Melanie-Maria Obermeier, MSc

Prospecting the Sphagnum microbiome for agriculture and medicine

DOCTORAL THESIS

to achieve the university degree of

Doktorin der Naturwissenschaften

submitted to

Graz University of Technology

Supervisor

Univ.-Prof. Dipl.-Biol. Dr.rer.nat. Gabriele Berg

Institute of Environmental Biotechnology

Graz, August 2019

AFFIDAVIT

I declare that I have authored this thesis independently, that I have not used other than the declared sources/resources, and that I have explicitly indicated all material which has been quoted either literally or by content from the sources used. The text document uploaded to TUGRAZonline is identical to the present doctoral thesis.

Date

Signature

The globally expanding threat of antimicrobial resistance (AMR) towards medicine and agriculture urges the development of novel antibiotics and fungicides, as well as a more comprehensive understanding of AMR distribution. Native plants are rarely explored to either end, despite their great potential for identifying novel drug leads and the existing knowledge gap about the plant resistome, in particular the intrinsic resistome.

To address all these needs a comprehensive prospecting approach was conceived aimed to: i) identify novel, antimicrobial lead compounds for the development of much needed antibiotics, ii) establish a high-throughput screening method for the identification of antimicrobial volatile organic compounds (VOCs) as ideal substances for biofumigation and in general surface sterilisation, and iii) elucidate the resistome of native plants. The study was directed towards the microbiome associated with *Sphagnum* mosses from pristine peat bogs. Previous work showed this evolutionary old plant species to harbour an extraordinary taxonomic diversity, a rich metabolism and a high share of antimicrobial traits. To get access to unexplored biosynthetic pathways and discover bioactives with a novel mode of action and new antibiotic resistance genes, functional metagenomics was employed. The elucidation of the *Sphagnum* resistome was additionally coupled with the screening of a culture collection and *in silico* analysis of a metagenomic dataset resulting in a multiphasic and first of its sort resistome analysis.

The bioprospecting strategy led to the identification of a novel biosynthetic gene cluster exerting an antifungal compound active against *Candida albicans*, the most important fungal, opportunistic human pathogen. Interestingly, the genes involved in the biosynthesis of the antimicrobial compound, which is proposed to be a fatty acid hydroperoxide metabolite, encode a new type of organic hydrogen peroxide resistance protein and a peroxidase. These proteins are known as crucial players in the oxidative stress response but have so far rarely been described with regard to antagonistic activity. Due to their functionality in reducing organic hydrogen peroxides, preferably fatty acid hydroperoxides, the antifungal compound may belong to a new class of antifungals and thus possess a potentially novel mode of action. Furthermore, a high-throughput screening method for the identification of antimicrobial VOCs from large clone libraries was successfully established. The new assay accelerates the screening by at least four-fold as compared to existing ones. Through functional screening of a metagenomic clone library, the vast potential of functional metagenomics for identification of antimicrobial VOCs from not-yet cultivable microorganisms was demonstrated. Lastly, the assessment of the *Sphagnum* resistome showed that plants naturally harbour a versatile, but also novel set of AMRs as exemplified by the discovery of the novel class A β-lactamase Mm3. The manifold and diverse resistance genes, thereby, play important roles for the taxonomic diversity and the bog ecosystem plasticity. Furthermore, *K*-selection was proposed as a strategy for microbial community management designed to contain exposure to resistances.

The undertaken prospecting approach highlights the potential of native plants for the identification of novel antibiotics and fungicides as well as their importance within the One Health concept. This study proves the unprecedented value of native plants for product discovery and AMR risk assessment, a topic which should be more emphasised in future research.

Die weltweit zunehmende Verbreitung von antimikrobiellen Resistenzen (AMR) bedroht Medizin als auch Landwirtschaft und drängt zur Entwicklung neuer Antibiotika und Fungizide sowie einem umfassenderen Verständnis über AMR Verbreitung. In beiderlei Hinsicht sind indigene Pflanzen nur wenig untersucht, trotz ihres großen Potentials für die Identifizierung neuer Wirkstoffe und der bestehenden Wissenslücke über das intrinsische Pflanzen-Resistom.

Um diese Fragestellungen anzugehen, wurde ein umfassendes Untersuchungskonzept entwickelt, welches zum Ziel gesetzt hatte: i) neue, antimikrobielle Wirkstoffe für die Entwicklung der dringend benötigten Antibiotika zu identifizieren, ii) ein Hochdurchsatz-Screening zu entwickeln, um antimikrobielle, flüchtige organische Stoffe (volatile organic compounds, VOCs) zu finden, da diese ideal für die Anwendung in Biofumigation oder generell zur Oberflächensterilisierung sind und iii) das Resistom einer heimischen Wildpflanze aufzuklären. Funktionelle Metagenomik wurde angewandt, um unerforschte Biosynthesewege auszuschöpfen und somit neue bioaktive Substanzen mit neuen Wirkmechanismen wie auch neue Resistenzgene zu finden. Zur Resistom-Aufklärung wurde zusätzlich eine Stammsammlung durchgemustert und eine *in silico* Datenanalyse eines Metagenoms durchgeführt; eine umfassende und bisher einzigartige Vorgehensweise. Die Untersuchungen drehten sich dabei um das Mikrobiom von *Sphagnum* Moosen aus den unberührten alpinen Hochmooren. Vorhergehende Forschungsarbeit zeigte, dass die evolutionär alten *Sphagnum* Moose eine außerordentliche taxonomische Artenvielfalt, einen reichhaltigen Metabolismus und starke antimikrobielle Eigenschaften besitzen.

Dies führte zur Entdeckung einer bislang unbekannten antifungalen Gengruppe, die das Wachstum von *Candida albicans* inhibiert, einem der wichtigsten, opportunistischen, humanpathogenen Pilze. Die biosynthetischen Gene kodieren ein neues organisches Wasserstoff Resistenz Protein und eine neue Peroxidase, die möglicherweise zur Bildung eines antimikrobiell wirkenden Fettsäurehydroperoxid-Abbauprodukts beitragen. Diese Enzyme, wohl bekannt als fundamentaler Bestandteil der oxidativen Stressantwort, wurden bisher selten mit Antagonismus in Verbindung gebracht. Aufgrund ihrer katalytischen Aktivität bei der Reduktion organischer Hydrogenperoxide, vorzugsweise Fettsäurehydroperoxide, könnten die entdeckten Gene und der neue Wirkmechanismus das Repertoire der antifungalen Wirkstoffgruppen erweitern. Des Weiteren wurde ein Hochdurchsatz-Screening zur Identifizierung von antimikrobiellen VOCs aus Klonbibliotheken entwickelt. Verglichen mit bestehenden Methoden, beschleunigt die neue Methode die Durchmusterung um mindestens das Vierfache. Ihre Anwendung bei der Durchmusterung einer metagenomischen Klonbibliothek demonstrierte das Potential von funktioneller Metagenomik neue, antimikrobielle VOCs von nicht kultivierbaren Mikroorganismen zu identifizieren. Die Aufklärung des *Sphagnum*-Resistoms führte zur Identifizierung der neuen Klasse A β-Laktamase Mm3 und zeigte darüber hinaus, dass Pflanzen eine angeborene und außergewöhnlich hohe Diversität an antimikrobiellen Resistenzgenen besitzen. Diese spielen eine wichtige Rolle für die taxonomische Vielfalt und die Plastizität des Moorökosystems. Zudem wurde *K*-Selektion als Strategie für das Management mikrobieller Gemeinschaften vorgeschlagen, um die Aussetzung gegenüber AMR einzudämmen.

Die durchgeführten Untersuchungsarbeiten betonen das Potential von Wildpflanzen für die Entdeckung neuer Antibiotika und Fungizide, sowie deren Bedeutung innerhalb des One Health Konzepts. Die vorliegende Studie bekräftigt den außerordentlichen Wert von indigenen Pflanzen für Naturstoffsuche, sowie AMR Risikoabschätzungen; Themen, die in künftiger Forschungsarbeit mehr in den Fokus gestellt werden sollten.

An African proverb says,

„If you want to go fast, go alone. If you want to go far, go together."

It is quite evident that doing a Ph.D. constitutes a journey one cannot master by walking it quickly and on one´s lonesome. There were many people who walked with me and whose help, guidance and encouragement got me where I am today. Here, I want to honour their part in this journey.

Firstly, I would like to express my gratitude to my supervisor Prof. Gabriele Berg. In view of the slow progress of prospecting approaches, her patience and continuous, positive support were essential to keep up my spirit. Although often demanding due to my different background, the insightful discussions we shared widened my research and view. I especially want to acknowledge her open-mindedness to promote my interest in communicating science beyond the academic boundaries.

Besides my advisor, I would like to thank my mentor Prof. Rolf Breinbauer for all the fruitful discussions, his input and ideas. Furthermore, I would like to thank Prof. Christine Moissl-Eichinger and Prof. Fröhlich for succeeding to the task of examining my Ph.D. thesis and defence.

My sincere thanks also go to my project leader Dr. Christina Müller Bogotá whose office door was always open and whose keen interest and unbreakable motivation were a driving force for the success of this project. Her feedback and guidance helped me throughout my Ph.D. study, for which I could not imagine a better project leader.

I thank Dr. Henry Müller and Dr. Tomislav Cernava for their help with data analysis as well as Barbara Fetz, Angelika Schäfer and Monika Schneider-Trampitsch for their technical assistance. Special thanks go to Isabella Wrolli whose technical help and attentive mind were indispensable for my research. I would like to thank all the project lab, bachelor´s and master students involved in this study for their contributions: Stephanie Hollauf, Franz Stocker, Sarah Prinz, Katharina Resch, Silvia Ferrario, Bettina Semler, and Harald Blasl.

I thank my peers, permanent and visiting ones, for their help and interesting discussions, for cheering me up in frustrating moments and for all the fun we have had. Especially the office - Julian Taffner, Alessandro Bergna and Manuel Reisinger - who are not only office mates but have become family.

Zu guter Letzt möchte ich meinen Eltern danken. Sie haben mich in dem Wissen aufgezogen nicht alles können zu müssen und ermutigten mich meine Stärken zu erkennen und zu verfolgen. Stets haben sie mich meine Wege frei gehen lassen, selbst wenn diese große Distanzen zwischen uns gebracht haben und sie es sich das eine oder andere Mal sicherlich anders gewünscht hätten.

This work has been supported by the Austrian BMWD, BMVIT, SFG, Standortagentur Tirol, Government of Lower Austria and Business Agency Vienna through the Austrian FFG-COMET- Funding Program.

Table of content

Table of content

Prospecting the *Sphagnum* microbiome

for medicine and agriculture

Exposé

Background

With antimicrobials being an integral part of human medicine, live-stock and crop farming alike, the ever-increasing emergence and spread of antimicrobial resistances (AMRs) threaten the advancements of modern medicine and pose a risk to food security (Bengtsson-Palme, Kristiansson and Larsson 2018; Fisher *et al.* 2018).

As AMR escalates globally so does the high morbidity and mortality associated with it. For instance, the estimated 700 000 AMR-related deaths in 2014 are foreseen to reach 10 Mio. by 2050 (O'Neill 2014). Simultaneously, antibiotics currently in the pipeline cannot meet the demand for new therapeutic agents with respect to the necessary number and especially the much-needed novel modes of action (Fernandes and Martens 2017). The wide-spread occurrence of resistances against antimicrobials is, thus, recognised as a major health crisis; of particular concern are resistance genes of clinical relevance such as extended spectrum β-lactamases (ESLBs) conferring resistance to the widely administered 3rd generation cephalosporins (World Health Organization 2014). These enzyme-coding genes, transmitted to the clinics via horizontal gene transfer, originated from the environment in which the diverse and ubiquitous family of β-lactamases has been around for millions of years (Bush 2018). In fact, many resistances predate the antibiotic era (D'Costa *et al.* 2011; Bhullar *et al.* 2012; Perron *et al.* 2015) and are widely spread in the environment, both in human controlled and pristine locations (Pal *et al.* 2016). While intra-community signalling and metabolic processes may constitute the original function of these genes, their ability to counteract antibiotics renders the environment a reservoir for novel antibiotic resistance genes (ARGs) from where they can transit to the clinics (Martínez 2008). Ensuing this realisation, the environmental ARG-distribution was elucidated in order to retrace their origin, understand their dissemination and perform risk assessments, which so far mainly focused on soil, water and air (Pal *et al.* 2016). Plants have been rarely investigated to this end, albeit more studies are required to determine the intrinsic plant resistome (Chen *et al.* 2019). This is a key for evaluating the role that plants play in ARG dissemination. Moreover, comprehensive analyses of native plants are missing entirely, can however unveil unexplored resistance mechanisms and advance our understanding of the ecology, evolution and functioning of plants in the context of AMR development and spread.

The focus regarding AMR centres around clinically relevant bacteria with lesser notion given to fungi (O'Neill 2014; World Health Organization 2014). Yet, fungal inflicted sepsis has been continuously increasing in the last three decades, whereby the genus *Candida,* specifically *C. albicans,* accounts as the leading causative agent of invasive fungal infections (Guinea 2014). Similarly to bacterial infections, invasive candidiasis is associated with high mortality rates of around 40% and an increase in resistance development against therapeutic drugs (Perfect 2017). Particularly striking in this respect is the limited amount of only four classes of antifungals available to combat this opportunistic yeast, but also other emerging human pathogenic fungi (Geddes-McAlister and Shapiro 2018). AMR in fungi is further of concern in the agricultural sector. While some taxa such as *Fusarium spec.* intrinsically tolerate many antifungals, resistance develops rapidly and today´s monoculture landscape

represents the perfect ground for resistant cultivars to spread readily (Fisher *et al.* 2018) which comes with an enormous economic burden. *Fusarium culmorum* for instance, produces the economically most relevant mycotoxin for the wheat sector and additionally to raising concerns regarding food safety lowers grain quality and yield (Khaneghah *et al.* 2018). For *Verticillium longisporum,* one of the major phytopathogens infesting oilseed rape, yield losses range between 10-50% (Depotter *et al.* 2016). The overall annual yield loss due to plant pathogenic fungi amounts to about 20% with an additional ~10% accounting to postharvest loss (Fisher *et al.* 2018). In contrary to ARGs, most clinically and agriculturally relevant antifungal resistances may have arisen by simultaneous evolution, but do as well cover every available drug class (Fisher *et al.* 2018). Hence, a great demand exists for new antifungal compounds, particularly those with a novel mode of actions, for the medical as well as the agricultural sector. For the latter, volatile organic compounds (VOCs) appear as a promising target class. These long distance communication molecules have a high vapour pressure at ordinary temperature and can exert antifungal activities (Kanchiswamy, Malnoy and Maffei 2015), making them suitable candidates for application in biofumigation.

In view of the escalating and rapid development of AMR in man-made environments important scientific questions arise, which need immediate action.

THE SCIENTIFIC PROBLEMS AT HAND:

1. There exists a great need for truly novel bioactive compounds for the development of new antimicrobial agents to combat the growing threat posed to society by resistant bacteria and fungi.
2. There exists a knowledge gap about the intrinsic plant resistome and overall about the intertwined circuits of environmental and clinical ARGs, specifically regarding native plants.

Natural products (NPs), the countless secondary metabolites naturally produced by living organisms, represent ideal scaffolds for the development of new antibiotics and fungicides. They constitute a diverse pool of bioactive compounds with an unprecedented structural and steric complexity that cannot be met by synthetic and combinatorial chemistry (Henkel *et al.* 1999; Feher and Schmidt 2003). Due to their selective bioactivity, making them evolutionary optimised lead compounds, and with their discovery rate correlated to drug approval rate, NPs are as relevant as ever for the development of pharmaceuticals and agrochemicals (Shen 2015).

However, traditional NP mining, faced by the problem of frequently re-isolating already known compounds, struggles to identify truly novel bioactives. This directs the focus to untapped bioresources like yet-uncultivable microorganisms or endophytes (Li and Lou 2018). Interestingly, nearly half of the NPs identified from microorganisms exhibit bioactivity (Bérdy 2012) indicating the great potential of microbial metabolites in general. In addition to exploring new sources, alternative prospecting methodologies are paving the way for successful NP discovery such as functional metagenomics. This cultivation independent methodology allows accessing and exploiting the biosynthetic pathways of the many yet-uncultivable bacteria by heterologous expression of their DNA in clone libraries (Handelsman 2004).

Great power lies in the combination of sourcing untapped bioresources and alternative prospecting methodologies, as detailed in **Chapter 1 'Prospects for biotechnological application of endophytes using functional metagenomics'**. Endophytic, secondary

metabolites span a diverse array of chemical structures including aliphatic compounds, peptides, steroids, and terpenoids (Gao *et al.* 2018). Their functionality is often associated with the symbiotic life style of these plant inhabiting microorganisms, including nutrient supply, plant growth promotion and pathogen defence (Hardoim *et al.* 2015). As outlined by Hardoim *et al.,* 23 different phyla colonise the plant interior – *Actinobacteria* among the most prevalent ones. Although *Actinobacteria* are well-known antimicrobial producers, endophytic *Actinobacteria* remain mostly unexplored (Qin *et al.* 2010).

The potential is, however, not constrained to endophytes only. Hundreds of species colonise the plant´s interior and exterior, which as a collective produce a rich and versatile, yet nearly untapped pool of secondary metabolites as described in **Chapter 2 'Bioprospecting plant-associated microbiomes'** in more detail. Orchestrated by the plant´s metabolites, microbial community composition differs from plant to plant and as a consequence thereof the microbial, metabolic profiles as well (Berg and Smalla 2009). Bioprospecting approaches should, therefore, integrate the functional and structural diversity of microbiota when selecting a target plant. As many cultivable and particularly not-yet cultivable microorganisms may cause the bioactivity linked to a certain plant (Li and Lou 2018), another valuable guiding principle resides in the ethnobotanical approach that is based on the traditional use of plants (Cox 2007). As depicted in chapter 2, this was already successfully pursued for the bryophyte species *S. magellanicum,* leading to the identification of *inter alia* putative antibiotic producing gene clusters (Müller *et al.* 2015).

The use of *Sphagnum* mosses and other bryophytes in traditional medicine due to their antifungal properties is well documented (Frahm 2004). They, thereby possess a nearly unexplored (Li and Lou 2018), but rich and specific metabolite pool, like in the case of *S. magellanicum* (Opelt *et al.* 2007b). Both the antifungal activity and the diverse metabolites may in part derive from the inhabiting microorganisms. For one, *S. magellanicum* hosts an extraordinarily high share of species with antagonistic activity (Opelt *et al.* 2007b; Müller *et al.* 2015). Secondly, its highly specific and stable microbial community belongs to 18 different phyla, with predominance of *Proteobacteria*, next to the phyla *Acidobacteria*, *Actinobacteria*, *Bacteroidetes* and *Verrucomicrobiota* but also including a high share of yet undescribed taxa (Bragina *et al.* 2012, 2014). The community composition is thereby stable across geographic location (Opelt *et al.* 2007a; Bragina *et al.* 2012); probably as a result of *Sphagnum´s* ecology. This native plant species forms the main component of raised bogs, which belong to the oldest vegetation forms on earth (Page and Baird 2016). In the prevailing harsh conditions of these water-saturated, acidic and low-nutrient environments and throughout a long history of co-evolution, *S. magellanicum* has built strong symbiotic interactions with its microbiota. Remarkable for plant resistome research, the *S. magellanicum* resistome can be regarded as intrinsic since its microbiota is unaffected by soil microorganisms given that *Sphagnum* moss grows on peat; accumulated, partly degraded plant material mostly stemming from the host plant itself.

Objectives

Because of the outstanding qualities of the microbiome associated with *S. magellanicum* to address all the above outlined scientific problems, a comprehensive prospecting approach was conceived.

Using functional metagenomics, the metabolically rich and diverse *S. magellanicum* microbiome sampled at the Austrian Alps was mined for novel antimicrobial NPs against clinically and agriculturally relevant (opportunistic) pathogens, including *C. albicans, Bacillus cereus, Staphylococcus aureus, F. oxysporum* and *V. longisporum.* Compounds of interest

were antimicrobial compounds, including VOCs, for application in biofumigation or for surface sterilisation in clinics and as promising lead molecules for the development of antibiotics. A 3.6 Gbps large metagenomic library constructed in *E. coli* during an earlier project (Müller *et al.* 2015) was used to this end. To make the screening of the large number of metagenomic clones feasible, efficient high-throughput screenings (HTS) were established to reliably and rapidly identify clones which inhibit the growth of the test organisms. Based on the well documented antimicrobial properties of the *Sphagnum* microbiota, the presence of versatile antimicrobial genes and consequently a diverse array of antifungal and antibacterial compounds was presumed.

The microbiome of this evolutionary old and native plant was further analysed for antibiotic resistances. To get a comprehensive insight, a unique combinatorial approach was pursued by screening a culture collection, a metagenomic data set and the metagenomic library. Due to the pristine origin of the *S. magellanicum* microbiome, its phylogenetic and functional diversity and stability it was expected to contain predominantly resistances against natural antibiotics, an overall versatile and evenly distributed pool of ARGs and yet-unknown ARGs.

THE OBJECTIVES OF THE PROJECT:

Employing the microbiome associated with *S. magellanicum*

1. identify novel drug candidates against *C. albicans, B. cereus,* and *S. aureus* by functional metagenomics.
2. establish and subsequently employ a HTS method for the identification of antimicrobial VOCs against *C. albicans, S. aureus, B. cereus, F. culmorum* and *V. longisporum* through functional metagenomics.
3. acquire a comprehensive understanding of a native plant resistome using a holistic screening approach including culture collection, *in silico* data mining and functional metagenomics.

Screening for novel drug candidates

About 90 000 fosmid clones from the *S. magellanicum* metagenomic library were screened against *C. albicans, B. cereus* and *S. aureus.* The HTS used to this end based on the agar overlay method. Library clones were pre-cultivated on different culture media and then overlaid with soft agar infused with the test organism. **Chapter 3 'Novel organic hydroperoxide resistance enzymes from uncultivated bacteria involved in antifungal activity against *Candida albicans'*** describes the identification of one active clone derived from the moss metagenome, which inhibited *C. albicans* and was denominated *E. coli* EPI300 pCC2FOS-CC (CC). CC contained a 44.8 kb large DNA insert holding 41 open reading frames (ORFs) as revealed by *de novo* sequencing and gene annotation. In a multi-step subcloning approach the number of genes potentially involved in the antagonistic activity was reduced to four, novel genes. The identified regulon consists of a putative sulfiredoxin and an organic hydroperoxide resistance (Ohr) gene cluster containing a putative OhrR, OhrB and a peroxidoredoxin homologue (Prx). *In silico* characterisation of OhrB and Prx by amino acid sequence and protein structure alignments confirmed the gene annotation.

The identification of the ohr regulon as the active gene cluster represents one of the first reports that ties peroxidase proteins to antagonistic activity. Known as important detoxification system that reduces organic hydroperoxides to their respective alcohols

(Fuangthong *et al.* 2001; Lesniak, Barton and Nikolov 2002), the produced inhibitory molecule may be a hydroxyl-compound. While any organic hydroperoxide could serve as potential substrate, fatty acid hydroperoxides were proposed as the natural substrates for Ohr peroxidases (Alegria *et al.* 2017) pointing towards hydroxy fatty acids as the inhibitory compound. Currently available antifungal drugs belong to polyenes, azoles, pyrimidine analogues and the cyclic lipopeptides echinocandins (Geddes-McAlister and Shapiro 2018). The discovered organic hydroperoxide regulon, thus, is highly interesting for further investigations representing a new type of antifungal system.

Screening for antimicrobial VOCs

As stated in **Chapter 4 'A new high-throughput screening method to detect antimicrobial volatiles from metagenonic clone libraries'**, antimicrobial VOCs were successfully isolated from yet-not cultivable microorganisms. For this purpose, an efficient HTS was established to make the screening of large clone libraries feasible. The strategy, thereby, built on the Two Clamps VOCs Assay (TCVA) that was developed to identify bacterial isolates producing antifungal VOCs (Cernava *et al.* 2015). The assay, a set-up of two 6-, 12- or 24-well plates separated by perforated silicon, creates a chamber in which the proliferation of fungal mycelium on one side is monitored as measure for the inhibitory efficacy of the bacterial isolate growing on the opposite side. For the high-throughput TCVA (htTCVA) the utilisation of 96-well plate was envisaged, which due to size constraint did not allow the usage of plaques of mycelium. Instead, macroconidia and spores were harvested from *F. culmorum* and *V. longisporum* and suspended in water used to inoculate the agar in the wells. This accelerated the screening process by four-fold as compared to 24-well plates, easily allowing the screening of 13000 clones per week.

To screen for antimicrobial VOCs against *C. albicans, B. cereus,* and *S. aureus* cell material of grown colonies was used to prepare suspensions. That way the target organism was infused into soft agar which was then pipetted into the wells of 96-well plates for the htTCVA.

In a two-step screening approach 18 000 moss metagenomic clones were screened in duplicate for antagonistic activity against *C. albicans, B. cereus, S. aureus, F. culmorum* and *V. longisporum* using the htTCVA as an initial step. Positive tested clones were subjected to the TCVA assay to verify their antagonistic activity. This resulted in the identification of six active clones; all inhibiting *F. culmorum*. As determined by a petri dish-based VOCs assay, their inhibition efficacy ranged from 3-9% and 20-30% for mycelium and spores, respectively.

For the six clones their VOCs were profiled using solid phase microextraction gas chromatography mass spectrometry. By comparing their volatilomes to that of the empty vector control strain nine potentially antifungal VOCs were determined. Eight VOCs could be identified, which included valeric acid, 1-decanol, 1-undecanol, 1-dodecanol, 2-tridecanone, phenoxyethylacetate, α-bisabolene and 2,2,4,6,6-pentamethylheptane. Their antifungal effect was verified *in vitro* through application of the purchased compounds as pure substance in the petri dish-based VOCs assay. The inhibitory capacity was thereby as high as 64% in the case of valeric acid.

The successful identification of antifungal VOCs confirmed the conceived approach, which accounts as the first example of a functional metagenomics screening for VOCs and demonstrates that functional metagenomics allows the identification of antimicrobial VOCs from uncultivable microorganisms. The htTCVA presents, thereby, an efficient HTS method to effectively screen large clone libraries for antimicrobial VOCs.

Elucidating the *S. magellanicum resistome*

A first comprehensive insight into the resistome of native plants was gained as highlighted in detail in **Chapter 5 'Unravelling native plant resistomes – The *Sphagnum* microbiome harbours versatile and novel antimicrobial resistance genes'**. The uncovered resistome is thereby regarded as intrinsic since *Sphagnum* mosses do not have contact to soil from where they could recruit external bacteria and their associated ARGs. The results, thus, address the still unanswered question regarding the intrinsic plant resistome and, thereby, emphasise that it presents an inherent and highly diverse repertoire of genetic signatures. This is affirmed by the identification of the novel class A β-lactamase Mm3.

The microbiome associated with *S. magellanicum* was analysed for AMR by subjecting a culture collection to a resistance screening using ten antibiotics; including agents of critical importance for therapeutic treatment (Collignon *et al.* 2016). Multi-resistant bacteria were frequently observed, including species commonly known to colonise *Sphagnum* mosses, such as *Paraburkholderia, Serratia* and *Rouxiella* (Belova, Pankratov and Dedysh 2006; Opelt, Berg and Berg 2007; Fléche-Matéos *et al.* 2017), but also for *Sphagnum* newly reported bacteria belonging to *Pandoraea* spp. Interestingly, *Pandoraea* spp. are considered emerging opportunistic pathogens (Green and Jones 2018). This highlights the importance of environmental resistomes in risk assessment as they can serve for early-onset identification of potential, novel opportunistic pathogens. Furthermore, resistance against all antibiotics was encountered, whereby against initial expectations, the bacterial isolates showed predominantly resistance to (semi)synthetic compounds. Due to the vast amount of metabolites exerted both by the competitive, highly plant-adapted *Sphagnum* microbiota and the moss itself (Bragina *et al.* 2014), microorganisms living within this community developed resistance mechanisms that naturally equip them against natural and (semi)synthetic antibiotics alike.

The great resistance capacity stems from specific and general resistance mechanisms as indicated by the conducted *in silico* analysis of a *Sphagnum* metagenome, which revealed a low abundant but evenly distributed and versatile assemblage of antibiotic resistance determinants spanning all antibiotic classes. This included 667 assigned ARGs and 220 assigned efflux pump determinants. A main driver for ARG diversity may reside in the highly abundant efflux pumps, which dominate the *Sphagnum* resistome to an extraordinarily high share of 80%. As a general resistance mechanism detoxifying a broad range of molecules (Martinez *et al.* 2009), they ensure co-existence within complex, metabolically active communities. Thus, they drive and maintain taxonomic diversity and in doing so foster ARG diversity and contribute, thereby, to ecosystem plasticity.

While efflux pumps shape diversity, evenness may be driven by *K*-selection. In contrast to the evenly distributed ARG pools found in natural environments, those of human-controlled environments comprise highly abundant, but less diverse ARGs (Pal *et al.* 2016). Environments under anthropogenic influence show, furthermore, lower bacterial diversity and simultaneously are linked to enrichment of opportunistic pathogens (Pal *et al.* 2016; Mahnert *et al.* 2019). It is, hence, interesting that the microbiota of *Sphagnum* is dominated by *Alphaproteobacteria* (Bragina *et al.* 2012, 2014), while that of arugula from urban gardening by *Gammaproteobacteria* such as multi-resistant *Enterobacteriacaea* (Cernava *et al.* 2019), which comprise many opportunistic pathogens. According to the ecological concept of *r*- and *K*-selection, microorganisms tending towards a copiotrophic lifestyle, like *Gammaproteobacteria*, thrive in carbon-rich and species-poor environments, while more oligotrophic organisms, such as *Alphaproteobacteria*, are favoured in nutrient poor, competitive environments (Kurm *et al.* 2017). By maintaining oligotrophy, *K*-Selection may stabilise community composition and as ARGs are often associated to certain taxa (Forsberg

et al. 2014; Goethem *et al.* 2018; Mahnert *et al.* 2019) maintain ARG evenness. *K*-selection, thus, represents a promising strategy for microbial community management which was suggested to reduce exposure to highly abundant ARGs (Mahnert *et al.* 2019).

THE KEY FINDINGS OF THE PROJECT:

1. Organic hydroperoxide resistance proteins were described, as one of the first reports, to be involved in the production of an antimicrobial compound representing a potentially novel class of antifungals.
2. A HTS method to identify antimicrobial VOCs through functional metagenomics was successfully established and demonstrated the potential to discover novel antimicrobial VOCs from yet uncultivable microorganisms.
3. A comprehensive insight into the intrinsic plant resistome was gained which beyond a more in-depth understanding of the ecology of ARGs gives new perspectives for risk assessment and management:
   - Efflux pumps and *K*-selection drive and maintain ARG diversity and evenness
   - Resistome evaluation can aid in early-onset identification of potential opportunistic pathogens
   - Microbial community management to reduce ARG exposure could be

Implications for medicine and agriculture

Great power lies in (functional) metagenomics and other omics-technologies which have transformed the field of microbiology. Their impact reached medicine, pharmacology and agriculture likewise. **Chapter 6 'Neues aus der Mikrobiomforschung: Von verbesserter Wirkstoffsuche, neuen Therapieansätzen und Resistenz-Management'** outlines the implications of the omics tool box for the pharmaceutical sector. In principle, omics-technologies can boost the success rate of NP discovery and drug development: i) they allow improved cultivation techniques to stimulate biosynthesis of certain NPs, ii) they facilitate the identification of new, promising antibiotic targets towards which drug development can be oriented (Pulido *et al.* 2016), iii) they enable a more targeted mining of resources through pre-evaluation. The value of the latter is herein demonstrated. The previous work on the *Sphagnum* microbiome gave a detailed understanding of its taxonomic diversity, rich metabolism and potential for the identification of especially antifungal compounds. This potential was successfully accessed by the identification of a metagenomic clone (*E. coli* EPI300 pCC2FOS-CC), which produces an antifungal compound with a potentially novel mode of action. Interestingly, the employment of functional metagenomics revealed the involvement of a foreign gene cluster in antifungal activity, the Ohr regulon, which otherwise may have been overlooked in this context due to its fundamental role in the oxidative stress response. The implications are, thereby, not restricted to pharmacology. The agricultural sector can benefit from omics-technologies the same way as the pharmaceutical sector does as exemplified by the screening for antifungal VOCs. Although generally promising for surface sterilisation, antimicrobial VOCs are particularly ideal for biofumigation. They are rarely applied to this end, meaning that a vast pool of new compounds is still to be sourced and benefitted from. Because of that, the identification of molecules with a novel mode of action is very likely. That VOCs identification through functional metagenomics is possible, was shown herein by the identification of eight antifungal volatiles inhibiting *F. culmorum*. The established, highly efficient HTS opens the doors towards the biosynthetic pathways of yet-not cultivable, environmental

microorganisms, a largely unexplored bioresource. That way it facilitates the discovery of truly novel VOCs and additionally allows a faster gene identification as compared to isolates. This adds additional value to the identification of VOCs via functional metagenomics, as not much is known about the mechanisms and genes involved in VOCs biosynthesis. The assay can however be employed not only for large clone libraries, but also for large culture collections accelerating the screening by several fold as compared to existing assays.

The exploited *Sphagnum* microbiome possessing a rich metabolism and a high share of antimicrobial properties is presumed to harbour a versatile, bioactive metabolites of which only a fraction was discovered during the employed bioprospecting approach. Thus, the metabolic pool is yet-not fully explored. Most functional metagenomics campaigns to date employ *E. coli* as host for heterologous gene expression. To source the metabolic potential of yet-not cultivable microorganisms to a wider extent, the usage of metagenomic libraries in alternative hosts is crucial. In this way a different cellular machinery gives access to a different range of metabolites. This approach, however, is still hampered by the lack of protocols and kits developed specifically for hosts. An obstacle which impaired the generation of a second moss metagenomic library in this study (data not shown). Further prospecting approaches, especially those employing alternative hosts, possess high chances of uncovering novel NPs from the manifold, still undiscovered bioactive metabolites.

Beyond the identification of novel producer strains and NPs, omics-technologies play an important role in the fight against AMR. Comprehensive elucidations of resistomes can be used not only for risk assessment of certain ARGs, but also to monitor and determine the efficacy of cleaning and sterilisation procedures, which are especially crucial in clinical environments and during pharmaceutical manufacturing processes (refer to chapter 6). By giving insight in the presence and abundance of ARGs, resistome analysis further provides a means to understand which antibiotics are most successful in a certain patient or a given environment. By understanding the ecological role and interplay of AMR, new ways to contain or even combat ARG spread can be found. One possibility is the herein proposed *K*-selection as a resistance management strategy to establish diverse, stable and beneficially designed microbiomes, foreseen to lower risk exposure (Mahnert *et al.* 2019). With the growing popularity of the One Health concept which acknowledges the interdependencies and circuits connecting humans, animals and the environment (McEwen and Collignon 2017), elucidation in form of resistome analyses will play an increasingly important role which includes human, animal and environment sources (Aenishaenslin *et al.* 2019). As described earlier, there exists a knowledge gap around the plant resistome (Chen *et al.* 2019). Plants too, as shown here by the first comprehensive analysis of the native *Sphagnum* moss resistome, harbour versatile AMR. The conducted analysis demonstrated that plants naturally possess a versatile set of innate ARGs, which might be propagated along with the host microbiome to the next generation via seeds (Bragina *et al.* 2012). Especially intrinsic ARGs that are vertically transmitted will play an important role within the frame work of One Health. As these ARGs are tied to the plant, they will be inherently transmitted through it. In terms of the One Health concept, the elucidation of native plant resistomes can provide valuable understanding of the potential of AMR found in nature. Through functional metagenomics novel and clinically relevant ARGs can be identified, but also emerging pathogens, which was demonstrated for the moss resistome. Such early-onset identification for risk assessment represents a crucial aspect, if we want to stay ahead of the ever-increasing AMR rise.

With the continuously decreasing costs associated with omics-technologies, they will play a more and more important part towards the identification of novel, antimicrobial NPs, the development of new drugs and risk assessment within the One Health framework. Natural and especially native plants should not remain exempt from such strategies. We might

otherwise miss valuable antimicrobials as well as overlook an important key for understanding AMR emergence and spread and for the development of containment strategies.

