Chapter 5
Metagenomics and Metatranscriptomics for the Exploration of Natural Products from Soil Fungi

Irshad Ul Haq and Jan Dirk van Elsas

Introduction

Microorganisms constitute rich sources of diverse biologically active metabolites. These metabolites have already found a broad spectrum of applications, for instance as antiparasitics, antibiotics, anticancer agents, immunosuppressants as well as agrochemicals [1]. A wide range of niches on Earth are occupied by microorganisms, ranging from deep rock sediments and marine environments to deserts, alpine, Arctic, and Antarctic regions, and even to thermal vents [1]. In terms of microbial diversity, soil is a remarkable site, which contains a hitherto largely unexplored microbiota. For instance, in as small as 1 g of soil, several thousands of bacterial species exist, the majority of which are uncultivable under standard microbiological conditions [2]. In parallel to prokaryotes, there is a substantial number of eukaryotic microorganisms hosted by soil, which contribute to the microbial biomass [3]. In the light of the enormous diversity of microorganisms in soil, only a handful of bacterial (less than 1 %) and fungal species (less than 5 %) are known at present. Hence, millions of microbial species out there need to be unearthed [4]. Isolation and in vitro growth of most prokaryotic and eukaryotic microorganisms is, however, difficult or impossible due to their general lack of cultivability. To address this obstacle, new experimental approaches, such as metagenomics, have been used to assess the true functional diversity and activities of microorganisms in soil. In the next sections, we will explore how the power of molecular tools can be harnessed to explore the wealth of genetic and functional information that exists right underfoot.

J. D. van Elsas (✉) · I. U. Haq
Microbial Ecology, Groningen Institute for Life Sciences (GELIFES),
University of Groningen, Groningen, The Netherlands
e-mail: j.d.van.elsas@rug.nl

I. U. Haq
e-mail: i.u.haq@rug.nl

S. Zeilinger et al. (eds.), *Biosynthesis and Molecular Genetics of Fungal Secondary Metabolites, Volume 2,* Fungal Biology, DOI 10.1007/978-1-4939-2531-5_5

Natural Product Exploration Using Metagenomics and Other "Omics" Tools

Natural environments, such as soil, are a great reservoir of genes involved in different biosynthetic pathways that are difficult to explore using cultivation techniques. This reservoir of genes can be unlocked using metagenomics; i.e., the study of the collective genomes of the microbial community. Metagenomics offers very powerful strategies that allow us to unearth both the functional potential and taxonomic diversity of microorganisms at the community level [5, 6]. This approach has already yielded a wealth of novel data. However, there are pitfalls in the approach, as will be discussed in the following section.

Analysis of Function of Soil—Amplicon Sequencing and Metagenomics

The function of soil can be studied using either a gene-centered or a genome-centered approach. In the former approach, the polymerase chain reaction (PCR) is used to amplify single target genes and the amplification products (amplicons) are used for sequencing to analyze the occurrence of the different orthologs of that gene in the whole community. In the second approach, random metagenomic sequencing is used in which total microbial community DNA is isolated from a soil sample and shotgun sequenced, resulting in an outline of all genes that are present in the community [7]. Next to these direct approaches, the DNA extracted from the sample can also be used to generate metagenomic DNA libraries, which are subsequently screened for function (Fig. 5.1).

The latter two approaches result in a wealth of information that is stored in the genomes of microorganisms, which occupy various niches in the soil environment [8]. The metagenomic libraries have potential applications in both applied and basic research. Several studies over the years have used metagenomics for purposes such as bioprospection for novel amylases [9], beta-agarases, cellulases and lipases [10]. Moreover, Schirmer et al. [11], Courtois et al. [12] and Gillespie et al. [13] reported on new polyketide synthase genes and their expressed compounds and two colored triaryl cation antibiotics. Other studies revealed information about important physiological processes of microorganisms after extensive sequencing of metagenomic libraries [14–16]. All these studies focused on prokaryotic microorganisms, thus excluding eukaryotic microorganisms, which was possibly due to their relative scarcity or because of their physical discrimination through filtration or centrifugation on density gradients before DNA extraction [3]. Eukaryotic metagenomics has faced certain constraints over the years, such as the giant genome sizes of most eukarya compared to the smaller bacterial genomes. (Micro)eukaryote genomes range from 13.8 Mbp for the yeast *Schizosaccharomyces pombe* [17] to 69 Mbp for the ciliate *Paramecium tetraurelia* [18]. The large sizes compromise seizing—to a sufficient extent—the eukaryotic microbial community gene content. Moreover, the

