis approach such as the co-extraction of humic substances, interfering contaminant DNA and, hence, it increases the efficiency of subsequent screening (Robe et al., 2003; Singh and Macdonald, 2010). Other methods for environmental DNA extraction include whole-genome amplification (Bodelier et al., 2009) and multiple displacement amplification (Abulencia et al., 2006 and 2008) that can be used in some circumstances when the DNA yield is very low. However, it remains clear that better extraction efficiencies need to be developed for the recovery of good-quality DNA from difficult samples, such as high humic content soils or low nutrient environments (Yun et al., 2004; Singh and Macdonald, 2010).

Genetic diversity in metagenomic libraries

Size fractionation of the metagenomic DNA has a major influence on the screening throughput of metagenomic libraries. Thus, recovery of metagenomic DNA material of high quality and high molecular size is a key step for successful discovery of new metabolite pathways. These complete pathways are fundamental for the majority of metabolite pathways that require several genes for their functional expression. Nevertheless, the generation of large inserts also can lead to a decrease in the statistical population of clones to be tested (Lefevre et al., 2007 and 2008). Metagenomic libraries of more than 150,000 clones can be routinely obtained with standardized insert sizes ranging from 40 to 50 kb using fosmid vectors (Torsvik et al., 2002). Such libraries are considered as a partial representation of the actual genetic diversity of an environmental sample. Results from molecular ecology studies have demonstrated that the initial bacterial diversity contained in a soil sample, for instance, is very important that at least two million such clones would be necessary to statistically consider that all the initial genomes have been included in a metagenomic library (Ginolhac et al., 2004; Lefevre et al., 2008). However, the actual diversity of sequences obtained in a metagenomic library is much higher than that obtained using gene discovery technologies. For instance, bioinformatic analysis of 44 clones randomly chosen among 139 genes encoding polyketide synthase genes recovered from a soil-derived metagenomic library has shown that no clone has a similarity level higher than 67% with all the existing gene databases (Figure 4), (Pettit, 2004; Lefevre et al., 2008).

Selection of the environmental samples

Bioactive compounds that can be discovered using a metagenomic approach directly depend on the microorganisms present in the environmental samples. Hence, it is important that the environmental samples used for metagenomics should be targeted from particular sites to increase the success rate (Wang et al., 2006; Lefevre et al., 2008). For example, if the target is an anti-tumour drug, screening of metagenomes from marine systems should be targeted because terrestrial samples show a low anti-tumour potential (0.01%) compared to those from marine systems (1%) (Pettit, 2004). Similarly, if a drug for detoxification is the target, samples with prior exposure to such toxicant might substantially increase the hit rate (Ginolhac et al., 2004; Singh, 2009).

Several environments have been prospected through metagenomic approaches; including a recent study of the human digestive system (Borruel et al., 2002; Manichanh et al., 2006). This biological niche was very poorly studied previously due to constraints related to the nature of this ecosystem, its difficult accessibility and the physiological specificities of the intestinal microbiota, such as the obligatory anaerobia (Lefevre et al., 2008). However, it was shown using molecular approaches that primary deterioration of human fecal microbiota composition is