1 Prospects for biotechnological exploitation of endophytes using functional metagenomics

Melanie-Maria Obermeier[1,2] and Christina A. Müller Bogotá[1]

[1] *Institute of Environmental Biotechnology, Graz University of Technology, Graz, Austria;*

[2] *Austrian Centre of Industrial Biotechnology, Graz, Austria*

Abstract

The usage of natural products, especially in the treatment of diseases, has a long history. While natural products used to be administered directly, they today serve as lead compounds and structural scaffolds for the development of new drugs and other market products. The success of combinatorial approaches to develop new products strongly depends on natural product-likeness. This exemplifies the importance of natural products as structural leads during product development and demonstrates natural product discovery to be as important as ever. This chapter highlights endophytes as a rich bio-resource for the identification of novel natural compounds and emphasises functional metagenomics as a promising method to source the endophytic potential.

With the majority of microorganisms not readily cultivable under laboratory conditions, a vast number of natural products synthesised by endophytes remains inaccessible. Functional metagenomics circumvents current cultivation limitations by direct cloning of bacterial community DNA. This procedure is, however, rarely performed exclusively on endophytes. This chapter outlines the procedures underlying this methodology with focus on its application to endophytes.

Contribution in:

Shouten, A. (eds) *Endophyte Biotechnology: Potential for Agriculture and Pharmacology*. CABI Publishing, Wallingford, UK, pp. 164-179

Introduction

Plants, animals and microorganisms naturally produce innumerable metabolites and metabolic by-products of medical or industrial value. This value has been recognised long-since. The molecules produced by living organisms, generally referred to as natural products (NPs), find application in the medical, agricultural and industrial sectors as pharmaceuticals, agrochemicals, food additives and preservatives (David, Wolfender and Dias 2015; Schmitt, Hoepfner and Krastel 2016; Sparks, Hahn and Garizi 2017; Lorenzo *et al.* 2018; Ribes *et al.* 2018). They either are commercialised directly or serve as lead compounds and structural templates for effective product development. Today NPs and their derivatives constitute a high share among market products (Newman and Cragg 2016; Patridge *et al.* 2016). Due to their great structural diversity and manifold and specific bioactivity, they constitute highly promising resource for the identification of novel drugs, pesticides or preservatives.

Novel therapeutics and pesticides are required in the wake of the ever-increasing emergence of resistances to treat infections and ensure food security. Rising numbers of cancer patients and those suffering from chronic diseases urges the development of new treatments. Industries have been attempting to meet this demand for new lead compounds and scaffolds primarily by combinatorial chemistry since the 1990s. By parallel synthesis utilising several building blocks simultaneously, this approach enables the synthesis of large compound collections consisting of diverse variants of the starting material. Combinatorial chemistry, hence, represents until today a rapid way to generate molecular diversity. With the simultaneously expanding high-throughput screening methodologies, which allowed for fast screening of large compound libraries, this approach was believed to lead to compound discovery much quicker as compared to identifying a bioactive compound of interest from the complex NP mixture present in biological samples. This caused among other things the decline of NP discovery programs. Over the last decades various synthesis methods were established, and while success stories were reported, combinatorial chemistry could not yet meet the high expectations that were placed on this approach (Ortholand and Ganesan 2004; Lindell, Pattenden and Shannon 2009; Liu, Li and Lam 2017). The limited structural diversity of the generated molecules counts as one potential reason. Two studies conducted early on found significant differences in the chemical structures of NPs and those of combinatorial compounds. Both evaluations found NPs to span a wider range in molecular weight and to generally comprise a higher amount of oxygen atoms but fewer nitrogen, sulphur, and halogen atoms. Furthermore, they showed that chiral centres, which add to steric complexity and account as an important determinant for selectivity, are prominent in NPs and mostly missing in synthetic compounds. NPs were generally determine to be of more complex steric structures (Henkel *et al.* 1999; Feher and Schmidt 2003). Especially, Henkel *et al.* took a strong stand emphasising the importance of NPs with their vast chemical diversity as source for lead discovery. By comparing compounds from a NP and a synthetics database, they estimated that synthetic compounds only represent 60% of the NPs structural diversity (Henkel *et al.* 1999).

As the generation of compound libraries continued and evolved to increase the number of synthesised molecules, the concept of the ‘chemical space’ arose. Seen as the possible structural diversity of molecules, estimates range in several orders of magnitude. Depending on the chosen parameters such as the number and types of atoms, or whether it is a chemical or a peptide, calculations determined the possible number of molecules between 10^{8} and 10^{390}, whereby only a fraction might show bioactivity (Medina-Franco *et al.* 2008). Chemoinformatic analyses to compare and evaluate the chemical space of NPs, combinatorial libraries, pharmaceuticals, and other marketed molecules unroll ‘NP-likeness’ as an important criterion to be considered when expanding the current available chemical

space by designing and generating new compound collections (Medina-Franco 2012). This necessity is exemplified in the latest survey of Newman and Cragg, which states that the antitumor drugs sorafenib and vemurafenib and the pharmaceutical ataluren account as the only three *de novo* synthesised combinatorial molecules entering the pharmaceutical market (Newman and Cragg 2016). However, NPs, their derivatives and NP-like molecules contribute to more than half of the market products and represent, especially as lead compounds, a valuable asset for product development (Newman and Cragg 2016; Sparks, Hahn and Garizi 2017). Along with the trend of generating more NP-like compound libraries, NP discovery itself is, hence, living up a revival (McChesney, Venkataraman and Henri 2007; Shen 2015).

Similarly, novel enzymes, which are generally not considered NPs, are in demand for the food and pharmaceutical industries (Coughlan *et al.* 2015). As outlined in detail by Coughlan *et al.* microbial enzymes find *inter alia* application in food processing, flavouring, dairy products, brewing and baking and novel bio-catalysts for the various processes are required: lipases for milk fat hydrolysis, esterases for flavour production in the beverage industry, α-amylases for starch-modification in the baking industry, among other things. Another class of enzymes which caught considerable attention over the last years is non-ribosomal peptide synthetases and polyketide synthases. As producers of secondary metabolites, many bioactive compounds synthesised by these enzymes have been identified (Nikolouli and Mossialos 2012).

Only a fraction of all the natural resources are exploited (Henkel *et al.* 1999; Bérdy 2012). The diversity of microbial NPs is mostly not yet investigated. According to estimates, the diversity of bacterial compounds smaller than 1 kDa reaches 10^9 molecules (Davies 2007) of which merely 60 000 - 80 000 metabolites are identified (Bérdy 2012). Interestingly, Bérdy (2012) reported that almost half of the known microbial NPs showed bioactivity, exemplifying the great potential of microbial NPs in general. This goes well in hand with an evaluation by Patridge *et al.* showing that half of the Food and Drug Administration (FDA)-approved NPs originate from bacteria and fungi (Patridge *et al.* 2016). While bacteria and fungi regain importance as bio-resource after a drop-down following the golden era in microbial drug discovery from 1960 to 1980, screenings focus primarily on microbial communities in soil, sludge and rhizosphere. The plant associated and especially plant inhabiting bacteria and fungi are widely unexplored (Strobel and Daisy 2003; Müller, Obermeier and Berg 2016).

Each plant is inhabited by a diverse and specific community of microorganisms which fulfil important functions for the host including nutrient supply, plant development, pathogen defence and stress resilience (Hardoim, van Overbeek and Elsas 2008; Brader *et al.* 2014; Hardoim *et al.* 2015). It was estimated that between 250 000 and 500 000 higher plant species populate the planet (McChesney, Venkataraman and Henri 2007), of which only a fraction has been investigated for their endophytes (Strobel and Daisy 2003). The species richness and diversity of endophytes is staggering. According to estimates, 1 million fungal endophytes can be found (Dreyfuss and Chapela 1994). Drivers for the endophytic community composition are: 1) biotic factors such as plant genotype, plant physiology, the microorganisms present in the surrounding bulk soil and 2) abiotic factors like soil type, nutrient availability, temperature (Hardoim *et al.* 2015; Kandel, Joubert and Doty 2017). Interestingly, plants actively recruit certain endophytes from the bulk soil (Lemanceau *et al.* 1995; Kloepper, Ryu and Zhang 2004; Rudrappa *et al.* 2008) and even propagate them to their offspring (Truyens *et al.* 2015; Frank, Saldierna Guzmán and Shay 2017; Shahzad *et al.* 2018). Similarly, mosses and other lower plants comprise endophytic communities, whereby the community structure underlies the same principles (Shcherbakov *et al.* 2013). These highly adapted microbial consortia represent widely unexplored bio-resources which comprise a unique set of metabolic pathways and a tremendous amount of NPs and

enzymes (Strobel and Daisy 2003; Gunatilaka 2006; Brader *et al.* 2014; Müller, Obermeier and Berg 2016).

This chapter focuses on the discovery of NPs and enzymes of microbial origin, highlighting endophytes as a plenteous and yet less explored bio-resource. Thereby, we assess the potential of sourcing endophytes by functional metagenomics, a methodology which facilitates the access to unexplored NPs and enzymes of yet uncultivable microorganisms.

Functional Metagenomics of Endophytes

The biotechnological potential of endophytes has long been recognised and sourced for NPs, which led to the isolation and identification of many novel compounds as summarised in various reviews (Tan and Zou 2001; Strobel and Daisy 2003; Gunatilaka 2006; Zhang, Song and Tan 2006; Chen *et al.* 2014; Deshmukh, Verekar and Bhave 2014; Newman and Cragg 2015; Martinez-Klimova and Rodríguez-Peña 2017; Nalini and Prakash 2017; Gao *et al.* 2018). The therein described work exploited the endophytic potential exclusively by cultivation-dependent techniques. Such screenings, however, allow only a glimpse into the endophytic treasure chest of metabolites, metabolic by-products and enzymes. The great plate count anomaly (Stewart 2012), the fact that most microorganisms are not readily cultivable under laboratory conditions, endophytes are not exempt from. Generally, about 99% of all microorganisms account as not yet cultivable by standard methods (Strobel and Daisy 2003). Following this, NPs and enzymes of the majority of the species inhabiting a plant are not accessible by cultivation-dependent techniques.

Cloning of microbial community DNA from the sample directly into a surrogate host represents a promising methodology to circumvent cultivation dependency. The great potential lies therein, that this can be applied to any microbial community (Ravin, Mardanov and Skryabin 2015). As heterologous expression of such DNA facilitates access to hidden genes and biosynthetic gene clusters, this method allows identification of novel NPs and enzymes from yet uncultivable microorganisms (Coughlan *et al.* 2015; Katz, Hover and Brady 2016), including endophytes. This methodology generally belongs to the metagenomics approaches which commonly investigate the collective genomes of all members of a microbial community, the metagenome (Handelsman *et al.* 1998). Mainly grouped into sequence- and function-based approaches, metagenomics comprises different methods as outlined by different reviews (Handelsman 2004; Simon and Daniel 2011; Coughlan *et al.* 2015; Ravin, Mardanov and Skryabin 2015). Summarily, metagenomics in its broadest interpretation also comprises amplicon sequencing of conserved gene regions, mostly the 16S rRNA marker gene for bacteria or the ITS region for fungi. Inferred from such species-specific marker genes the phylogenetic community composition can be analysed and based on knowledge about functional traits of single phyla the functional potential of a microbiome can in part be evaluated. In comparison, shotgun sequencing of adapter-ligated metagenomic DNA fragments or metagenome clone libraries allow a detailed analysis of functional properties based on gene homologies. Annotated genes can then be cloned for heterologous expression. With neither approach, however, can new genes and consequently novel NPs or enzymes be identified, which is where functional metagenomics comes into play. This methodology can lead to the identification of truly novel NPs and enzymes. Functional metagenomics is based on direct cloning of metagenomic DNA and therefore requires the isolation of high-quality metagenomic DNA, fragmentation and cloning of the metagenomic DNA into a vector and transformation into a host. The generated metagenomic library can then either be explored by polymerase chain reaction (PCR)-based screenings using degenerated primers which bind to conserved regions of the gene target of

interest or by phenotypic screenings for a desired activity. As this chapter assesses the exploitation of the endophytic potential by phenotypic-based functional metagenomics, this approach will be elaborated on in the following.

Historically, the first mile stone for functional metagenomics was laid by Pace *et al*. After proposing direct cloning of metagenomic DNA in 1986 (Pace *et al.* 1986), they constructed the first metagenomic library 5 years later for 16S rRNA gene sequencing and phylogenetic analysis of picoplankton-associated microorganisms (Schmidt, DeLong and Pace 1991). Almost another 5 years later, the first successful phenotypic screening of a metagenomic library resulted in the identification of four novel cellulases (Healy *et al.* 1995). Functional metagenomics has since then led to the discovery of many novel enzymes and bioactives from uncultivable microorganisms (Coughlan *et al.* 2015). The identification process generally involves two steps: the generation and subsequently the functional screening of metagenomic libraries. The process can be further broken down into several steps: plant selection and sampling, surface sterilisation including the indispensable usage of a DNA-degrading agent such as sodium hypochlorite (as an additional step when metagenomic DNA of only endophytes is desired), microbial enrichment, DNA isolation, cloning, host transformation and the screening (Fig. 1.1). Once a metagenomic clone has been identified, biosynthesis genes contained on the metagenomic DNA insert are more readily identified as compared to identifying the gene clusters within the whole genome of isolated microorganisms.

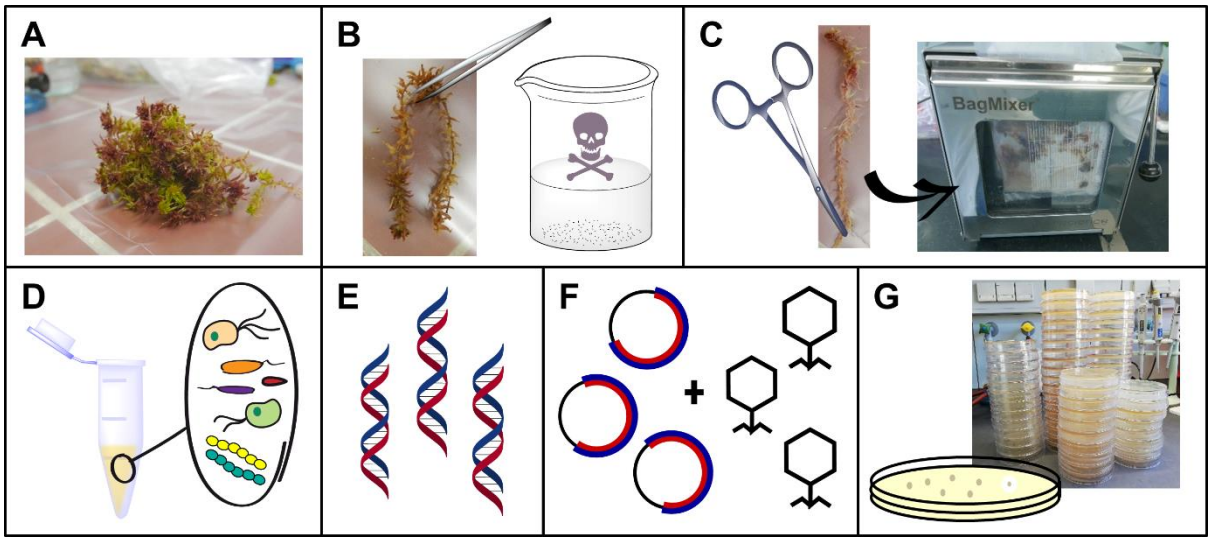


Figure 1.1: Generating of a metagenomic clone library from endophytes. The following steps are necessary for the construction of a clone library, here exemplarily shown for moss: (A) Plant selection and sampling. (B) Surface sterilisation of the fresh plant material including a DNA-degrading step using for instance sodium hypochlorite. (C) The plant material is treated by mechanical or enzymatic methods to enrich the microbial fraction (e.g. cutting, homogenisation, treatment with salt or detergents, bag-mixing, centrifugation). (D and E) The enriched endophytic microbiome is used for isolation of metagenomic DNA. (F) The DNA is ligated into an appropriate vector system and cloned into the library host, e.g. using fosmids and packaging them into phages prior to transfection of the host. (G) The generated clone library undergoes an activity- or sequence-based screening, which is designed to identify a desired activity or specific genetic traits.

The selection of a plant species for constructing of a metagenomic clone library (Fig. 1.1A) certainly depends on the molecule or the activity that is targeted for screening. A very promising way of selecting a plant as source of endophytes, not only for isolation of microorganisms but for isolation of metagenomic DNA, is the so called ethnobotanical approach (Cox 2007). Thus, the ethnobotanical knowledge from native people or the traditional use of plants in herbal medicine for treatment of diseases is employed to find

interesting candidates. Once the selected plant is collected, the treatment of the plant material for construction of the library should proceed without delay to avoid contamination or loss of microbial diversity.

Surface sterilisation of the plant material is the most commonly employed technique to obtain solely endophytic microorganisms (Hallmann, Berg and Schulz 2006) (Fig. 1.1B). Furthermore, high-quality DNA must be obtained from the selected material. In general, different isolation protocols and commercial kits are available, but mainly for soil and water samples (Leis, Angelov and Liebl 2013). When working with endophyte communities a big challenge resides in the extraction of endophyte DNA in enough quality and quantity for library construction. A study by Gabor *et al.* reports that direct DNA isolation, mechanically, enzymatically or using detergents, yields higher DNA amounts than indirect approaches for which the microbial cells are extracted and enriched prior to cell lysis. When evaluating the isolated DNA, they showed that lysates obtained by direct methods contained considerably higher amounts of eukaryotic DNA (>50%) as compared with indirect methods which yielded lysates containing more than 90% bacterial DNA (Gabor, Vries and Janssen 2003). As metagenomic libraries contain several thousand clones of which only a small amount will exhibit the desired phenotype (Handelsman 2004), it is desirable that metagenomic libraries contain as much endophytic DNA as possible. As mentioned above, extracting enough high-quality DNA from endophytes for library production is challenging. The high yield of plant DNA obtained by standard methods and the fact that the plant DNA interferes with further processing and analysis steps make usually an enrichment of the microbial fraction indispensable (Jiao *et al.* 2006) (Fig. 1.1C). Due to this restraint, only a few examples of library construction from endophyte DNA are available.

A few methods for enriching the plant microbiome, e.g. from stems, and simultaneously reducing or eliminating the plant DNA have been described (Wang *et al.* 2008; Ikeda *et al.* 2009). Wang *et al*. extracted and enriched microbial DNA from stem bark material using a combined treatment with salt and detergent (0.9% NaCl, 0.063% SDS) for disrupting plastids and eliminating plant DNA (Wang *et al.* 2008). A similar method was proposed by Ikeda *et al.* for enrichment of the bacterial fraction in soybean stems (Ikeda *et al.* 2009). In this case Triton X-100 was used as a mild detergent for disrupting the membranes of chloroplasts, in combination with a density gradient centrifugation using the non-ionic medium Nycodenz. Other microbial enrichment methods, for example, for sugarcane stems (Dos-Santos *et al.* 2017) or for leaves from the *Maytenus hookeri* tree (Jiao *et al.* 2006), have been reported; the latter involves the enzymatic hydrolysis of plant cell walls and differential centrifugation. The method of choice will ultimately depend on the plant morphology and composition, and has to be adapted and evaluated for each individual case.

For library establishment (Fig. 1.1D-F), several considerations have to be taken into account. Choice of the vector host system is crucial, as successful expression is a necessity for the later screening. Heterologous gene expression depends on various factors. Different codon usage and the metabolic background of the surrogate host in comparison to that of the species from which the metagenomic DNA insert derives from may differ greatly and can hamper gene expression (Liebl *et al.* 2014). The most widely used host strain is *E. coli* for which many protocols and commercial kits for DNA extraction and library generation exist, aiding in the procedure and making it more efficient (Simon and Daniel 2011). Yet, the predicted potential of *E. coli* to heterologously express metagenomic DNA varies greatly ranging from as low as 7% up to 73%, whereby expression of one-third of the genes depends on the vector promoter (Gabor, Alkema and Janssen 2004). This drives the interest towards different species such as *Pseudomonas* or *Streptomyces* strains for library generation (Liebl *et al.* 2014). Yeast could be theoretically employed as an alternative host for the directed expression of eukaryotic metagenomic DNA. Cloning vectors containing broad-host replicon

such as the RK2-based plasmid have been developed (Aakvik *et al.* 2009). This vector system can be in principle transferred to gram-negative and gram-positive bacteria and also eukaryotic hosts like yeast. Nevertheless, there are no reports available detailing the successful application of yeast as a metagenomic library host. Although being more efficient producers and more promising hosts for the identification of novel NPs, the establishment of metagenomic libraries in alternative hosts remains a laborious procedure. Another important question is that of the insert size to be used. With increasing insert size, the likelihood rises that complete biosynthesis pathways and gene clusters are cloned. This is interesting when the screenings targets metabolites or big enzyme complexes such as nonribosomal peptide synthetases. Cosmid or fosmid libraries holding high-molecular-weight DNA up to 40 kb or bacterial artificial chromosomes (>40 kb) are generated to this end (Simon and Daniel 2017). Yet, high-molecular-weight inserts affect the cloning efficiency limiting the resulting number of library clones. Small DNA fragments are cloned more readily, which favours libraries with low-molecular-weight inserts that are generally used for screenings of different enzymes classes like hydrolases (Simon and Daniel 2017).

As mentioned earlier metagenomic libraries comprise several thousand clones, whereby only a small fraction will exert the activity of interest (Handelsman 2004). Therefore, a high-throughput screening to process many clones simultaneously for the desired phenotype needs to be in place to make the screening of large clone collections feasible. Such screenings, however, come with the drawback of potentially missing clones of interest (Coughlan *et al.* 2015). As reviewed by Leis *et al.* as well as by Simon and David, screenings may involve the detection of enzymes by supplementing the growth medium with indicator reagents specific for tracing the desired enzymatic activity (called phenotypic detection) or the use of reporter genes for which expression is triggered only once the compound of interest is present. A further approach represents heterologous complementation of the host by the gene of interest. Only with the target gene being present will growth be observed under selection pressure (Simon and Daniel 2011; Leis, Angelov and Liebl 2013). For instance, phenotypic detection led to the discovery of six novel polyesterases from a moss metagenomic library (Müller *et al.* 2017). The underlying high-throughput screening procedure used tributyrin containing agar plates where hydrolytic activity could easily be spotted by halo formation. Novel 4´-phosphopantetheinyl transferases (PPT) from a metagenomic library in *E.coli* and *Streptomyces albus* were identified by a coupled *bpsA* reporter gene PPTase complementation approach (Owen *et al.* 2012; Bitok *et al.* 2017). Thereby, the PPT was deleted and pigment production by the PPT dependent *bpsA* gene only restored upon expression of a functional PPT. Furthermore, Fluorescens Activates Cell Sorting (FACS)-based screening methods have been established. The two methods, substrate-induced gene expression screening (SIGEX) (Uchiyama *et al.* 2005) and metabolite-regulated expression screening (METREX) (Williamson *et al.* 2005), are based on the induction of *gfp* reporter gene expression. While the latter employs a quorum-sensing promoter upstream of the *gfp* gene, for SIGEX the reporter gene is promoter-less. Hence, for SIGEX reporter gene expression depends on the presence of a promoter on the metagenomic DNA insert. Uchiyama *et al.* used this approach to identify clones that express enzymes which convert a substrate of interest - in their case, hydrocarbons. By adding the substrate of interest, the desired promoters get activated and drive the expression of not only the downstream enzyme, but also the reporter gene. METREX, on the contrary, requires the synthesis of the promoter activating molecules and facilitates the identification of signal molecules. However, both methodologies employing FACS for cell sorting have the advantage of a very efficient high-throughput screening.

A metagenomic analysis targeting the community composition and functions of endophytes in rice roots was reported for the first time by Sessitsch et al. (2011). Using the

metagenome data, the authors predicted main microbial adaption mechanisms supporting an endophytic lifestyle, for example, the availability of plant-polymer-degrading enzymes, iron acquisition and storage, protein secretion systems, among others. Later on, the endophyte community of other plant species and plant compartments including *Arabidopsis thaliana* roots (Bulgarelli *et al.* 2012), grapevine branches (Campisano *et al.* 2014), sugar beet (Shi *et al.* 2014), *Aloe vera* root, stem, and leaves (Akinsanya *et al.* 2015), tomato roots (Tian, Cao and Zhang 2015), or floating fern (*Azolla filliculoides*) (Dijkhuizen *et al.* 2018) has been evaluated. However, functional metagenomics, i.e. the screening of clone libraries for identification of new NPs, has not been in the focus of research yet. To the best of our knowledge, there is only one published study by Nikolic et al. thus far, targeting exclusively the endophytic microbial community for clone library generation and subsequent screening (Nikolic, Schwab and Sessitsch 2011). Here, using a few selected examples, we will highlight the potential of endophytes for the discovery of new natural products and enzymes.

Natural Products of Endophytes

The production of functional secondary metabolites by endophytes is linked to the improvement of plant fitness (Tan and Zou 2001). Besides plant-growth promotion, one major function of microbial metabolites is protecting the plant against biotic and abiotic stress, e.g. by inducing resistance against pathogens (Bailly and Weisskopf 2012). The structural and chemical diversity of NPs is extensive, including alkaloids, steroids, terpenoids, peptides and aliphatic compounds, among others (Tan and Zou 2001; Gao *et al.* 2018). A main group of interesting secondary metabolites is composed of high-molecular compounds, such as peptides and polyketides, encoded by nonribosomal peptide synthetases and polyketide synthases. This type of metabolites display complex structural diversity, along with a broad range of biological activities and functions, such as antibacterial, antifungal and cytotoxic activity, or acting as metal chelators (siderophores) (Cane and Walsh 1999). In this way they also support the lifestyle of endophytes in association with the host. Promising sources for this type of NPs are actinobacteria and fungi. Especially endophytic actinobacteria are regarded as a nearly unexplored reservoir of bioactive secondary metabolites (Qin *et al.* 2010). Most clinically relevant antibiotics used today have, for instance, their origin in actinomycetes (Baltz 2007). Comprehensive reviews on the discovery of antibiotics from endophytes were published by Deshmukh *et al.* (2014) and Martinez-Klimova and Rodrígez-Peña (2017).

The enhanced acquisition of iron has been generally hypothesized as a central aspect in the life cycle of endophytes (Reinhold-Hurek and Hurek 2011), in particular for nitrogen-fixing bacteria since this process is iron-dependent. Siderophores, which are mostly nonribosomal peptides (Crosa and Walsh 2002), have high chelating affinity for iron and they contribute to the nutritional requirement of microorganisms and the plant host. A new class of siderophores, the so called serobactins, were identified in the grass endophyte *Herbaspirillum seropedicae* Z67, a bacterium of interest due to its nitrogen-fixation ability (Rosconi *et al.* 2013). Another type of siderophore, epichloënin A, was discovered as a product of a fungal endophyte *Epichloë festucae*, which lives in perennial ryegrass (Koulman *et al.* 2012). This type of fungal endosymbiosis in temperate grasses not only improves the herbivore resistance of the plant (Lane, Christensen and Miles 2000). It was also shown that production of epichloënin A is required for maintaining of a mutual beneficial interaction between the fungus and its host (Johnson *et al.* 2013).

The cyclic depsipeptide FR900359 is a further example of a non-ribosomal peptide from an uncultivable endosymbiont, *Candidatus* Burkholderia crenata, living in the tropical plant *Ardisia crenata* (Crüsemann *et al.* 2018). The depsipeptide produced mainly in the leaf

nodules by the endosymbiotic partner functions probably as a protective defence chemical against plant herbivores like insects and nymphs. In medicinal applications, this peptide is an indispensable tool for pharmacological studies of cellular signalling processes, being a potent inhibitor of guanine nucleotide binding proteins. For more detailed examples, the review by Abdalla and Matasyoh gives a good overview of different peptide classes isolated from endophytes (Abdalla and Matasyoh 2014).

Polyketides, which are synthetized by large and iterative multifunctional proteins, so-called polyketide synthases, are also interesting bioactive NPs. Several new polyketides have been reported in fungal endophytes. For example, six novel bicyclic polyketides, the so called preussilides, were isolated from the fungus *Preussia similis*, an endosymbiont of *Globularia alypum*, and showed antiproliferation activity on eukaryotic cell lines. Similarly, different types of compounds belonging to the family of oblongolides were isolated from *Phomopsis oblonga*, and endophyte, from wild banana. Some of the isolated *oblongolides* displayed cytotoxic activity (Bunyapaiboonsri *et al.* 2010). Recently, another group of cytotoxic polyketides was found in a related fungus, *Phomopsis* sp. A818, isolated from mangrove (Zhang *et al.* 2017). A dimeric anthraquinone called skyrin was identified as a pigment of the fungal endophyte *Cyanodermella asteris* (Jahn *et al.* 2017). *C. asteris* was isolated from the plant *Aster tataricus*, which has been employed in traditional Chinese medicine as expectorant, and showing anti-inflammatory properties as well (Yu *et al.* 2015). Through in silico analysis of the *C. asteris* genome, putative biosynthetic pathways for production of skyrin were elucidated, suggesting the involvement of a non-reducing polyketide synthase (Jahn *et al.* 2017).

Another group of metabolites produced by the plant microbiota are volatile organic compounds (VOCs). VOCs enable chemical inter- and intra-species communication, over longer distances than non-volatile compounds (Kanchiswamy, Malnoy and Maffei 2015). Many microbial volatiles, or mixtures thereof, have been investigated for their bioactivity, and especially for the ability to antagonise plant pathogens (Berg 2009). One of the most prominent and first discovered examples of VOCs-producing microbes is the endophytic fungus *Muscodor albus*, isolated by Strobel *et al.* in the late 1990s from a cinnamon tree in a botanical garden in Honduras (Sears *et al.* 2001). The mixture of VOCs produced by *M. albus* contained different classes of organic substances (esters, alcohols, lipids, ketones and acids), which showed antibiotic effect on several plant and human pathogenic bacteria and fungi. While the single volatiles only inhibited the growth of the test organisms, the mixture thereof showed a potent lethal activity, being the most active single volatile isoamyl acetate.

Since this first discovery, efforts have been undertaken to isolate new VOC producers and evaluate their bioactivity and biotechnological potential. For instance, volatiles and semi-volatiles from endophytic fungi like *Hypoxylon anthochroum*, *Gleosporium* sp. and *Geotrichum candidum* PF005 have been studied, showing growth inhibition of important plant pathogens like *Fusarium oxysporum*, *Phytophthora palmivora, Rhizoctonia solani* and other fungi (Schaible *et al.* 2015; Ulloa-Benítez *et al.* 2016; Medina-Romero, Roque-Flores and Macías-Rubalcava 2017; Mookherjee *et al.* 2018). Not only fungi but also bacteria are capable of producing VOCs, many of them having plant-modulating properties and disease-suppressing activities (Weisskopf 2013). New volatiles identified from bacterial endophytes are scarcer than those reported for fungi. However, bacteria from the genera *Pseudomonas*, *Bacillus*, *Serratia* and *Stenotrophomonas* are well-known VOCs producers (Bailly and Weisskopf 2012), many of them being capable of an endophytic lifestyle. This is the case for *Pseudomonas putida* BP25, an isolated endophyte from the root of black pepper (Sheoran *et al.* 2015). This bacterium showed production of some well-known antimicrobial VOCs like 1-undecene and different types of pyrazines.

In general, sourcing endophytes for bioactive compounds attributed to plants represents a promising alternative to exploiting the plant itself. This is of interest when the commercial supply cannot be maintained, for instance due to the compound being isolated from slow growing or rare plants and its molecular structure being highly complex so that chemical synthesis is not a suitable option (Strobel and Daisy 2003; McChesney, Venkataraman and Henri 2007). A well-known example for this dilemma is the anticancer drug paclitaxel (Taxol®), a diterpenoid originally isolated from the bark of *Taxus brevifolia*, Western Yew (Cragg 1998). The search for alternative sources for paclitaxel extraction revealed not only other plant species to produce this compound, but also several plant endophytes, mainly fungi (Stierle, Strobel and Stierle 1993; Kharwar *et al.* 2011). Other examples include the anticancer compound camptothecin which was first isolated from a *Nyssaceae* (Wall *et al.* 1966) but is also produced by the endophytic fungi *Fasurium solani* (Kusari, Zühlke and Spiteller 2009); and the potentially antidepressant hypericin, first isolated from St. John´s wort (Aly, Debbab and Proksch 2013) but later found to be produced by the fungal endophyte *Thielavia subthermophila* (Kusari *et al.* 2008). These and several other examples are reviewed in detail by Aly *et al.* (2013).

Enzymes from Endophytes

Microorganisms living in close association with plants often produce a wealth of different (extracellular) enzymes for the degradation of plant polymers and oligomers. For example, the production of an endoglucanase, a cell wall-degrading enzyme, was reported as a key factor for the initial and active bacterial colonization of internal plant tissues (Reinhold-Hurek *et al.* 2006).

A metagenomic investigation of the root gall-associated microbiome in tomato plants showed a high abundance of oligosaccharide-degrading genes; however, only a lower frequency of genes coding for enzymes acting on full-length polymers, like cellulases or hemicellulases, was detected (Tian, Cao and Zhang 2015). Moreover, several endophytic fungi isolated from medicinal plants, mangrove or the shrub *Brucea javanica* were tested positive for extracellular enzyme activities, like cellulases, lipases, amylases, laccases, pectinases or proteases (Choi *et al.* 2005; Maria *et al.* 2005; Sunitha, Nirmala Devi and Srinivas 2013).

Ligninolytic enzymes, in particular laccases, are a group of enzymes commonly found in wood-decomposing fungi (Singh Arora and Kumar Sharma 2010). Compared to fungi, reports on laccases from bacterial origin are rare. Endophytes might represent a new source for this type of bacterial enzymes. A new bacterial laccase, showing lignin degradation, dye decoloration and acid-stable properties, was found in the rice endophyte *Pantoea ananatis* Sd-1 (Shi *et al.* 2015).

Chitin is a major constituent of the fungal cell wall. Chitinases and chitin-modifying enzymes are of biotechnological interest. One of the first reported enzymes was an extracellular chitobiosidase from *Bacillus cereus*, an endophyte isolated from mustard. The presence of extracellularly produced proteins from this bacterial strain decreased the rate of germination of spores from the plant pathogen *Fusarium oxysporum* and supported the idea, that the production of extracellular enzymes might protect the plant from fungal infection (Pleban, Chernin and Chet 1997). The availability of chitin-modifying enzymes in fungal endophytes was also investigated by Govinda et al. A high genetic diversity of this class of enzymes was found for several fungal isolates (Govinda Rajulu *et al.* 2011). In a further study, a new chitin deacetylase from the endophytic fungi *Pestalotiopsis* sp. was isolated and characterized. It was shown that the enzymatic deacetylation of the chitin

oligomers could be part of the survival strategy of the fungus inside the plant, since the modified oligomers were no longer recognized by the plant´s immune system (Cord-Landwehr *et al.* 2016).

Studies have also focused on the investigation of beneficial enzymes for the host plant, like those derived from plant-growth-promoting bacteria. This is the case for 1-aminocyclopropane 1-carboxylic acid (ACC) deaminases. ACC deaminases are important catalysts for regulating the level of plant-produced ethylene. ACC is a precursor of ethylene and its cleavage reduces the level of this stress-induced hormone and increases the stress resistance of the plant (Glick 2005). Nikolic et al. analysed the abundance and diversity of ACC deaminase genes (*acd*S) from bacterial endophytes colonising field-grown potato plants. One complete *acd*S operon was identified and analysed, showing the presence of a transcriptional regulator (*acd*R), which may be exclusive for the phyla of *Alpha-* and *Betaproteobacteria* (Nikolic, Schwab and Sessitsch 2011).

Similarly, the population of endophytic bacteria from the nickel (Ni) hyperaccumulator plant *Thlaspi gosingense* was analyzed by means of total DNA extraction from shoot-associated DNA. The division of *Proteobacteria* dominated the bacterial endophyte population, showing clear differences to the bacterial community from the plants rhizosphere. The presence of genes or genetic treats responsible for a higher Ni resistance was analysed using bacterial isolates but not the uncultivable bacterial fraction. All endophytic isolates were positive for the production of siderophores, and some showed ACC deaminase activity (Idris *et al.* 2004). This study underpins the potential of finding new types of siderophores and other heavy metal resistance determinants in endophytes from hyper-accumulating plants.