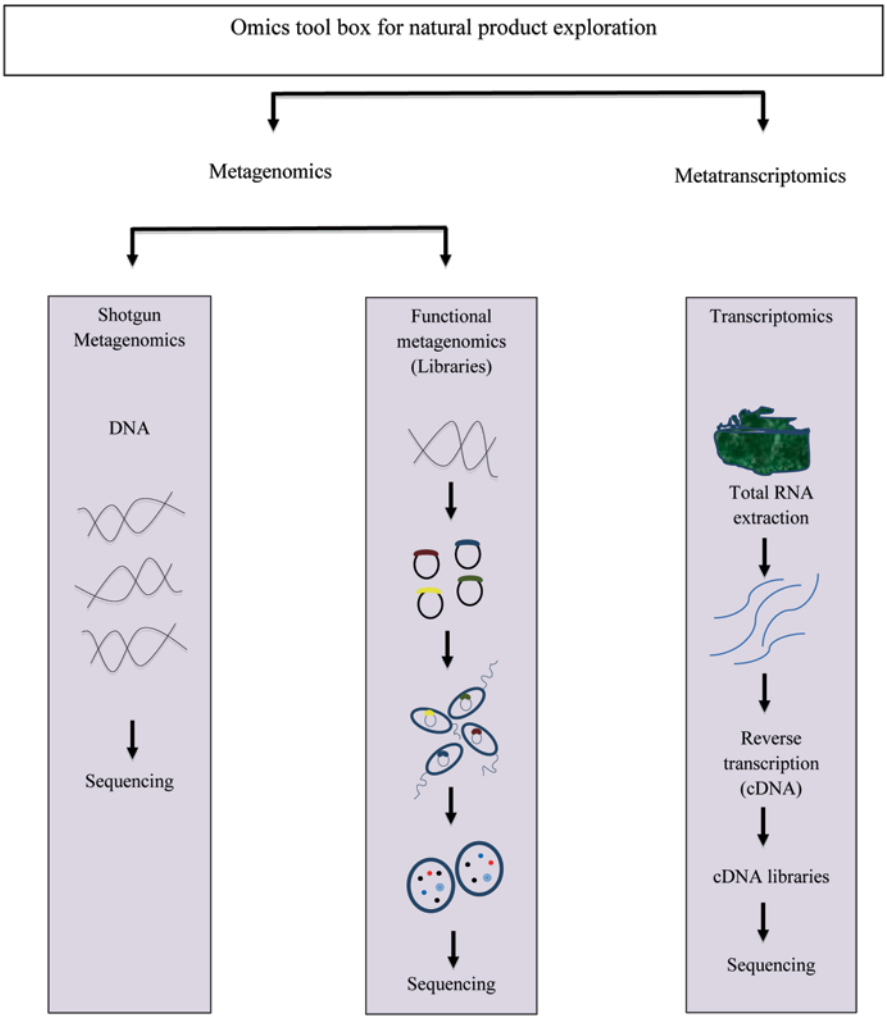

Fig. 5.1 A schematic workflow of 'omics' tools for screening of natural products

detection of the expression of eukaryotic protein-encoding genes is impeded by the existence of introns and the absence of a conserved motif in promoter sequences [3].

Construction and Screening of Metagenomic Libraries

Metagenomic libraries are most often constructed inside cloning vectors that are replicated in the common host *Escherichia coli*. Different cloning vectors can host DNA fragments ranging from up to 30 to 300 kb in size [19–23], which allows a wealth of possibilities in the cloning step. Single-gene traits can be picked up in

small-insert (up to several kb) vectors, whereas traits that are encoded by larger stretches of DNA (gene clusters) require larger insert vectors (see later). Gene clusters that encode natural products are mainly in the 30–300 kb range, which would allow cloning of whole intact gene clusters inside a single large-insert vector clone [24]. Recently, expression of whole gene clusters in suitable host organisms has been achieved, allowing biosynthesis of the natural product in question [25–27]. If gene clusters spread over multiple clones represent a target pathway, recombinogenic cloning could be used to streamline the whole metabolic pathway, as has been reported for heterologous expression of the tubulysin gene cluster [28].

A prerequisite for the construction of a metagenomic library is the selection of an appropriate vector, into which sheared and fragmented DNA isolated from an environmental sample can be inserted. For many gene clusters, vectors that can accommodate inserts up to 40 kb, such as fosmids [22] and cosmids [20], could be used. However, as discussed above, some gene clusters have sizes well above 40 kb. This has motivated researchers to use vectors that can harbor up to 300 kb fragments [19], such as P1-derived vectors [21] and bacterial artificial chromosomes (BACs) [23]. Consequently, an appropriate host organism should host vectors that carry random fragments of particular sizes. This results in a plethora of clones that are then screened to assess if DNA fragments of interest are successfully expressed or are detectable by genetic screens. Those screens that pinpoint interesting DNA fragments are often subjected to sequencing, resulting in information that is useful to mine the gene clusters of interest. This strategy has been used in a number of recent studies, i.e., to identify siderophore biosynthetic genes [26], to detect secondary metabolites from soil bacteria [29], and to discover novel natural products [30].

Shotgun Metagenomics or Direct Sequencing

The current affordability of next-generation sequencing (NGS) tools has revolutionized the direct shotgun metagenomics of the microbiota from environmental habitats. In this approach, microbial community DNA isolated from environmental samples is directly sequenced without first constructing clone libraries. Shotgun metagenomics has been applied in various studies to discover enzymes responsible for the biodegradation of lignocellulosic matter from sources such as cow rumen [31] and compost [32]. Furthermore, recent studies featured shotgun metagenomics and metagenomic library construction, exploring the microbiomes of the marine tunicate *Lissoclinum patella* as well as of coral reefs [33–35]. Other studies used pyrosequencing to investigate the marine sponge *Arenosclera brasiliensis* microbiome [36], and the *Ecteinascidia turbinata* tunicate metagenome [37]. The latter study identified a biosynthetic gene cluster encoding the chemotherapeutic natural product denoted ‘ET-743’. However, there are potential constraints related to the data generated, as processing the data and extracting relevant information about the biosynthetic gene clusters of natural products is a daunting task. In particular, the massive data that we generate nowadays pose problems for downstream analyses such as cleaning up, binning and subsequent sequence-based and statistical analy-

ses. Computer power is increasingly limiting and so is bioinformatics. Clearly, one needs advanced (better) bioinformatics tools and pipelines to pave the way for improvement in the analysis of the data generated by shotgun metagenomics.