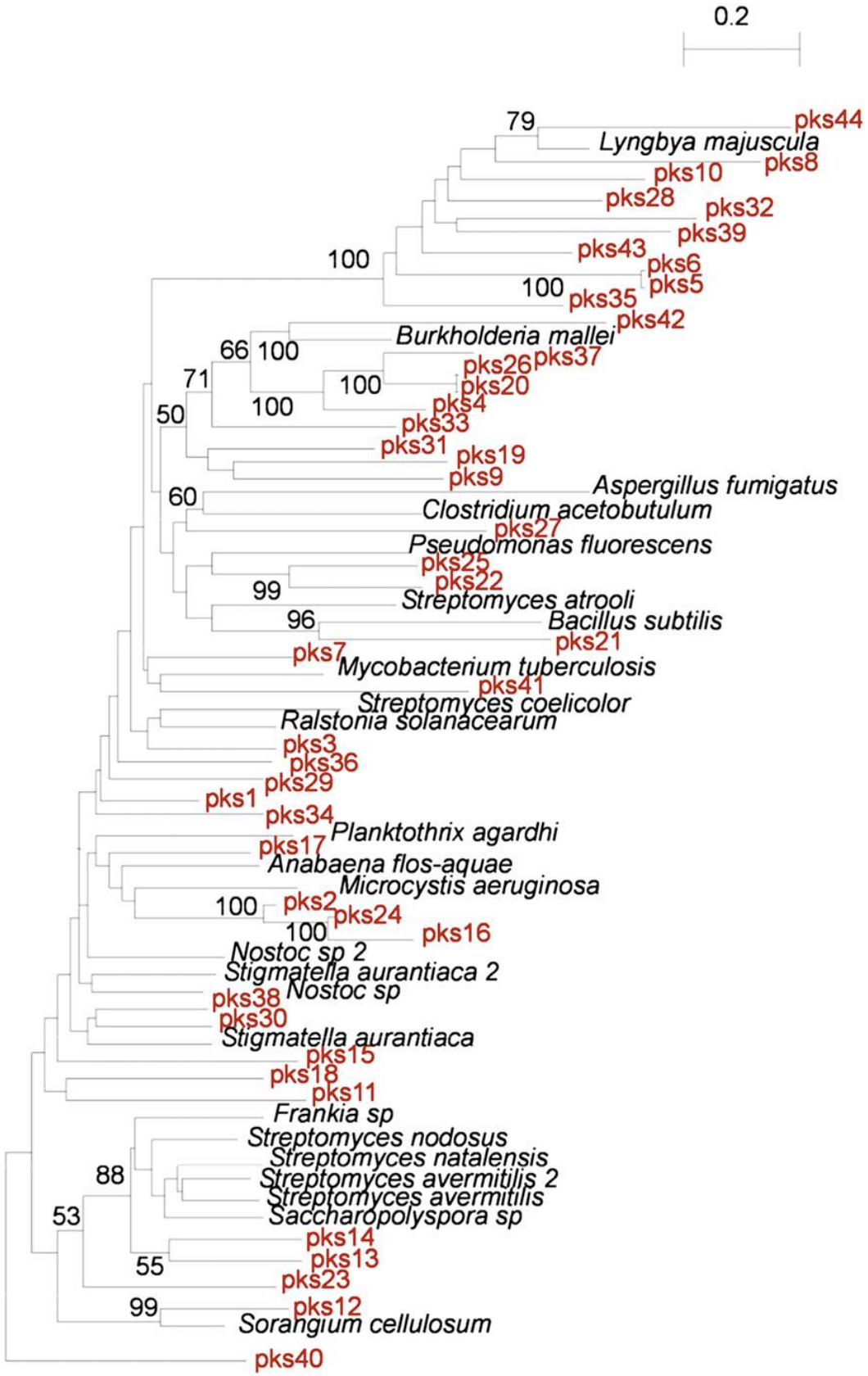

Figure 4. Phylogenetic tree created with sequences of the ketosynthase domain of 46 soil-derived metagenomic clones compared with public ketosynthase domains (in blue and red, metagenomic PKS (Ginolhac et al., 2004; Lefevre et al., 2008).

associated with Crohn's disease (Manichanh et al., 2006). Gut bacteria have been demonstrated to have a central role in the pathogenesis of Crohn's disease and colon cancer (Martin et al., 2004), down regulation of production of tumor necrosis factor (TNF)-alpha (a major cytokines in the mucosal interface) (Borruel et al., 2002), while others may drastically increase the incidence of cancer (Chu et al., 2004), production of activators of protein kinase C, which is involved in the regulation of colonic mucosal proliferation (Vulevic et al., 2004). Therefore, natural interactions between microbes and cells in several specific environments like the human digestive system, skin and vaginal mucosa are particularly promising environments for identification of new bioactive molecules, using appropriate discovery technologies such as metagenomics (Lefevre et al., 2008; Li et al., 2009). In addition, uncultured microorganisms that adapt to live in inhospitable environments such as deep sea hydrothermal vents, the Dead Sea, the Antarctic, alkaline lakes, and volcano soils are considered as a potential source of new chemical entities, and can be tapped using a metagenomic approach (Horikoshi, 1999; Kerkhof and Goodman; 2009; Horikoshi et al., 2011).

Construction and expression of metagenomic libraries

The vectors used in construction of metagenomic libraries have significant impact on the efficiency of the libraries. Plasmid-based vectors are the first choice for small DNA fragments (0.5 to 5 kbp). However, to target entire functional operons or complete pathways, it is desirable to clone larger fragments (>35 kbp) in cosmid or fosmid (both 35 to 45 kbp) vectors or to use BAC vectors (typically 100 kbp). However, because cosmid libraries can be unstable in maintaining prokaryotic or eukaryotic inserts, fosmid vectors are sometimes preferred (Daniel et al., 2004; Long et al., 2005; Bodelier et al., 2009; Singh and Macdonald, 2010). Expression within a heterologous host is another important parameter affecting the efficiency of the metagenomic libraries. *E. coli* is a convenient expression host because of the relative ease of transformation and genetic manipulations, and it is commonly used in industrial fermentations (Gillespie et al., 2002; Knietsch et al., 2003). Other hosts include *Streptomyces* spp., *Pseudomonas* spp. and *Bacillus* spp. (Martinez et al., 2004; Richard et al., 2004). *Streptomyces* spp. genomes synthesize a considerable amount of secondary metabolites and, thus, provide a surrogate organism capable of providing secondary metabolites precursors without the need for further genetic manipulation. Hence, *Streptomyces* spp. are particularly suitable hosts for the detection of new secondary metabolites (Mocali and Benedetti, 2010). However, these alternatives hosts are slow growing in comparison to *E. coli*, which has hampered their application in fermentation industries. Therefore, further improvements in expression vectors and expression in surrogate organisms need to be addressed on a priority basis to increase the hit rate for new products from the metagenomic libraries and natural product production at the industrial scale (Singh and Macdonald, 2010).

Screening of metagenomic libraries

For identification of clones in the metagenomic libraries that produce bioactive metabolites, both expression-dependent (functional) and sequence-based (homology) screening strategies have been used (Kerkhof and Goodman, 2009; Banik and Brady, 2010; Monier et al., 2011). In functional metagenomic studies, DNA libraries are screened in high-throughput assays designed to identify clones that have phenotypes traditionally associated with the production of small bioactive molecules, while in homology-based studies, libraries are probed to identify clones that contain conserved sequences traditionally associated with secondary metabolite biosynthesis. Hits identified in these initial high-throughput assays are subsequently examined for the ability to confer the production of small molecules to model cultured heterologous hosts (Van Lanen and Shen, 2006; Sing, 2009; Singh and Macdonald, 2010)

Functional metagenomic library screening strategy

Direct recombinant expression is the first choice for detecting bioactive secondary metabolites expressed in metagenomic libraries. Each recombinant *E. coli* clone is cultivated in 96-well or 384-well plates and tested for its ability to express an active metabolite compared to the non-transformed reference *E. coli* strain. High-throughput anti-infection assays can be run by spotting *E. coli* recombinant clones on a *Bacillus* spp.-seeded agar plates and appearance of the growth inhibition halo around the positive clones can be detected using automated image analysis of each plate (Figure 5) (Lamprecht et al., 2007; Lefevre et al., 2008). Other tests like anti-proliferative or anti-inflammatory bioassays can be achieved by preparing an extract from each *E. coli* recombinant clone. Then, each extract is tested using classical cell-based assays like cancer cell lines or inflammation reporter cell lines to test its potential activity versus reference compounds (Lamprecht et al., 2007; Lefevre et al., 2008). Using such approach, Gloux et al. (2007) were able to identify new compounds with modulatory activity toward a colon cancer cell line, from clones derived from a gut microbiome.