Concluding remarks

There is clearly an urgent need for new bioactive compounds in medicine, agriculture and industry, like antibiotics, anticancer drugs or pesticides. The discovery of novel pharmaceuticals by combinatorial chemistry has, however, not delivered the expected results in regard to the number and the bioactivity of new developed compounds. Growing interest in the search for new NPs, their corresponding biosynthetic pathways and enzymes as biocatalysts has now awaken. In particular, plants have historically been a source of NPs. Likewise, endophytes, through a tight synergism with their host, have evolved a specialized metabolism for the production of certain bioactive metabolites. Hence, endophytes represent an uttermost promising and yet less explored source for this type of molecules.

While cultivation dependent techniques provide an enormous share of bioactive molecules, a myriad of other compounds from not-yet cultivable microorganisms remains undiscovered. The application of functional metagenomics can aid in the search for those unexplored molecules from uncultivable microbes. In this chapter we described the methodology and highlighted important considerations for construction of metagenomic clone libraries from endophytes. Several examples illustrate the potential for the discovery of new enzymes and NPs using the aforementioned strategy. This is a promising and new field of study, since endeavours for using functional metagenomics from sole endophytic microbiomes are still limited. It remains to be seen, how the field of endophyte biotechnology in combination with functional metagenomics develops for the discovery of novel bioactive molecules.

2 Bioprospecting plant-associated microorganisms

Christina Andrea Müller[1,2], Melanie-Maria Obermeier[2] and Gabriele Berg[1]

[1] *Institute of Environmental Biotechnology, Graz University of Technology, Graz, Austria;*

[2] *Austrian Centre of Industrial Biotechnology, Graz, Austria*

Abstract

There is growing demand for new bioactive compounds and biologicals for the pharmaceutical, agro- and food industries. Plant-associated microbes present an attractive and promising source to this end but are nearly unexploited. Therefore, bioprospecting of plant microbiomes is gaining more and more attention. Due to their highly specialized and co-evolved genetic pool, plant microbiomes host a rich secondary metabolism. This article highlights the potential detection and use of secondary metabolites and enzymes derived from plant-associated microorganisms in biotechnology. As an example, we summarize the findings from the moss microbiome with special focus on the genus Sphagnum and its biotechnological potential for the discovery of novel microorganisms and bioactive molecules. The selected examples illustrate unique and yet untapped properties of plant-associated microbiomes, which are an immense treasure box for future research.

Published in Journal of Biotechnology:

Müller C.A., Obermeier M.M., Berg, G. (2016) Bioprospecting plant-associated microorganisms. *J Biotechnol,* 235, 171-180, doi:10.1016/j.jbiotec.2016.03.033

Introduction

Bioprospecting is the process of discovery and commercialization of new products based on biological resources (Strobel and Daisy 2003). A growing need for new bioactive compounds in the pharmaceutical (e.g. antibiotics against multi-resistant pathogens) and the agro- and food industries (e.g. agrochemicals, biocontrol products, food additives) stresses the importance of prospecting for novel bio-resources (Berg *et al.* 2013; Woolhouse and Farrar 2014). Since the chemical diversity of compounds as comprised in biological resources is higher than synthetic chemistry achieves, bio-resources have great potential to hold a manifold of promising compounds for biotechnological application (Bérdy 2012; Nováková and Farkašovský 2013). Plants have been described as one of the richest sources of valuable bioactive natural products (McChesney, Venkataraman and Henri 2007). Progress in the -omics era, including next-generation sequencing (metagenomics, metatranscriptomics) and microscopic advances, has changed our view on eukaryotic hosts and the role of microbial diversity and microbial functions (Jansson *et al.* 2012; Cardinale 2014; Mendes and Raaijmakers 2015). Nowadays, we consider plants, like humans and other eukaryotic hosts, as meta-organisms that undergoes a tight symbiotic relationship with their microbiome (Bulgarelli *et al.* 2012; Hirsch and Mauchline 2012; Lundberg *et al.* 2012).

Each plant is colonized by more than 1000 microbial species, which are to a high degree not cultivable. In addition, plants are divided into specific niches for microorganisms, where biotic and abiotic factors shape specific microbial communities. While the rhizosphere has been well-studied for more than 100 years (Philippot *et al.* 2013), the phyllosphere and especially the endosphere are currently under intense investigation (Vorholt 2012; Hardoim *et al.* 2015). Moreover, each plant microbiome has due to the plant-specific secondary metabolism and physiology a specific composition (Berg and Smalla 2009). This depends on the plant family, for instance dicotyledonous plants have developed a richer secondary metabolism than monocotyledonous plants. Recently it was shown that breeding has a strong impact on the plant microbiome as well, and cultivar-specific effects were identified (Schlaeppi *et al.* 2014; Cardinale *et al.* 2015; Pérez-Jaramillo, Mendes and Raaijmakers 2016). Thereby, the microbiome fulfils multiple functions for the host health, like pathogen defence and contribution to stress tolerance under adverse environmental conditions and further supports growth and nutrient supply (Berg and Smalla 2009; Berg *et al.* 2013; Bragina *et al.* 2014; Grube *et al.* 2015).

While the potential of the microbiome to influence the host is now well-recognized (Blaser 2014; Berg *et al.* 2016), the diversity of metabolites that are synthetized by the microbiota is largely unexplored (Berg *et al.* 2014; Nunes-Alves 2014). During the last decades it was shown that single cultivable microorganisms represent a plenteous source for novel enzymes and bioactive compounds: around 70,000 natural products derived from microorganisms were reported with almost half of them showing bioactive traits (Bérdy 2012). Altogether, these facts underline the high diversity of indigenous microbial populations in plant microbiomes, and supports their exploitation for biotechnological purposes, for biocatalytic processes and plant protection, or in the pharmaceutical industry (Harvey 2008; Duke *et al.* 2010).

Metagenomics is one means that facilitates examination of the entire genetic pool comprised by (plant) microbiomes, thereby providing access to the potential of the high share of uncultivable microbes. The field of metagenomics has been intensively reviewed in recent time, which illustrates the great interest within the scientific community spanning a wide range of research areas from ecology over medicine to biotechnology (Banerjee, Mishra and Dhas 2015; Coughlan *et al.* 2015; Cowan *et al.* 2015; Faust *et al.* 2015; Garza and Dutilh 2015; Ravin, Mardanov and Skryabin 2015; Roossinck, Martin and Roumagnac 2015;

Wang *et al.* 2015). Briefly, metagenomics is the study of the combined genomes of all the organisms present at site without culturing them first (Handelsman 2004). Upon extraction of the total environmental DNA it can either be examined at the sequence level or function driven. Sequence based analysis obviously requires the DNA to be sequenced, which commonly involves random shot gun sequencing (Abbasian *et al.* 2015). In contrast, when expressed heterologous, the metagenome can be screened for certain, desired activities by subjecting the metagenomic library to specifically designed screening assays (Gabor *et al.* 2007). Bioprospecting towards novel enzymes and bioactive compounds has been frequently performed using environmental samples originated from soil, marine environments and microbiota associated to mammals (e.g. the human gut microbiome), but rarely employing plant-associated microorganisms.

Here we present examples that highlight the enormous microbial and functional diversity, as well as the biosynthetic potential of plant-associated microbiomes for bioprospecting approaches.

Functional and structural insights in plant microbiomes from metagenomics

While the structural diversity of plant microbiomes is well-studied now by amplicon sequencing, it is difficult to find general genetic markers to analyze functional diversity. For this purpose, metagenomic datasets have to be compared and studied in depth. Interestingly, microbiome functions are more similar across different environments than previously thought. For example, Ramírez-Puebla *et al.* (2013) discussed extensively on similar functional traits in the gut and root microbiome. Based on deeper insights obtained by omics technologies, Mendes and Raaijmakers (2015) recently presented their concept that the structure and function of rhizosphere and gut microbiomes show cross-kingdom similarities. This concept was extended by Berg *et al.* (2016) to ecological rules. To date only a few studies have deciphered the functional diversity of plant microbiota employing shotgun next generation sequencing approaches (rev. in Knief, 2014). For instance, in depth analysis of the *Sphagnum* moss metagenome revealed high functional diversity within the microbiome and a higher structural diversity as previously detected with other methods (amplicon sequencing or cultivation-dependent methods) (Bragina *et al.* 2014).

In this section we focus on functional and structural diversity of plant metagenomes as summarised in Table 3.1, which are important criteria when prospecting for novel bioactive compounds. Based on phylogenetic analysis of plant metagenomic data it was demonstrated that bacterial diversity in higher plants (rhizosphere, phyllosphere, endosphere) is remarkably high, even though it is in general lower than the microbial diversity in the surrounding bulk soil (Delmotte *et al.* 2009; Bulgarelli *et al.* 2015). High abundance of bacterial taxa in the rhizosphere or phyllosphere of rice, barley, moss, lettuce and soybean belong predominantly to the phyla *Proteobacteria*, *Actinobacteria*, *Bacteroidetes*, *Firmicutes*, and *Acidobacteria.* Less abundant but highly diverse phyla with functionally outstanding properties are also present, for example *Cyanobacteria*, *Deinococcus*, and *Chloroflexi* (Delmotte *et al.* 2009; Lelie *et al.* 2012; Bragina *et al.* 2014; Kröber *et al.* 2014; Bulgarelli *et al.* 2015). The functional diversity of the aforementioned plant-microbiomes has been explored in relation to the plant compartment. In barley roots and rhizosphere, the authors found significant enrichment of biological functions coding for adhesion, stress response, secretion, host-pathogen, microbe–microbe and phage-microbe interactions, as well as iron mobilization and sugar transport (Bulgarelli *et al.* 2015). In contrast, phyllosphere bacterial communities in clover, soybean and *Arabidopsis thaliana* plants were characterized by high

expression of outer membrane proteins (porins, TonB receptors) and ABC-transport systems for carbohydrates (maltose, glucose, sucrose) and amino acids, underlining the role of the phyllosphere microbiota for carbon utilization on the plant leaves (Delmotte *et al.* 2009).

Other groups have focused on the investigation of highly specialized microbial communities and their metabolic functions, for example on the distribution and abundance of phototrophic bacteria on the plant phyllosphere, with putative functions for light sensing and utilization (Atamna-Ismaeel *et al.* 2012a, 2012b). Here the authors employed metagenome data mining to elucidate the relative abundance of rhodopsin-based phototrophs and anoxygenic phototrophs on leaves of different plants (tamarisk, clover, rice, soybean, and *A. thaliana*). Rhodopsin coding sequences (Atamna-Ismaeel *et al.* 2012b), as well as phototrophy-related genes (*pufM*, *bchY*, *pufL*) (Atamna-Ismaeel *et al.* 2012a), were used as genetic markers for in silico analysis. Several novel rhodopsin sequences were identified, especially in the tamarisk phyllosphere. These findings suggest the presence of highly diverse and phyllosphere-specific phototrophic species in the studied plant microbiomes. Similarly, van der Lelie *et al.* (2012) investigated the occurrence and enrichment of specialized indigenous microbial communities from poplar hard wood that are capable of decomposing or modifying lignin under anaerobic conditions. The biomass degrading community was dominated by species belonging to *Clostridia*, *Bacteroidetes* and *Methanomicrobiales,* but also by a novel biomass degrader bacterium displaying similarity to the genus *Magnetospirillum,* which may be involved in detoxification of aromatic metabolites. The poplar wood metagenome showed high abundance of carbohydrate active enzymes and also the presence of putative cellobiose dehydrogenases and new bacterial genes showing homology to fungal lignin oxidases.

Bioactiva from plant-microbiomes: microbes, enzymes and secondary metabolites

Microbes

The use of isolated microorganisms or microbial consortia as biological control agents (BCAs) is an outstanding alternative to conventional agricultural techniques (Berg *et al.* 2013). While in the past mainly microbes of soil or crop origin were selected, now native and endemic plants or mosses are in the focus of bioprospecting (Opelt, Berg and Berg 2007; Zachow, Tilcher and Berg 2008). Colonisation and plant-growth promotion were observed for several endophytic bacterial isolates from *Sphagnum* moss, belonging mainly to the genus *Burkholderia*, *Pseudomonas*, *Flavobacterium*, *Serratia* and *Collimonas* (Shcherbakov *et al.* 2013). Bacteria displaying stress protection and growth-promotion of economically relevant crops (e.g. maize and sugar beet) were isolated from alpine mosses, prime rose, and lichens (Zachow *et al.* 2013). A possible application of endophytic bacteria for bioremediation has also been discussed (Stępniewska and Kuźniar 2013). It was further shown that microbes assist plants in pathogen defence. When *A. thaliana* gets infected with *Pseudomonas syringae* pv. *tomato* it increasingly secretes malic acid which attracts *Bacillus subtilis BF-17*. The colonisation of root tissue triggers an immune response that aids the plant in antagonising the pathogen (Rudrappa *et al.* 2008). Raymond (2016) reported the production of ice-binding proteins by epiphytic bacteria in the moss *Bryum argenteum* grown in the Antarctica. This class of proteins are found mainly in cold-adapted microorganisms, protecting the plant against freezing damage, a property that could be exploited for crop protection strategies as well. These are only a few examples for bioactive microbes from plant microbiomes and a much longer list of studies exists (Berg 2009; Berg *et al.* 2013).

Enzymes

Enzymes with novel activities are classical targets in microbial screenings that aim for an industrial application (Schmeisser, Steele and Streit 2007). However, most enzyme discovery screenings to date have focused on analysing metagenomic libraries from soil and marine environments or the human gut (Simon and Daniel 2011). In the case of plant-associated microbiomes, detection of new enzymes has mainly been centred in the investigation of microbial activities that are directly related to antagonism or plant-growth promotion. Cell-wall degrading or lytic enzymes, like cellulases, glucanases, proteases, xylanases and chitinases have been detected in plant-associated microorganisms, especially in endophytic fungi and bacteria (Lumyong *et al.* 2002; Gayathri *et al.* 2010; Suryanarayanan *et al.* 2012; Robl *et al.* 2013) or rhizobacteria (Inbar and Chet 1991; Frankowski *et al.* 2001; Ghyselinck *et al.* 2013; Castaneda-Alvarez *et al.* 2016). In this studies microbiota adapted to extreme conditions received much attention. Reactions as carried out during industrial processes require enzymes to possess not only high functionality and specificity, but also high stability at particular conditions. Hence, tolerance to high salt content, cold or elevated temperature and low or increased pH are desired traits (Elleuche *et al.* 2015). Recent reports describe the discovery (functional screenings) and characterization of a novel salt-tolerant chitibiosidase for potential degradation of seafood waste, a new alkaline and thermostable esterase Est3K and lipase Lip3K, as well as the esterase EstGX2 that exhibits high stability at elevated temperatures and in the presence of organic solvents (Cretoiu *et al.* 2015; Gu *et al.* 2015; Kim *et al.* 2015). As in these studies, functional metagenomic screens are often performed on microbiota derived from various kinds of soil and ground (Coughlan *et al.* 2015), rather than on plant microbiomes. Plant associated microbes, however, are likely to hold new enzymes for industrial applications. As for polysaccharides, like starch and cellulose, that are widely used in food, textile, paper industry and bioethanol production among others, plant microbiomes most likely harbour microbes that express degrading enzymes with interesting features. Through activity based screening of a metagenomic library constructed from sugarcane bagasse the thermophilic endoglucanase GH9 and xylanase GH11 were isolated. Both enzymes, derived from the indigenous microbial fraction, were characterized as thermophilic with highest activity under slightly acidic conditions, which are interesting features for industrial application in cellulose degradation (Kanokratana *et al.* 2015). Likewise, several novel glycoside hydrolase genes (GHases) were identified in naturally occurring microbial biomass decay communities from poplar wood chips, after enrichment in an anaerobic bioreactor. By combination of metagenomic gene mining and fosmid library screening the authors succeeded in the isolation and cloning of 4 metagenomic GHases (Li *et al.* 2011). Despite the increasing number of novel enzymes detected in metagenomes in the last decades, only a very small fraction has been expressed and fully characterized to evaluate further utilization in industrial processes (Schmeisser, Steele and Streit 2007). Meanwhile, the demand for functionally characterized biocatalysts is still growing.

Secondary metabolites

Plants were suggested to select competent endophytes from the environment for their own ecological benefit (Hardoim *et al.* 2015). In accordance, a plethora of diverse and highly specialized secondary metabolites and enzymes are produced by the inhabiting microorganisms. Most bioprospecting campaigns have therefore focused on the exploitation of plant endophytes that produce manifold bioactive metabolites through cultivation-dependent methods, evaluating mainly endophytic fungi and *Actinomycetes* (Tan and Zou 2001; Schulz *et al.* 2002; Strobel and Daisy 2003; Gunatilaka 2006) or *Actinobacteria* (rev. in Qin et al., 2010). To a lesser extent root-associated microorganisms have been prospected for secondary metabolites (Gunatilaka 2006). Microbial production of manifold

Table 2.1: Structural and functional diversity of plant-associated metagenomes.

Plant species and habitat	Taxonomic structure: Bacteria/ Eukaryota/ Archaea (%)	Dominant bacterial taxa (%)	Characteristics (location, habitat, sampling, treatments)	Functional diversity	Reference
Soybean phyllosphere[a] *(Glycine max Gallec)*	99.07/ 0.58/ n.a.	α-*Proteobacteria* (42.8), *Bacteroidetes* (12.1), ß-*Proteobacteria* (10.3), *Actinobacteria* (9.8), γ-*Proteobacteria* (5.3)	Effretikon (Switzerland); Agricultural practice: field-grown plants, sampling at beginning of bean development.	The analysed enzymatic functions focused on carbon and nitrogen cycling: High abundance of outer membrane proteins (OprF), one-carbon compound metabolism (methanol-DH), transport systems, carbon storage (phasin) and stress-related proteins (chaperons GroEL) were detected.	Delmotte *et al.* 2009
Rice endosphere (Oryza sativa 'APO `)	n.a./ n.a./ 0.4	ß- & γ-*Proteobacteria* (61), α-*Proteobacteria* (5.5), *Firmicutes* (12), *Planctomycetes* (1.3)	Los Baños, International Rice Research Institute (Philippines); Plants grown on experimental field, sampling at flowering stage: endophytes extracted and enriched from roots.	Focus on metabolic systems for microbiome-host interactions: high abundance of genes for protein secretion (cytoplasmic and outer membrane proteins, type VI secretion system), plant-polymer-degrading enzymes (cellulases, xylanases, cellobiohydrolases, pectinases, cellulose-binding proteins), motility and adhesion (flagellins), detoxification of reactive oxygen species (catalase, superoxide dismutase, glutathione peroxidase), siderophores biosynthesis, quorum sensing systems (AI-2, homoserine lactone), and nitrogen cycling; Availability of other enzymes involved in butane-diol fermentation, aerobic degradation of aromatic, halogenated and aliphatic compounds, and synthesis/degradation of polyhydroxyalkanoates was detected.	Sessitsch *et al.* 2011
Rice phyllosphere (P) and rhizosphere (R)[b] *(Oryza sativa subsp. indica cv. IR-72)*	99.5 P, 97.4 R/ n.a./ 0.5 P, 2.6 R	*Actinobacteria* (38 P, 8.5 R), α-*Proteobacteria* (35 P, 14 R), ß-*Proteobacteria* (5 P, 17 R), δ-*Proteobacteria* (1.6 P, 11 R), *Chloroflexi* (0.6 P, 4.6 R)	Los Baños (Philippines); Agricultural practice: rice-field, sampling at 59 to 76 days after seedling transplantation.	Abundant functions in the phyllosphere are: substrate uptake (porins, ABC transport system), stress response, cell adhesion (fasciclin domain), one-carbon compound metabolism (methanol-DH), invasion-associated locus B-family protein; in the rhizosphere: methanogenesis and methane oxidation, dinitrogen fixation, chemotaxis and motility.	Knief *et al.* 2012
Tamarisk phyllosphere (Tamarix nilotica)	n.a./ n.a./ n.a.	n.a.	Oasis by the Dead See (Israel); sampling of leafs from *T. nilotica* tree.	Study focused on diversity of microbial rhodopsins: a higher abundance of sensory rhodopsins over rhodopsin proton pumps was detected; mostly novel rhodopsin sequences with uncertain phylogenetic affiliation were detected.	Atamna-Ismaeel *et al.* 2012b
Poplar wood[d] *(Populus tremula)*	99.2/ 0.04/ 0.74	*Firmicutes* (46), *Proteobacteria* (32), *Bacteroidetes* (10), *Cyanobacteria* (4)	USA; Non-sterile yellow poplar saw dust (indigenous microbiota), humidified; enrichment of microbial communities by incubation at 30 °C in the dark for 3 and 12 months. Sampling from bottom fraction (anaerobic zone).	The study focused on detection of enzymatic functions for the anaerobic degradation of recalcitrant plant biomass (lignin, cellulose, hemicellulose) and carbohydrate active enzymes, using a BlastP homology search in CAZy and FOLy databases: most abundant families are glycosyl transferases (GT2, GT4) and glycoside hydrolases (GH13, GH3, GH2); present but less abundant are bacterial genes with lower homology to fungal oxidase families (cellobiose-DH, dihydrolipoamide-DH, peroxidases and multi-copper oxidases).	Van der Lelie *et al.* 2012
Tomato phyllosphere (Solanum lycopersicum)	n.a.	*Proteobacteria* (n.a.), *Firmicutes* (n.a.), *Actinobacteria* (n.a.), *Cyanobacteria* (n.a.), *Chloroflexi* (n.a.)	Eastern Shore Agricultural Research and Extension Centre, Virginia Tech (USA); Leaves and fruits, untreated control.	Study focused on the effects of various culturing conditions (enrichment treatments) rather than on functional characteristics of the tomato phyllosphere.	Ottesen *et al.* 2013

Plant species and habitat	Taxonomic structure: Bacteria/ Eukaryota/ Archaea (%)	Dominant bacterial taxa (%)	Characteristics (location, habitat, sampling, treatments)	Functional diversity	Reference
Lotus japonicus rhizosphere (*L. japonicus* MG20)	90.74 NT, 87.23 F/ 5.45 NF, 9.03 F/ 3.03 NF, 3.54 F	*Proteobacteria* (43 NF; 33 F), *Acidobacteria/Fibrobacteres* (13 NF; 20 F), *Actinobacteria* (10 NF; 13 F), *Bacteroidetes/Chlorobi* (10 NF; 6 F), *Firmicutes* (4.4 NF; 5.8 F), *Chloroflexi* (3.4 NF; 4.2 F)	Hokkaido University (Japan); Plant from agricultural practice: surface sterilized seeds, plants grown in green-house, analysis of plants with addition of phytic acid (sampling at flowering stage, F) or without (not flowering, NF).	Significantly increased abundance of functions for phosphorus-utilization in plants supplemented with phytic acid (F) (alkaline phosphatase, *myo*-inositol 1-monophosphatase, *epi*-inositol hydrolase; glutamine-, glutamate-, aspartate- and asparagine-biosynthesis; glyoxylate synthesis), as well as production of secondary metabolites (antibiotic compounds, phytohormones) were observed; Other genes coding for outer membrane proteins, citrate synthase, and glycosyl-transferase were detected.	Unno and Shinano 2013
Soybean rhizosphere (*Glycine max* Merril cv. M-SOY 8866)	96/ 3/ <1	*Proteobacteria* (47), *Actinobacteria* (23) , *Firmicutes* (6), *Acidobacteria* (5)	University of Sao Paulo (Brazil); Soybean grown in mesocosm (green-house) using bulk soil from Amazonian agricultural fields, sampling after 80 days of growth, data for 1 and 5 years of soybean cultivation.	Functional cores were analysed: the most representative metabolic systems involve membrane transport (secretion system type IV), as well as acquisition of nitrogen, phosphorus (P uptake, alkylphosphonate utilization), potassium and iron (heme, hemin uptake), which are overrepresented in the rhizosphere as compared to bulk soil. Protein metabolism was most abundant after 1 year, while regulation and cell signalling systems were more abundant after 5 years of cultivation.	Mendes *et al.* 2014
Lettuce rhizosphere[c] (*Lactuca sativa*)	98.2/ n.a./ n.a.	*Proteobacteria* (64), *Actinobacteria* (24), *Bacteroidetes* (5)	Golzow (Germany); Agricultural practice: field-grown plants (untreated), 0, 2 and 5 weeks old plants after planting from peat blocks.	Focus on functional subclasses of carbohydrate and nitrogen metabolism: changes in saccharides, organic acids, amino acids and nitrogen metabolism are dependent on growth state and substrate (peat or soil).	Kröber *et al.* 2014
Moss phyllosphere/ endosphere (*Sphagnum magellanicum*)	95.1/ 4.4 / n.a.	*Proteobacteria* (65.8), *Acidobacteria* (11.4), *Actinobacteria* (5.6), *Bacteroidetes* (4.2), *Verrucomirobia* (2.0)	Pirker Waldhochmoor (Austria); Extreme abiotic conditions: acidic pH, repetitive desiccation, high oxidative stress, high temperature fluctuations.	Functions supporting genetic exchange (phages, type IV secretion), resistance to oxidative stress (GSH/MSH/Rr reactions), and motility (flagella, rhamnolipids) are highly abundant. High diversity of genes coding for quorum sensing molecules (AI-2, homoserine lactones, γ-butyrolactones), biofilm formation and adhesion (EPS, adhesins), toxin-antitoxin systems, antibiotics and toxins (multidrug efflux pumps, Co-Zn-Cd resistance), DNA repair, NRPs and PKs (toxins, siderophores, antibiotics) were also detected.	Bragina *et al.* 2014, Müller *et al.* 2015
Barley rhizosphere *(Hordeum vulgare)*	94.04/ 5.9/ 0.054	ß-*Proteobacteria* (7.7), γ-*Proteobacteria* (5.5), *Actinobacteria* (n.a.), *Bacteroidetes* (n.a.)	Golm (Germany); Agricultural practices: green-house cultivation of surface-sterilized seeds, sampling at early stem elongation stage.	Significant enrichment of protein coding sequences for adhesion, stress response, and secretion systems was observed; Functions support host-pathogen interactions (type III secretion system T3SS), microbe-microbe interactions (type VI secretion system, T6SS), microbe-phage interactions (transposable elements, bacteriophage integration), iron mobilization (siderophores), and sugar transport; Plant effector proteins (T3SEs) and phage protection systems (CRISPR-*cas*) were identified.	Bulgarelli *et al 2015*

Abbreviations: n.a.: not analysed; n.d.: not detected; GSH: glutathione; MSH: mycothiol; Rr: rubrerythrin; AI-2: autoinducer-2; EPS: extracellular polysaccharides; NRPs: non-ribosomal peptides; PKs: polyketides; CRISPR: Clustered regularly interspaced short palindromic repeat system; *cas*: CRISPR-associated genes; R: rhizosphere; P: phyllosphere; DH: dehydrogenase; F: flowering; NF: not flowering; CAZy: Carbohydrate-Active enZYmes database; FOLy: Fungal Oxidative Lignin enzYmes database.

[a] Functional diversity analysed using metaproteome data.[b] Functional diversity analysed using a combination of metagenome and metaproteome data.[c] Data for alpha-diversity and taxonomic composition is given for 5 weeks old plants after planting.[d] Taxonomic composition was analysed by pyrotag sequencing, while functional diversity was analysed using the metagenomic dataset.

molecules like phytohormones, antibiotics and quorum-sensing molecules has been reported in plants. Cytokinins and auxins are plant growth regulators expressed by a vast amount of microbes, especially bacteria, allowing them to influence cell elongation and division in plants (Hayat *et al.* 2010). Particularly, the auxin indole-3-acetic acid, its biosynthesis and mode of action received much attention. Microbial indole-3-acetic acid production can be of different consequence for plants, either detrimental as in crown gall formation induced by *Agrobacterium tumefaciens* or beneficial by promoting root development (rev. in Duca *et al.*, 2014), for instance.

Penicillin, the historical and classical example for microbial derived antibiotics is by far not the only microbial antibiotic produced and found. The production of such compounds is common among microorganisms. Examples of antimicrobial compounds synthesized by plant associated microbes include cyclosporine that is produced by the endophytic fungus *Tolypocladium inflatum* and shows antifungal activity (Borel *et al.* 1976). More recently discovered microbial antibiotics are antifungal ecomycins derived from the plant endophyte *Pseudomonas viridiflava* or turbomycin A and B, which were identified from a soil derived metagenomic library (Miller *et al.* 1998; Gillespie *et al.* 2002). Bioactive compounds produced by plant associated microbes not only display antimicrobial but also antiviral activity. Xiamycin from *Streptomyces GT2002/1503*, for instance, was reported to have anti-HIV properties (Ding *et al.* 2010).

Signalling compounds such as *N*-acylhomoserine-lactones (AHL) and cyclic peptides play an important role in bacterial communication and bacterial-plant interactions, e.g. for regulation of bacterial swarming or biosynthesis of antibiotics and exopolysaccharides (Brelles-Mariño and Bedmar 2001). Quorum sensing molecules have been well-characterized in plant-associated bacteria such as *Rhizobium* spp. (Downie 2010) and *A. tumefaciens* (Zhang, Wang and Zhang 2002). In addition to secondary metabolites, pathways or genes that confer resistance to heavy metals, salts and acids are of industrial interest. Using the rhizosphere metagenome of *Erica andevalensis*, a plant adapted to acid mine drainage at the banks of the Tinto River in Spain, Mirete *et al.* (2007) discovered novel nickel resistance genes and proposed mechanisms involved in metal resistance.

Another class of secondary metabolites are volatile organic compounds (VOCs) that are gaining more and more attention. In contrast to enzymes, VOCs and their possible biotechnological application have hardly been investigated, despite their beneficial effect on plant growth and health and their potential for biological control (Ryu *et al.* 2003; Müller *et al.* 2009). In fact, only little is known in regard to the enormous diversity of VOCs from microbial sources and their mode of action (Kanchiswamy, Malnoy and Maffei 2015). New volatiles were detected in plant-associated microorganisms (Strobel and Daisy 2003; Kai *et al.* 2007; Bailly and Weisskopf 2012). Bacterial volatiles can interact directly with plants, but they can also have a strong effect on fungal growth (Weisskopf 2013). The observation that prokaryotes produce volatile compounds that are able to inhibit the growth of fungi is relatively new. In 2007, Kai and colleagues compared the volatile-mediated inhibition of a range of phytopathogenic fungi by different bacterial strains. Since then, bacterial strains emitting antifungal volatiles, the chemical nature of those volatiles and their effect as pure compounds have been reported (recently rev. by Effmert *et al.*, 2012 and Weisskopf, 2013). However, despite of a long lists of VOCs produced by plant-associated bacteria, much less is known about the synthesis pathways. Widespread classes of volatiles emitted by bacteria are fatty acid derivatives (aliphatic volatiles e.g. butan-1-ol), terpenoids (e.g. geosmin, sesquiterpenes), and aromatic (e.g. 2-phenylethanol) or halogenated compounds (methyl iodine), as well as compounds containing nitrogen (e.g. ammonia, pyrazines) or sulphur (e.g. 3-dimethylsulfoniopropionate) (Schulz and Dickschat 2007; Kai *et al.* 2009). One of the most prominent anti-fungal volatiles is hydrogen cyanide, a respiratory poison, which is

inappropriate for agricultural applications due to its high toxicity and lacking specificity (Blom *et al.* 2011). In the last years, the study of microbial volatile emissions has been confined to single cultivable bacterial species. Given the fact that more than 98% of the microbial diversity is not readily cultivable under current lab conditions (Handelsman 2004), a search in metagenomes – bypassing the need of cultivation – is a promising approach.

Interesting industrial applications for microbial VOCs are bio-fumigation processes for developing crop welfare and sustainable agricultural practices (e.g. agricultural pre- and postharvest treatment) (Kanchiswamy, Malnoy and Maffei 2015) as well as for control of microbial hazards in human environments. The identification and analysis of novel secondary metabolites and their producers is yet another encouraging field, with an ever-growing demand for microbial and biochemical biopesticides. Moreover, VOCs can be applied as natural aroma chemicals, such as flavours and fragrances in the food and health care production industry (Krings and Berger 1998). Some classes of VOCs produced by plant-associated bacteria are active against multi-resistant bacteria (Berg, pers. comm.), which can lead to the development of promising and novel concepts against the world-wide antimicrobial resistance challenge (Woolhouse and Farrar 2014).

Exploiting the first moss metagenome – a successful example for bioprospecting

Our investigations on moss have revealed that the plant-associated microbiota represent a nearly untapped source for novel microbial functions, originating from species that are adapted to diverse biotic and abiotic conditions (Bragina *et al.* 2014). Mosses are considered model organisms for higher plants (Cove *et al.* 2009) and are therefore well suited for studying plant-microbiome interactions. Peat bogs are considered extreme habitat where mosses face extreme environmental conditions, such as high temperature fluctuations, repetitive desiccation, nutrient deficiency and high oxidative stress. The colonizing microorganisms are therefore highly specialised and display specific genetic features that distinguish them significantly from microbiomes of higher plants and peat soils and support the symbiotic nature of the ecosystem. In silico analysis of the *Sphagnum* moss metagenome revealed increased abundance of functional groups that are responsible for stress tolerance and higher bacterial interaction through quorum sensing (e.g. autoinducer-2, AHL, and gamma-butyrolactones signalling pathways), biofilm formation, nutrient exchange (e.g. nitrogen cycling) and enhanced bacterial motility (Bragina *et al.* 2014). The occurrence of biosynthetic genes belonging to non-ribosomal peptide synthetases (NRPS) and polyketide synthases (PKS) was also investigated. Contigs containing NRPS and PKS sequences contributed 0.06% of the assembled metagenome, leading to detection of putative gene sequences coding for antibiotics, siderophores and phytotoxins (Müller *et al.* 2015). Bioprospecting for novel enzymes in the moss-associated microbiome has also been successful, yielding novel esterase, decarboxylase and phosphatase genes with potential application in biocatalytic processes (Müller, pers. comm.).

Concluding remarks

The phylogenetic and metabolic diversity hidden in plant-associated microorganisms is undeniable. This biosynthetic diversity can be regarded as highly diverse yet specialised and adapted to biotic and abiotic environmental conditions, making the discovery and exploitation of functional genes and secondary metabolites very promising. The full potential is hardly explored and should continue to be exploited.

Acknowledgements

This manuscript was supported by different grants from the Austrian Science Foundation FWF (I183-B16; I882-B16) to G.B. The cooperation of M.M.O. was funded by a project in the Austrian Centre of Industrial Biotechnology, supported by the Austrian BMWFW, BMVIT, SFG, Standortagentur Tirol, Government of Lower Austria and ZIT through the Austrian FFG-COMET- Funding Program. We also gratefully acknowledge support from NAWI Graz.

3 A novel *ohr* regulon from uncultivated bacteria involved in antifungal activity against *Candida albicans*

Melanie-Maria Obermeier[1,2], Sarah Prinz[1,2], Stephanie Hollauf[1], Anna Komor[3], Gabriele Berg[1,2], and Christina Andrea Müller Bogotá[1,2]

[1]Institute of Environmental Biotechnology, Graz University of Technology, Graz, Austria

[2]Austrian Centre of Industrial Biotechnology, Graz, Austria

[3]Department of Biomolecular Chemistry, Leibniz Institute for Natural Product Research and Infection Biology – Hans-Knöll-Institute, Jena, Germany

Abstract

Candida albicans accounts as the leading causative agent for invasive fungal infections which are on the rise worldwide. The high mortality rate, emerging resistances and limited drug availability urge the development of novel therapeutic agents. We, therefore, prospected the microbial community associated with *Sphagnum* mosses, which is characterised by a rich and diverse secondary metabolism and a high share of antifungal properties. To capture the metabolic potential of the whole microbiota, including the not-yet cultivable microorganisms, we pursued functional metagenomics; a cultivation independent methodology. The screening of a metagenomic library in *E. coli* led to the identification of one fosmid clone with antifungal activity towards *C. albicans* and further to the identification of an organic hydroperoxide resistance (Ohr) gene cluster driving the antagonistic effect. The discovered regulon consists of four novel putative genes, *ohrR, ohrB,* a peroxidoredoxin homologue (*prx*) and a sulfiredoxin (*srx*). We hypothesise that these genes act on a metabolite from the *E. coli* host background, which then gains antifungal activity and lay out its involvement in the antagonistic effect. We further suspect the inhibitory molecule to be a fatty acid hydroperoxide metabolite, which represents a potentially novel mode of action. These findings open new possibilities for the search of novel drug candidates against fungal infections.