Metatranscriptomics

Metagenomics studies do not address questions regarding the expression state of genes, hampering conclusions about the functional role of particular genes in the environment under study. Therefore, new technology-coined metatranscriptomics strategies have come into being. Metatranscriptomics analyzes the collective set of messenger RNAs that are present in an environmental sample, in an all-at-once manner. However, RNA extracted from the natural microbiota is often dominated by ribosomal RNA (rRNA). Hence, in the early studies messenger RNA (mRNA) has been enriched by rRNA depletion (for bacteria) or poly-A tailed mRNA enrichment (for eukaryotes) to allow the investigation of overall gene expression profiles in the environments under study [3, 38]. Later on, massive parallel (pyro)sequencing, following a reverse transcription step, was adopted to analyze bacterial and archaeal mRNA from environmental (marine) samples, giving a much larger scale as compared to the previous studies. In one study, an in vitro amplification step was included to keep sample size small and preparation fast [39]. In another study in soil, the total RNA was analyzed all at once. This allowed the assessment of the limitations of earlier approaches (linking phylogeny to function) by the simultaneous determination of soil microbial community structure (rRNA) and function (mRNA) through metatranscriptomics [40]. However, the amount of mRNA that could be analyzed in the study was disappointingly low, and so the analysis of in situ gene expression was limited. On the other hand, in spite of the technical difficulties, analyses based on metatranscriptomics are very useful, as they provide clues concerning the in situ gene expression and point us to conditions under which key genes (e.g., those involved in the production of natural compounds) are expressed.

The Rhizosphere: A Potential Hotspot for Natural Products

The biologically active region in the immediate vicinity of plant roots, which is inhabited by soil microorganisms (in particular bacteria and fungi), is termed the rhizosphere [41–43]. The rhizosphere is under the direct influence of plant roots and their exuded products, such as secreted compounds, cell lysates, mucilage and gases such as respiratory CO_2 [44]. Although recalcitrant compounds are also present, the availability of easily degradable nutrients makes the rhizosphere a dreamland for microorganisms. It is also a playground for complex interactions among microorganisms, such as cooperation through cross-feeding or competition for nutrients, using, for instance, antagonism through chemical warfare [45–47]. The diversity and complexity of the rhizosphere in terms of microbial life, in addition to the selection

for interaction-proficient microbes, makes it a potentially important source of natural products. A number of natural products from fungi associated with plants have already been isolated and characterized, as described in the following examples. Penicillic acid and two new natural products (6-methoxy-5, 6-dihydropenicillic acid and 4*R*, 5*S* dihydroxy-3-methoxy-5-methylcyclohex-2-enone) have been isolated from *Aspergillus cervinus* associated with *Anicasanthus thurberi* [48]. Similarly, *Aspergillus terreus*, which is associated with the roots of *Opuntia versicolor*, is a producer of compounds such as betulinan, quadrone, terricyclic acid A, asterriquinone C-1, asterriquinone D and asterredione, among others [49, 50]. These few examples suggest that the rhizosphere is a rich source of natural products. The diverse microorganisms that inhabit the rhizosphere have apparently learned to deal with the ecology of the niches present and developed key genetic systems allowing survival by chemical warfare accordingly. On the other hand, the studies only reported on fungal strains that can be grown in vitro. In the light of the lack of culturability of many microorganisms, it is likely that there is much more potential for finding and exploring natural product producers. Thus, techniques such as metagenomics, metatranscriptomics and even metabonomics should be applied to overcome the limitation of microorganism culturing in vitro.

Endophytic Fungi as Sources of Natural Products

Microorganisms that reside inside plants without causing disease symptoms have been coined endophytes [51]. This unique plant–microbe interaction is established entirely inside plant tissues [52] and is defined by the fact that the two partners do not affect each other lethally. Endophytes offer great biotechnological potential in terms of the biosynthesis of natural products and bioactive metabolites for application such as therapeutics for a number of diseases [52–56]. Some studies have already reported the finding of key therapeutically important secondary metabolites produced by fungi, such as taxol [57], deoxypodophyllotoxin [58], podophyllotoxin [59, 60], hypericin and emodin [61, 62], azadirachtin [63] and camptothecin [64–67]. The production of these metabolites, as well as many others, makes endophytes very important microorganisms for studying from an ecological as well as biochemical standpoint. In order to investigate and explore new secondary metabolites, it is imperative that such plant-interactive microorganisms are exploited in the best possible way, which should include an assessment on how genes for the biosynthesis of key metabolites are regulated.

Bacterial–Fungal Interactions and Natural Product Discovery

Recently, microbial interactions were found to be drivers the of the production of particular metabolites in bacteria and fungi. The interactions between microorganisms, especially the bipartite ones (e.g., between bacteria and fungi) were deemed

important, as natural product formation can indeed be induced by bacteria that occur in the vicinity of fungi. For instance, the soil-dwelling bacterium *Streptomyces rapamycinicus* physically interacts with *Aspergillus nidulans*, and, in this interaction, activates a silent polyketide biosynthesis gene cluster [68]. Therefore, including microbial "neighbors" in studies for exploration of natural products is the way to go in further screens that eventually will include meta-omics techniques. The former study showed the production of polyphenols—i.e., cathepsin K inhibitors and lecanoric acid—derived from orsellinic acid [68]. In other studies on non-endophytic fungi, *Variovorax paradoxus* strain HB44 was found to be selected in the mycosphere of *Laccaria proxima* [69]. The organism was able to grow on compounds released by a close relative of *L. proxima*, i.e., *Lyophyllum* sp. strain Karsten, particularly glycerol. The study also reported the release of other compounds, i.e., acetic acid and formic acid, by the fungus [69]. Recent work in our laboratory shows that *Lyophyllum* sp. strain Karsten releases glycerol-rich exudates, which may be due to a stimulatory effect of the fungal-interactive *Burkholderia terrae* strain BS001. This mechanism may be of great significance for strain BS001 in an ecological context. The stimulation of the fungal release of glycerol could be considered as a strategy to acquire easily degradable carbonaceous food, allowing a better survival in the mycosphere [70].