A variety of new long-chain N-acylated amines (1), as well as a new isonitrile functionalized indole antibiotic (2), were identified by isolation of clone-specific metabolites produced by antibacterially active clones that were screened using top agar overlay assays (Figure 6), (Brady and Clardy, 2000 and 2005). Other small molecule antibiotics have also been detected by examining

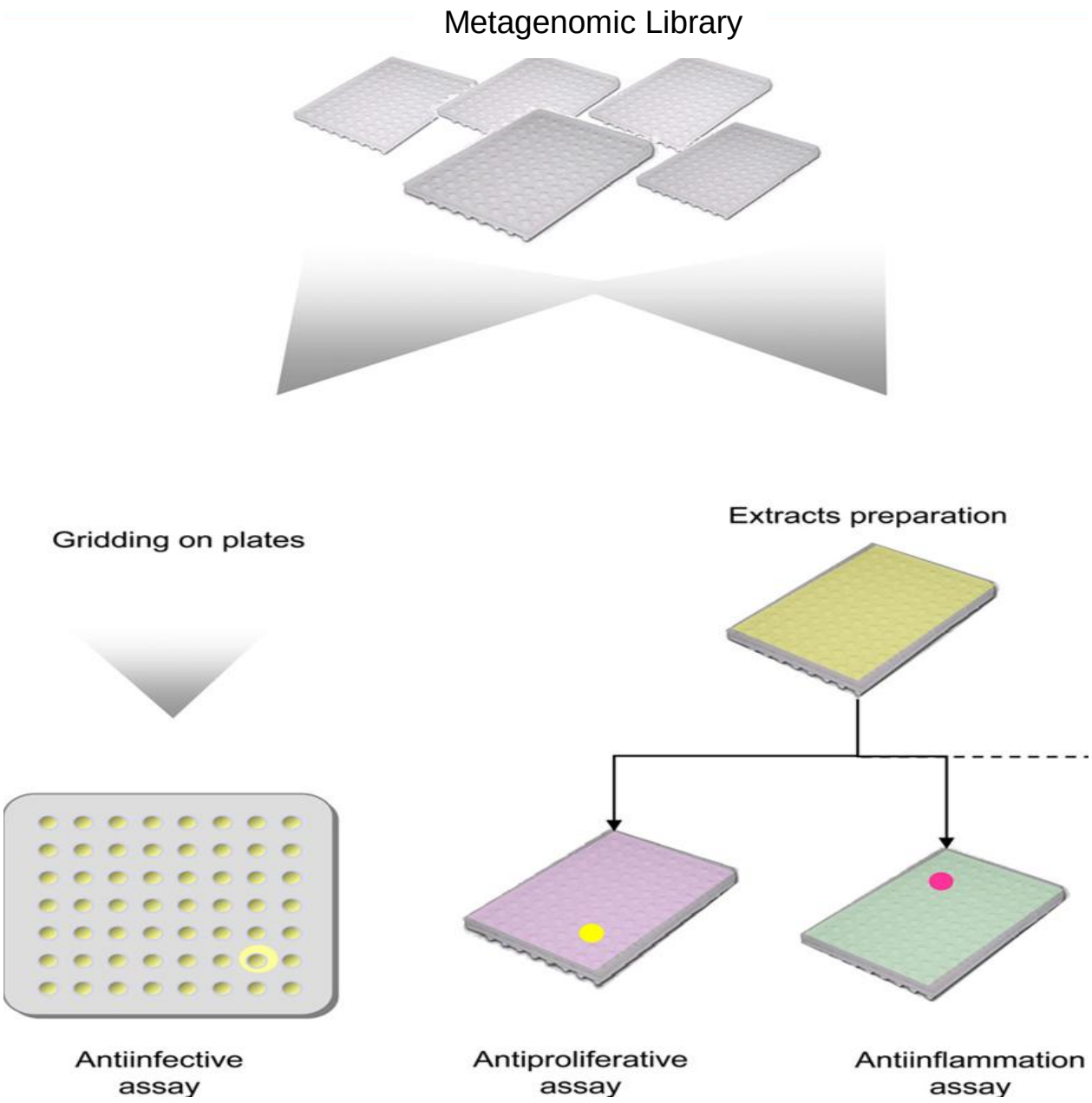

Figure 5. Functional screening of metagenomic libraries for new active metabolites (Lamprecht et al., 2007; Lefevre et al., 2008).

pigmented DNA clones, as well as through the direct examination of culture broth extracts from randomly selected clones. Bioactive compounds identified from these types of studies include the antibacterially active pigments violacein, indigo (three) and the turbomycins (four), all recovered from soil libraries, as well as the known cyclic peptides patellamide D (five) and nocardamine (six), isolated from marine sponge and soil libraries, respectively (Wang et al., 2000; Brady et al., 2000; MacNeil et al., 2001; Gillespie et al., 2002; Lim et al., 2005; Long et al., 2005; Banik and Brady, 2010).

Most functional metagenomic studies focused on small molecule discovery have been carried out in *E. coli*. However, it is unlikely that all of the biosynthetic diversity present in an environmental sample can be functionally accessed using a single heterologous host cell. A computational analysis of promoters and ribosomal binding sites of various sequenced bacteria has indicated that only 40% of the enzymatic activities present within a typical metagenomic sample could be accessed using *E. coli* as a heterologous host (Gabor et al., 2004). Thus, vector–host pairs that allow for the introduction and screening of