MANUSCRIPT

Introduction

Increasing fungal infections associated with the growing number of immunocompromised or -suppressed patients and drug related, pervasive dysbiosis have become a growing concern for modern medicine (Geddes-McAlister and Shapiro 2018). *Candida spec.* account as the leading causative agents for invasive fungal infections with high mortality rates worldwide, whereby most cases are attributed to *C. albicans* (Guinea 2014). This common human commensal colonising skin, mouth, the gastrointestinal and the female reproductive tract (Noble, Gianetti and Witchley 2016) can orchestrate infections intrinsically as part of the patient´s microbiome or cause exogenous infections as contaminant of intravenous catheters (Guinea 2014). With increasing resistance and only four classes of antifungal drugs available new therapeutic agents for the treatment of invasive fungal infection are in great demand; especially compounds with a new mechanism of action (Perfect 2017).

Natural products which surpass the structural diversity of synthetic compounds (Henkel *et al.* 1999) represent a versatile source for drug development, serving as novel lead molecules. Very promising in this regard are in particular the unexplored bioactive products from plant-associated microorganisms, like for instance of peatland-forming *Sphagnum* mosses (Müller, Obermeier and Berg 2016). The use of *Sphagnum* moss and other bryophytes as disinfectants in traditional medicine due to their antifungal properties is well documented (Frahm 2004). *Sphagnum spec.* are characterised by a specific and rich metabolic profile (Opelt *et al.* 2007b) to which its taxonomically and metabolically versatile microbiome likely contributes to (Bragina *et al.* 2014). In the harsh conditions of the bog ecosystem, these bryophytes share close symbiotic interactions with their inhabiting microbial community which fulfils important host functions such as nutrient supply and protection from fungal plant pathogens (Opelt, Berg and Berg 2007; Bragina *et al.* 2014). Further, the *Sphagnum* microbiota comprises an extraordinarily high share of species with antifungal traits (Opelt, Berg and Berg 2007).

Due to the distinctive antimicrobial properties of the *Sphagnum*-associated microbiome, we previously assessed a *S. magellanicum* metagenome from an Austrian Alpine peat bog by *in silico* data mining and PCR-based *in vitro* analysis of a fosmid clone library for the presence of non-ribosomal peptide synthetases and polyketide synthases (Müller *et al.* 2015). The study highlighted a versatile and novel assemblage of these antimicrobial production modules within the *S. magellanicum* microbiome. Here, to harness the antifungal potential, we employed a functional metagenomics approach using a previously generated moss metagenomic library. As this powerful, cultivation independent approach allows access to the metabolic pathways of the 97% not-yet cultivable *Sphagnum*-associated species (Bragina *et al.* 2012), we expected to identify entirely novel genes and metabolites. Here we demonstrate that a novel organic hydroperoxide resistance (Ohr) regulon in *E. coli*, originating from the *Sphagnum* metagenome, is involved in the antifungal activity acquired by *E. coli* as library host against the opportunistic pathogen *C. albicans*.

Results

Activity-based screening of 90 000 fosmid clones, collectively holding 3.6 Gbps of moss metagenomic DNA, was carried out in a high-throughput fashion using the agar overlay method. The library was, thereby, screened with 3 × coverage. This led to the identification of one unique clone effectively inhibiting *C. albicans*, termed *E. coli* EPI300 pCC2FOS-CC (CC) (Figure 3.1 a). Interestingly, clone CC exhibited its antagonistic activity towards *C. albicans* only when cultivated on Terrific Broth agar (TB). No effect was observed for Müller-Hinton and neither for Luria-Bertani (LB) agar, although the latter has a similar

composition to TB (data not shown). Therefore, the assay was performed using modified TB and LB media, where single medium components and concentrations were exchanged. This showed that yeast extract was the single, necessary component for the antagonistic effect of clone CC (Figure 3.1 a-d). It further prompted the idea of amending the medium with casein hydrolysate, which proved to increase the inhibition zone (data not shown).

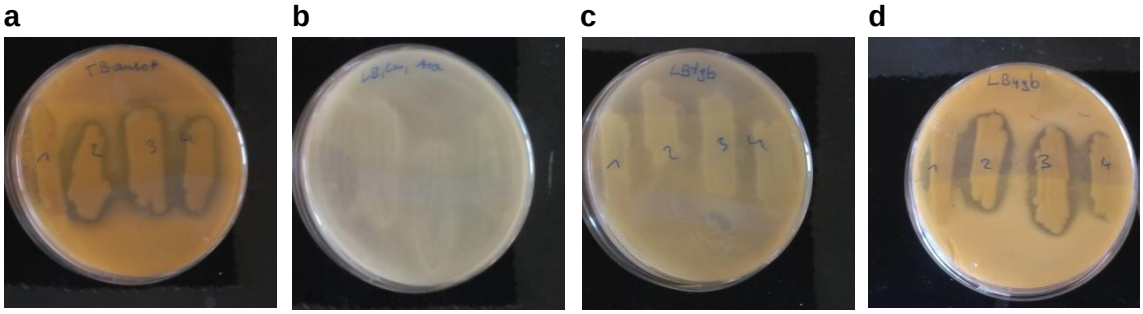


Figure 3.1: *C. albicans* inhibition. Three re-transformants of *E. coli* EPI300 pCC2FOS-CC (2-4) and their inhibitory effect onto *C. albicans* compared with the empty vector strain (1) on TB agar (a), LB agar (b), LB agar amended with glycerol and potassium phosphate buffer and increased concentration of tryptone (c), and LB agar amended with glycerol and potassium phosphate buffer and increased concentration of yeast extract (d).

De novo sequencing of pCC2FOS-CC and gene annotation resulted in a 44.8 kb large DNA fragment holding 41 open reading frames (ORFs). As no synthetase gene, e.g. non-ribosomal peptide synthetases or polyketide synthases, nor a biosynthetic gene cluster were present, we hypothesised that the bioactive compound is a metabolite from the host background, which processed by the metagenome insert gene(s) acquires antifungal activity.

To identify the gene(s) of interest, a multi-step subcloning approach was pursued (Figure 3.2). By restriction digest of pCC2FOS-CC (Figure 3.2 a), five different subclones were generated (CC-3, 5, 9, 17, and 18) of which *E. coli* EPI300 pCC2FOS-CC-18 retained the antifungal phenotype (Figure 3.2 b). This decreased the metagenomic DNA insert to 18.7 kb with 19 ORFs. It was then further subcloned using AQUA cloning (Beyer *et al.* 2015), an assembly cloning technique based on *in vivo* homologous recombination. Primers were designed on a rational basis based on gene annotations, such that the four identified clusters of four to six ORFs present on CC-18 were subcloned as one unit (Figure 3.2 b). This resulted in the four strains *E. coli* 10-beta pET17b-W, X, Y, and Z. With clone Z still inhibiting the yeast, the number of potentially active genes was successfully downsized to six. These included, two putative deaminases, a putative sulfiredoxin (*srx*), and the organic hydroperoxide regulon (Ohr) consisting of a transcription repressor (*ohrR*), and two peroxidases, an organic hydroperoxide resistance protein *ohrB* and a peroxiredoxin homologue (*prx*) (Figure 3.2 c). We speculated that the deaminases were not involved in the antagonistic activity. Thus, two subclones were generated in another AQUA cloning step. One of which contained *ohrR, ohrB* and *prx* (subclone S), while in addition to these three genes the other one held the *srx* gene (subclone T) (Figure 3.2 d). Activity screening confirmed our venture and proved the antifungal activity to result from the Ohr regulon in combination with *srx*, as present in subclone *E. coli* 10-beta pET17b-T. A clone holding only *srx* did not exhibit the antifungal effect against *C. albicans* (data not shown). Due to sequence similarity, *srx* was initially annotated as dibenzothiophene desulfurization enzyme (SoxC) by the prokka tool (Seemann 2014). Based on literature research and its importance for the antifungal activity in relation to the Ohr regulon, we propose that this protein is a novel type of prokaryotic sulfiredoxin.

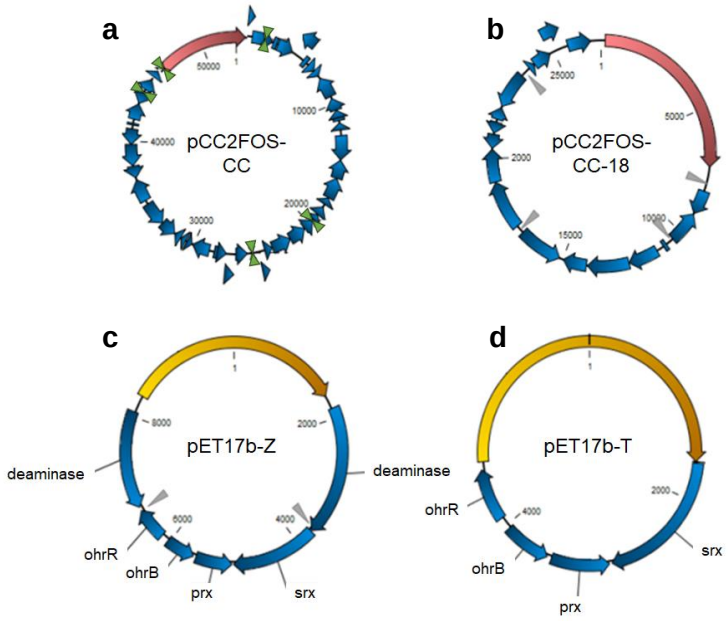


Figure 3.2: Subcloning. Restriction sites are marked in green on the fosmid and primer sites indicated by grey arrows next to the fosmid. a) The identified pCC2FOS-CC, which confers antifungal activity against *C. albicans*, holding a metagenomic insert with 41 ORFs. b) Subclone pCC2FOS-CC-18 conferring the antifungal phenotype. c) Subclone pET17b-Z. d) Subclone pET17b-T.

The Ohr regulon consisting of the putative, transcriptional regulator OhrR and the two putative, catalytically active peroxidases OhrB and Prx was next analysed in more depth. The peroxidoredoxin and *ohrB* genes encode a 180 and a 143 amino acid protein of an estimated molecular weight of 20.04 kDa and 14.54 kDa, respectively. The family of Ohr and Prx proteins are both Cys-based, thiol-dependent peroxidases, which differ in their biochemical and structural properties, but share similar catalytic mechanisms (Lesniak, Barton and Nikolov 2002; Alegria *et al.* 2017). While Ohr contains a unique α/β-fold (Lesniak, Barton and Nikolov 2002), the Prx family is characterised by a thioredoxin-fold (Hall *et al.* 2011). In both protein families many peroxidases contain two catalytic active cysteine residues (2-Cys), the peroxidatic (C_p) and the resolving (C_R) cysteine (Perkins, Poole and Karplus 2014; Alegria *et al.* 2017).

The novel OhrB gene from the *Sphagnum* moss metagenome

The novel SmMOhrB (OhrB from the *Sphagnum magellanicum* Metagenome) shares the highest sequence similarity with several *Burkholderia sp.* derived OhrB proteins (85.9-88.7%). Elucidation of the phylogenetic relation of SmMOhrB with sequences from different bacterial phyla showed OhrB of the *Sphagnum* metagenome to cluster together with those from *Bulkholderia sp.* (Figure 3.3). Furthermore, SmMOhrB together with the other OhrB proteins and those of the OhrA type form overall two distinct clusters.

A multiple amino acid sequence alignment of the newly identified SmMOhrB (OhrB from the *Sphagnum magellanicum* Metagenome) with other Ohr proteins showed the catalytically important, highly conserved amino acids reported for this protein family to be present (Figure 3.4). This includes two cysteines separated by 64 amino acids (Cys60 and C125), one arginine (Arg18) and one glutamic acid (Glu50) (Lesniak, Barton and Nikolov 2002; Cussiol *et al.* 2003; Oliveira *et al.* 2006). Furthermore, SmMOhrB contains highly conserved glycines (Gly16/17) and a tyrosine (Tyr7) near the amino terminus, and the 17CysArg (CR motif), 48AsnProGluLysLeu (NPExL motif), 60CysPhe (CF motif) and 125CysProTyr (CPY motif) (Völker *et al.* 1998; Meireles *et al.* 2017).

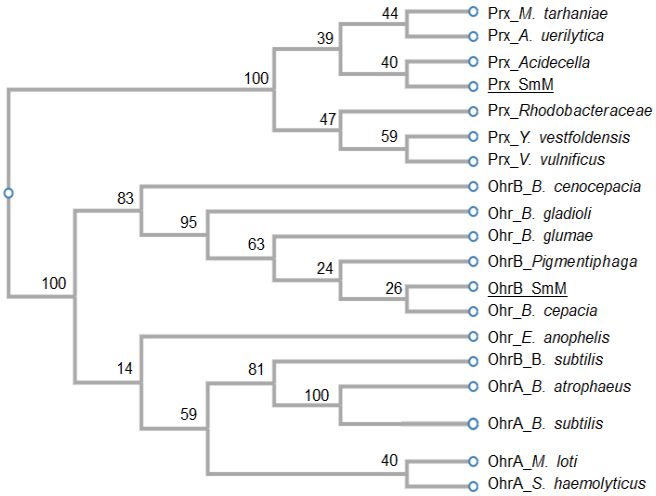


Figure 3.3: Phylogenetic relationship of SmMOhrB and SmMPrx with other Ohr proteins and peroxiredoxins. The phylogenetic relation of SmMOhrB and SmMPrx was elucidated via a Maximum Likelihood tree using PhyML with bootstrap values shown next to the branches. The used sequences for Prx are: *Vibrio vulnificus* (A0A1Z0YU25), *Yoonia vestfoldensis* (A0A1Y0E978), *Rhodobactereae* (A0A1B6YJ88), *Methylobacterium tarhaniae* (A0A0J6SKP2), *Aureimonas ureilytica* (A0A175 RPG0), *Acidocella sp.* (K5YHF6). The selected Ohr sequences are: *Staphylococcus haemolyticus* (Q4L4R0), *Bacillus subtilis* (P80242, O34762), *Bacillus atrophaeus* (A0A080UD17), *Mesorhizobium loti* (NZ_CP016079.1), *Elizabethkingia anophelis* (A0A077EDR1), *Burkholderia cenocepacia* (B4EPK2), *Burkholderia gladioli* (F2LQG7), *Pigmentiphaga sp.* (A0A3P4AWU6), *Burkholderia glumae* (A0A2J9HM40), *Burkholderia cepacia* (WP_043187524.1

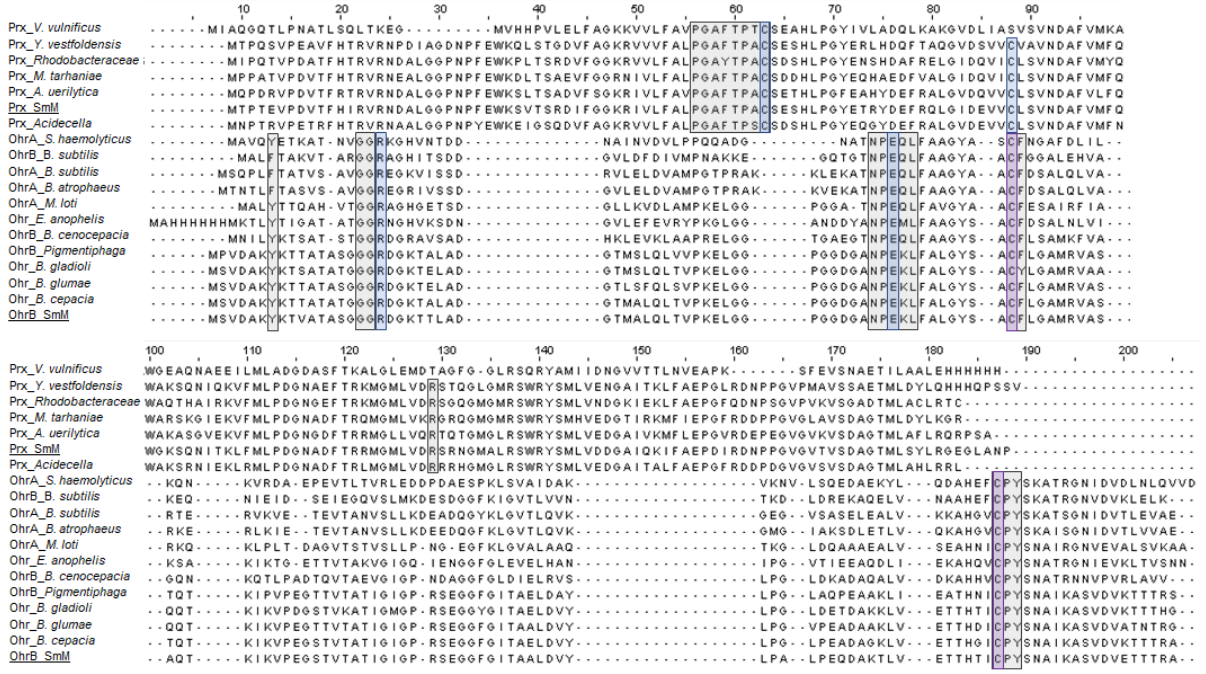


Figure 3.4: Multiple amino acid sequence alignment of Ohr and Prx enzymes. The *Sphagnum magellanicum* metagenome OhrB and Prx (SmMOhrB and SmMPrx) were aligned against sequences from different bacterial phyla using Clustal Omega. Catalytic cysteins as well as catalytic arginine and glutamic acid are highlighted by a purple or blue background, respectively. Other conserved amino acids or motives are highlighted by a grey background. The employed Prx sequences are: *Vibrio vulnificus* (A0A1Z0YU25), *Yoonia vestfoldensis* (A0A1Y0E978), *Rhodobactereae* (A0A1B6YJ88), *Methylobacterium tarhaniae* (A0A0J6SKP2), *Aureimonas ureilytica* (A0A175RPG0), *Acidocella sp.* (K5YHF6). The selected Ohr sequences are: *Staphylococcus haemolyticus* (Q4L4R0), *Bacillus subtilis* (P80242, O34762), *Bacillus atrophaeus* (A0A080UD17), *Mesorhizobium loti* (NZ_CP016079.1), *Elizabethkingia anophelis* (A0A077EDR1), *Burkholderia cenocepacia* (B4EPK2), *Burkholderia gladioli* (F2LQG7), *Pigmentiphaga sp.* (A0A3P4AWU6), *Burkholderia glumae* (A0A2J9HM40), *Burkholderia cepacia* (WP_043187524.1).

Next, a homology model was constructed and aligned with other OhrB structures (Figure 3.4). OhrB was reported to be a homodimer, with nearly structurally identical monomers, as depicted in our homology model (Lesniak, Barton and Nikolov 2002). OhrB and the template structure shared 41.9% sequence identity and homology modelling resulted in a global model quality estimate of 0.76 and a QMEAN Z-score of -0.5. In the quaternary structure of the OhrB model, the conserved cysteine residues are located each on one of the two α-helices, the conserved arginine on the loop between the first two β-sheets, and the conserved glutamic acid on the first α-helix (Figure 3.4 a,c). The cysteines on each monomer are located such that they can form an intermolecular disulphide bond in the homodimer (C60-S-S-C125) and adopt a highly similar orientation compared with the crystallised OhrB structures. Glutamic acid shows a highly similar orientation for all aligned structures as well, while arginine displays more freedom in its orientation due to its location on the less conserved loop (Figure 3.4 b,d,e).

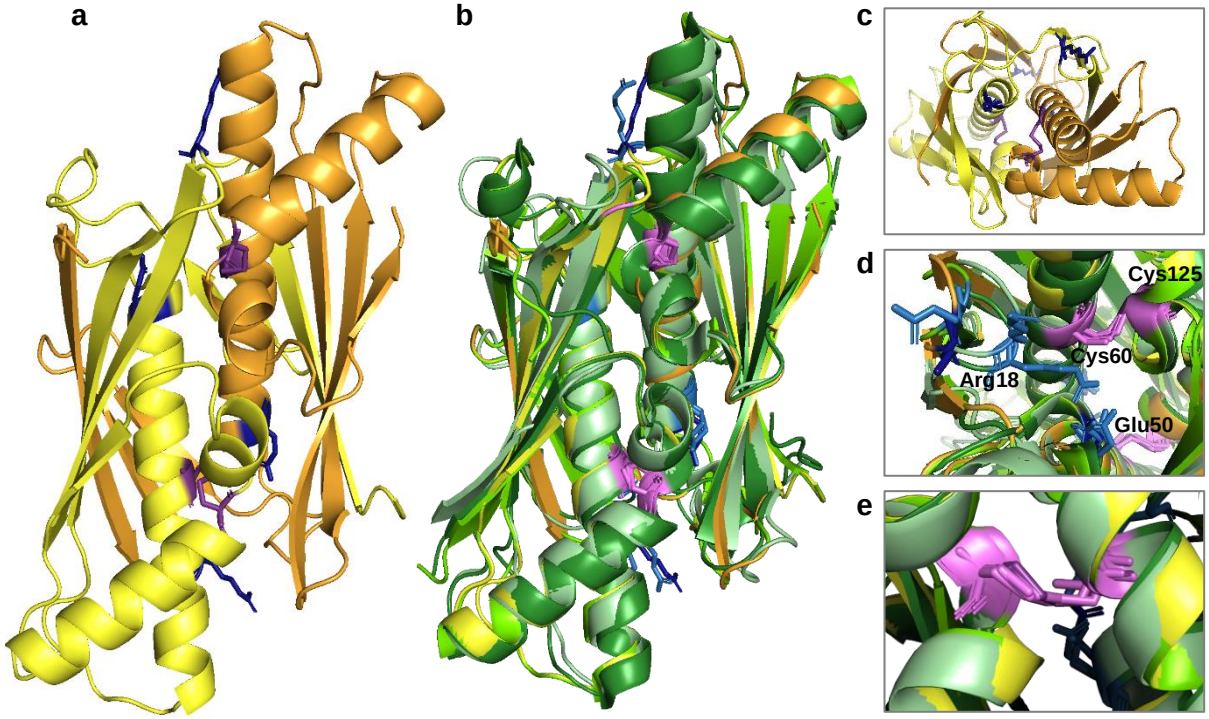


Figure 3.5: Homology model of SmMOhrB. a) Structure of SmMOhrB homodimer based on homology modelling against OhrB from *Elizabethkingia anophelis* (6D9N). b) Structure alignment of the modelled SmMOhrB and crystallised OhrB structures of *E. anophelis*, *Pseudomonas aerigunosa* (1N2F) and *Bacillus subtilis* (2BJO). c) Side view of SmMOhrB homodimer. d) Active site of the OhrB alignment displaying the catalytically active and conserved cysteine residues (purple), arginine and glutamic acid residues (blue). e) The two catalytically active cysteine residues.
Cysteine residues (purple), arginine and glutamic acid (blue) are depicted for the novel OhrB (dark colour) and the crystallised structures (bright colour).

The novel Prx gene from the *Sphagnum* moss metagenome

The new Prx from the *Sphagnum magellanicum* Metagenome (SmMPrx) shares the highest sequence similarity with uncharacterized peroxidoredoxins from *Acidocella sp.* (76.3%), *Aureimonas ureilytica* (75.1%), and several *Methylobacterium sp.* derived OhrB (~74%). Elucidation of the phylogenetic relation of SmMPrx to other peroxidoredoxins showed SmMPrx to cluster together with its next neighbour sequence from *Acidocella* and in closer proximity to Prx from *Aureimonas* and *Methylobacterium sp.* than to *Vibrio vulnificus*, *Yoonia vestfoldensis*, and *Rhodobactereae* (Figure 3.3).

SmMPrx contains the highly conserved C_P (Cys55) and C_R (Cys80) as well as the conserved arginine (Arg121) and 48ProGlyAlaPheThrProAlaCys (Pxxx(T/S)xxC motif) (Hall *et al.* 2011) as depicted in the multiple sequence alignment (Figure 3.4). No modelled quaternary structure was generated as no suitable bacterial 2-Cys peroxiredoxin template was available.

Discussion

Functional screening of a moss metagenomic library for the discovery of novel drug leads to combat the growing number of fungal invasive infections led to the identification of one active clone, CC. Subcloning revealed the Ohr regulon, encoding the putative, transcriptional regulator OhrR and the two putative peroxidases Prx and OhrB, together with the putative sulfiredoxin Srx, to cause the antagonistic phenotype. Their involvement in antagonism has yet not been described. Interestingly, the Ohr regulon is known as an oxidative stress response involved in the reduction of organic hydroperoxides (OHPs), in particular fatty acid hydroperoxides and peroxinitrite (Fuangthong *et al.* 2001; Lesniak, Barton and Nikolov 2002; Alegria *et al.* 2017). Sulfiredoxin was first reported in connection to the hydroperoxide stress resistance in eukaryotes and especially to the catalytic activity of Prx (Biteau, Labarre and Toledano 2003). Srx recovers hyperoxidized peroxidases and reactivates peroxidase activity (Perkins, Poole and Karplus 2014).

The alignment of the primary and a modelled quaternary structure of OhrB showed the highly conserved, catalytically active cysteines, the arginine residue and glutamic acid to be located at the active site. Its quaternary structure, thereby, shares high similarity with that of other OhrB structures, showing the characteristic α/β-fold (Lesniak, Barton and Nikolov 2002). The multiple sequence alignment showed also that the two catalytic cysteins are present in Prx. Based on *in silico* analysis, the identified *ohrB* and *prx* genes encode a new 2-Cys dependent peroxidase of the Ohr subfamily and the Prx family, respectively, which are herein reported for the first time to be involved in inhibition.

The produced inhibitory molecule, which could no be isolated for analysis thus far (data not shown), may be a hydroxy fatty acid (HFA). Fatty acid hydroperoxides and peroxinitrite were proposed as natural substrates for peroxidases of the Ohr subfamily, which they reduce to their respective alcohols (Alegria *et al.* 2017). The substrate(s) for the Ohr regulon may originate from the host background, specifically from the secondary metabolism. This seems likely, as the inhibitory effect depends on medium composition and prolonged incubation time (data not shown). Upon reduction, the metabolised substrate acquires antagonistic activity and can then inhibit *C. albicans.* Alternative, the yeast itself could provide the substrate. Ohr proteins are widely spread among bacteria including many pathogens (Atichartpongkul *et al.* 2001). The release of highly toxic organic hydroperoxides such as fatty acid hydroperoxides as innate immune response to combat pathogens in eukaryotic cells (Clifford and Repine 1982; Prost *et al.* 2005; Perkins, Poole and Karplus 2014) and the detoxification of OHPs via Ohr proteins and peroxiredoxins as a response and important driver for virulence have been described (Atichartpongkul *et al.* 2010; Perkins, Poole and Karplus 2014). It is hence conceivable that such a response is triggered in *C. albicans* by the presence of *E. coli*^{ohr}. The yeast could produce OHPs, like the suggested fatty acid hydroperoxides, as part of an oxidative attack against *E. coli,* which is warded off through the acquired peroxidase proteins.

However, the antifungal compound potentially being a hydroxy fatty acid or any hydroxyl-compound derived from OHPs for that matter is highly interesting from the drug development point of view. Current therapeutic agents against *C. albicans* belong to only four classes, namely polyenes, azoles, pyrimidine derivatives and echinocandins (Perfect

2017; Geddes-McAlister and Shapiro 2018). HFAs or the broader group of hydroxy compounds derived from organic hydroperoxides, thus, represent a novel chemical class. Based on that, the antifungal compound may possess a new mode of action. Further analysis of the produced inhibitory molecule should give valuable insight into the potential of the compound as a new drug lead.

Methods

Strains and culture conditions. Strains and vectors used in this study are listed in Table 1. *C. albicans* was cultured on Nutrient Broth (NB) II (Sifin diagnostics, Germany) agar plates at 37˚C. Unless otherwise stated, all *E. coli* strains were cultivated at 37˚C using Luria-Bertani (LB) medium amended with chloramphenicol (12.5 µg ml^{-1}, Carl Roth, Germany) or ampicillin (100 µg ml^{-1},Carl Roth, Germany), respectively, for vector maintenance. For high copy number induction of the CopyControl fosmid pCC2FOS medium was additionally supplemented with L-arabinose (0.01% w/v, Sigma-Aldrich, Germany).

Metagenomic library construction. Metagenomic DNA from the moss *S. magellanicum* sampled at the Alpine peat bog Pirker Waldhochmoor (N46˚37´38.66´´, E14˚26´5.66´´) was used for the previously generated 3.6 Gbps clone library, which was established in *E. coli* EPI300 pCC2FOS (Epicentre, Wisconsin, USA) employing the CopyControl fosmid library production kit (Epicentre, Wisconsin, USA) as previously described (Müller *et al.* 2017). The 90 000 fosmid clones, holding a ~40 kb DNA fragment each, were stored as amplified library pools in LB medium supplied with 20% (v/v) glycerol for long term storage at -70˚C.

Screening for antimicrobial activity via the agar overlay method. For the initial screening of the fosmid library, the metagenomic clones were revived for 3 h at 37˚C with shaking at 130 rpm. They were then plated onto selective Terrific Broth agar (TB, 12 g l^{-1} tryptone, 24 g l^{-1} yeast extract, 0.4% glycerol (v/v), 0.017 M KH_2PO_4, 0.072 M K_2HPO_4, 15 g l^{-1} agar) to ~500 colony forming units per plate. After 72 h of incubation at 37˚C, the plates were overlaid with 3 ml of *C. albicans*-infused NB soft agar (0.07% w/v). Therefore, yeast cell material was suspended in 0.85% (w/v) sodium chloride solution to a turbidity of McFarland 0.5 and diluted 1:20 into the agar. Plates were incubated at 20˚C for three days followed by a 24 h incubation at 37˚C and then evaluated for halo formation. Positive clones were recovered, and their activity verified by re-transformation into *E. coli* EPI300 followed by functional screening.

During the subcloning the assay was modified. The medium was additionally amended with 1% CAS amino acids (w/v). The turbidity of the yeast suspension was adjusted to McFarland 1 and diluted 1:5 into the soft agar to achieve a more distinct halo formation. Furthermore, *S. cerevisiae* was used as alternative target organism in addition to *C. albicans* due to a more stable phenotype during the activity screenings. Transformants and active subclones were otherwise subjected to the assay as described above.

***De novo* sequencing of pCC2FOS-CC.** Sequencing was performed as previously described (Obermeier *et al.,* unpublished). Briefly, the 300 bp reads generated using the Illumina MiSeq technology (Illumina, California, USA) were quality filtered with Trimmomatic v0.36 (Bolger, Lohse and Usadel 2014) and assembled with SPAdes v3.10.0 (Bankevich *et al.* 2012), resulting in a 53 kb contig. Prokka (Seemann 2014) was employed for gene annotation.

Identification of antimicrobial genes. Active genes were identified in a step wise subcloning approach (Figure 2). Enzymes and kits used during the procedure were purchased from Thermo Fisher Scientific or New England Biolabs and used according to the manufacturer´s instructions unless otherwise stated.

First, pCC2FOS-CC was digested using HindIII generating five insert fragments which were separately ligated into the HindIII digested pCC2FOS vector using T4 DNA ligase and transformed into chemically competent *E. coli* EPI300. Transformants were subjected to the agar overlay based activity screening, as described above, to identify active clones.

All further subclones were established in purchased, chemically competent *E. coli* 10-beta cells using the pET17b vector. AQUA cloning (Beyer *et al.* 2015) was employed to this end. Insert primers were designed as described by Beyer *et al.* with flanking 32 bp extensions homologous to the pET17b vector and *vice versa.* Primers and PCR conditions are listed in Table 2. Amplified insert and backbone fragments were mixed in a 1:1 ratio, incubated at room temperature for 1 h and 2.5 µl of the mixture added to 25 µl of competent cells for subsequent transformation. Clones were validated via restriction enzyme digestion of isolated plasmids using XhoI and in case of subclone pCC2FOS-CC-T and S, additionally with Eco32I. Correct clones were confirmed by sequencing and activity screening.

Phylogenetic analysis of SmMOhrB and SmMPrx. The phylogenetic analysis for inferring the evolutionary relationship of SmMOhrB and SmMPrx with other Ohr proteins and peroxiredoxins was conducted using the software ClustalW. First alignment and phylogenetic reconstructions were generated using the function "build" of ETE3 v3.0.0b32 (Huerta-Cepas, Serra and Bork 2016), as implemented on GenomeNet (https://www.genome.jp/tools/ete/). Using the generated multiple sequence alignment, a Maximum Likelihood tree was inferred using PhyML v20160115 at default settings (Guindon *et al.* 2010). Branch supports are computed out of 100 bootstrapped trees.

Sequence alignment and homology modelling. Protein sequences were aligned using Clustal Omega v1.2.4 with Clustal W default settings and visualised using Jalview v2.10.5. The structure of OhrB was modelled using the automated modelling server Swiss-Model (Waterhouse *et al.* 2018), for which the most similar, non-ligand bound structure was used as template. The obtained protein model was then aligned with Ohr structures using PyMol v.2.3.2.

Table 3.1: Strains or vectors used in this study

Strains or vectors	Properties	Sources or references
Candida albicans H5	Wildtype	Institute of Environmental Biotechnology, TU Graz
Saccharomyces cerevisiae	Wildtype	Institute of Environmental Biotechnology, TU Graz
E. coli		
EPI300	F¯ *mcrA Δ(mrr-hsdRMS-mcrBC)* (Str^R) *φ80dlacZΔM15 ΔlacX74 recA1 endA1 araD139Δ(ara, leu)7697 galU galK λ¯rpsL nupG trfA tonA dhrf*	Epicentre
10-beta	*Δ(ara-leu) 7697 araD139 fhu ΔlacX74 galK16 galE15 e14-φ80dlacZΔM15recA1 relA1 endA1 nupG rpsL* (Str^R) *rph spoT1 Δ(mrr-hsdRMS-mcrBC)*	New England Biolabs
Vectors		
pCC2FOS	Cloning and maintainance vector, *ori2* and *oriV*, Cm^R,8.1 kb	Epicentre
pCC2FOS-CC	pCC2FOS holding a XX kb insert of the *S. magellanicum* metagenome active against *C. albicans*	This study
pCC2FOS-CC-18	Active pCC2FOS-CC subclone comprising a 18 kb fragment with XX genes of the pCC2FOS-CC insert	This study
pET17b	Expression vector, T7 promoter, Amp^R, 3.3 kb	Novagen
pET17b-Z	pET17b with incorporated fragment Z of the pCC2FOS-CC clone, comprises XX genes	This study
pET17b-Z4	pET17b derivative containing a small, junk DNA insert	This study
pET17b-T	pET17b holding fragment T of pCC2FOS-CC with *soxC, ohrB, ohrR* and putative *ohrA*	This study

Table 3.2: Primers and PCR reactions for AQUA cloning

Primer	Oligonucleotide sequence 5´- 3´	PCR programme
pET17b-Wfwd	CCCGACGGGCAAATCACCCAATCTCTGTCCAGCTCGAGCAGATCCGGCTGCTAAC	2 step PCR using the Q5 High-Fidelity DNA Polymerase: 98°C – 30 s; 5 cycles of 98° - 10 s, 71°C – 20 s, 72°C – 2 min with 15 µl reactions for each primer; 25 cycles of 98°C – 10 s, 71°C – 20 s, 72°C – 2 min with combined reactions of primer pairs; 72°C - 5 min
pET17b-Wrev	AAGCCCTTTCAGTCGAGGGTTTTAGTCATATCCATATGTATATCTCCTTC	
pET17b-Xfwd	CCGCCGGCACGGGATAATACTTTGTGGGTGAACTCGAGCAGATCCGGCTGCTAAC	
pET17b-Xrev	ATCCCACCAGGAAACAGCCTAGGAACACGTGGCATATGTATATCTCCTTC	
pET17b-Yfwd	GTGAACAATGTCACTACCCCCTCTGTGCCAACCTCGAGCAGATCCGGCTGCTAAC	
pET17b-Yrev	TTCGGACGTTCACCCACAAAGTATTATCCCGTCATATGTATATCTCCTTC	
pET17b-Zfwd	ATTGTGCTGACGCCCGATATGACTAAAACCCTCTCGAGCAGATCCGGCTGCTAAC	
pET17b-Zrev	CCCTGGTCGCGTTGGCACAGAGGGGGTAGTGACATATGTATATCTCCTTC	
pET17b-Tfwd	GACGCCAGTGGCGATTGTTCCCAGACGCTCGTCTCGAGCAGATCCGGCTGCTAAC	2 step PCR using Phusion High-Fidelity DNA Polymerase: 98°C – 30 s; 5 cycles of 98°C – 10 s, 71°C – 20 s, 72°C – 2 min with 15 µl reactions for each primer; 25 cycles of 98C° - 10s, 71°C – 20 s, 72°C – 2 min with combined reactions of primer pairs; 72°C - 5 min
pET17b-Trev	ATATCCATCGGAGCTATCCTGTCACACTTCGACATATGTATATCTCCTTC	
pET17b-Sfwd	GACGCCAGTGGCGATTGTTCCCAGACGCTCGTCTCGAGCAGATCCGGCTGCTAAC	
pET17b-Srev	CCAATCCGTAAAACAAATCATCGAGCATCCAGCATATGTATATCTCCTTC	
pCC2FOS-Wfwd	TTGTTTAACTTTAAGAAGGAGATATACATATGGATATGACTAAAACCCTCGACTG	Using Q5 or Phusion High-Fidelity DNA Polymerase: 98°C – 3 min; 34 cycles of 98°C – 10s, 55°C – 45 s, 72°C – *; 72°C – 10 min; *W: 2:50 min, X: 5 min, Y: 3 or 5 min, Z: 2 min
pCC2FOS-Wrev	TCGGGCTTTGTTAGCAGCCGGATCTGCTCGAGCTGGACAGAGATTGGGTGATTT	
pCC2FOS-Xfwd	TTGTTTAACTTTAAGAAGGAGATATACATATGCCACGTGTTCCTAGGCTGTTTCC	
pCC2FOS-Xrev	TCGGGCTTTGTTAGCAGCCGGATCTGCTCGAGTTCACCCACAAAGTATTATCCCGT	
pCC2FOS-Yfwd	TTGTTTAACTTTAAGAAGGAGATATACATATGACGGGATAATACTTTGTGGGTGAA	
pCC2FOS-Yrev	TCGGGCTTTGTTAGCAGCCGGATCTGCTCGAGGTTGGCACAGAGGGGGTAGTGA	

pCC2FOS-Zfwd	TTGTTTAACTTTAAGAAGGAGATATACATATGTCACTACCCCCTCTGTGCCAAC	
pCC2FOS-Zrev	TCGGGCTTTGTTAGCAGCCGGATCTGCTCGAGAGGGTTTTAGTCATATCGGGCGTC	
pCC2FOS-Tfwd	TTGTTTAACTTTAAGAAGGAGATATACATATGTCGAAGTGTGACAGGATAGC	Using Q5 High-Fidelity DNA Polymerase: 98 °C – 3 min; 34 cycles of 98°C – 10 s, 55°C – 45 s, 72°C – 1:30 min (S)/ 72.5°C – 2:30 min (T); 72°C – 10
pCC2FOS-Trev	TCGGGCTTTGTTAGCAGCCGGATCTGCTCGAGACGAGCGTCTGGGAACAAT	
pCC2FOS-Sfwd	TTGTTTAACTTTAAGAAGGAGATATACATATGCTGGATGCTCGATGATTTGT	
pCC2FOS-Srev	TCGGGCTTTGTTAGCAGCCGGATCTGCTCGAGACGAGCGTCTGGGAACAAT	

Flanking homologous primer extensions are underlined. Enzymes and reagents were purchased from New England Biolabs and Thermo Fisher Scientific and used according to their manuals, if not otherwise stated.