With respect to bacterial–fungal interactions, complex interplays of events have been shown, in which the toxic compound rhizoxin was produced. Rhizoxin is the causative agent of rice seedling blight. Until recently, it was believed that the rice-pathogenic fungus *Rhizopus microsporus* was the producer of this rhizoxin. However, this turned out not to be the case, as the bacterium *Burkholderia rhizoxinica*, which inhabits the fungal cytosol, was revealed to be the producer [71]. Recently, it has been reported that *R. microsporus* also contributes to the potency of the (phytotoxic) rhizoxin, as the rhizoxin produced in co-cultures with *B. rhizoxinica* contained two bis-epoxide moieties compared to the one that is solely produced by the bacterium, clearly indicating that there is synergism in production of natural products [72].

In another study, fungi interacting with lichens were reported to produce natural products such as bis-naphtopyrones and lichenicolins A and B, both of which have activity against Gram-positive bacteria [73]. Co-culturing an unidentified bacterium with the fungus *Libertella* sp., diterpenoids- including libertellenone A–D—were discovered. The compounds were induced by the presence of the bacterium, as they did not show up in cultures without the bacterium [74]. The marine fungus *Emericella* sp., when grown in co-culture with *Salinispora arenicola*, produced two depsipeptides (emericellamides A and B) that were shown to exert antibacterial activity against methicillin-resistant *Staphylococcus aureus* [75], and are thus interesting as antibiotics against this dangerous bacterium. Similarly, formyl-xanthocillin analogs were shown to be synthesized when *Streptomyces peucetius* was grown with *Aspergillus fumigatus* [76].

All these studies on the biosynthesis of natural products have been carried out in relatively "simple" circumstances, in which basic (co-)culturing techniques were applied. However, the complexities of their natural environment, including fluctuating and often harsh circumstances, are rarely included in such experiments

and the full potential of natural compound production in nature is still cryptic. Therefore, to unlock the wealth of natural products hidden in the natural microbiota, it is imperative to dig deeply into natural habitats, focusing on the interacting microorganisms and to "eavesdrop" on the cross-talk between them using meta-omics tools. In such strategies, a combination of the current and newly emerging advanced molecular tools, including RNA sequencing and metabonomics, with traditional cultivation-based efforts needs to be applied as such an approach will take advantage of the complementary strong points of both types of analyses.

Conclusion

Outlook

There is a progressively increasing feeling that the scientific community, and society as a whole, is close to exhausting the capability of finding novel natural products that serve society, if we continue to bioexplore our natural environments by the traditional (cultivation-based) and advanced molecular methods. The natural products include the dearly needed novel antibiotics, which are often produced by fungi from soil or other natural environments, that allow us to treat dangerous bacterial or fungal diseases. The reasons for this contention of reaching the "limit" are the simple facts that (1) increasingly we reencounter organisms and the natural compounds they produce via the traditional way of cultivation and assessing bioantagonism, and (2) the novel molecular tools often stop short of telling us the complete story on the expression of genes/operons for natural products in the light of the absence of the suitable conditions that trigger gene expression.

To tackle both types of problems, it is imperative that a better focus is placed on mimicking the conditions that govern the life of the target microbiota in its natural environment. And, included in such conditions, we need to consider the biotic component of it (i.e., the presence of other organisms), as microorganisms such as fungi in nature have most likely "learned" to express their key antagonistic compounds when other organisms (that might present ecological threats) are in their vicinity. Hence, we here reviewed the available literature with respect to the effect of microbial "neighbors" on the expression of (otherwise silent) genes that underlie microbial antagonism and thus might yield novel antibiotic compounds. Moreover, we strongly advocate the inclusion of such organisms, or consortia of organisms, in screens for the production of novel compounds. Then, the power of meta-omics techniques might be harnessed to improve our screens and get at the natural products and their underlying genes in the most efficient way possible. This may include (in that order) metabonomics, metatranscriptomics and metagenomics, leading to the identification and isolation of genes/operons responsible for the biosynthesis.

However, in the end it might be very useful to also attempt to isolate the producer organism, allowing production by the natural organism, as incited by neighbors. In this isolation effort, the availability of molecular tools will be a great asset.

In other words, if fragments of interesting genes are discovered, probes and primers might be generated that enable the monitoring of the organisms in enrichments and allow the guidance of a directed isolation effort, leading to the availability of novel "nature-derived" producer organisms.