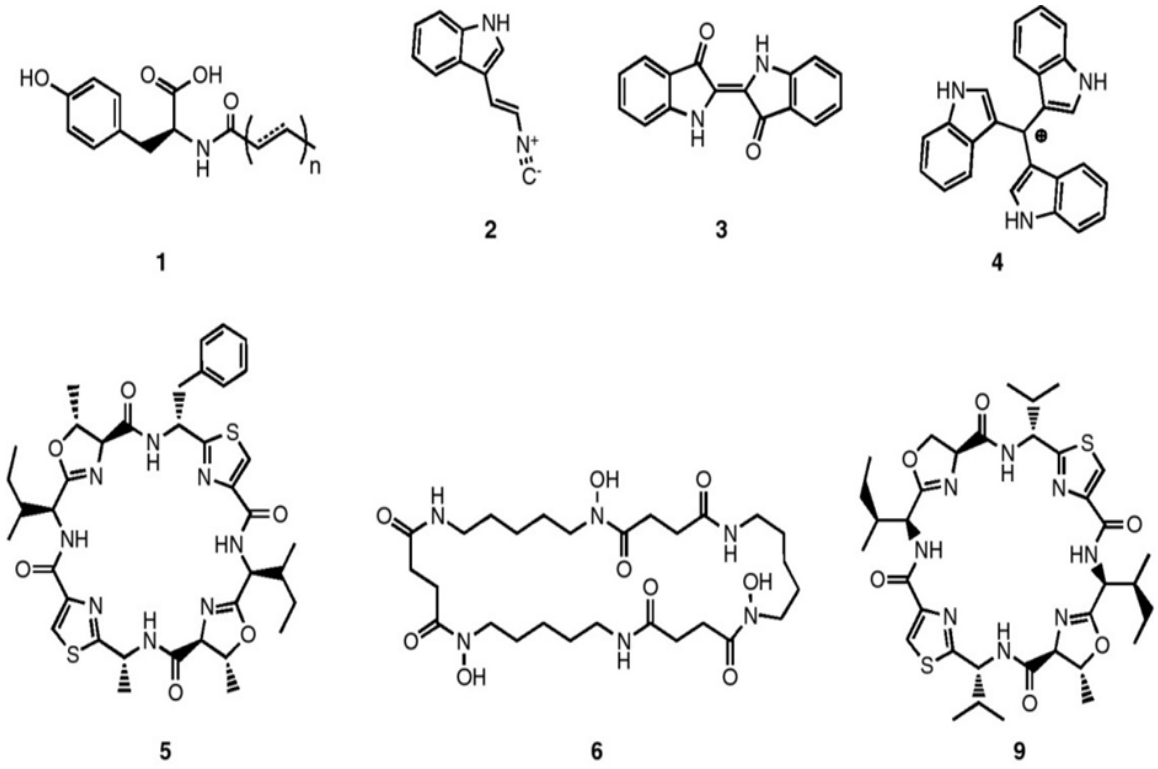

Figure 6. Representative natural products that were discovered using functional metagenomic library screening strategy. 1, N-acyltyrosine; 2, isocyanide functionalized indole; 3, indigo; 4, turbomycin A; 5, cyanobactinpatellamide D; 6, nocardamine and cyanobactin patellamide A (Banik and Brady, 2010).

metagenomic libraries in diverse bacteria have the potential to expand the number and type of bioactive compounds that can be discovered from metagenomic analysis (Aakvik et al., 2009; Craig et al., 2009 and 2010).

Sequence-based metagenomic library screening

Sequence-based strategies or sequence mining includes sequencing of inserts in the whole metagenomic library (or without prior cloning) to identify clones that contain conserved sequences traditionally associated with secondary metabolite biosynthesis. It is advantageous in that, it does not rely on heterologous expression (Singh and Macdonald, 2010). Recent development of automated and high-throughput sequencing technologies and bioinformatics have enabled sequencing of entire clone libraries in short time periods at reasonable cost (Wang et al., 2006). When the entire metagenome sequence is available, searches for particular functions or proteins can be targeted by analysis of the sequence data, and once putative homologous sequences are detected, exact sequence determination and expression of genes can be achieved by polymerase chain reaction (PCR) amplification and expression in suitable host cells (Gullo et al., 2005; Lefevre et al., 2007 and 2008; Singh and Macdonald, 2010). This approach has enhanced the discovery of several novel products including antibiotics and other therapeutic agents via detection of open reading frames (ORF) with sequence homology to known functions. For example several multigenic pathways responsible for synthesis of the antibiotic violacein have been successfully isolated using this approach (Brady and Clardy, 2005; Yang et al., 2009). Piel (2002) reported recovery of the biosynthetic gene cluster for pederin (an anticancer agent originally isolated from the beetle *Paederus fuscipes*) from a beetle-derived metagenomic library, and was shown to originate from an uncultured symbiotic Pseudomonad. Since the pederin-like structures onnamide and psymberin (seven) (Figure 7), had also been isolated from marine sponge extracts, it was hypothesized that these molecules might originate from bacterial symbionts rather than the sponge (Piel, 2002; Fisch et al., 2009; Banik and Brady, 2010; Craig et al., 2010). Zimmermann et al. (2009) recently reported the use of a recombinant O-methyltransferase, PedO, from the pederin biosynthetic gene cluster to site-specifically methylate mycalamide A, resulting in the production of a derivative (eight) that exhibits enhanced antitumor activity (Figure 7). Biosynthetic gene clusters for a number of patellamides, cytotoxic cyclic peptides originally isolated from sponge extracts (five and nine), were identified from a metagenomic library constructed from uncultured cyano-