4 A new high-throughput screening method to detect antimicrobial volatiles from metagenomic clone libraries

Melanie-Maria Obermeier[1,2], Franz Stocker[1,2], Katharina Resch[1,2], Christina Andrea Müller Bogotá[1,2], and Gabriele Berg[1]

[1] *Institute of Environmental Biotechnology, Graz University of Technology, Graz, Austria;*

[2] *Austrian Centre of Industrial Biotechnology, Graz, Austria*

Abstract

The agricultural sector records high yield losses every year due to fungal infestation. Combating fungal phytopathogens, however, becomes increasingly difficult as resistances to the limited number of available fungicide classes expands worldwide. A promising group of chemicals to meet the great need for novel fungicides are microbial volatile organic compounds (VOCs). These rarely exploited molecules can exhibit antimicrobial functions and their high vapour pressure makes them ideal for the application in biofumigation.

We previously developed the Two Clamps VOCs Assay (TCVA) for the detection of antimicrobial VOCs and describe herein its adaption into a high-throughput screening method (htTCVA), which allows the efficient screening of large culture collections and clone libraries. In both assays two multi-well plates assembled top to top and separated by a perforated silicon foil create self-contained chambers in which the microbial activity of bacterial isolates or clones is assessed against a test organism. By using fungal spores instead of mycelium for the htTCVA in 96-well plates, the screening process is accelerated by at least four-fold. The robustness and feasibility of the htTCVA was demonstrated by the successful, first functional metagenomics screening for antifungal VOCs. In a multi-step screening approach using the htTCVA in the initial screening round, 20 000 fosmid clones from a *Sphagnum* moss derived metagenomic library were tested against *F. culmorum*. This led to the identification of eight antifungal VOCs, which inhibited the growth of spores and mycelium *in vivo* up to 30% and 9%, respectively, and using the reference compounds *in vitro* 100%.

The results demonstrate that antimicrobial VOCs can be successfully identified from not-yet cultivable microorganisms via functional metagenomics, which grants access to completely new and unexplored biosynthetic pathways and their secondary metabolites. The developed htTCVA accounts, thereby, as an efficient high-throughput screening method to detect antimicrobial VOCs not only from metagenomic libraries, but also from large culture collections. Beyond biofumigation, newly discovered antimicrobial VOCs can further find application in surface sterilisation in clinical environments.

Manuscript in preparation

5 Unravelling native plant resistomes – The *Sphagnum* resistome comprises versatile and novel antimicrobial resistance genes

Melanie-Maria Obermeier[1,2], Julian Taffner[1], Alessandro Bergna[1,2], Anja Poehlein[3], Tomislav Cernava[1], Christina Andrea Müller Bogotá[1,2], and Gabriele Berg[1]

[1]*Institute of Environmental Biotechnology, Graz University of Technology, Graz, Austria;*

[2]*Austrian Centre of Industrial Biotechnology, Graz, Austria*

[3]*Goettingen Genomics Laboratory, Georg-August-University Goettingen, Geottingen, Germany*

Abstract

The expanding antibiotic resistance crisis calls for a more in depth understanding of the importance of antimicrobial resistance genes (ARGs) in pristine environments. We, therefore, studied the microbiota associated with *Sphagnum* forming the main vegetation in undomesticated, evolutionary old bog ecosystems. In our complementary analysis of a culture collection, metagenomic data and a fosmid library, we identified a low abundant but highly diverse pool of resistance determinants, which targets an unexpected broad range of antibiotics including natural and synthetic compounds. This derives both, from the extraordinarily high abundance of efflux pumps (80%), and the unexpectedly versatile set of ARGs underlying all major resistance mechanisms. The overall target spectrum of detected resistance determinants spans 21 antibiotic classes, whereby β-lactamases and vancomycin resistance appeared as the predominant resistances in all screenings. Multi-resistance was frequently observed among bacterial isolates, e.g. in *Serratia*, *Pandorea, Paraburkholderia* and *Rouxiella*. In a search for novel ARGs we identified the new class A β-lactamase Mm3. The native *Sphagnum* resistome comprising a highly diversified and partially novel set of ARGs contributes to the bog ecosystem´s plasticity. Our results shed light onto the antibiotic resistance background of non-agricultural plants and highlight the ecological link between natural and clinically relevant resistomes.

MANUSCRIPT

Published as preprint at biorxiv:

Obermeier, M.-M., Taffner, J., Bergna, A., Poehlein, A., Cernava, T., Müller Bogotá, C.A., Berg, G. (2019) Unravelling native plant resistomes – The *Sphagnum* microbiome harbours versatile and novel antimicrobial resistance genes, doi:10.1101/695973

Introduction

The risk posed to modern medicine by increased morbidity and mortality associated with antibacterial resistance continues to escalate globally and has reached a stage where a post-antibiotic era is not unthinkable anymore (Kåhrström 2013; World Health Organization 2014). Many of the clinically relevant antimicrobial resistance genes (ARGs) originate from the environment, wherein they may act in intra-community signalling and metabolic processes; in the presence of selective pressure they can adapt antibiotic resistance as primary function (Martínez 2008). In order to retrace the origin and habitat transitions of resistant microorganisms, a detailed understanding of native resistomes is crucial (Martínez 2008). So far, such elucidations focused on soil, water and air (Pal *et al.* 2016). Limited work has been performed on plants and thereby mostly evolved around fresh produce to assess the risk potential of crops in serving as gateway of ARGs to humans (Blau *et al.* 2019; Cernava *et al.* 2019; Chen *et al.* 2019). The resistome of native plants from pristine vegetation was neglected so far. It can, however, provide the missing ecological link to understand the evolution and functioning of native resistomes as well as their role as pools of unexplored resistance mechanisms (Berendonk *et al.* 2015). Since the resistome reflects the continuous co-evolution of small bioactive molecules and microbial genomes within an environment (Wright 2007), native plants, which provide an extraordinarily diversified secondary metabolism, are expected to possess a diversified intrinsic resistome as well.

Sphagnum magellanicum Brid. covering peatlands, was selected as a model plant to study ARGs in a representative pristine as well as evolutionary old ecosystem (Kostka *et al.* 2016; Page and Baird 2016). *Sphagnum*-dominated peatlands constitute balancing and productive ecosystems, in which the prevailing harsh conditions fostered symbiotic connections throughout a long plant-microbe co-evolution (Opelt *et al.* 2007b; Bragina *et al.* 2014). As a result, the *Sphagnum* microbiome is highly abundant and diverse with a specialised structure and function similar across geographic locations (Opelt *et al.* 2007a; Bragina *et al.* 2012). The microbiota fulfils important functions like nutrient supply and protection against biotic and abiotic stress; its metagenome is characterised to a remarkably high extend by signatures indicating horizontal gene transfer and communication systems thought to facilitate the balance between plasticity and stability within the bog ecosystem (Opelt *et al.* 2007b). Moreover, the highly stable microbiome is not affected by soil microbiota, given that the rootless *Sphagnum* moss grows on peat; accumulated, partly degraded plant material mostly stemming from the plant itself and forming the largest terrestrial carbon sink on Earth (Kostka *et al.* 2016). *Sphagnum* mosses harbour specific and rich metabolite profiles (Opelt *et al.* 2007b) and their associated microbiota is characterised by a high proportion of antimicrobial activity (Opelt, Berg and Berg 2007). Altogether, *S. magellanicum* represents an ideal model to elucidate the antibiotic resistance background of plants which we expect to: i) comprise predominantly resistances against natural antibiotics due to the missing selective pressure by synthetic ones, ii) encompass versatile but evenly distributed ARGs due to the diverse and stable microbial community, iii) contain yet-unknown resistance genes. For our study, we pursued a unique approach combining analysis of a culture collection, *in silico* data mining of deep-sequenced metagenomic data, and functional metagenomics; the importance of combinatorial approaches for functional validation of *in silico* predictions was emphasised but rarely considered before (Allen *et al.* 2010; Chen *et al.* 2019).

Results

Sphagnum isolates display predominantly resistance against (semi)synthetic antibiotics

The culture collection was established using *Sphagnum* gametophytes from an Austrian Alpine bog, well known to host a highly abundant microbiota (Figure 5.1). Resistance assessment of the bacterial isolates included ten different antibiotics, comprising those ranked as critically important for medical applications (Collignon *et al.* 2016). Of the 264 isolated bacteria, 90% grew in the presence of at least one antibiotic, thereby, displaying 121 different resistance profiles (Table S1). With predominantly observed vancomycin, ampicillin, rifampicin, ciprofloxacin, and sulfadiazine resistance, resistance against semisynthetic and synthetic antibiotics was more prevalent than resistance to natural antibiotics (Figure 5.2). This contradicts our preliminary expectation of dominating resistances against natural antibiotics. Overall, resistance against all ten antibiotics was observed. Multi-resistance against eight and nine antibiotics was encountered for three isolates each (Table S1). The six isolates all displayed a distinct resistance profile, whereby all grew in the presence of erythromycin, vancomycin, rifampicin, ciprofloxacin and sulfadiazine. They were identified as the plant beneficial bacterium *Paraburkholderia phytofirmans*, formerly *Burkholderia phytofirmans* (Sawana, Adeolu and Gupta 2014), the potential nosocomial species *Serratia marcescens,* and as the newly described bacteria *Rouxiella chamberiensis*, *Pandoraea terrae* and *P. apista.*

Figure 5.1: The *Sphagnum*-dominated peat bog. a) The Austrian Alpine peat bog Pürgschachen Moor (N47°34'50.57" E14°20'29.29"). b) *S. magellanicum* gametophytes. c) Cross-section of a *Sphagnum* gametophyte displaying the highly abundant microbial colonisation (red spheres) of the moss´s hyaline cells (blue).

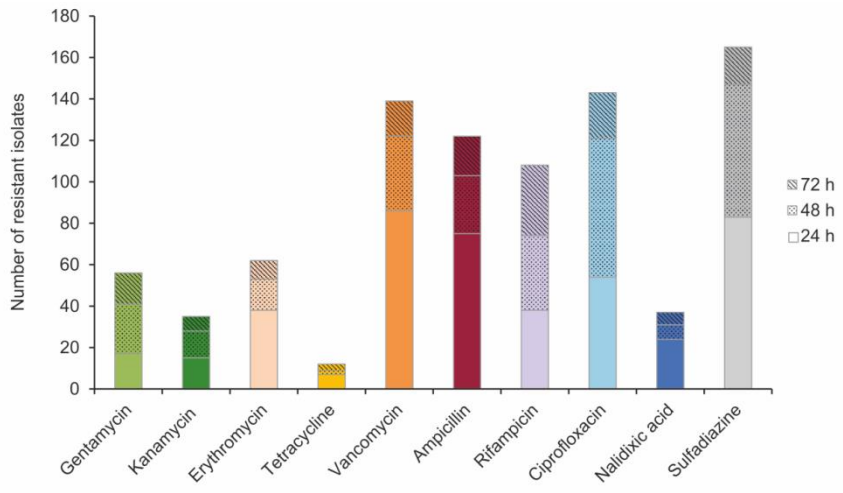


Figure 5.2: Antibiotic resistance profile of the *S. magellanicum* culture collection. Absolute number of resistant moss isolates which grew in the presence of different antibiotics.

The *Sphagnum* microbiome comprises a highly diverse resistome

An Illumina-sequenced *S. magellanicum* metagenome was aligned against Comprehensive Antibiotic Resistance Database (CARD) sequences using high stringency (90% threshold) yielding matches with ≥30 amino acids identity. This revealed a low abundant, but highly diverse pool of resistance determinants (Table S2). After curation and double normalisation of the generated hits (Table S3), 0.14% of all metagenomic reads were assigned to 887 ARGs with a collective antibiotic resistance abundance index (ARAI)(Elbehery, Aziz and Siam 2016) of 2.53 ppm. This included all major resistance mechanisms (Figure 5.3a): antibiotic target protection with 19 ARGs and 0.04 ppm (1.5%), antibiotic target replacement with 26 ARGs and 0.07 ppm (2.8%), antibiotic target alteration with 107 ARGs and 0.22 ppm (8.8%), antibiotic inactivation with 515 ARGs and 0.11 ppm (4.5%) and efflux-mediated resistance with 220 ARGs to an extraordinarily high share of collectively 2.07 ppm (82.4%).

To understand the strong contribution of efflux pumps towards resistance in more detail, we evaluated the extent of this resistance mechanism against antibiotics at class level (Figure 5.3b). The focus was, thereby, restricted to the antibiotic classes used during the screening of the culture collection. Glycopeptides were omitted as these act on the outer cell wall(Wright 2007). Efflux pumps, which export multiple antibiotics, were included in the abundance of each of the respective antibiotic classes. Although to a varying degree between 80% to almost 100%, efflux pumps constituted the most abundant resistance mechanism for all studied classes. Efflux-mediated resistance is more prevalent for macrolides, tetracyclines, β-lactams and fluoroquinolones than for aminoglycosides, rifamycins and sulphonamides based on the determined ARAI.

Next, the detected resistance determinants were grouped according to their antibiotic class to compare for all classes their distribution and abundance. However, efflux pumps were excluded entirely in this analysis as they often confer resistance to multiple antibiotics. The overall target spectrum of the detected 667 non-efflux pump determinants spans 21 antibiotic classes including synthetic antibiotics such as diaminopyrimidines, fluoroquinolones and sulphonamides and many classes ranked as critically important for human medicine (Collignon *et al.* 2016) like aminoglycosides, glycopeptides and β-lactams (Figure 5.3c). These results show a high degree of genetic diversity and an even distribution of the detected ARGs as expected, ranging from 8.3×10^{-6} to 1.5×10^{-2} ppm (Figure 3c, Table S2). Only two ARGs, *dfrE* and *parY*^{R}, mark an exception being the most prominent determinants with a considerable difference in abundance with 4.1×10^{-2} and 3.8×10^{-2} ppm, respectively. β-Lactams represent the most abundant class with more than 400 ARGs. In contrast, for aminocoumarins, mupirocin, nucleoside and elfamycin just one ARG was assigned to each (*parY*^{R}, *ileS, tmrB,* and *EF-Tu*). Almost all genetic determinants that were suggested by Berendonk *et al.* (Berendonk *et al.* 2015) as potential indicators to survey the antibiotic resistance status in environmental samples were detected in this metagenome as well, although at lower abundance. These include *sul1*, *sul2*, $bla_{\text{CTX-M}}$, bla_{TEM}, bla_{VIM}, bla_{KPC}, *qnrS*, *vanA*, *mecA*, *ermB*, *ermF*, *tetM* and *aph* (Table S2). Altogether, the data highlight the predominance of glycopeptide and β-lactam resistance determinants in the studied resistome, both in terms of abundance and versatility with 60 and 403 ARGs and 0.14ppm (32.6%) and 0.05 ppm (11.6%), respectively.

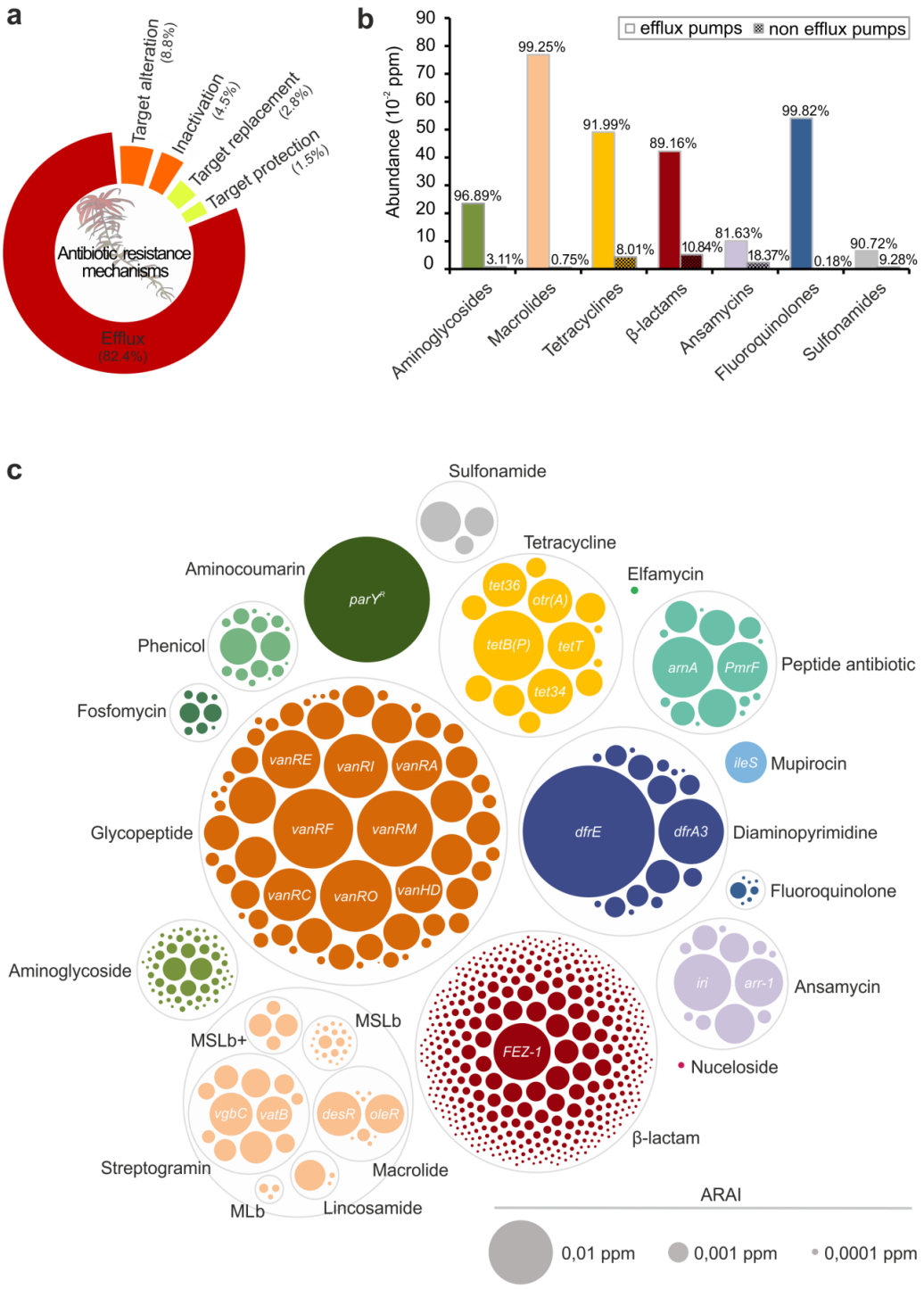


Figure 5.3: The *S. magellanicum* metagenome comprises a highly versatile resistome. The Illumina generated 41.8 Gbps moss metagenome was aligned against the CARD sequences. a) The five major resistance mechanisms presented by their relative abundance within the moss resistome. b) For a selected group of antibiotic classes, the extent of efflux pump mediated and non-efflux pump mediated resistance is compared. Abundance within the metagenome is given in absolute numbers by the ARAI in ppm (≙reads per million reads), while the abundance within antibiotic classes is given as proportion in percent. c) All detected non-efflux pump related ARGs grouped according to antibiotic classes. Each bubble represents one determinant with absolute abundance within the metagenome reflected by bubble size. The most abundant determinants are labelled with the gene names. MLb, macrolide and streptogramin. MSLb, MLb and streptogramin. MSLb+, MSLb and oxazolidione, pleuromutilin and phenicol. CARD, Comprehensive Antibiotic Resistance Database. ARAI, Antibiotic Resistance Abundance Index.

In silico analysis showed the predominance of β-lactam resistance determinants, and that β-lactamase diversity in the *Sphagnum* resistome covers every β-lactam class. The list of assigned β-lactamases included extended spectrum as well as metallo β-lactamases of environmental but also clinical origin, such as GIM-2, SHV-16 and TEM-102 (Table S2). Due to the relevance of extended-spectrum and metallo β-lactamases, which pose a problem to the still widely administrated β-lactams, a network analysis was conducted to assess the target spectrum of the 398 assigned β-lactamases (Figure 5.4). All six β-lactam classes, penams, penems, monobactams, cephalosporins, cephamycins and carbapenems, are represented in the constructed network. The majority of determinants (67.6%) cluster in groups acting on more than one β-lactam class. These clusters often connect to three, four and five β-lactam classes comprising 22.1%, 13.3% and 6.5% of the detected determinants, respectively. Overall, the determinants connect most frequently to cephalosporins and penams. However, connections to carbapenems, drugs of last resort, are also highly represented in the network analysis.

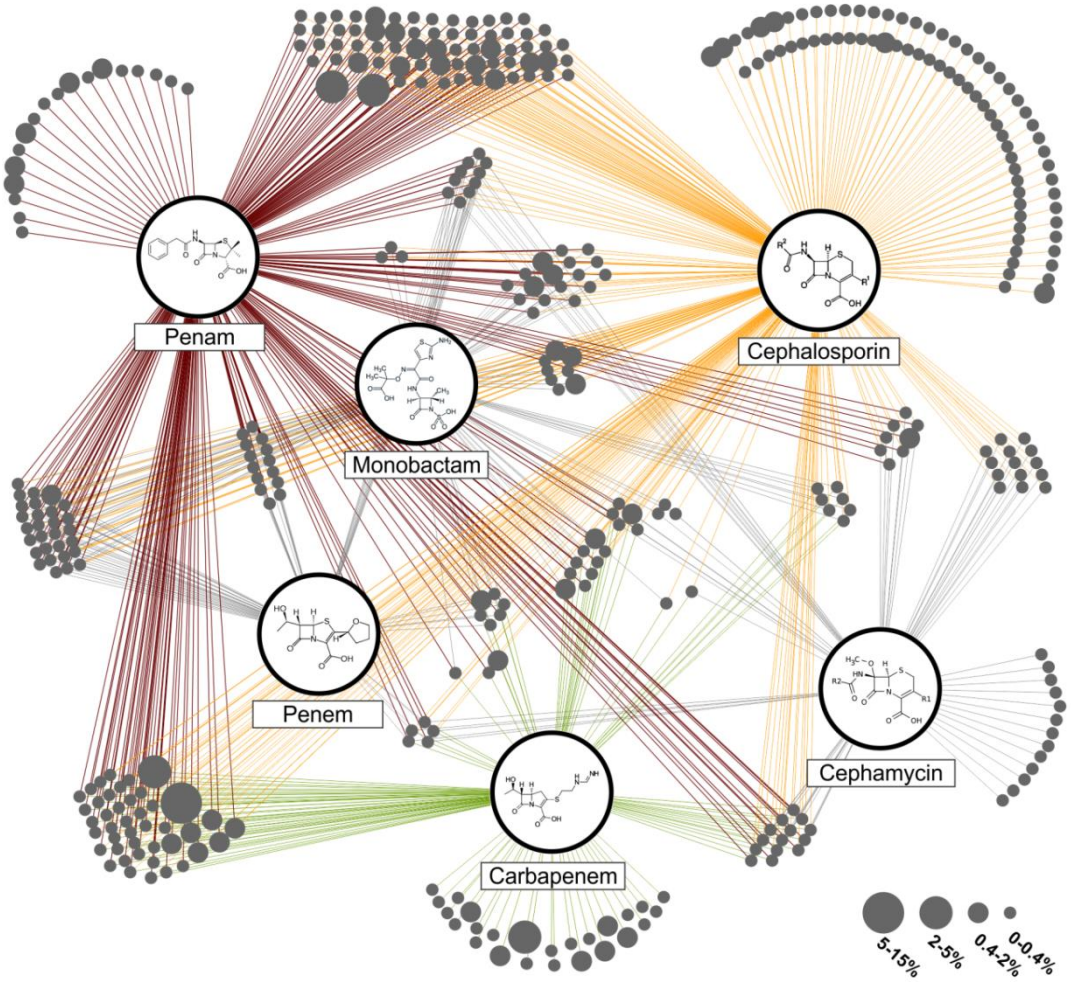


Figure 5.4: Substrate spectrum of detected β-lactamases. All assigned β-lactamases of the *S. magellanicum* metagenome represented as single bubbles (in dark grey) were grouped into clusters based on their reported substrate spectrum. Enzymes with the same substrate spectrum form one cluster. Connecting lines from the clusters to the β-lactam classes display the substrate specificity. Bubble size relates to the relative abundance of single enzymes within the whole β-lactamase pool. The three most abundant classes are penam (red), cephalosporin (orange) and carbapenem (green).

Identification of the novel class A β-lactamase Mm3 from the *Sphagnum* metagenome

A functional metagenomics approach was pursued to identify novel resistance genes. Therefore, a fosmid library comprising 3.6 Gbps of cloned moss metagenomic DNA was screened for ARGs against nine different antibiotics. The screening identified three unique resistant metagenomic clones (*E. coli* EPI300 pCC2FOS-Mm1, Mm2 and Mm3), all three conferring resistance against ampicillin. The initially determined MICs for ampicillin were 64 µg ml^{-1} for clones Mm1 and Mm2, and >512 µg ml^{-1} for Mm3, as compared to 32 µg ml^{-1} for the control strain (Supplementary Table 2). The clone *E. coli* EPI300 pCC2FOS-Mm3, exhibiting the highest MIC for ampicillin, was chosen for *de novo* sequencing. This revealed a novel β-lactamase gene encoding a 304 amino acid protein with an estimated weight of 32.8 kDa to be present on the 40.7 kb DNA insert. The gene was designated *blaMm3* (β-lactamase from Moss metagenome clone 3).

The novel β-lactamase gene *blaMm3* shares the highest sequence similarity with two annotated but not yet characterised β-lactamases from *Rhodanobacter* sp. (70.6%) and *Frateuria* sp. (66.8%). Both species belong to the family of *Rhodanobacteraceae* and order of *Xanthomonadales*. Together with reference sequences from characterised β-lactamases the phylogenetic relation of Mm3 was elucidated (Figure 5.5). Its clustering into defined groups of class A β-lactamases was evaluated according to the updated classification by Philippon *et al*. (Philippon *et al.* 2016). The Mm3 β-lactamase clustered, together with the next neighbour sequences from *Rhodanobacter* sp. and *Frateuria* sp., in closer proximity to members of the so called *Xanthomonas* (XANT) group, which contains β-lactamases from *Xanthomonas* sp., *Stenotrophomonas maltophilia* and *Pseudomonas aeruginosa*. Other clusters in the phylogenetic tree include members showing a lower degree of similarity (33 to 44% identity), like those belonging to the limited-spectrum (LSBL1 to 4) and extended-spectrum β-lactamases (ESBL1 and 3). Members of the LSBL2 and 3 clusters have been described as true carbenicillinases, while enzymes from the ESBL group hydrolyse cephalosporins like cefotaxime additionally to penicillins. In accordance with the phylogenetic analysis, the amino acid sequence of Mm3 harbours characteristic class A Ambler motifs (Ambler 1980) as follows: 70SerThrPheLys (SxxK motif), 130SerAspAsn (SDN motif), 234LysThrGly (KTG motif), Glu166 and 166GluProGluLeuAsn (ExxLN motif).

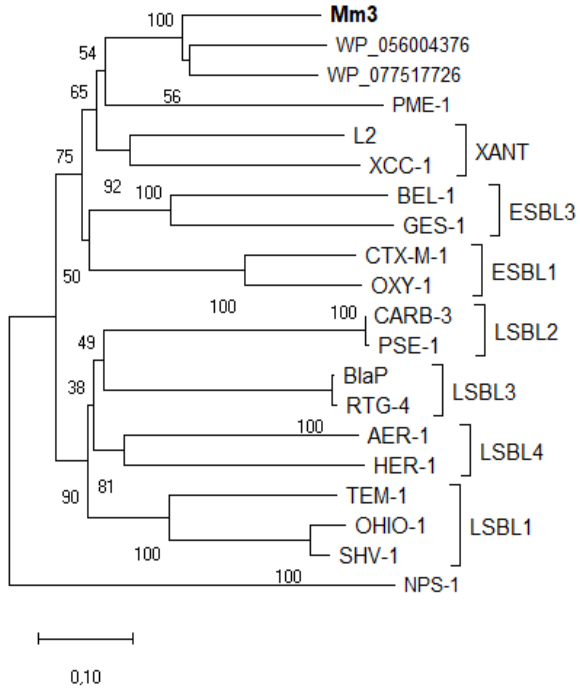


Figure 5.5: Phylogenetic relationship of Mm3 and other class A β-lactamases. The evolutionary analysis of aligned amino acid sequences was conducted using the neighbor-joining method. Bootstrap values are shown next to the branches. The scale bar indicates the number of amino acid differences per sequence. The reference sequences are: *Frateuria* sp. Soil773 (WP_056004376), *Rhodanobacter* sp. C03 (WP_077517726), PME-1 (*Pseudomonas aeruginosa,* E9N9H5), L2 (*Stenotrophomonas maltophilia*, P96465), XCC-1 (*Xanthomonas campestris* pv. *campestris*, O87643), BEL-1 (*Pseudomonas aeruginosa*, Q3SAW3), GES-1 (*Klebsiella pneumonia*, Q9KJY7), CTX-M-1 (*Escherichia coli*, P28585), OXY-1 (*Klebsiella oxytoca*, P22391), CARB-3 (*Pseudomonas aeruginosa*, P37322), PSE-1 (*Pseudomonas aeruginosa*, Q03170), AER-1 (*Aeromonas hydrophila*, Q44056), HER-1 (*Escherichia hermannii*, Q93FN7), BlaP (*Proteus mirabilis*, P30897), RTG-4 (*Acinetobacter baumannii*, ACJ61335), TEM-1 (*Shigella flexneri*, AAC97980), OHIO-1 (*Enterobacter cloacae*, P18251), SHV-1 (*Klebsiella pneumonia*, P0AD64). The tree was rooted with NPS-1, a class D β-lactamase from *Pseudomonas aeruginosa* (AAK1479).

To investigate the substrate spectrum of the novel β-lactamase, MICs for penam and cephalosporin antibiotics were determined (Table S4). Similarly, to the control strain, clone Mm3 showed no to little resistance to the tested cephalosporin concentrations with MICs of <0.5, 64 and 8 µg ml^{-1} for cefatoxime, cephalexin and cefalothin, respectively. On the contrary, Mm3 displayed a distinctly higher resistance against the penam antibiotics ampicillin (>512 µg ml^{-1}) and carbenicillin (>1024 µg ml^{-1}).

The identified *blaMm3* gene encoding a novel β-lactamase was cloned and expressed in *E. coli* BL21 (DE3) for subsequent purification and biochemical characterisation. After confirming solubility of the overexpressed enzyme by SDS-PAGE analysis, β-lactam-hydrolysing activity was verified by testing cell-free lysates on nitrocefin disks (data not shown). The N-terminally His-tagged enzyme was then purified by affinity chromatography to a purity of 90% as estimated by SDS-PAGE. Two prominent bands with a molecular weight of around 32 and 35 kDa were visible (Figure S1). LC-MS/MS analysis of both bands determined each of the respective proteins to comprise the right β-lactamase amino acid sequence (data not shown). However, the 35 kDa protein contained the N-terminal His-Tag while the lower weight protein did not (32 kDa), probably as a result from proteolytic activity during purification or SDS-PAGE analysis. Data obtained from the kinetic measurements of Mm3 for ampicillin were fitted using the Hill equation (n_H of 2.36). In the case of carbenicillin the kinetic data were fitted with the Michaelis-Menten equation (Figure S2). The kinetic analysis revealed a higher affinity of Mm3 for ampicillin (V_{max} = 179.2 ± 6.1 U mg^{-1}, K_M = 270.8 ± 16.4 µM) compared to carbenicillin (V_{max} = 264.6 ± 8.6 U mg^{-1}, K_M = 399.85 ± 42.69 µM).

Discussion

Our multi-faceted analysis uncovered a highly versatile resistome present in the evolutionary old and long-term stable bog ecosystem, and underlined the natural, strong resilience of *Sphagnum*-associated bacteria against antibiotics. Given the highly adapted plant-associated lifestyle, the strong microbial competition and the vast pool of microbial and plant-produced secondary metabolites (Bragina *et al.* 2014), the *Sphagnum* microbiota has developed general and also specific antimicrobial resistance mechanisms that naturally equip them against antibiotics. (Semi)synthetic drugs were not exempt from this as demonstrated in the present study. Contrary to our initial expectation, a predominance of resistances against synthetic antibiotics was observed in the culture collection , despite its pristine origin(Page and Baird 2016). Stemming to a large part from efflux pumps as indicated by our *in silico* analysis, the ability to combat these compounds may not exclusively result from extrusion. For instance, resistance of environmental bacteria against synthetics has been tied to high sequence variations in target genes (D'Costa *et al.* 2006). Furthermore, a high level of multi-resistances was encountered in the culture-dependent analysis. The six isolates with the highest level of multi-resistance belonged to *Burkholderiales* and *Enterobacteriales,* which are typical and dominant orders within the bacterial community of *Sphagnum* mosses (Bragina *et al.* 2012, 2014). *S. marcescens, P. phytofirmans,* and *R. chamberiensis* were isolated from this habitat before (Belova, Pankratov and Dedysh 2006; Opelt, Berg and Berg 2007; Fléche-Matéos *et al.* 2017). For *Pandoraea* spp. an association with *Sphagnum* was not described so far. Interestingly, this bacterium has been mostly isolated from the sputum of cystic fibrosis patients (Coenye *et al.* 2000) and is considered as emerging opportunistic pathogen (Green and Jones 2018). Clinically isolated *Pandoraea* spp. are known to be highly resistant, including resistances to last defence antibiotics of the carbapenem class (Schneider, Queenan and Bauernfeind 2006; Green and Jones 2018). The two multi-resistant isolates described here show that these species naturally possess a high level of resistance

and they might easily transit to clinical environments as they are well equipped with ARGs. *S. marcescens*, a common plant-associated bacterium, also exhibits various antibiotic resistances and opportunistic traits (Mahlen 2011). Less is known to this end for the identified *Paraburkholderia* and *Rouxiella* species, especially for *P. phytofirmans*, which is a promising plant growth promoting agent (Mitter *et al.* 2017). Interestingly, plants in general and *Sphagnum* in particular constitute reservoirs for plant growth promoting bacteria with antifungal and antibacterial activity (Vandamme *et al.* 2007), while simultaneously hosting species known as opportunistic human pathogens (Berg, Eberl and Hartmann 2005; Opelt, Berg and Berg 2007). According to our results, assessment of the environmental resistome in a given habitat can be used to predict emerging opportunistic pathogens and thereby help to counteract bacterial infections, which are considered a serious public health issue world-wide (Troeger *et al.* 2017).