References

1. Gunatilaka AAL (2006) Natural products from plant-associated micro-organisms: distribution, structural diversity, bioactivity, and implications of their occurrence. J Nat Prod 69(3):509–526
2. Rappe MS, Giovannoni SJ (2003) The uncultured microbial majority. Annu Rev Microbiol 57:369–394
3. Bailly J, Fraissinet-Tachet L, Verner MC, Debaud JC, Lemaire M, Wésolowski-Louvel M, Marmeisse R (2007) Soil eukaryotic functional diversity, a metatranscriptomic approach. ISME J 1(7):632–642
4. Young P (1997) Major microbial diversity initiative recommended. ASM News 63:417–421
5. Cowan D, Meyer Q, Stafford W, Muyanga S, Cameron R, Wittwer P (2005) Metagenomic gene discovery, past, present and future. Trends Biotechnol 23(6):321–329
6. Schloss PD, Handelsman J (2003) Biotechnological prospects from metagenomics. Curr Opin Biotechnol 14(3):303–310
7. Gilbert JA, Dupont CL (2011) Microbial metagenomics: beyond the genome. Ann Rev Mar Sci 3:347–371
8. Rondon MR, August PR, Bettermann AD, Brady SF, Grossman TH, Liles MR, Loiacono KA, Lynch BA, MacNeil IA, Minor C, Tiong CL, Gilman M, Osburne MS, Clardy J, Handelsman J, Goodman RM (2000) Cloning the soil metagenome: a strategy for accessing the genetic and functional diversity of uncultured microorganisms. Appl Environ Microbiol 66(6):2541–2547
9. Yun J, Kang S, Park S, Yoon H, Kim MJ, Hew S, Ryu S (2004) Characterization of a novel amylolytic enzyme encoded by a gene from a soil derived metagenomic library. Appl Environ Microbiol 70(12):7229–7235
10. Voget S, Leggewie C, Uesbeck A, Raasch C, Jaeger KE, Streit WR (2003) Prospecting for novel biocatalysts in a soil metagenome. Appl Environ Microbiol 69(10):6235–6242
11. Schirmer A, Gadkari R, Reeves CD, Ibrahim F, DeLong EF, Hutchinson CR (2005) Metagenomic analysis reveals diverse polyketide synthase gene clusters in micro-organisms associated with the marine sponge *Discodermia dissoluta*. Appl Environ Microbiol 71(8):4840–4849
12. Courtois S, Cappellano CM, Ball M Francou FX, Normand P, Helynck G, Martinez A, Kolvek SJ, Hopke J, Osburne MS, August PR, Nalin R, Guérineau M, Jeannin P, Simonet P, Pernodet JL (2003) Recombinant environmental libraries provide access to microbial diversity for drug discovery from natural products. Appl Environ Microb 69:49–55
13. 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(9):4301–4306
14. Rusch DB, Halpern AL, Sutton G, Heidelberg KB, Williamson S, Yooseph S , Wu D, Eisen JA, Hoffman JM, Remington K, Beeson K, Tran B, Smith H, Baden-Tillson H, Stewart C, Thorpe J, Freeman J, Andrews-Pfannkoch C, Venter JE, Li K, Kravitz S, Heidelberg JF, Utterback T, Rogers YH, Falcón LI, Souza V, Bonilla-Rosso G, Eguiarte LE, Karl DM, Sathyendranath S, Platt T, Bermingham E, Gallardo V, Tamayo-Castillo G, Ferrari MR, Strausberg RL, Nealson K, Friedman R, Frazier M, Venter JC (2007) The sorcerer II global sampling expedition: Northwest Atlantic through eastern tropical pacific. PLoS Biol 5(3):e77