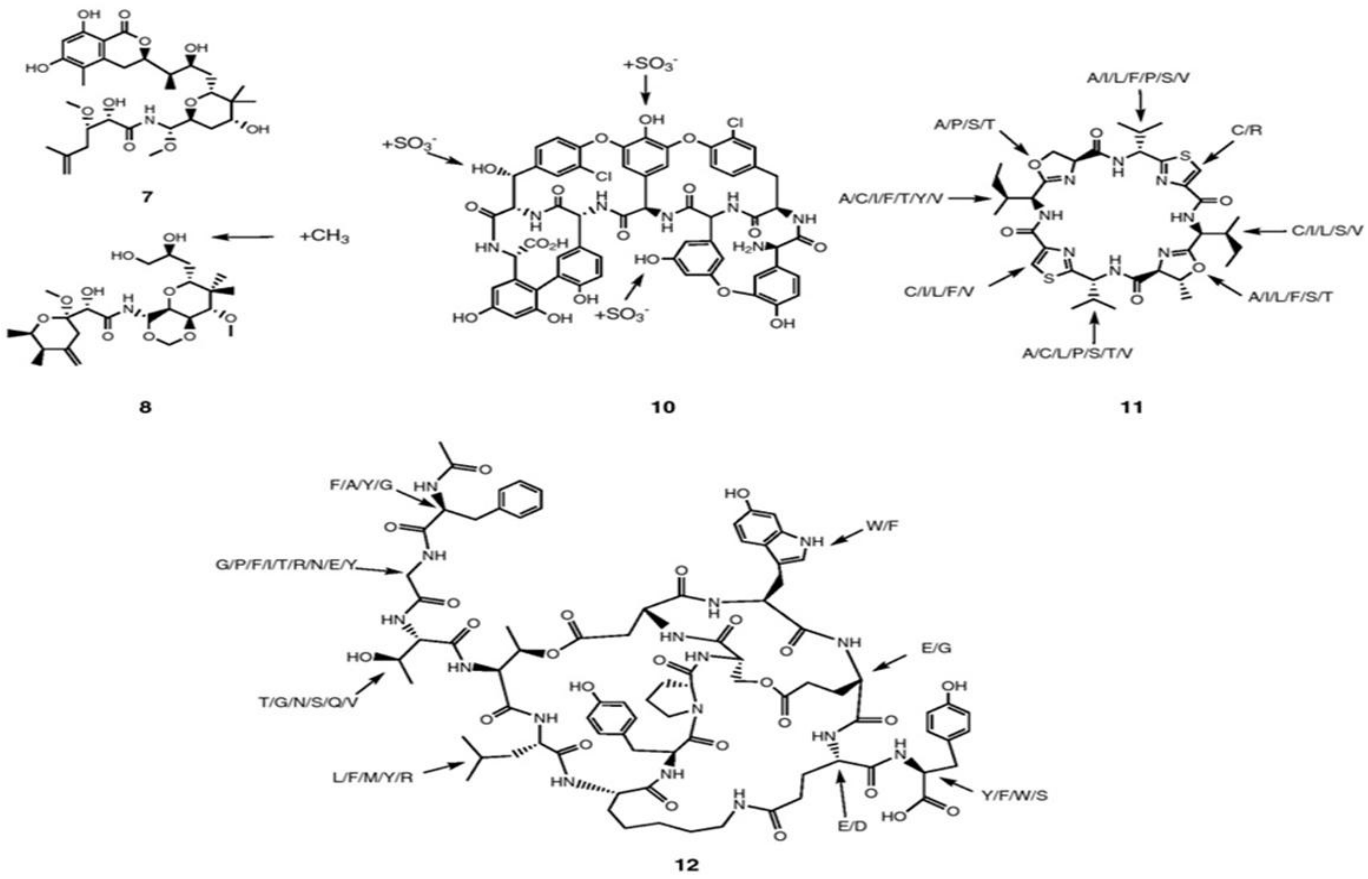

Figure 7. Natural product families that have been discovered using a sequence-based metagenomic library screening strategy. 7, Psymberin; 8, methylated mycalamide A; 10, glycopeptides; 11, patellamides; 12, microviridins.

bacterial symbionts associated with marine Didemnidae sponges (Figure 6) (Schmidt et al., 2005). Recently, Banik and Brady (2008) screened a soil DNA cosmid library for clones containing genes associated with the biosynthesis of teicoplanin-and vancomycin-like glycol-peptides antibiotic and were able to identify a new glycol-peptide biosynthetic cluster that contains unique genes encoding three sulfotransferases. Using the teicoplanin aglycone as a substrate, seven new anionic glycopeptide congeners (10) were generated *in vitro* using the metagenomic DNA-derived sulfotransferases (Figure 7) (Banik and Brady, 2008 and 2010). The cyanobactin and microviridin precursor peptide diversity found in metagenomic studies is displayed on patellamide A and microviridin B, respectively (Banik and Brady, 2010).

The Schmidt group recently reported the PCR amplification of 30 genes encoding novel patellamide-like precursor peptides from uncultured *Prochloron* spp. Symbionts living in consortia with marine sponges (11), (Figure 7) (Schmidt et al., 2009). In another PCR-based study, Ziemert et al. (2010) was able to identify 15 new variants of the gene that encodes for the precursor to the microviridin peptide (12), an important protease inhibitor, using DNA isolated from uncultured freshwater cyano-bacteria of the genera *Microcystis* (Figure 7).

Despite such success in some cases using direct sequencing of clone libraries, it remains costly because a large number of clones need to be screened. Current estimates suggest that millions of clones would be needed to access all of the genomes present within some environmental samples (Henne et al., 2000; Ginolhac et al., 2004). Furthermore, although this approach is not dependent on expression, it relies on the design of pro-bes based on available databases derived from already known genes and, hence, it is unlikely to discover novel bioactive metabolites using such an approach (Singh and Macdonald, 2010; Banik and Brady, 2010).

INTEGRATION OF METAGENOMICS WITH OTHER EMERGING TECHNOLOGIES

Recently, Beloqui et al. (2009) presented a new sensitive metabolite array, called the “reactome array”. This approach reports the direct isolation of enzymes from environ-mental samples without the need for culturing or creating metagenomic libraries. Here, the substrate is