The high microbial diversity of the *Sphagnum* moss microbiome was reflected by the corresponding resistome, which displayed high versatility and evenness. Given the high stringency applied to our analysis, the observed coverage on the functional and chemical level with more than 800 ARG-like genes covering 21 antibiotic classes is staggering. A key driver for this vast diversity resides in the presence of a great repertoire of efflux pumps. They are considered as an evolutionary ancient and general resistance mechanism against toxic molecules like heavy metals, solvents, and plant-produced antimicrobials(Martinez *et al.* 2009), and may be a missing link in understanding resistome composition in natural environments. For the *Sphagnum* resistome efflux-mediated resistance comprised more than 80%. An extraordinarily high share when compared to other CARD-based analyses, which attributed a relative proportion of 20-50% of the resistances to efflux pumps for various natural and human controlled environments (Xiao *et al.* 2016; Cernava *et al.* 2019; Mahnert *et al.* 2019). Yet, the high share of efflux pumps is a common attribute for other plant-associated resistomes in *Sphagnum*-dominated peatlands (unpublished data). As their typically high taxonomic diversity is inherently a driving force for chemical diversity, efflux pumps, thus, constitute a pivotal point in ensuring co-existence within this highly complex community. This general resistance mechanism fosters the diverse pool of ARGs present in the moss resistome and in doing so contributes to the great plasticity found within the peat bog ecosystem.

Since *Sphagnum* mosses are rootless plants that within *Sphagnum*-dominated peat bogs do not have soil contact, the elucidated resistome is regarded as intrinsic. As such, it addresses the need to understand the extent to which the plant resistome is intrinsic or recruited from soil, which represents still an unanswered question (Chen *et al.* 2019). In addition, we expect that it will be vertically transmitted with the core microbiome from the gametophyte to the sporophyte and *vice versa* (Bragina *et al.* 2012). Our data highlight that the plant microbiome naturally comprises a versatile, intrinsic resistome. This is reinforced by the identification of a novel class A β-lactamase. Notably, Mm3 shares low sequence similarity to characterised β-lactamases. This can be explained by the fact that most moos-associated microorganisms are not yet cultivable and not much is known about their origin and genetic content (Bragina *et al.* 2014). The metagenome derived β-lactamase Mm3 is phylogenetically closest related to β-lactamases of the XANT group and the uncharacterised β-lactamases from the environmental isolates *Rhodanobacter* (soil) and *Frauteuria* (rhizosphere), both belonging to the order of *Xanthomonadales* (Bai *et al.* 2015; Lycus *et al.* 2017), *Xanthomonadales spp.* constitute common colonisers of *Sphagnum* mosses (Bragina *et al.* 2013; Bragina, Berg and Berg 2015). Interestingly, many of the other β-lactamase sequences clustering in close proximity were isolated from well-known nosocomial human pathogens such as *S. maltophilia*, *P. aeruginosa* or *K. pneumoniae* (Green and Jones 2018). The relatedness of *bla*Mm3 to these genes is not surprising, since β-lactamases account as

evolutionary old enzymes and are widely spread in nature (Wright 2007). The latter was clearly confirmed in the network analysis, displaying high abundance and an extraordinary diverse substrate range of the *in silico* detected β-lactamases. The isolated β-lactamase Mm3 showed a higher affinity for penam antibiotics, but no activity for the tested cephalosporins, exhibiting in this case a narrow substrate spectrum. With K_M values around 270 to 400 µM for the penam antibiotics, the activity of the new Mm3 is not outstanding and surpassed by the ones reported for β-lactamases from many facultative human pathogens, e.g. the plasmid-encoded MIR-1, CMY-1 or ACT-1 from *E. coli* (0.16 to 2.2 µM, ampicillin) (Papanicolaou, Medeiros and Jacoby 1990; Bauernfeind *et al.* 1996; Dahmen *et al.* 2012). Based on this finding, the presence of antibiotic-inactivating enzymes with a rather limited activity seems to be characteristic for the described natural environment as compared to the clinical settings.

Since ARGs are often associated with specific taxa (Forsberg *et al.* 2014; Goethem *et al.* 2018; Mahnert *et al.* 2019), we expected the taxonomically diverse and balanced *Sphagnum*-associated microbiome to comprise a resistome with evenly distributed ARGs at low abundance; a common observation for natural environments (Pal *et al.* 2016). In contrast, environments under anthropogenic influence are characterised by highly abundant ARGs (Pal *et al.* 2016; Mahnert *et al.* 2019), and further correlate with a loss in bacterial diversity and enrichment of opportunistic pathogens (Mahnert *et al.* 2019). The antimicrobial selective pressure exerted by our life style is without question the driving force for imbalance, leading to a shift in bacterial community composition that ensues the increase of opportunistic pathogens and their associated ARGs. However, overlooked in this context is the ecological concept of *K*- and *r*-selection favouring oligotrophic or copiotrophic taxa, respectively. We recently reported that the phyllosphere of arugula from urban gardening was dominated by *Gammaproteobacteria* and in particular by multi-resistant *Enterobacteriacaea*(Cernava *et al.* 2019). This class, which comprises many opportunistic pathogens, tends towards a copiotrophic lifestyle, displaying faster growth and substrate generalisation as compared to the more oligotrophic *Alphaproteobacteria* (Kurm *et al.* 2017). The *Sphagnum* microbiota in contrary is dominated by *Alphaproteobacteria,* such as the slow growing *Methylobacteria* (Bragina *et al.* 2012, 2014). We assume that *K*-selection maintains oligotrophy and stabilises the bacterial community in the nutrient poor, microbial rich ecosystem of *Sphagnum*-dominated bogs. In doing so it represents a driving force in shaping the observed evenness and diversity within the moss resistome. Microbial community management ensuing diverse, stable and beneficially designed microbiomes is foreseen to abate exposure to resistances (Mahnert *et al.* 2019). We propose that resistance management in form of microbial community management could be achieved through *K*-selection. The advantageous effects of such a strategy have already proven valuable in improving the larvae viability in aquaculture (Vadstein *et al.* 2018).

Based on our complementary screening strategy, the herein presented novel findings deliver a first comprehensive picture of a native plant resistome consisting of a highly diverse genetic pool and novel antibiotic resistance genes.

Methods

CARD-based *Sphagnum* resistome profiling. A 41.8 Gbps metagenome previously generated by Illumina HiSeq paired-end sequencing from the Alpine peat bog moss *S. magellanicum* (Bragina *et al.* 2014) was implemented for antibiotic resistance profiling. The 172 590 841 paired-end reads were aligned against sequences from the CARD (McArthur *et al.* 2013), retrieved in April 2017, using the diamond protein aligner v0.9.24 (Buchfink, Xie and Huson 2015). BLASTX (Altschul *et al.* 1990) was performed at high stringency with a similarity threshold of 90% over the full read length and otherwise default settings giving hits with ≥30 amino acids identity. The reads were assigned to their best BLASTX hit and the obtained dataset was manually curated for gene redundancy or in case of antibiotic target genes for known resistance conferring mutations (Supplementary Table 1). The reads were normalised by calculating the ARAI (number of reads assigned to one ARG per number of total reads and respective ARG length) (Elbehery, Aziz and Siam 2016). Abundance of ARGs or resistance mechanism within the metagenome is given as ppm (≙read per million reads), while their proportion among all assigned ARGs or resistance mechanisms is given as percentage. The detected non-efflux pump determinants were visualised using RAWGraphs (Mauri *et al.* 2017). The distribution and abundance network of assigned β-lactamases was constructed with Cytoscape v3.3.0 (Shannon *et al.* 2003).

Sampling and isolation of *S. magellanicum* associated bacteria. Gametophyte samples of the moss *S. magellanicum* were collected from the Austrian Alpine bog Pürgschachen Moor (N47°34'50.57" E14°20'29.29") in September 2017 (Figure 1). Fluorescent *in situ* hybridization and confocal laser scanning microscopy were performed on *Sphagnum* gametophytes as described previously using the reported probes (Cy3-labeled ALF968 for *Alphaproteobacteria*, Cy5-labeled EUB338, EUB338II and EUB338III for *Eubacteria*) (Bragina *et al.* 2011).

The cleaned and fractionated plant material was shaken in a Stomacher laboratory blender (BagMixer, Interscience, France) twice for 120 s batch-wise in sterile plastic bags to 20 g in 50 ml of chilled 0.85% NaCl solution. After straining the suspension through double-layered gauze and a sterile analysis sieve (mesh size 63 µm, Retsch, Germany), the undiluted suspension as well as serial dilutions thereof were plated on R2A agar (Roth, Germany) containing nystatin (25 µg ml^{-1}, Duchefa Biochemie, Netherlands) and incubated at 20°C for four days. Isolates were subcultured until purity and liquid cultures grown from single, isolated colonies in Nutrient Broth II (Sifin diagnostics, Germany) supplied with glycerol to 20% (v/v) for long term storage at -70°C.

Antibiotic resistance screening of the *Sphagnum* culture collection. Bacterial isolates were screened against ten different antibiotics as listed in Supplementary Table 3. Concentrations were based on those used in other studies (Wardwell *et al.* 2009; Cernava *et al.* 2019), which followed the guidelines by the Clinical Laboratory Standards Institute. All 264 isolates were transferred to R2A agar plates of up to 50 isolates per plate and incubated at 20°C for four days. The colonies were then replica printed onto Müller-Hinton agar plates supplemented with the different antibiotics and incubated at 20°C. The plates were monitored every 24 hours for three days. This was done in duplicate and only isolates which had grown to visible colonies in both screenings were considered resistant.

Annotation of multi-resistant isolates. Cells were mechanically disrupted by bead-beating (6 m s^{-1}, 20 s) and the lysates incubated at 95˚C for 10 minutes followed by centrifugation at 5000 rpm for 5 min. The 16S rRNA gene was amplified using 2 µl of the supernatant and the universal bacterial primer pair 27f (5'AGAGTTTGATCMTGGCTCAG) and 1492r (5'TACGGYTACCTTGTTACGACTT), 0.5 µM each, in a 50 µl PCR reaction with 1x Taq-&GO Ready-to-use PCR Mix (MP Biomedicals, Germany) (98˚C - 4 min; 25 cycles of 98˚C - 30 s, 48˚C - 30 s, 72˚C - 90 s; 72˚C - 5 min). The 1400 bp long DNA fragments were purified (Wizard®SV Gel and PCR Clean-Up System, Promega, Germany) and sequenced using the same 27f and 1492r primer pair. Sequences were annotated to their best NCBI hit.

Metagenomic library construction. The 3.6 Gbps metagenomic fosmid library was established in *E. coli* EPI300 pCC2FOS (Epicentre, Wisconsin, USA), incorporating ~40 kb metagenomic DNA from the Alpine peat bog moss *S. magellanicum* as previously described (Müller *et al.* 2017). The generated 90 000 metagenomic clones were pooled by resuspending them in LB medium supplied with 20% glycerol for long term storage at -70˚C.

Screening for novel antibiotic resistance genes. The functional screening of the *Sphagnum* metagenomic library was carried out on LB agar plates containing chloramphenicol (12.5 µg ml^{-1}) for fosmid maintenance and arabinose (0.01% w/v) to induce high-copy number. Metagenomic clones were screened against nine different antibiotics as listed in Supplementary Table 3. Concentrations were chosen according to those used in other studies employing the CopyControl system with *E. coli* EPI300 (Allen *et al.* 2009; Vercammen *et al.* 2013). Alternatively, the minimal inhibitory concentration (MIC) was determined as described below. Cells of the pooled library stock were revived in LB broth containing chloramphenicol (12.5 µg ml^{-1}) at 37˚C for 3 h with shaking at 130 rpm. The library was screened with at least 3× coverage by plating 50 000 to 100 000 CFU per plate. Colonies that had formed after 16 h of incubation at 37˚C were re-cultivated under the same conditions to confirm the phenotype. Resistant clones were evaluated by restriction digest and unique clones were retransformed to confirm the presence of the resistance phenotype on the fosmid insert.

***De novo* sequencing of pCC2FOS-AmpR3.** Extracted DNA was used to generate Illumina shotgun paired-end sequencing libraries, which were sequenced with a MiSeq instrument and the MiSeq reagent kit version 3, as recommended by the manufacturer (Illumina, California, USA). Quality filtering using Trimmomatic version 0.36 (Bolger, Lohse and Usadel 2014) resulted in paired-end reads with an average read length of 301 bp. The assembly was performed with the SPAdes genome assembler software version 3.10.0 (Bankevich *et al.* 2012), resulting in a 50.2 kb contig with a 9.2-fold coverage. The assembly was validated, and the read coverage determined with QualiMap v2.1 (García-Alcalde *et al.* 2012). Automatic gene prediction was performed using the software tool Prokka v1.12 (Seemann 2014).

Phylogenetic analysis of *blaMm3*. The phylogenetic analysis for inferring the evolutionary relationship of Mm3 with other β-lactamases was conducted using the software MEGA X v10.0.2(Kumar *et al.* 2018). The amino acid sequences were aligned using MUSCLE (Edgar 2004) and the tree was constructed by the neighbour-joining method (Saitou and Nei 1987)

with a bootstrap test of 2000 replicates, using the p-distance method (Masatoshi Nei and Sudhir Kumar 2000) for computing evolutionary distances.

Minimum inhibitory concentrations. Minimum inhibitory concentrations (MICs) were determined according to the guidelines of the European Committee for Antimicrobial Susceptibility Testing using the broth microdilution method(European Committee for Antimicrobial Susceptibility Testing (EUCAST) of the European Society of Clinical Microbiology and Infectious Diseases (ESCMID) 2003). The assays were conducted in triplicate. MICs for the functional metagenomics screening were determined using the empty vector library host *E. coli* EPI300 pCC2FOS. MICs for the ampicillin resistant clones were determined for ampicillin and for *E. coli* EPI300 pCC2FOS-Mm3 additionally for cefotaxime, cephalothin, cephalexin, carbenicillin (Sigma-Aldrich, Germany) using *E. coli* EPI300 pCC2FOS as control strain.

Subcloning *blaMm3*. The *blaMm3* gene was cloned into the pET28a(+) expression vector (Novagen, USA) with N-terminal His Tag and inducible T7 promoter using the NdeI and EcoRI restriction sites. With primers comprising the respective restrictions sites (underlined) (F: 5'-3' TGCAGACATATGAACCCCAACCACTCTG, R: 5'-3' TACTAGAATTCCTAGACGCTCGAT GTCGCC, Sigma-Aldrich, Germany), the full ORF was amplified from pCC2FOS-Mm3 by a standard PCR reaction using the Phusion DNA polymerase (New England BioLabs, Germany) at 72°C annealing temperature. The vector ligated gene was transformed into high efficiency *E. coli* DH5α (New England BioLabs) for selection of the recombinant pET28a-*blaMm3* plasmid, which was then introduced into *E. coli* BL21(DE3) (Thermo Scientific, Germany) for overexpression.

Expression and purification of the β-lactamase Mm3. LB broth (400 ml) with kanamycin (50 µg ml^{-1}) was inoculated with 2% (v/v) of an overnight culture of *E. coli* BL21(DE3) pET28a-*blaMm3* which was then grown at 37°C under shaking at 130 rpm to an OD_{600} of 0.8. The culture was supplemented with isopropyl β-D-1-thiogalactopyranoside to 0.4 mM end concentration and further incubated for 4 hours. Harvested cells were resuspended in 50 ml binding buffer (20 mM sodium phosphate buffer, 500 mM NaCl, 20 mM imidazole, pH 7.4) containing 0.8 g l^{-1} lysozyme and disrupted by sonication with a digital sonifier (pulses of 2 s and 4 s pause, 5 min, 70% amplitude; Branson, Emerson, Missouri, USA). The His-tagged protein was isolated from the centrifuged lysate (12000 x g, 10 min) using a 1-ml HisTrap column (GE health Care, Illinois, USA) and an elution gradient (1 ml min^{-1}, 20 min) up to 500 mM imidazole. Two fractions containing active β-lactamase, as judged by application on nitrocefin disks (Sigma Aldrich, Germany) and SDS-PAGE were mixed together. The buffer was exchanged with 20 mM sodium phosphate buffer (200 mM NaCl, pH 7.4) through multiple dilution and centrifugation steps using Amicon Ultra-15 centrifugal filters (10 kDa cut off, Merck Millipore, Germany). Protein purity was estimated with SDS-PAGE and the concentration determined with the Pierce BCA assay kit (Thermo Scientific) using bovine serum albumin as reference. The purified enzyme was shock-frozen in liquid nitrogen and stored at -70°C.

Kinetic characterisation of the β-lactamase Mm3. To determine kinetic values (V_{max} and K_m) the activity of the purified β-lactamase Mm3 was measured spectrophotometrically (U-2001, Hitachi, Tokyo, Japan). Initial hydrolysis rates for ampicillin and carbenicillin were recorded at 235 nm and 30°C in 450 µl reaction buffer (20 mM sodium phosphate buffer, 0.2 M NaCl, pH 7.4) upon addition of 50 µl substrate at different concentrations (1 mM to 100 mM in

H_2O). The kinetic data was fitted for ampicillin with the Hill and for carbenicillin with the Michaelis-Menten equation, respectively, using the software Origin 9.0.0G (Origin Lab Corporation, Massachusetts, USA).

Data accessibility. The complete *S. magellanicum* metagenome is stored at the MG-RAST server under the accession no. 4533611.3. The nucleotide sequence of the β-lactamase *blaMm3* is deposited in Genbank under the accession no. MK831000 and the 16S rRNA sequences from the multi-resistant bacteria under the accession numbers MK801238-MK801243.

Acknowledgements

We thank Christian Berg for sampling of *S. magellanicum,* and Stephanie Hollauf, Franz Stocker and Angelika Schäfer for helping with the preparation of moss samples for downstream processing, Harald Blasl and Bettina Semler for their help with the screenings and Silvia Ferrario for helping with subcloning and protein expression. Special thanks go to Isabella Wrolli for her indispensable technical help throughout the entire study as well as to Henry Müller (all Graz) for his technical guidance. This work was supported by the Federal Ministry of Science, Research and Economy (BMWFW), the Federal Ministry of Traffic, Innovation and Technology (bmvit), the Styrian Business Promotion Agency SFG, the Standortagentur Tirol, the Government of Lower Austria and ZIT - Technology Agency of the City of Vienna through the COMET-Funding Program managed by the Austrian Research Promotion Agency FFG.

Contributions

M.-M.O., C.A.M. and G.B. conceived the study. M.-M.O. and C.A.M. designed and performed the experiments. M.-M.O, J.T., T.C. and A.B. analysed and/or interpreted the data. A.P. performed *de novo* sequencing and annotation. M.-M.O., C.A.M., and G.B. wrote the paper. All authors revised the manuscript and approved the final version.

6 Neues aus der Mikrobiom-Forschung: Von verbesserter Wirkstoffsuche, neuen Therapieansätzen und Resistenz-Management

Melanie-Maria Obermeier[1,2], Christina Andrea Müller Bogotá[1,2], and Gabriele Berg[1,2]

[1] *Institute of Environmental Biotechnology, Graz University of Technology, Graz, Austria;*

[2] *Austrian Centre of Industrial Biotechnology, Graz, Austria*

Zusammenfassung

Mikrobiom-Forschung erlebte in den letzten zwanzig Jahren einen enormen Boom. Dabei schuf dieses Forschungsfeld eine komplett neue Sichtweise auf Mikroorganismen und deren Bedeutung für Mensch und Natur. Mit weitreichenden Folgen. Die gewonnenen Erkenntnisse und verwendeten Methoden eröffnen innovative Perspektiven für die Nutzung von Mikroorganismen und den Umgang mit diesen. Ein Potential, das die Pharmazie in vielen Bereichen signifikant verändern kann. Dies betrifft insbesondere neue Möglichkeiten der pharmazeutischen Wertschöpfung von Mikroorganismen, wie der Entwicklung neuer Antibiotika, Biologicals und Mikrobiom-basierter Therapien. Über die pharmazeutische Mikrobiologie hinaus wird ein verbessertes Management von Keimen und antimikrobiellen Resistenzen ermöglicht.

MANUSKRIPT

Verfasst für Pharmind, Editio Cantor Verlag

Einleitung

Das Mikrobiom. Seit zwei Jahrzehnten ein wichtiger Fachbereich innerhalb der Mikrobiologie, erlangte in den letzten Jahren auch einen Bekanntheitsgrad in der Öffentlichkeit. Dort prägen vor allem Berichte über eine gesunde Darmflora, als wichtiges Kriterium für Gesundheit und Wohlbefinden, eine veränderte Wahrnehmung gegenüber Mikroorganismen. Diese werden heutzutage nicht mehr lediglich als Krankheitserreger angesehen und gefürchtet, sondern mittels Prä- und Probiotika gehegt und gepflegt. Darmmikroben sind allerdings nur ein Teil dessen was als Mikrobiom beschrieben wird. Was also versteht man nun unter dem Mikrobiom genau? Es handelt sich dabei um die komplexe Gemeinschaft aller Mikroorganismen in einem Habitat, welches zum Beispiel der Mensch als Ganzes oder der Darm im Speziellen sein kann, aber auch eine Pflanze oder Bodenprobe. Zusammen mit den Habitat-spezifischen Umweltbedingungen sind alle zellulären Elemente der vorkommenden Mikroorganismen wie DNA, Proteine, Lipide und die von diesen produzierten Metaboliten ebenfalls Bestandteil des Mikrobioms (Abbildung 1) (Berg unpublished).

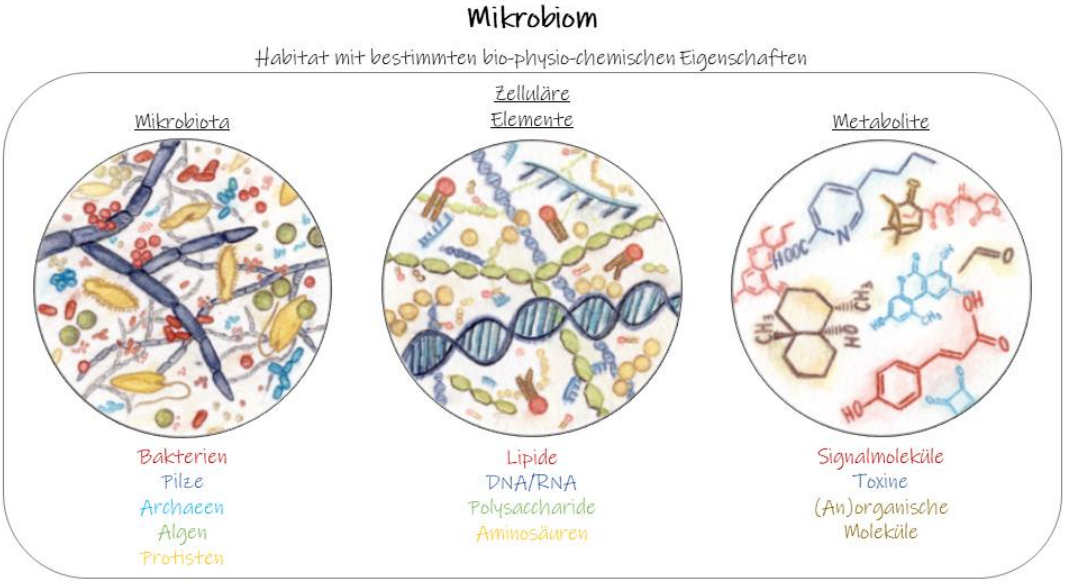
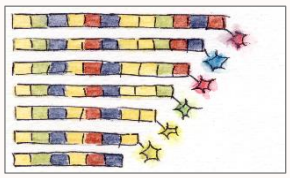


Abbildung 6.1: Das Mikrobiom. Die Bakterien, Pilze, Archaeen, Algen und Protisten (=Mikrobiota) in einem bestimmten Habitat bilden zusammen mit ihren zellulären Elementen (Lipide, DNA, RNA, Polysaccharide, Aminosäuren) und die von ihnen produzierten Metabolite (Signalmoleküle, Toxine, (An)organische Moleküle) das Mikrobiom. Weiterer Bestandteil sind die Habitat-spezifischen biologischen, physikalischen und chemischen Eigenschaften.

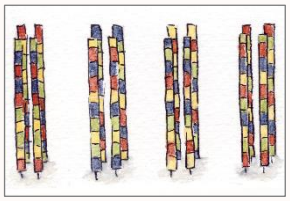
Ein komplexes Zusammenwirken mehrerer Faktoren, das eine wesentliche Rolle für die Gesundheit von Mensch, Tier und Pflanze, aber auch für die Integrität von Böden spielt. Die Erkenntnis darüber welche große Bedeutung diese kleinen Organismen innehaben, schnellte vor allem in den letzten zwanzig Jahren empor. Maßgeblich verantwortlich dafür war die rasante Entwicklung auf dem Gebiet der Sequenzierungstechnologien (Box 1). Neue, immer bessere Sequenziermethoden mit höheren Durchsatzraten und den damit einhergehenden, sinkenden Kosten führten zu einem regelrechten Boom auf dem Gebiet der Mikrobiom-Forschung. Das genetische Material ganzer mikrobieller Gemeinschaften kann heutzutage simultan entschlüsselt und studiert werden. Dies gibt Aufschluss darüber, welche Arten vorkommen und welche Gene und damit Funktionen vorhanden sind. Während Metagenomik fortwährend Schlüsseldisziplin ist, operiert moderne Mikrobiom-Forschung auf der gesamten Bandbreite der Omics-Technologien, wie Metatranskriptomik, Metaproteomik und Metabolomik (Box 2). Erst in ihrer Kombination kann ein umfassendes Verständnis über die mikrobiellen Aktivitäten in einem Habitat und die Rolle der Mikroorganismen für dieses erlangt werden.

BOX 1 Sequenziertechnologien
Um die Basenabfolge von DNA zu bestimmen, wurden seit den 1980er Jahren drei Generationen an Sequenziertechnologien entwickelt.
(rezensiert in Heather and Chain 2016).

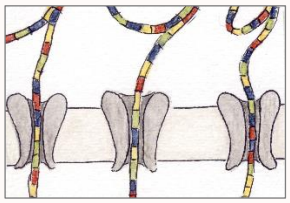
Sangersequenzierung. Hierbei wird die DNA zunächst mittels PCR vervielfältigt. Es folgt ein zweiter PCR-Schritt, bei welchem unterschiedlich lange DNA-Fragmente generiert werden. Dies geschieht durch das zufällige Integrieren von speziell zugegebenen, modifizierten Basen. Diese Basen beenden nicht nur die Duplikation des jeweiligen DNA-Moleküls, sondern emittieren zusätzlich ein für die Base spezifisches Lichtsignal. Nachdem die DNA-Fragmente ihrer Größe nach aufgetrennt worden sind, wird das Lichtsignal detektiert und so die DNA-Sequenz bestimmt.



Sequenzierung der nächsten Generation. Die darunterfallenden Methoden, Illumina, Roche 454 und Ion Torrent, beruhen alle auf dem gleichen Prinzip. Sie arbeiten mit vervielfältigten DNA-Fragmenten, welche parallel sequenziert werden. Schrittweise werden dabei die Basen nacheinander zugegeben und zwischen den einzelnen Schritten weggewaschen. Der Einbau der Basen wird in Echtzeit verfolgt, entweder durch Lichtsignale, die beim Integrieren der Basen emittiert werden oder durch das Aufzeichnen der beim Einbau entstehenden Ionenschwankung. Durch das gleichzeitige Sequenzieren hunderter DNA-Fragmente, schaffen diese Methoden einen enormen Durchsatz.



Sequenzierung der dritten Generation. Die von Pacific Bioscience and Oxford Nanopore entwickelten Technologien zeichnen sich dadurch aus, dass es keiner aufwendigen, vorherigen DNA-Vervielfältigung oder -Bearbeitung bedarf. PacBio und Nanopore Technologien arbeiten mit einem einzigen DNA-Molekül. Das parallele Sequenzieren multipler Einzel-DNAs ist zentraler Bestandteil. Die Replikation der einzelnen DNA-Stücke wird dabei wieder durch Lichtemission oder mittels Ionenstärke in Echtzeit verfolgt; mit dem Unterschied, dass die Basen nicht mehr schrittweise zugegeben und entfernt werden, was den Sequenziervorgang zusätzlich beschleunigt.



Neue Wege für die Wirkstoffsuche

Omics-Technologien, die den Mikrobiomforschern dazu dienen das Wechselspiel zwischen Mikrobiom und Wirt auf genetischer, enzymatischer und metabolischer Ebene zu verstehen, können der Pharmazie helfen neue Angriffspunkte für Antibiotika zu identifizieren (Pulido *et al.* 2016). Mittels (Meta-)Transkriptomik und Metabolomik lässt sich zum Beispiel herausfinden welche bakteriellen Gene und Metabolite während Infektionen besonders stark aktiviert sind bzw. produziert werden. Genau diese Gene und Metabolite könnten eine essentielle Rolle für die Infektiosität spielen und stellen damit ein vielversprechendes Angriffsziel für neue Antibiotika dar.

Darüber hinaus ermöglichen Omics-Technologien eine zielgerichtete Wirkstoffsuche indem das metabolische Potential von Mikrobiomen analysiert werden kann. Durch den gewonnenen Einblick in das für eine Probe spezifische Repertoire an produzierten Metaboliten, Peptiden und Enzymen mittels Metabolomik und Metaproteomik können Bioressourcen entsprechend ihres charakteristischen Potentials genutzt werden. Die Quantität einzelner Stoffe kann ebenfalls ermittelt werden und gestattet damit eine Abschätzung über die Erfolgsrate eben diese Stoffe aus der komplexen Mischung an zellulären Produkten zu isolieren.

Durch die Kombination verschiedener Omics-Technologien lässt sich das Mikrobiom zusätzlich im ökologischen Kontext studieren. Solche Analysen geben wertvollen Aufschluss darüber, wann Mikroorganismen welche Gene anschalten. Wissen, welches in der Pharmazie genutzt werden kann, um durch Simulation der nötigen Bedingungen im Labor die gewünschten Gene anzuwerfen; denn oft liegt eine Diskrepanz in der Genaktivität zwischen Labor und dem ursprünglichen Habitat vor. Zusammen mit der Anwendung von neuen Kultivierungsmethoden, steigen so die Chancen neue Naturstoffproduzenten zu finden und wirken damit einem großen Problem der Naturstoffsuche entgegen: Der niedrigen Erfolgsrate.

BOX 2 Omics-Technolgien.
(rezensiert in Aguiar-Pulido *et al.* 2016)

Metagenomik bezeichnet das Arbeiten mit DNA einer komplexen Gemeinschaft von Mikroorganismen mittels gen-technischer und bioinformatischer Methoden. Dazu wird DNA aus den sich in der Probe befindlichen Mikro-organismen direkt isoliert, ohne diese vorher zu kultivieren. Damit verschafft Metagenomik Zugang zum genetischen Material von bisher nicht kultivierbaren Mikroorganismen, welche über 90% der Vielfalt betragen. Mit der isolierten DNA gewinnt man Erkenntnis darüber welche Spezies vorhanden sind, welches genetische Potential und welche Funktionen diese besitzen.

Metatranskriptomik beschäftigt sich entgegen der Metagenomik nicht mit der DNA eines Mikrobioms, sondern mit dessen RNA; einem DNA-Transkript. Da nur aktive Gene in RNA überschrieben werden, liefert Metatranskriptomik ein Aktivitätsprofil. Anhand dieses Profils lässt sich der Zusammenhang zwischen Genaktivität und Umweltfaktoren studieren – anders gesagt, wie Mikro-organismen unter bestimmte Beding-ungen agieren. Denn es sind nie alle Gene zugleich aktiv. Durch die vorherrschenden Umwelteinflüsse wer-den einzelne Gene an- und ab-geschaltet.

Metaproteomik ist das Studium der Proteine, die zu einem bestimmten Zeitpunkt und unter bestimmten Bedingungen von einer mikrobiellen Gemeinschaft produziert werden. Diese werden ebenfalls direkt aus einer Probe isoliert und anhand von Massen-spektrometrie analysiert und per Datenbasenabgleich identifiziert. Dies erlaubt unter anderem einen Rück-schluss auf die enzyma-tischen und metabolischen Aktivitäten innerhalb eines Mikrobioms.

Metabolomik befasst sich mit der Analyse der metabolischen Aktivität von mikrobiellen Gemeinschaften. Dabei werden alle zu einer bestimmten Zeit, unter bestimmten Bedingungen produ-zierten Metabolite erfasst. Solche Analysen geben Aufschluss darüber wie Mikrobiome auf ihre Umwelt reagieren, aber vor allem wie sie mit dieser wechselwirken. Damit lässt sich ver-stehen wie das Mikrobiom mit seinem Wirt bzw. Habitat interagiert. Im Gegensatz zu Rückschlüssen auf das metabolische Repertoire eines Mikro-bioms basierend auf Metaproteomik, können hierbei auch bisher nicht bekannte Metabolite erfasst werden.

Naturstoffe spielen bis heute eine essentielle Rolle in der Pharmazie; vor allem als „lead compounds“ für die Entwicklung neuer Pharmazeutika. Unter diesen machen Naturstoffe und deren Derivate einen beträchtlichen Anteil aus. Der Naturstoffs Campto-thecin fungierte zum Beispiel als „drug lead“ für die Entwicklung der Antikrebsmedikamente Campto und Hycamtin. Als eines der erfolgreichsten Arzneistoffe innerhalb der Krebstherapie gilt fortwährend Taxol; ein von Pflanzen und Mikroorganismen gebildetes Alkaloid. Die geringen Erfolgschancen haben Naturstoffsuche jedoch zu einer nur noch selten verfolgten Disziplin werden lassen. Dabei ist erst ein Bruchteil der von Mikroorganismen stammenden Naturstoffe identifiziert. Dies liegt mitunter an der sogenannten Plate-Count-Anomalie; der Tatsache, dass gerade einmal 5% aller Bakterien mit heutigen Methoden kultivierbar sind.

Um das Kultivierungsproblem zu umgehen entwick-elten Forscher in den 80er Jahren das Konzept der funktionellen Metagenomik, welches seither erfolg-reich Zugang zu ungenutzten Ressourcen ermöglicht. Das gesamte genetische Material einer mikrobiellen Gemeinschaft, das Metagenom, wird dazu direkt aus einer komplexen Probe isoliert. In Form kleiner Fragmente wird die metagenomische DNA mittels eines Carriers in ein Wirtsbakterium kloniert und so eine Sammlung von Tausenden von Klonen generiert. Aufgrund der großen Zahl an Klonen, von denen nur ein geringer Teil die relevante metagenomische DNA trägt und somit den gewünschten Effekt aufzeigt, werden diese mittels Hochdurchsatz-Screenings durch-gemustert. Dabei kann nach alternativen Genvarianten, sowie nach neuen Genen, Enzymen und Metaboliten gesucht werden. Erstere lassen sich anhand der für ihre Genklasse charakteristischen Gensequenzen in einem PCR- (Polymerase Ketten Reaktion) Screening identifizieren. Für letzteres werden Screening-methoden entwickelt, bei denen Klone auf eine bestimmte Aktivität untersucht werden. Oft werden diese Screenings mit leicht auszuwertenden Methoden kombiniert. Zum Beispiel durch die Zugabe von Substraten, die eine sichtbare visuelle Reaktion auslösen, sobald ein Klon die gewünschte Aktivität ausübt. Wirtsbakterien können zudem für ein Screening speziell optimiert werden. Vor allem das Einbringen von Reportergenen erleichtert dabei nicht nur die Identifikation, sondern auch das Durchmustern der vielzähligen Klone. Löst das Reportergen Fluores-zenz aus, kann die Auswertung und Sortierung der Klonsammlung maschinell mittels Fluoreszenz-basierter Zellsortierungs-geräte hocheffizient durch-geführt werden.