15. Venter JC, Remington K, Heidelberg JF, Halpern AL, Rusch D, Eisen JA, Wu D, Paulsen I, Nelson KE, Nelson W, Fouts DE, Levy S, Knap AH, Lomas MW, Nealson K, White O, Peterson J, Hoffman J, Parsons R, Baden-Tillson H, Pfannkoch C, Rogers YH, Smith HO (2004) Environmental genome shotgun sequencing of the Sargasso Sea. Science 304(5667):66–74
16. Yooseph S, Sutton G, Rusch DB, Halpern AL, Williamson SJ, Remington K, Eisen JA, Heidelberg KB, Manning G, Li W, Jaroszewski L, Cieplak P, Miller CS, Li H, Mashiyama ST, Joachimiak MP, van Belle C, Chandonia JM, Soergel DA, Zhai Y, Natarajan K, Lee S, Raphael BJ, Bafna V, Friedman R, Brenner SE, Godzik A, Eisenberg D, Dixon JE, Taylor SS, Strausberg RL, Frazier M, Venter JC (2007) The Sorcerer II global ocean sampling expedition: expanding the universe of protein families. PLoS Biol 5(3):e16
17. Wood V, Gwilliam R, Rajandream MA, Lyne M, Lyne R, Stewart A, Sgouros J, Peat N, Hayles J, Baker S, Basham D, Bowman S, Brooks K, Brown D, Brown S, Chillingworth T, Churcher C, Collins M, Connor R, Cronin A, Davis P, Feltwell T, Fraser A, Gentles S, Goble A, Hamlin N, Harris D, Hidalgo J, Hodgson G, Holroyd S, Hornsby T, Howarth S, Huckle EJ, Hunt S, Jagels K, James K, Jones L, Jones M, Leather S, McDonald S, McLean J, Mooney P, Moule S, Mungall K, Murphy L, Niblett D, Odell C, Oliver K, O'Neil S, Pearson D, Quail MA, Rabbinowitsch E, Rutherford K, Rutter S, Saunders D, Seeger K, Sharp S, Skelton J, Simmonds M, Squares R, Squares S, Stevens K, Taylor K, Taylor RG, Tivey A, Walsh S, Warren T, Whitehead S, Woodward J, Volckaert G, Aert R, Robben J, Grymonprez B, Weltjens I, Vanstreels E, Rieger M, Schäfer M, Müller-Auer S, Gabel C, Fuchs M, Düsterhöft A, Fritzc C, Holzer E, Moestl D, Hilbert H, Borzym K, Langer I, Beck A, Lehrach H, Reinhardt R, Pohl TM, Eger P, Zimmermann W, Wedler H, Wambutt R, Purnelle B, Goffeau A, Cadieu E, Dréano S, Gloux S, Lelaure V, Mottier S, Galibert F, Aves SJ, Xiang Z, Hunt C, Moore K, Hurst SM, Lucas M, Rochet M, Gaillardin C, Tallada VA, Garzon A, Thode G, Daga RR, Cruzado L, Jimenez J, Sánchez M, del Rey F, Benito J, Domínguez A, Revuelta JL, Moreno S, Armstrong J, Forsburg SL, Cerutti L, Lowe T, McCombie WR, Paulsen I, Potashkin J, Shpakovski GV, Ussery D, Barrell BG, Nurse P (2002) The genome sequence of *Schizosaccharomyces pombe*. Nature 415(6874):871–880
18. Aury JM, Jaillon O, Duret L, Noel B, Jubin C, Porcel BM, Ségurens B, Daubin V, Anthouard V, Aiach N, Arnaiz O, Billaut A, Beisson J, Blanc I, Bouhouche K, Câmara F, Duharcourt S, Guigo R, Gogendeau D, Katinka M, Keller AM, Kissmehl R, Klotz C, Koll F, Le Mouël A, Lepère G, Malinsky S, Nowacki M, Nowak JK, Plattner H, Poulain J, Ruiz F, Serrano V, Zagulski M, Dessen P, Bétermier M, Weissenbach J, Scarpelli C, Schächter V, Sperling L, Meyer E, Cohen J, Wincker P (2006) Global trends of whole-genome duplications revealed by the ciliate *Paramecium tetraurelia*. Nature 444(7116):171–178
19. Alduina R, Gallo G (2012) Artificial chromosomes to explore and to exploit biosynthetic capabilities of actinomycetes. J Biomed Biotechnol. doi:10.1155/2012/462049
20. Collins J, Hohn B (1978) Cosmids: a type of plasmid gene-cloning vector that is packagable in vitro in bacteriophage lambda heads. Proc Natl Acad Sci U S A 75(9):4242–4246
21. Ioannou PA, Amemiya CT, Garnes J, Kroisel PM, Shizuya H, Chen C, Batzer MA, de Jong PJ (1994) A new bacteriophage P1-derived vector for the propagation of large human DNA fragments. Nat Genet 6(1):84–89
22. Kim UJ, Shizuya H, de Jong PJ, Birren B, Simon MI (1992) Stable propagation of cosmid sized human DNA inserts in an F factor based vector. Nucleic Acids Res 20(5):1083–1085
23. Shizuya H, Birren B, Kim UJ, Mancino V, Slepak T, Tachiiri Y, Simon M (1992) Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in *Escherichia coli* using an F-factor-based vector. Proc Natl Acad Sci 89(18):8794–8797
24. Schofield MM, David HS (2013) Meta-omic characterization of prokaryotic gene clusters for natural product biosynthesis. Curr Opin Biotechnol 24(6):1151–1158
25. 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
26. Fujita MJ, Kimura N, Sakai A, Ichikawa Y, Hanyu T, Otsuka M (2011) Cloning and heterologous expression of the vibrioferrin biosynthetic gene cluster from a marine metagenomic library. Biosci Biotechnol Biochem 75:2283–2287