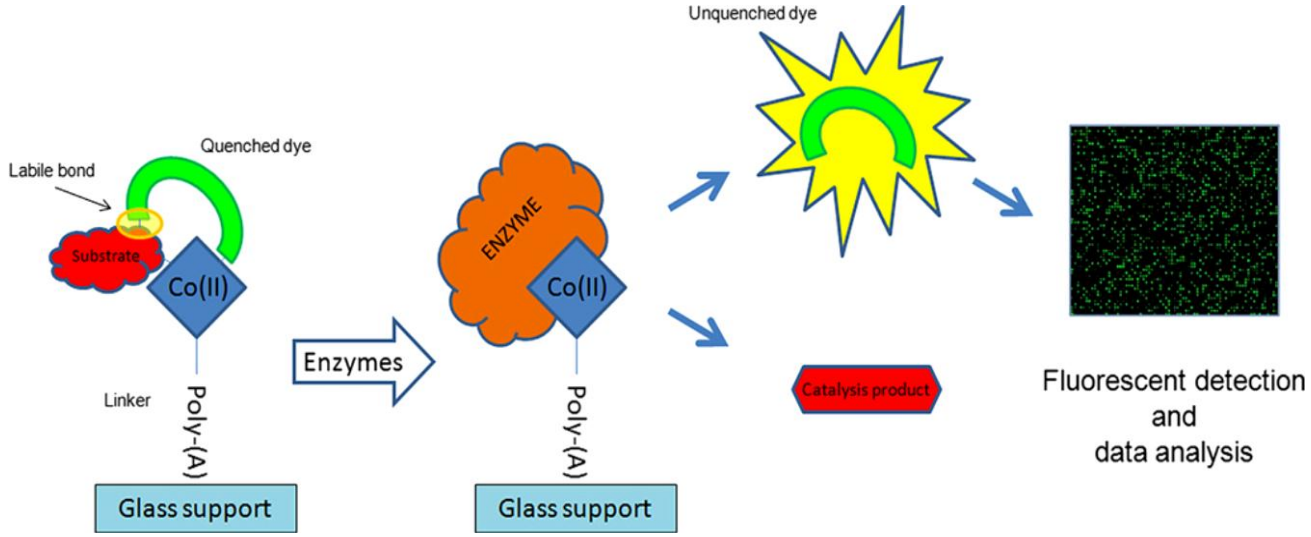

Figure 8. The reactome strategy. The process initially involves three linked molecules: the enzyme substrate-metabolite, the quenched dye (Cy3) and the linker.

linked with a dye and the degradation of a compound is recorded with concomitant release of the dye signal. The enzyme responsible for degradation is captured on nanoparticles coated with the cognate metabolite (Figure 8). Potentially a metagenomic approach could be combined with a reactome array as a novel method to discover novel products. For example, where heterologous expression of gene is a problem but metagenomic sequence analysis suggests the presence of novel microbial products in a particular environmental sample, a reactome array could be used to directly capture that particular protein. Conver-sely, if a reactome array suggests the presence of a parti-cular protein in an environmental sample but the quantity of the product being an issue, protein sequencing can be done to discover the amino acid composition, which can be used to design PCR primers for amplification of the target gene for subsequent cloning and expression in a surrogate host (Beloqui et al., 2009; Mocali and Benedetti, 2010).

The substrate-metabolite is linked to the quencher through a labile nitrogen bond and both are anchored to the Co(II)-containing poly(A) linker by hystidine tags. Enzymatic recognition of the substrate induces a reaction which causes a chemical change and rupture of the labile bond with the release of quenched dye. The consequent release of the reaction product exposes the active cobalt cation that ligates and immobilizes the enzyme on the glass support (array spot). The released dye is no longer quenched and gives a fluorescent signal (Beloqui et al., 2009; Mocali and Benedetti, 2010).

CONCLUSION AND FUTURE PROSPECTS

The huge potential of metagenomics to access environmental microbial diversity, the widest biodiversity on the entire planet, could be a great resource for natural products and drug discovery. Industrialization of this mature technology now offers great opportunities to pharmaceutical companies to enable access to the untapped reservoir of secondary metabolites from previously uncultured bacteria and to converts this genetic diversity into new sources of bioactive molecules and novel chemical entities. Metagenomics also enables detection of new molecules from already known clusters of active compounds, and production of the discovered compounds in well characterized heterologous hosts. Finally, a metagenomic strategy can provide an efficient tool for identifying and following new diseases markers among human microbiomes (Lefevre et al., 2007 and 2008). The discovery of bioactive small molecules using metagenomic methods will undoubtedly benefit greatly from future advances in sequencing technology that allow for the comprehensive sequencing of complex microbiomes, as well as from increasing our understanding of the expression barriers encountered by foreign DNA in model laboratory grown bacterial hosts. Although only a small number of compounds have been characterized to date using culture-independent methods, these initial studies indicate that as-yet-uncultured bacteria are likely to be rich sources of previously unstudied biologically active small molecules.

ACKNOWLEDGEMENT

The authors extend their appreciation to the Deanship of Scientific Research at King Saud University for funding the work through the research group project No RGPVPP-045.

REFERENCES

Aakvik T, Degnes KF, Dahlsrud R, Schmidt F, Dam R, Yu L, Volker U, Ellingsen TE, Valla S (2009). A plasmid RK2-based broadhost-range cloning vector useful for transfer of metagenomic libraries to a variety of bacterial species. FEMS Microbiol. Lett. 296:149-158.

Abulencia CB, Wyborski DL, Garcia JA, Podar M, Chen W, Chang SH, Chang HW, Watson D, Brodie EL, Hazen HC, Keller M (2006). Environmental whole-genome amplification to access microbial populations in contaminated sediments. Appl. Environ. Microbiol. 72:3291-3301.

Abulencia CB, Wyborski DL, Garcia JA, Podar M, Chen W, Chang SH, Chang HW, Watson D (2008). Environmental whole-genome amplification to access microbial populations in contaminated sediments. Appl. Environ. Microbiol. 72(5):3291-3301.

Amann RI, Ludwig W, Schleifer KH (1995). Phylogenetic identification and in-situ detection of individual microbial-cells without cultivation. Microbiol. Rev. 59:143-169.

Baltz RH (2007). Antimicrobials from Actinomycetes: Back to the Future. Microbe 2(3):125-132.

Banik JJ, Brady SF (2008). Cloning and characterization of new glycopeptide gene clusters found in an environmental DNA megalibrary. Proc. Natl. Acad. Sci. USA. 105:17273-17277.

Banik JJ, Brady SF (2010). Recent application of metagenomic approaches toward the discovery of antimicrobials and other bioactive small molecules. Curr. Opin. Microbiol. 13:603-609.

Beloqui A, Guazzaroni ME, Pazos F, Vieites JV, Godoy M, Golyshina OV, Chernikova, TN, Waliczek A (2009). Reactome array: forging a link between metabolome and genome. Science 326:252-257.

Berrue F, Ibrahim AS, Boland P, Kerr RG (2009). Newly isolated marine *Bacillus pumilus* (SP21): A source of novel lipoamides and other antimicrobial agents. Pure Appl. Chem. 81(6):1027-1031.

Bertrand H, Poly F, Van VT, Lombard N, Nalin R, Vogel TM, Simonet P (2005). High molecular weight DNA recovery from soils prerequisite for biotechnological metagenomic library construction. J. Microbiol. Methods 62:1-11.

Bodelier P, Kamst M, Meima-Franke M, Stralis-Pavese N, Bodrossy L. (2009). Whole community genome amplification (WCGA) leads to compositional bias in methane-oxidizing communities as assessed by pmoA-based microarray analyses and QPCR. Environ. Microbiol. Rep. 1:434-441.