Viele bereits generierte Klonbibliotheken lagern in den Tiefkühlern von Universitäten, wie der Rockefeller University, der Georg-August-Universität Göttingen oder der Technischen Universität Graz. In Zusammenarbeit könnten diese auf spezielle Meta-bolite und Enzyme durchgemustert werden, denn wenn einmal erfolgreich generiert, kann eine Klonsammlung generell verschiedenen Screenings unterzogen werden. Institutionen wie das österreichische Kompetenzzentrum acib (Austrian Centre of Industrial Biotechnology) fungieren dabei erfolgreich als Vermittler und schaffen seit Jahren vielver-sprechende Kooperationen zwischen Universitäten und Industrie. Vor allem da die Erstellung einer qualitativ hochwertigen Klonsammlung immer noch einen limitierenden Faktor darstellt, kann so zum Vorteil beider Seiten das an den Universitäten schlummernde Potential ausgeschöpft werden.

Mikrobiom-informierte und –basierte Pharmazie

Omics-Technologien und Hochdurchsatz-Screenings kommen nicht nur bei der Identifikation neuartiger Produzenten, Gene und Wirkstoffe zum Einsatz. Forscher kombinieren diese ebenfalls, um das Potential Mikrobiom-basierter Krankheitstherapien zu studieren. Zum Beispiel im Rahmen des künstlichen Darmsystems. Bestehend aus kleinen Bioreaktoren ermöglicht dieses das kontrollierte Arbeiten mit der Darmflora. Diese spielt eine zentrale Rolle innerhalb der Mikrobiom-basierten Medizin, da es einen Dreh- und Angelpunkt für eine Vielzahl an Erkrankungen darstellt.

Dickdarmkrebs und entzündliche Darmerkrankungen wie Morbus Crohn stehen mit dem Darmmikrobiom in Zusammenhang. Kennzeichnend für diese Erkrankungen ist eine deregulierte Darmflora, die sich neben der Dominanz von einigen wenigen Spezies vor allem durch eine reduzierte Artenvielfalt auszeichnet. Die Wiederherstellung eines gesunden Darmmikrobioms als Behandlungstherapie besitzt erfolgreiche Heilungschancen, wie Stuhltransplantationen bei an Morbus Crohn erkrankten Patienten oder bei *Clostridium difficile* Infektionen zeigen. Forscher arbeiten daher nicht nur an der Identifikation jener Spezies, die innerhalb der Arten-armen Mikrobiome aufblühen und Krankheiten verursachen. Sie identifizieren mehr und mehr Gesundheitsfördernde Mikroorganismen, die Krankheitssymptomen entgegenwirken. Sogenannte Pharmabiotika - Probiotika mit therapeutischer Wirkung (Lee *et al.* 2018). Diese wurden bereits mit vielversprechenden Ergebnissen zur vorbeugenden und heilenden Behandlung im Darm von Patienten gegen Antibiotika-assoziierte Diarrhö und das Reizdarmsyndrom angereichert. Das Potential der therapeutischen Behandlung des Darmmikrobioms ist allerdings weitaus größer. Durch unter anderem der Interaktion mit dem Immun- und Hormonsystem reicht dessen Einfluss über den Wirkort Darm hinaus. So wurde das Darmmikrobiom mit Fettleibigkeit, Diabetes und Depression in Verbindung gebracht und rückt Pharmabiotika damit als Mittel gegen diese Krankheiten in den Fokus.

Pharmabiotika könnten zusätzlich als Begleittherapie zur verbesserten Wirkeffizienz von Medikamenten Anwendung finden. Denn die Vielzahl an verschiedenen Spezies, die uns besiedeln, nimmt direkt und indirekt Einfluss auf die Prozesse im Körper und damit auf die medikamentöse Wirkung (Doestzada *et al.* 2018). Dabei beeinflusst das Mikrobiom die biologische Verfügbarkeit, wie auch die Aktivität von Medikamenten, in dem diese verstoffwechselt werden oder das chemische Milieu am Wirkort verändert wird. Vor allem aber für die individuelle Therapie besitzen Pharmabiotika großes Potential. Kein Mikrobiom gleicht nämlich dem Anderen. Durch diese individuelle Zusammensetzung sprechen einzelne Patienten unterschiedlich auf Behandlungen an, so dass sich das Mikrobiom, neben dem Genom und der Lebensweise als essentielles Kriterium für die personalisierte Medizin einreiht. Ist die Anzahl bestimmter Spezies, welche für die Wirkeffizienz von Medikamenten

wichtig sind, zu gering oder fehlen solche Mikroorganismen ganz, könnten diese als Pharmabiotika ergänzend verabreicht werden.

Ein weiterer Aspekt der Pharmakokinetik, die auf das Mikrobiom zurückzuführen ist, ist dessen Einwirken auf die Toxizität von Medikamenten. Aber auch umgekehrt haben Studien gezeigt, dass medikamentöse Behandlung einen Einfluss auf das mikrobielle Gleichgewicht nehmen kann. Am deutlichsten ist dieser Effekt bei Antibiotika (Ferrer *et al.* 2017), wurde aber bereits bei anderen Medikamenten beobachtet, wie gegen gastroösophageale Refluxerkrankungen verabreichte Protonenpumpeninhibitoren oder dem bei Diabetes Typ II verwendeten Metformin (Doestzada *et al.* 2018). Mit den Implikationen des Mikrobioms für Gesundheit und Krankheit kann das beeinflusste Patienten-Mikrobiom im weiteren Sinne zu Nebenwirkungen führen. Mikrobiom-verursachte und -basierte Nebenwirkungen könnten damit in Zukunft zu einem wichtigen Kriterium für die Verabreichung bzw. Zulassung von Medikamenten werden. Das künstliche Darmsystem, wie anfangs beschrieben, stellt hierbei eine hilfreiche Methodik für die Aufklärung solcher Effekte dar.

Resistenzmanagement

Antibiotikaresistenzen sind so allgegenwärtig wie ihre Träger. Während diese in ihrer natürlichen Umwelt in metabolische Prozesse involviert sind und der Zell-Zell-Kommunikation dienen, sind sie in Kliniken, sowie der Agrar- und Pharmaindustrie ein fortwährend wachsendes Problem. Dort stellt es eine immer größer werdende Herausforderung dar, allgemeine aber vor allem Antibiotika-resistenten Bakterien und Keimen durch Reinigungsmittel und antibakterielle Substanzen dauerhaft Einhalt zu gebieten. Dabei zeigen neueste Studien, dass das Reinigungsverhalten vor allem zur Reduktion der bakteriellen Vielfalt führt, aber die Einzeller nicht komplett von der Bildfläche verschwinden lässt. Die überlebenden Bakterien breiten sich verstärkt in der menschlich geschaffenen, Artenarmen Umgebung aus. Der regelmäßige Kontakt mit Reinigungsmitteln und antibakteriellen Substanzen selektiert folglich speziell wettkampfstarke Bakterien. Dabei sind diese zu einem höheren Anteil mit Virulenz, Krankheit und Antibiotikaresistenzen assoziiert (Mahnert *et al.* 2019).

Ein überwachtes Management von Keimen und Antibiotikaresistenzen gewinnt damit mehr und mehr an Bedeutung. Um entsprechende Strategien zu entwickeln, evaluieren Forscher mittels Omics-Technologien das Vorkommen und die Häufigkeit von Resistenzen auf Mikrobiomebene. Dieser Einblick erlaubt die Verbreitung von einzelnen Resistenzen zu verstehen und deren Risiken abzuwägen. Zudem lässt sich daraus erschließen, welche Antibiotikabehandlungen und Reinigungsstrategien die größeren Erfolgschancen versprechen. Analysen, wie sie zur Entwicklung und Überwachung von Resistenzmanagementstrategien oder zur Abschätzung von deren Wirksamkeit in der pharmazeutischen Industrie herangezogen werden können.

Fazit

Omics-Technologien ermöglichen einen enormen und rasanten Fortschritt in der Mikrobiologie und sind aus dieser nicht mehr wegzudenken. Der Informationsgehalt, den diese Technologien vor allem in ihrer Kombination liefern, ist allerdings nicht nur innerhalb der Mikrobiom-Forschung von großer Bedeutung. Durch ihr weitreichendes Potential haben diese bereits Fuß in der Start-Up Szene gefasst. Es ist also nur eine Frage der Zeit bis Omics-Technologien in der Industrie eine ebenso breite Anwendung finden wie in der Akademie.

References

Aakvik T, Degnes KF, Dahlsrud R *et al.* A plasmid RK2-based broad-host-range cloning vector useful for transfer of metagenomic libraries to a variety of bacterial species. *FEMS Microbiol Lett* 2009;**296**:149–58.

Abbasian F, Lockington R, Megharaj M *et al.* The integration of sequencing and bioinformatics in metagenomics. *Rev Environ Sci Bio/Technology* 2015;**14**:357–83.

Abdalla MA, Matasyoh JC. Endophytes as producers of peptides: an overview about the recently discovered peptides from endophytic microbes. *Nat Products Bioprospect* 2014;**4**:257–70.

Aenishaenslin C, Häsler B, Ravel A *et al.* Evidence needed for antimicrobial resistance surveillance systems. *Bull World Health Organ* 2019;**97**:283–9.

Aguiar-Pulido V, Huang W, Suarez-Ulloa V *et al.* Metagenomics, metatranscriptomics, and metabolomics approaches for microbiome analysis. *Evol Bioinform Online* 2016;**12**:5–16.

Akinsanya MA, Goh JK, Lim SP *et al.* Metagenomics study of endophytic bacteria in *Aloe vera* using next-generation technology. *Genomics Data* 2015;**6**:159–63.

Alegria TGP, Meireles DA, Cussiol JRR *et al.* Ohr plays a central role in bacterial responses against fatty acid hydroperoxides and peroxynitrite. *Proc Natl Acad Sci* 2017;**114**:E132–41.

Allen HK, Donato J, Wang HH *et al.* Call of the wild: antibiotic resistance genes in natural environments. *Nat Rev Microbiol* 2010;**8**:251–9.

Allen HK, Moe LA, Rodbumrer J *et al.* Functional metagenomics reveals diverse β-lactamases in a remote Alaskan soil. *ISME J* 2009;**3**:243–51.

Altschul SF, Gish W, Miller W *et al.* Basic local alignment search tool. *J Mol Biol* 1990;**215**:403–10.

Aly AH, Debbab A, Proksch P. Fungal endophytes – secret producers of bioactive plant metabolites. *Die Pharm - An Int J Pharm Sci* 2013;**68**:499–505.

Ambler RP. The structure of β-lactamases. *Philos Trans R Soc B Biol Sci* 1980;**289**:321–31.

Atamna-Ismaeel N, Finkel O, Glaser F *et al.* Bacterial anoxygenic photosynthesis on plant leaf surfaces. *Environ Microbiol Rep* 2012a;**4**:209–16.

Atamna-Ismaeel N, Finkel OM, Glaser F *et al.* Microbial rhodopsins on leaf surfaces of terrestrial plants. *Environ Microbiol* 2012b;**14**:140–6.

Atichartpongkul S, Fuangthong M, Vattanaviboon P *et al.* Analyses of the regulatory mechanism and physiological roles of *Pseudomonas aeruginosa* OhrR, a transcription regulator and a sensor of organic hydroperoxides. *J Bacteriol* 2010;**192**:2093–101.

Atichartpongkul S, Loprasert S, Vattanaviboon P *et al.* Bacterial Ohr and OsmC paralogues define two protein families with distinct functions and patterns of expression. 2001;**147**:1775–82.

Bai Y, Müller DB, Srinivas G *et al.* Functional overlap of the *Arabidopsis* leaf and root microbiota. *Nature* 2015;**528**:364–9.

Bailly A, Weisskopf L. The modulating effect of bacterial volatiles on plant growth. *Plant Signal Behav* 2012;**7**:79–85.

Baltz R. Antimicrobials from actinomycetes: back to the future. *Microbe* 2007;**2**:125–31.

Banerjee J, Mishra N, Dhas Y. Metagenomics: a new horizon in cancer research. *Meta Gene* 2015;**5**:84–9.

Bankevich A, Nurk S, Antipov D *et al.* SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 2012;**19**:455–77.

Bauernfeind A, Stemplinger I, Jungwirth R *et al.* Sequences of β-lactamase genes encoding CTX-M-1 (MEN-1) and CTX-M-2 and relationship of their amino acid sequences with those of other β-lactamases. *Antimicrob Agents Chemother* 1996;**40**:509–13.

Belova SE, Pankratov TA, Dedysh SN. Bacteria of the genus *Burkholderia* as a typical component of the microbial community of *Sphagnum* peat bogs. *Microbiology* 2006;**75**:90–6.

Bengtsson-Palme J, Kristiansson E, Larsson DGJ. Environmental factors influencing the development and spread of antibiotic resistance. *FEMS Microbiol Rev* 2018;**42**:68–80.

Bérdy J. Thoughts and facts about antibiotics: where we are now and where we are heading. *J Antibiot (Tokyo)* 2012;**65**:385–95.

Berendonk TU, Manaia CM, Merlin C *et al.* Tackling antibiotic resistance: the environmental framework. *Nat Rev Microbiol* 2015;**13**:310–7.

Berg G. Plant-microbe interactions promoting plant growth and health: perspectives for controlled use of microorganisms in agriculture. *Appl Microbiol Biotechnol* 2009;**84**:11–8.

Berg G, Eberl L, Hartmann A. The rhizosphere as a reservoir for opportunistic human pathogenic bacteria. *Environ Microbiol* 2005;**7**:1673–85.

Berg G, Grube M, Schloter M *et al.* Unraveling the plant microbiome: looking back and future perspectives. *Front Microbiol* 2014;**5**:1–7.

Berg G, Ryboakocva D, Fischer D, *et al.* Microbiome definition re-visited: old concepts and new recommendations. Unpublished

Berg G, Rybakova D, Grube M *et al.* The plant microbiome explored: implications for experimental botany. *J Exp Bot* 2016;**67**:995–1002.

Berg G, Smalla K. Plant species and soil type cooperatively shape the structure and function of microbial communities in the rhizosphere. *FEMS Microbiol Ecol* 2009;**68**:1–13.

Berg G, Zachow C, Müller H *et al.* Next-generation bio-products sowing the seeds of success for sustainable agriculture. *Agronomy* 2013;**3**:648–56.

Beyer HM, Gonschorek P, Samodelov SL *et al.* AQUA cloning: a versatile enzyme-free cloning approach. Chatterji D (ed.). *PLoS One* 2015;**10**:e0137652.

Bhullar K, Waglechner N, Pawlowski A *et al.* Antibiotic resistance is prevalent in an isolated cave microbiome. Aziz RK (ed.). *PLoS One* 2012;**7**:e34953.

Biteau B, Labarre J, Toledano MB. ATP-dependent reduction of cysteine–sulphinic acid by *S. cerevisiae* sulphiredoxin. *Nature* 2003;**425**:980–4.

Bitok JK, Lemetre C, Ternei MA *et al.* Identification of biosynthetic gene clusters from metagenomic libraries using PPTase complementation in a *Streptomyces* host. *FEMS Microbiol Lett* 2017;**364**, DOI: 10.1093/femsle/fnx155.

Blaser MJ. The microbiome revolution. *J Clin Invest* 2014;**124**:4162–5.

Blau K, Jacquiod S, Sørensen SJ *et al.* Manure and doxycycline affect the bacterial community and its resistome in lettuce rhizosphere and bulk soil. *Front Microbiol* 2019;**10**:725.

Blom D, Fabbri C, Eberl L *et al.* Volatile-mediated killing of *Arabidopsis thaliana* by bacteria is mainly due to hydrogen cyanide. *Appl Environ Microbiol* 2011;**77**:1000–8.

Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 2014;**30**:2114–20.

Borel JF, Feurer C, Gubler HU *et al.* Biological effects of cyclosporin A: a new antilymphocytic agent. *Agents Actions* 1976;**6**:468–75.

Brader G, Compant S, Mitter B *et al.* Metabolic potential of endophytic bacteria. *Curr Opin Biotechnol* 2014;**27**:30–7.

Bragina A, Berg C, Berg G. The core microbiome bonds the Alpine bog vegetation to a transkingdom metacommunity. *Mol Ecol* 2015;**24**:4795–807.

Bragina A, Berg C, Cardinale M *et al. Sphagnum* mosses harbour highly specific bacterial diversity during their whole lifecycle. *ISME J* 2012;**6**:802–13.

Bragina A, Cardinale M, Berg C *et al.* Vertical transmission explains the specific *Burkholderia* pattern in *Sphagnum* mosses at multi-geographic scale. *Front Microbiol* 2013;**4**:394.

Bragina A, Maier S, Berg C *et al.* Similar diversity of *Alphaproteobacteria* and nitrogenase gene amplicons on two related *Sphagnum* mosses. *Front Microbiol* 2011;**2**:275.

Bragina A, Oberauner-Wappis L, Zachow C *et al.* The *Sphagnum* microbiome supports bog ecosystem functioning under extreme conditions. *Mol Ecol* 2014;**23**:4498–510.

Brelles-Mariño G, Bedmar EJ. Detection, purification and characterisation of quorum-sensing signal molecules in plant-associated bacteria. *J Biotechnol* 2001;**91**:197–209.

Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. *Nat Methods* 2015;**12**:59–60.

Bulgarelli D, Garrido-Oter R, Münch PC *et al.* Structure and function of the bacterial root microbiota in wild and domesticated barley. *Cell Host Microbe* 2015;**17**:392–403.

Bulgarelli D, Rott M, Schlaeppi K *et al.* Revealing structure and assembly cues for *Arabidopsis* root-inhabiting bacterial microbiota. *Nature* 2012;**488**:91–5.

Bunyapaiboonsri T, Yoiprommarat S, Srikitikulchai P *et al.* Oblongolides from the endophytic fungus Phomopsis sp. BCC 9789. *J Nat Prod* 2010;**73**:55–9.

Bush K. Past and present perspectives on beta-lactamases. *Antimicrob Agents Chemother* 2018;**62**:e01076-18.

Campisano A, Antonielli L, Pancher M *et al.* Bacterial endophytic communities in the grapevine depend on pest management. *PLoS One* 2014;**9**:e112763.

Cane DE, Walsh CT. The parallel and convergent universes of polyketide synthases and nonribosomal peptide synthetases. *Chem Biol* 1999;**6**:R319–25.

Cardinale M. Scanning a microhabitat: plant-microbe interactions revealed by confocal laser microscopy. *Front Microbiol* 2014;**5**:94.

Cardinale M, Grube M, Erlacher A *et al.* Bacterial networks and co-occurrence relationships in the lettuce root microbiota. *Environ Microbiol* 2015;**17**:239–52.

Castaneda-Alvarez C, Prodan S, Rosales IM *et al.* Exoenzymes and metabolites related to the nematicidal effect of rhizobacteria on *Xiphinema index* Thorne & Allen. *J Appl Microbiol* 2016;**120**:413–24.

Cernava T, Aschenbrenner IA, Grube M *et al.* A novel assay for the detection of bioactive volatiles evaluated by screening of lichen-associated bacteria. *Front Microbiol* 2015;**6**:398.

Cernava T, Erlacher A, Soh J *et al. Enterobacteriaceae* dominate the core microbiome and contribute to the resistome of arugula (*Eruca sativa* Mill.). *Microbiome* 2019;**7**:13.

Chen L, Zhang Q-Y, Jia M *et al.* Endophytic fungi with antitumor activities: their occurrence and anticancer compounds. *Crit Rev Microbiol* 2014:1–20.

Chen Q-L, Cui H-L, Su J-Q *et al.* Antibiotic resistomes in plant microbiomes. *Trends Plant Sci* 2019;**24**:530–41.

Choi YW, Choi YW, Hodgkiss IJ *et al.* Enzyme production by endophytes of Brucea javanica. *J Agric TECHNOL* 2005;**1**:55–66.

Clifford DP, Repine JE. Hydrogen peroxide mediated killing of bacteria. 1982;**149**:143–4.

Coenye T, Falsen E, Hoste B *et al.* Description of *Pandoraea* gen. nov. with *Pandoraea apista* sp. nov., *Pandoraea pulmonicola* sp. nov., *Pandoraea pnomenusa* sp. nov., *Pandoraea*

sputorum sp. nov. and *Pandoraea norimbergensis* comb. nov. 2000;**50**:887–99.

Collignon PC, Conly JM, Andremont A *et al.* World Health Organization ranking of antimicrobials according to their importance in human medicine: a critical step for developing risk management strategies to control antimicrobial resistance from food animal production. Griffin PM (ed.). *Clin Infect Dis* 2016;**63**:1087–93.

Cord-Landwehr S, Melcher RLJ, Kolkenbrock S *et al.* A chitin deacetylase from the endophytic fungus *Pestalotiopsis* sp. efficiently inactivates the elicitor activity of chitin oligomers in rice cells. *Sci Rep* 2016;**6**:38018.

Coughlan LM, Cotter PD, Hill C *et al.* Biotechnological applications of functional metagenomics in the food and pharmaceutical industries. *Front Microbiol* 2015;**6**:1–22.

Cove DJ, Perroud P-F, Charron AJ *et al.* The moss *Physcomitrella patens*: a novel model system for plant development and genomic studies. *Cold Spring Harb Protoc* 2009;**2009**:pdb.emo115.

Cowan D, Ramond J-B, Makhalanyane T *et al.* Metagenomics of extreme environments. *Curr Opin Microbiol* 2015;**25**:97–102.

Cox PA. The ethnobotanical approach to drug discovery: strengths and limitations. *Sci Am* 2007;**270**:82–7.

Cragg GM. Paclitaxel (Taxol®):osa success story with valuable lessons for natural product drug discovery and development. *Med Res Rev* 1998;**18**:315–31.

Cretoiu MS, Berini F, Kielak AM *et al.* A novel salt-tolerant chitobiosidase discovered by genetic screening of a metagenomic library derived from chitin-amended agricultural soil. *Appl Microbiol Biotechnol* 2015;**99**:8199–215.

Crosa JH, Walsh CT. Genetics and assembly line enzymology of siderophore biosynthesis in bacteria. *Microbiol Mol Biol Rev* 2002;**66**:223–49.

Crüsemann M, Reher R, Schamari I *et al.* Heterologous expression, biosynthetic studies, and ecological function of the selective Gq-signaling inhibitor FR900359. *Angew Chemie Int Ed* 2018;**57**:836–40.

Cussiol JRR, Alves SV, De Oliveira MA *et al.* Organic hydroperoxide resistance gene encodes a thiol-dependent peroxidase. *J Biol Chem* 2003;**278**:11570–8.

D'Costa VM, King CE, Kalan L *et al.* Antibiotic resistance is ancient. *Nature* 2011;**477**:457–61.

D'Costa VM, McGrann KM, Hughes DW *et al.* Sampling the antibiotic resistome. *Science* 2006;**311**:374–7.

Dahmen S, Mansour W, Charfi K *et al.* Imipenem resistance in *Klebsiella pneumoniae* is associated to the combination of plasmid-mediated CMY-4 AmpC β-lactamase and loss of an outer membrane protein. *Microb Drug Resist* 2012;**18**:479–83.

David B, Wolfender J-L, Dias DA. The pharmaceutical industry and natural products: historical status and new trends. *Phytochem Rev* 2015;**14**:299–315.

Davies J. Small molecules: the lexicon of biodiversity. *J Biotechnol* 2007;**129**:3–5.

Delmotte N, Knief C, Chaffron S *et al.* Community proteogenomics reveals insights into the physiology of phyllosphere bacteria. *Proc Natl Acad Sci U S A* 2009;**106**:16428–33.

Depotter JRL, Deketelaere S, Inderbitzin P *et al. Verticillium longisporum*, the invisible threat to oilseed rape and other brassicaceous plant hosts. *Mol Plant Pathol* 2016;**17**:1004–16.

Deshmukh SK, Verekar SA, Bhave S V. Endophytic fungi: a reservoir of antibacterials. *Front Microbiol* 2014;**5**:715.

Dijkhuizen LW, Brouwer P, Bolhuis H *et al.* Is there foul play in the leaf pocket? The metagenome of floating fern *Azolla* reveals endophytes that do not fix N2 but may

denitrify. *New Phytol* 2018;**217**:453–66.

Ding L, Münch J, Goerls H *et al.* Xiamycin, a pentacyclic indolosesquiterpene with selective anti-HIV activity from a bacterial mangrove endophyte. *Bioorg Med Chem Lett* 2010;**20**:6685–7.

Doestzada M, Vila AV, Zhernakova A *et al.* Pharmacomicrobiomics: a novel route towards personalized medicine? *Protein Cell* 2018;**9**:432–45.

Dos-Santos CM, Souza DG de, Balsanelli E *et al.* A culture-independent approach to enrich endophytic bacterial cells from sugarcane stems for community characterization. *Microb Ecol* 2017;**74**:453–65.

Downie JA. The roles of extracellular proteins, polysaccharides and signals in the interactions of rhizobia with legume roots. *FEMS Microbiol Rev* 2010;**34**:150–70.

Dreyfuss MM, Chapela IH. Potential of fungi in the discovery of novel, low-molecular weight pharmaceuticals. *Discov Nov Nat Prod with Ther Potential* 1994;**26**:49–80.

Duca D, Lorv J, Patten CL *et al.* Indole-3-acetic acid in plant–microbe interactions. *Antonie Van Leeuwenhoek* 2014;**106**:85–125.

Duke SO, Cantrell CL, Meepagala KM *et al.* Natural toxins for use in pest management. *Toxins (Basel)* 2010;**2**:1943–62.

Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 2004;**32**:1792–7.

Effmert U, Kalderás J, Warnke R *et al.* Volatile mediated interactions between bacteria and fungi in the soil. *J Chem Ecol* 2012;**38**:665–703.

Elbehery AHA, Aziz RK, Siam R. Antibiotic resistome: improving detection and quantification accuracy for comparative metagenomics. *Omi A J Integr Biol* 2016;**20**:229–38.

Elleuche S, Schäfers C, Blank S *et al.* Exploration of extremophiles for high temperature biotechnological processes. *Curr Opin Microbiol* 2015;**25**:113–9.

European Committee for Antimicrobial Susceptibility Testing (EUCAST) of the European Society of Clinical Microbiology and Infectious Diseases (ESCMID). Determination of minimum inhibitory concentrations (MICs) of antibacterial agents by broth dilution. *Clin Microbiol Infect* 2003;**9**:1–7.

Faust K, Lahti L, Gonze D *et al.* Metagenomics meets time series analysis: unraveling microbial community dynamics. *Curr Opin Microbiol* 2015;**25**:56–66.

Feher M and, Schmidt JM. Property distributions: differences between drugs, natural products, and molecules from combinatorial chemistry. 2003;**43**:218–27.

Fernandes P, Martens E. Antibiotics in late clinical development. *Biochem Pharmacol* 2017;**133**:152–63.

Ferrer M, Méndez-García C, Rojo D *et al.* Antibiotic use and microbiome function. *Biochem Pharmacol* 2017;**134**:114–26.

Fisher MC, Hawkins NJ, Sanglard D *et al.* Worldwide emergence of resistance to antifungal drugs challenges human health and food security. *Science* 2018;**360**:739–42.

Fléche-Matéos A Le, Kügler JH, Hansen SH *et al. Rouxiella badensis* sp. nov. and *Rouxiella silvae* sp. nov. isolated from peat bog soil and emendation description of the genus *Rouxiella*. *Int J Syst Evol Microbiol* 2017;**67**:1255–9.

Forsberg KJ, Patel S, Wencewicz TA *et al.* Bacterial phylogeny structures soil resisomte across habitats. 2014;**509**:612–6.

Frahm J-P. Recent developments of commercial products from bryophytes. *Bryologist* 2004;**107**:277–83.

Frank AC, Saldierna Guzmán JP, Shay JE. Transmission of bacterial endophytes. *Microorganisms* 2017;**5**:70.

Frankowski J, Lorito M, Scala F *et al.* Purification and properties of two chitinolytic enzymes of *Serratia plymuthica* HRO-C48. *Arch Microbiol* 2001;**176**:421–6.

Fuangthong M, Atichartpongkul S, Mongkolsuk S *et al.* OhrR is a repressor of *ohrA* , a key organic hydroperoxide resistance determinant in *Bacillus subtilis*. 2001;**183**:4134–41.

Gabor E, Liebeton K, Niehaus F *et al.* Updating the metagenomics toolbox. *Biotechnol J* 2007;**2**:201–6.

Gabor EM, Alkema WBL, Janssen DB. Quantifying the accessibility of the metagenome by random expression cloning techniques. *Environ Microbiol* 2004;**6**:879–86.

Gabor EM, Vries EJ, Janssen DB. Efficient recovery of environmental DNA for expression cloning by indirect extraction methods. *FEMS Microbiol Ecol* 2003;**44**:153–63.

Gao H, Li G, Lou H-X *et al.* Structural diversity and biological activities of novel secondary metabolites from endophytes. *Molecules* 2018;**23**:646.

García-Alcalde F, Okonechnikov K, Carbonell J *et al.* Qualimap: evaluating next-generation sequencing alignment data. *Bioinformatics* 2012;**28**:2678–9.

Garza DR, Dutilh BE. From cultured to uncultured genome sequences: metagenomics and modeling microbial ecosystems. *Cell Mol Life Sci* 2015;**72**:4287–308.

Gayathri S, Saravanan D, Radhakrishnan M *et al.* Bioprospecting potential of fast growing endophytic bacteria from leaves of mangrove and salt-marsh plant species. *Indian J Biotechnol* 2010;**9**:397–402.

Geddes-McAlister J, Shapiro RS. New pathogens, new tricks: emerging, drug-resistant fungal pathogens and future prospects for antifungal therapeutics. *Ann N Y Acad Sci* 2018;**1435**:57–78.

Ghyselinck J, Velivelli SLS, Heylen K *et al.* Bioprospecting in potato fields in the Central Andean Highlands: screening of rhizobacteria for plant growth-promoting properties. *Syst Appl Microbiol* 2013;**36**:116–27.

Gillespie DE, Brady SF, Bettermann AD *et al.* Isolation of antibiotics turbomycin A and B from a metagenomic library of soil microbial DNA. *Appl Environ Microbiol* 2002;**68**:4301–6.

Glick BR. Modulation of plant ethylene levels by the bacterial enzyme ACC deaminase. *FEMS Microbiol Lett* 2005;**251**:1–7.

Goethem MW Van, Pierneef R, Bezuidt OKI *et al.* A reservoir of ´historical´antibiotic resistance genes in remote pristine Antarctic soils. 2018;**6**:40.

Govinda Rajulu MB, Thirunavukkarasu N, Suryanarayanan TS *et al.* Chitinolytic enzymes from endophytic fungi. *Fungal Divers* 2011;**47**:43–53.

Green H, Jones AM. Emerging Gram-negative bacteria: pathogenic or innocent bystanders. *Curr Opin Pulm Med* 2018;**24**:592–8.

Grube M, Cernava T, Soh J *et al.* Exploring functional contexts of symbiotic sustain within lichen-associated bacteria by comparative omics. *ISME J* 2015;**9**:412–24.

Gu X, Wang S, Wang S *et al.* Identification and characterization of two novel esterases from a metagenomic library. *Food Sci Technol Res* 2015;**21**:649–57.

Guindon S, Dufayard J-F, Lefort V *et al.* New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol* 2010;**59**:307–21.

Guinea J. Global trends in the distribution of *Candida* species causing candidemia. *Clin Microbiol Infect* 2014;**20**:5–10.

Gunatilaka AAL. Natural products from plant-associated microorganisms: distribution, structural diversity, bioactivity, and implications of their occurrence. *J Nat Prod* 2006;**69**:509–26.

Hall A, Nelson K, Poole LB *et al.* Structure-based insights into the catalytic power and

conformational dexterity of peroxiredoxins. *Antioxid Redox Signal* 2011;**15**:795–815.

Hallmann J, Berg G, Schulz B. Isolation procedures for endophytic microorganisms. In: Schulz BJE, Boyle CJC, Sieber TN (eds.). *Microbial Root Endophytes*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006, 299–319.

Han J, Sahin O, Barton Y-W *et al.* Key role of Mfd in the development of fluoroquinolone resistance in *Campylobacter jejuni*. *PLoS Pathog* 2008;**4**:e1000083.

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

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

Hardoim PR, van Overbeek LS, Berg G *et al.* The hidden world within plants: ecological and evolutionary considerations for defining functioning of microbial endophytes. *Microbiol Mol Biol Rev* 2015;**79**:293–320.

Hardoim PR, van Overbeek LS, Elsas JD van. Properties of bacterial endophytes and their proposed role in plant growth. *Trends Microbiol* 2008;**16**:463–71.

Harvey AL. Natural products in drug discovery. *Drug Discov Today* 2008;**13**:894–901.

Hayat R, Ali S, Amara U *et al.* Soil beneficial bacteria and their role in plant growth promotion: a review. *Ann Microbiol* 2010;**60**:579–98.

Healy FG, Ray RM, Aldrich HC *et al.* Direct isolation of functional genes encoding cellulases from the microbial consortia in a thermophilic, anaerobic digester maintained on lignocellulose. *Appl Microbiol Biotechnol* 1995;**43**:667–74.

Heather JM, Chain B. The sequence of sequencers: the history of sequencing DNA. *Genomics* 2016;**107**:1–8.

Henkel T, Brunne RM, Müller H *et al.* Statistical investigation into the structural complementarity of natural products and synthetic compounds. *Angew Chemie Int Ed* 1999;**38**:643–7.

Hirsch PR, Mauchline TH. Who's who in the plant root microbiome? *Nat Biotechnol* 2012;**30**:961–2.

Huerta-Cepas J, Serra F, Bork P. ETE 3: reconstruction, analysis, and visualization of phylogenomic data. *Mol Biol Evol* 2016;**33**:1635–8.

Idris R, Trifonova R, Puschenreiter M *et al.* Bacterial communities associated with flowering plants of the Ni *Hyperaccumulator Thlaspi goesingense*. *Appl Environ Microbiol* 2004;**70**:2667–77.

Ikeda S, Kaneko T, Okubo T *et al.* Development of a bacterial cell enrichment method and its application to the community analysis in soybean stems. *Microb Ecol* 2009;**58**:703–14.

Inbar J, Chet I. Evidence that chitinase produced by *Aeromonas caviae* is involved in the biological control of soil-borne plant pathogens by this bacterium. *Soil Biol Biochem* 1991;**23**:973–8.

Jahn L, Schafhauser T, Wibberg D *et al.* Linking secondary metabolites to biosynthesis genes in the fungal endophyte *Cyanodermella asteris*: the anti-cancer bisanthraquinone skyrin. *J Biotechnol* 2017;**257**:233–9.

Jansson JK, Neufeld JD, Moran MA *et al.* Omics for understanding microbial functional dynamics. *Environ Microbiol* 2012;**14**:1–3.

Jiao J-Y, Wang H-X, Zeng Y *et al.* Enrichment for microbes living in association with plant tissues. *J Appl Microbiol* 2006;**100**:830–7.

Johnson LJ, Koulman A, Christensen M *et al.* An extracellular siderophore is required to maintain the mutualistic interaction of *Epichloë festucae* with *Lolium perenne*. *PLoS Pathog* 2013;**9**:e1003332.

Kåhrström CT. Entering a post-antibiotic era? *Nat Rev Microbiol* 2013;**11**:146.

Kai M, Effmert U, Berg G *et al.* Volatiles of bacterial antagonists inhibit mycelial growth of the plant pathogen *Rhizoctonia solani*. *Arch Microbiol* 2007;**187**:351–60.

Kai M, Haustein M, Molina F *et al.* Bacterial volatiles and their action potential. *Appl Microbiol Biotechnol* 2009;**81**:1001–12.

Kanchiswamy CN, Malnoy M, Maffei ME. Chemical diversity of microbial volatiles and their potential for plant growth and productivity. *Front Plant Sci* 2015;**6**:151.

Kandel SL, Joubert PM, Doty SL. Bacterial endophyte colonization and distribution within plants. *Microorganisms* 2017;**5**:77.