27. McMahon MD, Guan C, Handelsman J, Thomas MG (2012) Metagenomic analysis of *Streptomyces lividans* reveals host-dependent functional expression. Appl Environ Microbiol 78(10):3622–3629
28. Chai Y, Shan S, Weissman KJ, Hu S, Zhang Y, Muller R (2012) Heterologous expression and genetic engineering of the tubulysin biosynthetic gene cluster using Red/ET recombineering and inactivation mutagenesis. Chem Biol 19(3):361–371
29. Feng Z, Kallifidas D, Brady SF (2011) Functional analysis of environmental DNA-derived type II polyketide synthases reveals structurally diverse secondary metabolites. Proc Natl Acad Sci U S A 108(31):12629–12634
30. Freeman MF, Gurgui C, Helf MJ, Morinaka BI, Uria AR, Oldham NJ, Sahl HG, Matsunaga S, Piel J (2012) Metagenome mining reveals polytheonamides as posttranslationally modified ribosomal peptides. Science 338(6105):387–390
31. Hess M, Sczyrba A, Egan R, Kim TW, Chokhawala H, Schroth G, Luo S, Clark DS, Chen F, Zhang T, Mackie RI, Pennacchio LA, Tringe SG, Visel A, Woyke T, Wang Z, Rubin EM (2011) Metagenomic discovery of biomass-degrading genes and genomes from cow rumen. Science 331(6016):463–467
32. Allgaier M, Reddy A, Park JI, Ivanova N, D'haeseleer P, Lowry S, Sapra R, Hazen TC, Simmons BA, Vander Gheynst JS, Hugenholtz P (2010) Targeted discovery of glycoside hydrolases from a switchgrass-adapted compost community. PLoS ONE 5(1):e8812
33. Donia MS, Fricke WF, Partensky F, Cox J, Elshahawi SI, White JR, Phillippy AM, Schatz MC, Piel J, Haygood MG, Ravel J, Schmidt EW (2011a) Complex microbiome underlying secondary and primary metabolism in the tunicate-*Prochloron* symbiosis. Proc Natl Acad Sci U S A 108(51):E1423–E1432
34. Donia MS, Fricke WF, Ravel J, Schmidt EW (2011b) Variation in tropical reef symbiont metagenomes defined by secondary metabolism. PLoS ONE 6(3):e17897
35. Kwan JC, Donia MS, Han AW, Hirose E, Haygood MG, Schmidt EW (2012) Genome streamlining and chemical defense in a coral reef symbiosis. Proc Natl Acad Sci U S A 109(50):20655–20660
36. Trindade-Silva AE, Rua C, Silva GG, Dutilh BE, Moreira AP, Edwards RA, Hajdu E, Lobo-Hajdu G, Vasconcelos AT, Berlinck RG, Thompson FL (2012) Taxonomic and functional microbial signatures of the endemic marine sponge *Arenosclera brasiliensis*. PLoS ONE 7(7):e39905
37. Rath CM, Janto B, Earl J, Ahmed A, Hu FZ, Hiller L, Dahlgren M, Kreft R, Yu F, Wolff JJ, Kweon HK, Christiansen MA, Håkansson K, Williams RM, Ehrlich GD, Sherman DH (2011) Meta-omic characterization of the marine invertebrate microbial consortium that produces the chemotherapeutic natural product ET-743. ACS Chem Biol 6(11):1244–1256
38. Poretsky RS, Bano N, Buchan A, LeCleir G, Kleikemper J, Pickering M, Pate WM, Moran MA, Hollibaugh JT (2005) Analysis of microbial gene transcripts in environmental samples. Appl Environ Microbiol 71(7):4121–4126
39. Frias-Lopez J, Shi Y, Tyson GW, Coleman ML, Schuster SC, Chisholm SW, Delong EF (2008) Microbial community gene expression in ocean surface waters. Proc Natl Acad Sci U S A 105(10):3805–3810
40. Urich T, Lanzén A, Qi J, Huson DH, Schleper C, Schuster SC (2008) Simultaneous assessment of soil microbial community structure and function through analysis of the meta-transcriptome. PLoS ONE 3(6):e2527
41. Hiltner L (1904) Ueber neuere Erfahrungen und Probleme auf dem Gebiete der Bodenbakteriologie und unter besonderer Berücksichtigung der Gründüngung und Brache. Arb Deut Landw Gesell 98:59–78
42. Lynch J (1982) The rhizosphere. In: Burns RG, Slater JH (eds) Experimental microbial ecology. Blackwell Scientific Publications, Oxford, pp 395–411
43. Singh BK, MillardP, Whiteley AS, Murrell JC (2004) Unravelling rhizosphere-microbial interactions: opportunities and limitations. Trends Microbiol 12(8):386–393
44. Lynch JM, Whipps M (1990) Substrate flow in the rhizosphere. Plant Soil 129:1–10