Borruel N, Carol M, Casellas F, Antolin M, de Lara F, Espin E, Naval J, Guarner F, Malagelada JR (2002). Increased mucosal tumor necrosis factor alpha production in Crohn's disease can be down regulated ex *vivo* by probiotic bacteria. Gut 51:659-664.

Brady SF, Clardy J (2000). Long-chain N-acyl amino acid antibiotics isolated from heterologously expressed environmental DNA. J. Am. Chem. Soc. 122:12903-12904.

Brady SF, Clardy, J (2005). Cloning and heterologous expression of isocyanide biosynthetic genes from environmental DNA. Angew Chem Int. Ed Engl. 44:7063-7065.

Chu FF, Esworthy RS, Chu PG, Longmate JA, Huycke MM, Wilczynski S, Doroshow JH (2004). Bacteria-induced intestinal cancer in mice with disrupted Gpx1 and Gpx2 genes. Cancer Res. 64:962-968.

Cowan D, Meyer Q, Stafford W, Muyanga S, Cameron R, Wittwer P (2005). Metagenomics gene discovery: past present and future. Trend. Biotechnol. 23:321-329.

Craig JW, Chang FY, Brady SF (2009) Natural products from environmental DNA hosted in *Ralstonia metallidurans*. ACS Chem. Biol. 4:23-28.

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. Appl. Environ. Microbiol. 76:1633-1641.

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

Fernandez-Arrojo L, Guazzaroni M, Lopez-Cortes L, Beloqui A, Ferrer M (2010). Metagenomic era for biocatalyst identification. Curr. Opin. Biotechnol. 21:725-733.

Fisch KM, Gurgui C, Heycke N, Van der Sar SA, Anderson SA, Webb VL, Taudien S, Platzer M, Rubio BK, Robinson SJ (2009). Polyketide assembly lines of uncultivated sponge symbionts from structure-based gene targeting. Nat. Chem. Biol. 5:494-501.

Gabor EM, Alkema WB, Janssen DB (2004). Quantifying the accessibility of the metagenome by random expression cloning techniques. Environ. Microbiol. 6:879-886.

Gillespie DE, Brady SF, Bettermann AD, Cianciotto NP, Liles MR, Rondon MR, Clardy J, Goodman RM, Handelsman J (2002). Isolation of antibiotics turbomycin a and B from a metagenomic library of soil microbial DNA. Appl. Environ. Microbiol. 68:4301-4306.

Ginolhac A, Jarrin C, Gillet B, Robe P, Pujic P, Tuphile K, Bertrand H, Vogel TM, Perriere G, Simonet P, Nalin R (2004). Phylogenetic analysis of polyketide synthase I domains from soil metagenomic libraries allows selection of promising clones. Appl. Environ. Microbiol. 70:5522-5527.

Gloux K, Leclerc M, Iliozer H, Haridon R, Manichanh C, Corthier G, Nalin R, Blottiere, H, Dore J (2007). Development of a high-throughput phenotyping of metagenomics clones from the human microbiome for modulation of eukaryotic cell growth. Appl. Environ. Microbiol. 71:3734-3737.

Gullo VP, Hughes DE (2005). Exploiting new approaches for natural product drug discovery in the biotechnology industry. Drug Discov. Today Technol. 2(3):281-286.

Henne A, Schmitz RA, Bömeke M, Gottschalk G, Daniel R (2000). Screening of environmental DNA libraries for the presence of genes conferring lipolytic activity on *Escherichia coli*. Appl. Environ. Microbiol. 66:3113-3116.

Hirsch PR, Mauchline TH, Clark IM (2010). Culture-independent molecular techniques for soil microbial ecology. Soil Biol. Biochem. 42:878-887.

Horikoshi K (1999). Alkaliphiles: some applications of their products for biotechnology. Microbiol. Mol. Biol. Rev. 63:735-750.

Horikoshi K, Antranikian G, Bull AT, Robb FT, Stetter KO (eds) (2011). Extremophiles handbook, vol. 1–2. Springer, Berlin.

Kerkhof LJ, Goodman RM (2009). Ocean microbial metagenomics. Deep-Sea Research II. 56:1824-1829.

Knietsch A, Waschkowitz T, Bowien S, Henne, A, Daniel R (2003). Metagenomes of complex microbial consortia derived from different soils as sources of novel genes conferring formation of carbonyls from short-chain polyols on Escherichia coli. J. Mol. Microbiol. Biotechnol. 5:46-56.

Koehn FE, Carter GT (2005). The evolving role of natural products in drug discovery. Nat. Rev. Drug Discov. 4:206-220.

Lamprecht MR, Sabatini DM, Carpenter AE (2007). Cell Profiler: free, versatile software for automated biological image analysis. Biotechniques 42(1):71-75.

Lefevre F, Jarrin C, Ginolhac A, Auriol D, Nalin R (2007). Environmental metagenomics: innovative resource for industrial biocatalysis. Biocatal. Biotrans. 25:242-250.

Lefevre F, Robe P, Jarrin C, Ginolhac A, Zago C, Auriol D, Vogel TM, Simonet P, Nalin R (2008). Drugs from hidden bugs: their discovery via untapped resources. Res. Microbiol. 159:153-161.

Li JW, Vederas JC (2009). Drug discovery and natural products: end of an era or an endless frontier. Science 325:161-165.

Lim HK, Chung EJ, Kim JC, Choi GJ, Jang KS, Chung YR, Cho KY, Lee SW (2005). Characterization of a forest soil metagenome clone that confers indirubin and indigo production on Escherichia coli. Appl. Environ. Microbiol. 71:7768-7777.

Long PF, Dunlap WC, Battershill CN, Jaspars M (2005). Shotgun cloning and heterologous expression of the patellamide gene cluster as a strategy to achieving sustained metabolite production. Chembiochem. 6(10):1760-1765.

MacNeil IA, Tiong CL, Minor C, August PR, Grossman TH, Loiacono KA, Lynch BA, Phillips T, Narula S, Sundaramoorthi R (2001). Expression and isolation of antimicrobial small molecules from soil DNA libraries. J. Mol. Microbiol. Biotechnol. 3:301-308.

Manichanh C, Rigottier-Gois L, Bonnaud E, Gloux K, Pelletier E, Frangeul L, Nalin R, Jarrin C, Chardon P, Marteau P, Roca J, Dore J (2006). Reduced diversity of faecal microbiota in Crohn's disease revealed by a metagenomic approach. Gut 55:205-211.

Martin HM, Campbell BJ, Hart CA, Mpofu C, Nayar M, Singh R, Englyst H, Williams HF, and Rhodes JM (2004). Enhanced *Escherichia coli* adherence and invasion in Crohn's disease and colon cancer.

Gastroenterology 127:80-93.
Martinez A, Kolvek SJ, Yip CLT, Hopke J, Brown KA, MacNeil IA, Osburne MS (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.
Mocali S, Benedetti A (2010). Exploring research frontiers in microbiology: the challenge of metagenomics in soil microbiology. Res. Microbiol. 161:497-505.
Monier J, Demaneche S, Delmont T, Mathieu A, Vogel TM, Simonet P (2011). Metagenomic exploration of antibiotic resistance in soil. Curr. Opin. Microbiol. 14:229-235.
Oren A (2004). Prokaryote diversity and taxonomy: current status and future challenges. Philosophical Transactions of the Royal Society B. Biol. Sci. 359:623-638.
Ortholand J, Ganesan A (2004). Natural products and combinatorial chemistry: back to the future. Curr. Opin. Chem. Biol. 8:271-280.
Pelaez F (2006). The historical delivery of antibiotics from microbial natural products--can history repeat? Biochem. Pharmacol. 71(7):981-990.
Pettit RK (2004). Soil DNA libraries for anticancer drug discovery. Cancer Chemother. Pharmacol. 54:1-6.
Piel J (2002). A polyketide synthase-peptide synthetase gene cluster from an uncultured bacterial symbiont of Paederus beetles. Proc. Natl. Acad. Sci. USA. 99:14002-14007.
Richard CL, Tandon A, Kranz RG (2004). *Rhodobacter capsulatus* nifHA1 promoter: high-GC-10 regions in high-GC bacteria and the basis for their transcription. J. Bacteriol. 186:740-749.
Robe P, Nalin R, Capellano C, Vogel TM, Simonet P (2003). Extraction of DNA from soil. Eur. J. Soil Biol. 39:183-190.
Roose-Amsaleg CL, Garnier-Sillam E, Harry M (2001) Extraction and purification of microbial DNA from soil and sediment samples. Appl. Soil Ecol. 18:47-60.
Rouhi AM (2003). Rediscovering natural products. Chem. Eng. News 81:77-91.
Schloss PD, Handelsman J (2004). Status of the microbial census. Microbiol. Mol. Biol. Rev. 68:686-691.
Schmidt EW, Donia MS (2009). Chapter 23. Cyanobactin ribosomally synthesized peptides — a case of deep metagenome mining. Methods Enzymol. 458:575-596.
Schmidt EW, Nelson JT, Rasko DA, Sudek S, Eisen JA, Haygood MG, Ravel J (2005). Patellamide A and C biosynthesis by a microcin-like pathway in Prochloron didemni, the cyanobacterial symbiont of *Lissoclinum patella*. Proc. Natl. Acad. Sci. USA. 102:7315-7320.
Singh BK (2009). Organophosphorus degrading bacteria: ecology and industrial application. Nat. Rev. Microbiol. 7:156-164.
Singh BK, Macdonald CA (2010). Drug discovery from uncultivable microorganisms. Drug Discov. Today 15:792-800.
Torsvik V, Ovreas L, Thingstad, TF (2002). Prokaryotic diversity magnitude, dynamics, and controlling factors. Science 296:1064-1066.
Tsai YL, Olson BH (1992). Rapid method for separation of bacterial DNA from humic substances in sediments for polymerase chain reaction. Appl. Environ. Microbiol. 58(7):2292-2295.
Van Lanen SG, Shen B (2006). Microbial genomics for the improvement of natural product discovery. Curr. Opin. Microbiol. 9:252-260.
Vulevic J, McCartney AL, Gee JM, Johnson IT, Gibson GR (2004). Microbial species involved in production of 1,2-sn-diacylglycerol and effects of phosphatidylcholine on human fecal microbiota. Appl. Environ. Microbiol. 70:5659-5666.
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. 2:2401-2404.
Wang J, Soisson SM, Young K, Shoop W, Kodali S, Galgoci A, Painter R, Parthasarathy, G, Tang YS, Cummings R, Ha S, Dorso K, Motyl M, Jayasuriya H, Ondeyka J, Herath K, Zhang C, Hernandez L, Allocco J, Basilio A, Tormo JR. Genilloud O, Vicente F, Pelaez, F, Colwell L, Lee SH, Michael B, Felcetto T, Gill C, Silver LL, Hermes JD, Bartizal K, Barrett J, Schmatz D, Becker JW, Cully D, Singh SB (2006). Platensimycin is a selective FabF inhibitor with potent antibiotic properties. Nature 441:358-361.
Yang Y, Adelstein SJ, Kassis AI (2009). Target discovery from data mining approaches. Drug Discov. Today 14:147-154.
Yun J, Kang S, Park S, Yoon H, Kim MJ, Heu S, Ryu S (2004). Characterization of a novel amylolytic enzyme encoded by a gene from a solid-derived metagenomic library. Appl. Environ. Microbiol. 70:7229-7235.
Zhou J, Bruns MA, Tiedje JM (1996). DNA recovery from soils of diverse composition. Appl. Environ. Microbiol. 62:316-322.
Ziemert N, Ishida K, Weiz A, Hertweck C, Dittmann E (2010). Exploiting the natural diversity of microviridin gene clusters for discovery of novel tricyclic depsipeptides. Appl. Environ. Microbiol. 76:3568-3574.
Zimmermann K, Engeser M, Blunt JW, Munro MH, Piel J (2009). Pederintype pathways of uncultivated bacterial symbionts: analysis of o-methyltransferases and generation of a biosynthetic hybrid. J. Am. Chem. Soc. 131:2780-2781.