45. Kent AD, Triplett EW (2002) Microbial communities and their interactions in soil and rhizosphere ecosystems. Annu Rev Microbiol 56:211–236
46. Morello JE, Pierson EA, Pierson LS III (2004) Negative cross-communication among wheat rhizosphere bacteria: effect on antibiotic production by the biological control bacterium *Pseudomonas aureofaciens* 30–84. Appl Environ Microbiol 70(5):3103–3109
47. Whipps JM (2001) Microbial interactions and biocontrol in the rhizosphere. J Exp Bot 52:487–511
48. He J, Wijeratne EM, Bashyal BP, Zhan J, Seliga CJ, Liu MX, Pierson EE, Pierson LS, Van Etten HD, Gunatilaka AA (2004) Cytotoxic and other metabolites of *Aspergillus* inhabiting the rhizosphere of Sonoran desert plants. J Nat Prod 67(12):1985–1991
49. Turbyville TJ, Wijeratne EMK, Whitesell L, Gunatilaka AA (2005) The anticancer activity of the fungal metabolite terrecyclic acid A is associated with modulation of multiple cellular stress response pathways. Mol Cancer Ther 4(10):1569–1576
50. Wijeratne EMK, Turbyville TJ, Zhang Z, Bigelow D, Pierson LS III, Whitesell L, Canfield LM, Gunatilaka AA (2003) Cytotoxic constituents of *Aspergillus terreus* from the rhizosphere of *Opuntia versicolor* of the Sonoran Desert. J Nat Prod 66(12):1567–1573
51. Bacon CW, White JF (2000.) Microbial endophytes. Marcel Deker Inc., New York
52. Kusari S, Spiteller M (2012) Metabolomics of endophytic fungi producing associated plant secondary metabolites: progress, challenges and opportunities. In: Roessner U (ed) Metabolomics. InTech, Rijeka, pp 241–66
53. Aly AH, Debbab A, Kjer J, Proksch P (2010) Fungal endophytes from higher plants: a prolific source of phytochemicals and other bioactive natural products. Fungal Divers 41(1):1–16
54. Kharwar RN, Mishra A, Gond SK, Stierle A, Stierle D (2011) Anticancer compounds derived from fungal endophytes: their importance and future challenges. Nat Prod Rep 28(7):1208–1228
55. Staniek A, Woerdenbag HJ, Kayser O (2008) Endophytes: exploiting biodiversity for the improvement of natural product-based drug discovery. J Plant Interact 3:75–93
56. Strobel GA, Daisy B, Castillo U, Harper J (2004) Natural products from endophytic microorganisms. J Nat Prod 67(2):257–268
57. Stierle A, Strobel GA, Stierle D (1993) Taxol and taxane production by *Taxomyces andreanae*, an endophytic fungus of Pacific yew. Science 260(5105):214–216
58. Kusari S, Lamshoft M, Spiteller M (2009a) *Aspergillus fumigatus* Fresenius, an endophytic fungus from *Juniperus communis* L. Horstmann as a novel source of the anticancer pro-drug deoxypodophyllotoxin. J Appl Microbiol 107(3):1019–1030
59. Eyberger AL, Dondapati R, Porter JR (2006) Endophyte fungal isolates from *Podophyllum peltatum* produce podophyllotoxin. J Nat Prod 69(8):1121–1124
60. Puri SC, Nazir A, Chawla R, Arora R, Riyaz-Ul-Hasan S, Amna T, Ahmed B, Verma V, Singh S, Sagar R, Sharma A, Kumar R, Sharma RK, Qazi GN (2006) The endophytic fungus *Trametes hirsuta* as a novel alternative source of podophyllotoxin and related aryl tetralin lignans. J Biotechnol 122(4):494–510
61. Kusari S, Lamshoft M, Zuhlke S, Spiteller M (2008) An endophytic fungus from *Hypericum perforatum* that produces hypericin. J Nat Prod 71(2):159–162
62. Kusari S, Zuhlke S, Kosuth J, Cellarova E, Spiteller M (2009b) Light-independent metabolomics of endophytic *Thielavia subthermophila* provides insight into microbial hypericin biosynthesis. J Nat Prod 72(10):1825–1835
63. Kusari S, Verma VC, Lamshoft M, Spiteller M (2012) An endophytic fungus from *Azadirachta indica* A. Juss. that produces azadirachtin. World J Microbiol Biotechnol 28(3):1287–1294
64. Kusari S, Zuhlke S, Spiteller M (2009c) An endophytic fungus from *Camptotheca acuminata* that produces camptothecin and analogues. J Nat Prod 72(1):2–7
65. Kusari S, Zuhlke S, Spiteller M (2011) Effect of artificial reconstitution of the interaction between the plant *Camptotheca acuminata* and the fungal endophyte *Fusarium solani* on camptothecin biosynthesis. J Nat Prod 74(4):764–775
66. Puri SC, Verma V, Amna T, Qazi GN, Spiteller M (2005) An endophytic fungus from *Nothapodytes foetida* that produces camptothecin. J Nat Prod 68(12):1717–1719

67. Shweta S, Zuehlke S, Ramesha BT, Priti V, Mohana Kumar P, Ravikanth G, Spiteller M, Vasudeva R, Uma Shaanker R (2010) Endophytic fungal strains of *Fusarium solani*, from *Apodytes dimidiate* E. Mey. ex Arn (Icacinaceae) produce camptothecin, 10-hydroxycamptothecin and 9-methoxycamptothecin. Phytochemistry 71(1):117–122
68. Schroeckh V, Scherlach K, Nutzmann HW, Shelest E, Schmidt-Heck W, Schuemann J, Martin K, Hertweck C, Brakhage AA (2009) Intimate bacterial-fungal interaction triggers biosynthesis of archetypal polyketides in *Aspergillus nidulans*. Proc Natl Acad Sci U S A 106(34):14558–14563
69. Boersma FGH, Otten R, Warmink JA, Nazir R, van Elsas JD (2010) Selection of *Variovorax paradoxus*-like bacteria in the mycosphere and the role of fungal-released compounds. Soil Biol Biochem 42(12):2137–2145
70. Nazir R, Warmink JA, Voordes DC, van de Bovenkamp HH, van Elsas JD (2013) Inhibition of mushroom formation and induction of glycerol release—ecological strategies of *Burkholderia terrae* BS001 to create a hospitable niche at the fungus *Lyophyllum* sp. Strain Karsten. Microb Ecol 65(1):245–254
71. Partida-Martinez LP, Hertweck C (2005) Pathogenic fungus harbours endosymbiotic bacteria for toxin production. Nature 437(7060):884–888
72. Scherlach K, Busch B, Lackner G, Paszkowski U, Hertweck C (2012) Symbiotic cooperation in the biosynthesis of a phytotoxin. Angew Chem Int Ed Engl 51(38):9615–9618
73. He H, Bigelis R, Yang HY, Chang LP, Singh MP (2005) Lichenicolins A and B, new bisnaphthopyrones from an unidentified lichenicolous fungus, strain LL-RB0668. J Antibiot 58(11):731–736
74. Oh DC, Jensen PR, Kauffman CA, Fenical W (2005) Libertellenones A–D: induction of cytotoxic diterpenoid biosynthesis by marine microbial competition. Bioorg Med Chem 13(17):5267–5273
75. Oh DC, Kauffman CA, Jensen PR, Fenical W (2007) Induced production of emericellamides A and B from the marine-derived fungus *Emericella* sp. in competing co-culture. J Nat Prod 70(4):515–520
76. Zuck KM, Shipley S, Newman DJ (2011) Induced production of N-formyl alkaloids from *Aspergillus fumigatus* by co-culture with *Streptomyces peucetius*. J Nat Prod 74(7):1653–1657