Kanokratana P, Eurwilaichitr L, Pootanakit K *et al.* Identification of glycosyl hydrolases from a metagenomic library of microflora in sugarcane bagasse collection site and their cooperative action on cellulose degradation. *J Biosci Bioeng* 2015;**119**:384–91.

Katz M, Hover BM, Brady SF. Culture-independent discovery of natural products from soil metagenomes. *J Ind Microbiol Biotechnol* 2016;**43**:129–41.

Khaneghah AM, Martins LM, von Hertwig AM *et al.* Deoxynivalenol and its masked forms: characteristics, incidence, control and fate during wheat and wheat based products processing - a review. *Trends Food Sci Technol* 2018;**71**:13–24.

Kharwar RN, Mishra A, Gond SK *et al.* Anticancer compounds derived from fungal endophytes: their importance and future challenges. *Nat Prod Rep* 2011;**28**:1208–1028.

Kim HJ, Jeong YS, Jung WK *et al.* Characterization of novel family IV esterase and family I.3 lipase from an oil-polluted mud flat metagenome. *Mol Biotechnol* 2015;**57**:781–92.

Kloepper JW, Ryu C-M, Zhang S. Induced systemic resistance and promotion of plant growth by Bacillus spp. *Phytopathology* 2004;**94**:1259–66.

Knief C. Analysis of plant microbe interactions in the era of next generation sequencing technologies. *Front Plant Sci* 2014;**5**:216.

Kostka JE, Weston DJ, Glass JB *et al.* The *Sphagnum* microbiome: new insights from an ancient plant lineage. *New Phytol* 2016;**211**:57–64.

Koulman A, Lee TV, Fraser K *et al.* Identification of extracellular siderophores and a related peptide from the endophytic fungus *Epichloë festucae* in culture and endophyte-infected *Lolium perenne*. *Phytochemistry* 2012;**75**:128–39.

Krings U, Berger RG. Biotechnological production of flavours and fragrances. *Appl Microbiol Biotechnol* 1998;**49**:1–8.

Kröber M, Wibberg D, Grosch R *et al.* Effect of the strain *Bacillus amyloliquefaciens* FZB42 on the microbial community in the rhizosphere of lettuce under field conditions analyzed by whole metagenome sequencing. *Front Microbiol* 2014;**5**:252.

Kumar S, Stecher G, Li M *et al.* MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol Biol Evol* 2018;**35**:1547–9.

Kurm V, van der Putten WH, de Boer W *et al.* Low abundant soil bacteria can be metabolically versatile and fast growing. *Ecology* 2017;**98**:555–64.

Kusari S, Lamshöft M, Zühlke S *et al.* An endophytic fungus from Hypericum perforatum that produces hypericin. *J Nat Prod* 2008;**71**:159–62.

Kusari S, Zühlke S, Spiteller M. An endophytic fungus from Camptotheca acuminata that produces camptothecin and analogues. *J Nat Prod* 2009;**72**:2–7.

Lane GA, Christensen MJ, Miles CO. Coevolution of fungal endophytes with grasses: the significance of secondary metabolites. In: Bacon CW, White J (eds.). *Microbial Endophytes*. New York: Marcel Dekker, 2000, 341–88.

Lee ES, Song EJ, Nam Y Do *et al.* Probiotics in human health and disease: from nutribiotics to

pharmabiotics. *J Microbiol* 2018;**56**:773–82.

Leis B, Angelov A, Liebl W. Screening and expression of genes from metagenomes. *Adv Appl Microbiol* 2013;**83**:1–68.

Lelie D van der, Taghavi S, McCorkle SM *et al.* The metagenome of an anaerobic microbial community decomposing poplar wood chips. *PLoS One* 2012;**7**:e36740.

Lemanceau P, Corberand T, Gardan L *et al.* Effect of two plant species, flax (*Linum usitatissinum* L.) and tomato (*Lycopersicon esculentum* Mill.), on the diversity of soilborne populations of fluorescent pseudomonads. *Appl Environ Microbiol* 1995;**61**:1004–12.

Lesniak J, Barton WA, Nikolov DB. Structural and functional characterization of the *Pseudomonas* hydroperoxide resistance protein Ohr. *EMBO J* 2002;**21**:6649–59.

Li G, Lou HX. Strategies to diversify natural products for drug discovery. *Med Res Rev* 2018;**38**:1255–94.

Li L-L, Taghavi S, McCorkle SM *et al.* Bioprospecting metagenomics of decaying wood: mining for new glycoside hydrolases. *Biotechnol Biofuels* 2011;**4**:23.

Liebl W, Angelov A, Juergensen J *et al.* Alternative hosts for functional (meta)genome analysis. *Appl Microbiol Biotechnol* 2014;**98**:8099–109.

Lindell SD, Pattenden LC, Shannon J. Combinatorial chemistry in the agrosciences. *Bioorg Med Chem* 2009;**17**:4035–46.

Liu R, Li X, Lam KS. Combinatorial chemistry in drug discovery. *Curr Opin Chem Biol* 2017;**38**:117–26.

Lorenzo JM, Munekata PES, Gómez B *et al.* Bioactive peptides as natural antioxidants in food products – a review. *Trends Food Sci Technol* 2018;**79**:136–47.

Lumyong S, Lumyong P, McKenzie EH. *et al.* Enzymatic activity of endophytic fungi of six native seedling species from Doi Suthep-Pui National Park, Thailand. *Can J Microbiol* 2002;**48**:1109–12.

Lundberg DS, Lebeis SL, Paredes SH *et al.* Defining the core *Arabidopsis thaliana* root microbiome. *Nature* 2012;**488**:86–90.

Lycus P, Lovise Bøthun K, Bergaust L *et al.* Phenotypic and genotypic richness of denitrifiers revealed by a novel isolation strategy. *ISME J* 2017;**11**:2219–32.

Mahlen SD. Serratia infections: from military experiments to current practice. *Clin Microbiol Rev* 2011;**24**:755–91.

Mahnert A, Moissl-Eichinger C, Zojer M *et al.* Man-made microbial resistances in built environments. *Nat Commun* 2019;**10**:968.

Maria G, Sridhar K, Agricultural NR-J of *et al.* Antimicrobial and enzyme activity of mangrove endophytic fungi of southwest coast of India. *ijat-aatsea.com* 2005;**1**:67–80.

Martinez-Klimova E, Rodríguez-Peña K. Endophytes as sources of antibiotics. *Biochem Pharmacol* 2017;**134**:1–17.

Martínez JL. Antibiotics and antibiotic resistance genes in natural environments. *Science* 2008;**321**:365–7.

Martinez JL, Sánchez MB, Martínez-Solano L *et al.* Functional role of bacterial multidrug efflux pumps in microbial natural ecosystems. *FEMS Microbiol Rev* 2009;**33**:430–49.

Masatoshi Nei, Sudhir Kumar. *Molecular Evolution and Phylogenetics*. New York: Oxford University Press, 2000.

Mauri M, Elli T, Caviglia G *et al.* RAWGraphs: a visualisation platform to create open outputs. *Proceedings of the 12th Biannual Conference on Italian SIGCHI Chapter*. New York, New York, USA: ACM Press, 2017, 28:1-28–5.

McArthur AG, Waglechner N, Nizam F *et al.* The comprehensive antibiotic resistance database. *Antimicrob Agents Chemother* 2013;**57**:3348–57.

McChesney JD, Venkataraman SK, Henri JT. Plant natural products: back to the future or into extinction? *Phytochemistry* 2007;**68**:2015–22.

McEwen SA, Collignon PJ. Antimicrobial resistance: a One Health colloquium. *Microbiol Spectr* 2017;**6**:1–26.

Medina-Franco J, Martinez-Mayorga K, Giulianotti M *et al.* Visualization of the chemical space in drug discovery. *Curr Comput Aided-Drug Des* 2008;**4**:322–33.

Medina-Franco JL. Interrogating novel areas of chemical space for drug discovery using chemoinformatics. *Drug Dev Res* 2012;**73**:430–8.

Medina-Romero YM, Roque-Flores G, Macías-Rubalcava ML. Volatile organic compounds from endophytic fungi as innovative postharvest control of *Fusarium oxysporum* in cherry tomato fruits. *Appl Microbiol Biotechnol* 2017;**101**:8209–22.

Meireles DA, Domingos RM, Gaiarsa JW *et al.* Functional and evolutionary characterization of Ohr proteins in eukaryotes reveals many active homologs among pathogenic fungi. *Redox Biol* 2017;**12**:600–9.

Mendes R, Raaijmakers JM. Cross-kingdom similarities in microbiome functions. *ISME J* 2015;**9**:1905–7.

Miller, Miller, Garton-Kenny *et al.* Ecomycins, unique antimycotics from *Pseudomonas viridiflava*. *J Appl Microbiol* 1998;**84**:937–44.

Mirete S, de Figueras CG, González-Pastor JE. Novel nickel resistance genes from the rhizosphere metagenome of plants adapted to acid mine drainage. *Appl Environ Microbiol* 2007;**73**:6001–11.

Mitter B, Pfaffenbichler N, Flavell R *et al.* A new approach to modify plant microbiomes and traits by introducing beneficial bacteria at flowering into progeny seeds. *Front Microbiol* 2017;**8**:11.

Mookherjee A, Bera P, Mitra A *et al.* Characterization and synergistic effect of antifungal volatile organic compounds emitted by the *Geotrichum candidum* PF005, an endophytic fungus from the eggplant. *Microb Ecol* 2018;**75**:647–61.

Müller CA, Oberauner-Wappis L, Peyman A *et al.* Mining for NRPS and PKS genes revealed a high diversity in the *Sphagnum* bog metagenome. *Appl Environ Microbiol* 2015;**81**:5064–72.

Müller CA, Obermeier MM, Berg G. Bioprospecting plant-associated microbiomes. *J Biotechnol* 2016;**235**, DOI: 10.1016/j.jbiotec.2016.03.033.

Müller CA, Perz V, Provasnek C *et al.* Discovery of polyesterases from moss-associated microorganisms. *Appl Environ Microbiol* 2017;**83**:e02641-16.

Müller H, Westendorf C, Leitner E *et al.* Quorum-sensing effects in the antagonistic rhizosphere bacterium *Serratia plymuthica* HRO-C48. *FEMS Microbiol Ecol* 2009;**67**:468–78.

Naas T, Nordmann P. Analysis of a carbapenem-hydrolyzing class A beta-lactamase from *Enterobacter cloacae* and of its LysR-type regulatory protein. *Proc Natl Acad Sci U S A* 1994;**91**:7693–7.

Nalini MS, Prakash HS. Diversity and bioprospecting of actinomycete endophytes from the medicinal plants. *Lett Appl Microbiol* 2017;**64**:261–70.

Newman DJ, Cragg GM. Endophytic and epiphytic microbes as “sources” of bioactive agents. *Front Chem* 2015;**3**:34.

Newman DJ, Cragg GM. Natural products as sources of new drugs from 1981 to 2014. *J Nat Prod* 2016;**79**:629–61.

Nikolic B, Schwab H, Sessitsch A. Metagenomic analysis of the 1-aminocyclopropane-1-carboxylate deaminase gene (*acdS*) operon of an uncultured bacterial endophyte colonizing *Solanum tuberosum* L. *Arch Microbiol* 2011;**193**:665–76.

Nikolouli K, Mossialos D. Bioactive compounds synthesized by non-ribosomal peptide synthetases and type-I polyketide synthases discovered through genome-mining and metagenomics. *Biotechnol Lett* 2012;**34**:1393–403.

Noble SM, Gianetti BA, Witchley JN. *Candida albicans* cell-type switching and functional plasticity in the mammalian host. *Nat Publ Gr* 2016;**15**:96–108.

Nováková J, Farkašovský M. Bioprospecting microbial metagenome for natural products. *Biologia (Bratisl)* 2013;**68**:1079–86.

Nunes-Alves C. Commensally sourced antibiotics. *Nat Rev Drug Discov* 2014;**13**:812–812.

O'Neill J. Antimicrobial resistance: tackling a crisis for the health and wealth of nations, 2014. *Rev Antimicrob Resist* 2014, DOI: 10.1038/510015a.

Oliveira MA, Guimarães BG, Cussiol JRR *et al.* Structural insights into enzyme–substrate interaction and characterization of enzymatic intermediates of organic hydroperoxide resistance protein from *Xylella fastidiosa*. *J Mol Biol* 2006;**359**:433–45.

Opelt K, Berg C, Berg G. The bryophyte genus *Sphagnum* is a reservoir for powerful and extraordinary antagonists and potentially facultative human pathogens. *FEMS Microbiol Ecol* 2007;**61**:38–53.

Opelt K, Berg C, Schönmann S *et al.* High specificity but contrasting biodiversity of *Sphagnum*-associated bacterial and plant communities in bog ecosystems independent of the geographical region. *ISME J* 2007a;**1**:502–16.

Opelt K, Chobot V, Hadacek F *et al.* Investigations of the structure and function of bacterial communities associated with *Sphagnum* mosses. *Environ Microbiol* 2007b;**9**:2795–809.

Ortholand J-Y, Ganesan A. Natural products and combinatorial chemistry: back to the future. *Curr Opin Chem Biol* 2004;**8**:271–80.

Owen JG, Robins KJ, Parachin NS *et al.* A functional screen for recovery of 4′-phosphopantetheinyl transferase and associated natural product biosynthesis genes from metagenome libraries. *Environ Microbiol* 2012;**14**:1198–209.

Pace NR, Stahl DA, Lane DJ *et al.* The analysis of natural microbial populations by ribosomal RNA sequences. In: Marshall KC (ed.). *Advances in Microbial Ecology*. Springer, Boston, MA, 1986, 1–55.

Page SE, Baird AJ. Peatlands and global change: response and resilience. *Annu Rev Environ Resour* 2016;**41**:35–57.

Pal C, Bengtsson-Palme J, Kristiansson E *et al.* The structure and diversity of human, animal and environmental resistomes. *Microbiome* 2016;**4**:1–15.

Papanicolaou GA, Medeiros AA, Jacoby GA. Novel plasmid-mediated β-lactamase (MIR-1) conferring resistance to oxyimino- and α-methoxy β-lactams in clinical isolates of *Klebsiella pneumoniae*. *Antimicrob Agents Chemother* 1990;**34**:2200–9.

Patridge E, Gareiss P, Kinch MS *et al.* An analysis of FDA-approved drugs: natural products and their derivatives. *Drug Discov Today* 2016;**21**:204–7.

Pérez-Jaramillo JE, Mendes R, Raaijmakers JM. Impact of plant domestication on rhizosphere microbiome assembly and functions. *Plant Mol Biol* 2016;**90**:635–44.

Perfect JR. The antifungal pipeline : a reality check. *Nat Rev Drug Discov* 2017;**16**:603–16.

Perkins A, Poole LB, Karplus PA. Tuning of peroxiredoxin catalysis for various physiological roles. *Biochemistry* 2014;**53**:7693–705.

Perron GG, Whyte L, Turnbaugh PJ *et al.* Functional characterization of bacteria isolated from ancient arctic soil exposes diverse resistance mechanisms to modern antibiotics. *PLoS*

One 2015;**10**:1–19.

Philippon A, Slama P, Dény P *et al.* A structure-based classification of class A β-lactamases, a broadly diverse family of enzymes. *Clin Microbiol Rev* 2016;**29**:29–57.

Philippot L, Raaijmakers JM, Lemanceau P *et al.* Going back to the roots: the microbial ecology of the rhizosphere. *Nat Rev Microbiol* 2013;**11**:789–99.

Pleban S, Chernin L, Chet I. Chitinolytic activity of an endophytic strain of *Bacillus cereus*. *Lett Appl Microbiol* 1997;**25**:284–8.

Prost I, Dhondt S, Rothe G *et al.* Evaluation of the antimicrobial activities of plant oxylipins supports their involvement in defense against pathogens. *Plant Physiol* 2005;**139**:1902–13.

Pulido MR, García-Quintanilla M, Gil-Marqués ML *et al.* Identifying targets for antibiotic development using omics technologies. *Drug Discov Today* 2016;**21**:465–72.

Qin S, Xing K, Jiang J-H *et al.* Biodiversity, bioactive natural products and biotechnological potential of plant-associated endophytic actinobacteria. *Appl Microbiol Biotechnol* 2010;**89**:457–73.

Ramírez-Puebla ST, Servín-Garcidueñas LE, Jiménez-Marín B *et al.* Gut and root microbiota commonalities. *Appl Environ Microbiol* 2013;**79**:2–9.

Ravin N V., Mardanov A V., Skryabin KG. Metagenomics as a tool for the investigation of uncultured microorganisms. *Russ J Genet* 2015;**51**:431–9.

Raymond JA. Dependence on epiphytic bacteria for freezing protection in an Antarctic moss, *Bryum argenteum*. *Environ Microbiol Rep* 2016;**8**:14–9.

Reinhold-Hurek B, Hurek T. Living inside plants: bacterial endophytes. *Curr Opin Plant Biol* 2011;**14**:435–43.

Reinhold-Hurek B, Maes T, Gemmer S *et al.* An endoglucanase is involved in infection of rice roots by the not-cellulose-metabolizing endophyte Azoarcus Sp. etrain BH72. *Mol Plant-Microbe Interact* 2006;**19**:181–8.

Ribes S, Fuentes A, Talens P *et al.* Prevention of fungal spoilage in food products using natural compounds: a review. *Crit Rev Food Sci Nutr* 2018;**58**:2002–16.

Robl D, Delabona P da, Mergel C *et al.* The capability of endophytic fungi for production of hemicellulases and related enzymes. *BMC Biotechnol* 2013;**13**:94.

Roossinck MJ, Martin DP, Roumagnac P. Plant virus metagenomics: advances in virus discovery. *Phytopathology* 2015;**105**:716–27.

Rosconi F, Davyt D, Martínez V *et al.* Identification and structural characterization of serobactins, a suite of lipopeptide siderophores produced by the grass endophyte *Herbaspirillum seropedicae*. *Environ Microbiol* 2013;**15**:916–27.

Rudrappa T, Czymmek KJ, Paré PW *et al.* Root-secreted malic acid recruits beneficial soil bacteria. *Plant Physiol* 2008;**148**:1547–56.

Ryu C-M, Farag MA, Hu C-H *et al.* Bacterial volatiles promote growth in *Arabidopsis*. *Proc Natl Acad Sci U S A* 2003;**100**:4927–32.

Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 1987;**4**:406–25.

Savery NJ. The molecular mechanism of transcription-coupled DNA repair. *Trends Microbiol* 2007;**15**:326–33.

Sawana A, Adeolu M, Gupta RS. Molecular signatures and phylogenomic analysis of the genus *Burkholderia*: proposal for division of this genus into the emended genus *Burkholderia* containing pathogenic organisms and a new genus *Paraburkholderia* gen. nov. harboring env. *Front Genet* 2014;**5**:429.

Schaible GA, Strobel GA, Mends MT *et al.* Characterization of an endophytic *Gloeosporium*

sp. and its novel bioactivity with "Synergistans." *Microb Ecol* 2015;**70**:41–50.

Schlaeppi K, Dombrowski N, Oter RG *et al.* Quantitative divergence of the bacterial root microbiota in *Arabidopsis thaliana* relatives. *Proc Natl Acad Sci U S A* 2014;**111**:585–92.

Schmeisser C, Steele H, Streit WR. Metagenomics, biotechnology with non-culturable microbes. *Appl Microbiol Biotechnol* 2007;**75**:955–62.

Schmidt TM, DeLong EF, Pace NR. Analysis of a marine picoplankton community by 16S rRNA gene cloning and sequencing. *J Bacteriol* 1991;**173**:4371–8.

Schmitt EK, Hoepfner D, Krastel P. Natural products as probes in pharmaceutical research. *J Ind Microbiol Biotechnol* 2016;**43**:249–60.

Schneider I, Queenan AM, Bauernfeind A. Novel carbapenem-hydrolyzing oxacillinase OXA-62 from *Pandoraea pnomenusa*. *Antimicrob Agents Chemother* 2006;**50**:1330–5.

Schulz B, Boyle C, Draeger S *et al.* Endophytic fungi: a source of novel biologically active secondary metabolites. *Mycol Res* 2002;**106**:996–1004.

Schulz S, Dickschat JS. Bacterial volatiles: the smell of small organisms. *Nat Prod Rep* 2007;**24**:814–42.

Sears J, Dirkse E, Markworth C *et al.* Volatile antimicrobials from *Muscodor albus*, a novel endophytic fungus. *Microbiology* 2001;**147**:2943–50.

Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 2014;**30**:2068–9.

Sessitsch A, Hardoim P, Döring J *et al.* Functional characteristics of an endophyte community colonizing rice roots as revealed by metagenomic analysis. *Mol Plant-Microbe Interact* 2011;**25**:28–36.

Shahzad R, Khan AL, Bilal S *et al.* What is there in seeds? vertically transmitted endophytic resources for sustainable improvement in plant growth. *Front Plant Sci* 2018;**9**:24.

Shannon P, Markiel A, Ozier O *et al.* Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 2003;**13**:2498–504.

Shcherbakov A V, Bragina A V, Kuzmina EY *et al.* Endophytic bacteria of *Sphagnum* mosses as promising objects of agricultural. *Microbiology* 2013;**82**:306–15.

Shen B. A new golden age of natural products drug discovery. *Cell* 2015;**163**:1297–300.

Sheoran N, Valiya Nadakkakath A, Munjal V *et al.* Genetic analysis of plant endophytic *Pseudomonas putida* BP25 and chemo-profiling of its antimicrobial volatile organic compounds. *Microbiol Res* 2015;**173**:66–78.

Shi X, Liu Q, Ma J *et al.* An acid-stable bacterial laccase identified from the endophyte *Pantoea ananatis* Sd-1 genome exhibiting lignin degradation and dye decolorization abilities. *Biotechnol Lett* 2015;**37**:2279–88.

Shi Y, Yang H, Zhang T *et al.* Illumina-based analysis of endophytic bacterial diversity and space-time dynamics in sugar beet on the north slope of Tianshan mountain. *Appl Microbiol Biotechnol* 2014;**98**:6375–85.

Simon C, Daniel R. Metagenomic analyses: past and future trends. *Appl Environ Microbiol* 2011;**77**:1153–61.

Simon C, Daniel R. Construction of small-insert and large-insert metagenomic libraries. *Methods Mol Biol* 2017;**1539**:1–12.

Singh Arora D, Kumar Sharma R. Ligninolytic fungal laccases and their biotechnological applications. *Appl Biochem Biotechnol* 2010;**160**:1760–88.

Sparks TC, Hahn DR, Garizi N V. Natural products, their derivatives, mimics and synthetic equivalents: role in agrochemical discovery. *Pest Manag Sci* 2017;**73**:700–15.

Stępniewska Z, Kuźniar A. Endophytic microorganisms—promising applications in

bioremediation of greenhouse gases. *Appl Microbiol Biotechnol* 2013;**97**:9589–96.

Stewart EJ. Growing unculturable bacteria. *J Bacteriol* 2012;**194**:4151–60.

Stierle A, Strobel G, Stierle D. Taxol and taxane production by *Taxomyces andreanae*, an endophytic fungus of Pacific yew. *Science* 1993;**260**:214–6.

Strobel G, Daisy B. Bioprospecting for microbial endophytes and their natural products. *Microbiol Mol Biol Rev* 2003;**67**:491–502.

Sunitha V, Nirmala Devi D, Srinivas C. Extracellular enzymatic activity of endophytic fungal strains isolated from medicinal plants. *World J Agric Sci* 2013;**9**:1–9.

Suryanarayanan TS, Thirunavukkarasu N, Govindarajulu MB *et al.* Fungal endophytes: an untapped source of biocatalysts. *Fungal Divers* 2012;**54**:19–30.

Tan RX, Zou WX. Endophytes: a rich source of functional metabolites (1987 to 2000). *Nat Prod Rep* 2001;**18**:448–59.

Tian B-Y, Cao Y, Zhang K-Q. Metagenomic insights into communities, functions of endophytes and their associates with infection by root-knot nematode, *Meloidogyne incognita*, in tomato roots. *Sci Rep* 2015;**5**:17087.

Troeger C, Forouzanfar M, Rao PC *et al.* Estimates of the global, regional, and national morbidity, mortality, and aetiologies of lower respiratory tract infections in 195 countries: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet Infect Dis* 2017;**17**:1133–61.

Truyens S, Weyens N, Cuypers A *et al.* Bacterial seed endophytes: genera, vertical transmission and interaction with plants. *Environ Microbiol Rep* 2015;**7**:40–50.

Uchiyama T, Abe T, Ikemura T *et al.* Substrate-induced gene-expression screening of environmental metagenome libraries for isolation of catabolic genes. *Nat Biotechnol* 2005;**23**:88–93.

Ulloa-Benítez Á, Medina-Romero YM, Sánchez-Fernández RE *et al.* Phytotoxic and antimicrobial activity of volatile and semi-volatile organic compounds from the endophyte *Hypoxylon anthochroum* strain Blaci isolated from *Bursera lancifolia* (Burseraceae). *J Appl Microbiol* 2016;**121**:380–400.

Vadstein O, Attramadal KJK, Bakke I *et al. K*-Selection as microbial community management strategy: a method for improved viability of larvae in aquaculture. *Front Microbiol* 2018;**9**:2730.

Vandamme P, Opelt K, Knochel N *et al. Burkholderia bryophila* sp. nov. and *Burkholderia megapolitana* sp. nov., moss-associated species with antifungal and plant-growth-promoting properties. *Int J Syst Evol Microbiol* 2007;**57**:2228–35.

Vercammen K, Garcia-Armisen T, Goeders N *et al.* Identification of a metagenomic gene cluster containing a new class A beta-lactamase and toxin-antitoxin systems. *Microbiologyopen* 2013;**2**:674–83.

Völker U, Andersen KK, Antelmann H *et al.* One of two osmC homologs in *Bacillus subtilis* is part of the sigmaB-dependent general stress regulon. *J Bacteriol* 1998;**180**:4212–8.

Vorholt JA. Microbial life in the phyllosphere. *Nat Rev Microbiol* 2012;**10**:828–40.

Wall ME, Wani MC, Cook CE *et al.* Plant antitumor agents. I. The isolation and structure of camptothecin, a novel alkaloidal leukemia and tumor inhibitor from *Camptotheca acuminata*. *J Am Chem Soc* 1966;**88**:3888–90.

Wang A, Xu Y, Gao Y *et al.* Chemical and bioactive diversities of the genera *Stachybotrys* and *Memnoniella* secondary metabolites. *Phytochem Rev* 2015;**14**:623–55.

Wang H-X, Geng Z-L, Zeng Y *et al.* Enriching plant microbiota for a metagenomic library construction. *Environ Microbiol* 2008;**10**:2684–91.

Wardwell LH, Jude BA, Moody JP *et al.* Co-selection of mercury and antibiotic resistance in

Sphagnum core samples dating back 2000 years. *Geomicrobiol J* 2009;**26**:351–60.

Waterhouse A, Bertoni M, Bienert S *et al.* SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res* 2018;**46**:W296–303.

Weisskopf L. The potential of bacterial volatiles for crop protection against phytophathogenic fungi. In: Méndez-Vilas A (ed.). *Microbial Pathogens and Strategies for Combating Them: Science, Technology and Education*. Badajoz: Formatex Research Center, 2013.

Williamson LL, Borlee BR, Schloss PD *et al.* Intracellular screen to identify metagenomic clones that induce or inhibit a quorum-sensing biosensor. *Appl Environ Microbiol* 2005;**71**:6335–44.

Woolhouse M, Farrar J. Policy: An intergovernmental panel on antimicrobial resistance. *Nature* 2014;**509**:555–7.

World Health Organization. *Antimicrobial Resistance: Global Report on Surveillance.*, 2014.

Wright GD. The antibiotic resistome: the nexus of chemical and genetic diversity. *Nat Rev Microbiol* 2007;**5**:175–86.

Xiao K-Q, Li B, Ma L *et al.* Metagenomic profiles of antibiotic resistance genes in paddy soils from South China. *FEMS Microbiol Ecol* 2016;**92**:3.

Yu P, Cheng S, Xiang J *et al.* Expectorant, antitussive, anti-inflammatory activities and compositional analysis of *Aster tataricus*. *J Ethnopharmacol* 2015;**164**:328–33.

Zachow C, Müller H, Tilcher R *et al.* Catch the best: novel screening strategy to select stress protecting agents for crop plants. *Agronomy* 2013;**3**:794–815.

Zachow C, Tilcher R, Berg G. Sugar beet-associated bacterial and fungal communities show a high indigenous antagonistic potential against plant pathogens. *Microb Ecol* 2008;**55**:119–29.

Zhang H-B, Wang L-H, Zhang L-H. Genetic control of quorum-sensing signal turnover in *Agrobacterium tumefaciens*. *Proc Natl Acad Sci U S A* 2002;**99**:4638–43.

Zhang HW, Song YC, Tan RX. Biology and chemistry of endophytes. *Nat Prod Rep* 2006;**23**:753–71.

Zhang W, Zhao B, Du L *et al.* Cytotoxic polyketides with an oxygen-bridged cyclooctadiene core skeleton from the mangrove endophytic fungus *Phomosis* sp. A818. *Molecules* 2017;**22**:1547.

Appendix

Supplementary tables

Table S 1: Isolate resistance profiles. Showing resistance (1) and sensitivity (0) to all tested antibiotic for each bacterial isolate.

Isolate No.	Ampicillin	Tetracycline	Erythromycin	Gentamycin	Rifampicin	Kanamycin	Nalidixic acid	Ciprofloxacin	Sulfadiazine	Vancomycin
					CULTURE COLLECTION					
1a1	0	0	0	0	0	0	0	0	1	1
1a2	1	0	0	0	1	0	0	1	1	1
1a3	1	0	0	1	1	1	0	0	1	1
1a4	1	0	1	1	1	0	0	0	1	1
1a5	1	0	1	1	1	1	1	1	1	1
1a6	1	0	0	1	1	0	0	0	1	1
1a7	0	0	0	0	0	0	0	0	1	1
1a8	1	0	1	0	1	0	0	1	1	1
1a9	1	0	0	0	0	0	0	1	1	1
1a10	1	0	0	0	0	0	0	0	1	1
1a11	1	0	0	1	0	0	0	0	1	1
1a12	1	0	0	0	1	0	0	1	1	0
1a13	1	0	1	1	0	0	0	1	1	0
1a14	1	0	0	0	0	0	0	0	1	0
1a15	0	0	0	0	1	0	0	1	1	1
1a16	0	0	1	0	0	0	0	0	0	1
1a17	1	0	0	0	0	0	0	0	0	0
1a18	1	1	0	1	1	1	1	1	1	1
1a19	0	0	0	1	1	1	0	1	1	1
1a20	1	0	0	0	0	0	0	0	1	0
1a21	1	0	0	0	0	0	0	1	1	1

Isolate No.	Ampicillin	Tetracycline	Erythromycin	Gentamycin	Rifampicin	Kanamycin	Nalidixic acid	Ciprofloxacin	Sulfadiazine	Vancomycin
CULTURE COLLECTION										
1a22	1	0	1	0	1	0	0	0	1	1
1a23	1	0	0	1	0	0	1	0	1	1
1a24	1	0	1	1	0	0	0	0	1	1
1a25	1	0	0	0	0	0	0	1	0	0
1a26	1	0	0	1	0	0	0	0	0	0
1a27	1	0	0	0	1	0	0	1	1	1
1a28	0	0	0	0	0	0	0	0	0	0
1a29	0	0	0	0	0	0	0	0	0	0
1a30	0	0	1	0	0	1	0	1	1	1
1a31	0	0	0	0	1	0	0	1	1	1
1a32	0	0	0	1	0	0	0	0	1	0
1a33	1	0	1	0	0	0	0	1	1	1
1a34	1	0	0	1	0	1	0	1	1	1
1a35	1	0	0	0	1	0	1	1	1	1
1a36	1	0	1	0	0	0	0	0	1	1
1a37	0	0	0	1	0	0	0	1	0	1
1a38	0	0	0	0	0	0	1	0	1	0
1a39	1	0	0	1	0	1	0	1	1	1
1a40	1	0	1	1	1	1	0	0	1	1
1a41	1	0	0	1	0	0	0	0	1	0
1a42	1	0	0	1	1	0	0	1	1	1
1a43	0	0	0	0	0	0	0	0	0	0
1a44	1	0	0	1	0	0	0	0	0	0
1a45	1	0	0	1	1	0	0	0	0	0
1a46	0	0	0	0	1	0	0	0	0	0
1a47	0	0	0	0	0	0	0	0	0	0
1a48	0	0	1	1	1	0	0	1	1	1
1a49	0	0	0	1	0	1	1	1	1	0
1a50	1	0	1	1	0	0	0	0	1	1

Isolate No.	Ampicillin	Tetracycline	Erythromycin	Gentamycin	Rifampicin	Kanamycin	Nalidixic acid	Ciprofloxacin	Sulfadiazine	Vancomycin
CULTURE COLLECTION										
1b1	1	0	1	0	0	1	1	1	1	1
1b2	0	0	0	0	0	0	0	0	1	1
1b3	0	0	0	1	0	0	0	0	0	1
1b4	1	0	0	1	0	0	0	1	1	1
1b5	1	0	1	1	0	1	0	1	1	1
1b6	0	0	0	0	1	0	0	1	1	0
1b7	0	0	1	0	1	0	0	1	1	0
1b8	1	0	0	0	0	0	0	0	1	1
1b9	0	0	0	0	0	0	0	0	0	0
1b10	0	0	0	0	1	0	0	0	0	0
1b11	1	0	0	0	0	1	1	1	1	1
1b12	0	0	0	0	0	0	0	0	1	0
1b13	0	0	0	1	0	0	0	1	1	0
1b14	1	0	1	1	1	0	0	0	1	1
1b15	1	0	0	0	0	0	0	1	1	1
1b16	0	0	0	0	0	0	0	0	1	0
1b17	0	0	0	0	0	0	0	0	0	0
1b18	0	0	0	0	0	0	0	0	0	0
1b19	0	0	0	1	0	0	0	0	0	0
1b20	0	0	0	0	1	0	0	1	1	0
1b21	0	0	0	0	0	0	0	0	1	0
1b22	0	0	0	0	0	0	0	1	1	0
1b23	0	0	0	0	0	0	0	0	0	0
1b24	1	0	1	1	1	0	0	0	1	1
1b25	0	0	0	0	0	0	0	0	0	0
1b26	1	0	1	1	1	0	0	0	1	1
1b27	1	0	0	0	0	0	1	1	1	0
1b28	1	1	1	1	0	0	0	0	1	1
1b29	0	0	0	0	1	0	0	1	1	0

Isolate No.	Ampicillin	Tetracycline	Erythromycin	Gentamycin	Rifampicin	Kanamycin	Nalidixic acid	Ciprofloxacin	Sulfadiazine	Vancomycin
CULTURE COLLECTION										
1b30	1	1	1	0	1	0	1	1	1	1
1b31	0	0	0	0	0	0	1	1	1	0
1b32	0	0	0	0	0	0	0	0	1	0
1b33	0	0	0	0	0	0	0	0	0	1
1b34	0	0	0	1	0	0	0	0	1	0
1b35	0	0	0	0	0	0	0	0	0	0
1b36	0	0	0	0	1	0	0	1	1	0
1b37	0	0	0	0	0	0	1	0	1	0
1b38	0	0	0	0	0	0	0	0	1	0
1b39	1	0	0	0	1	0	0	1	1	1
1b40	1	0	1	1	1	0	0	1	1	1
1b41	0	1	0	1	1	0	0	0	0	0
1b42	0	1	1	1	1	0	0	0	1	1
1b43	1	0	1	1	0	1	0	1	1	0
1b44	1	0	1	1	0	0	0	0	1	0
1b45	1	0	0	0	0	0	1	1	1	1
2a1	0	0	0	0	0	0	0	0	0	0
2a2	0	0	1	0	0	0	0	1	1	1
2a3	0	0	0	0	0	0	0	1	1	1
2a4	0	0	0	0	1	0	0	1	1	1
2a5	1	0	0	0	0	0	1	1	1	1
2a6	0	0	0	0	0	0	0	1	1	1
2a7	0	0	0	0	0	0	0	0	0	0
2a8	0	0	0	0	0	0	0	1	1	1
2a9	0	0	0	0	0	0	0	0	0	0
2a10	0	0	1	1	1	0	0	0	0	0
2a11	0	0	1	0	1	0	1	1	1	1
2a12	1	0	0	0	1	0	0	1	0	0
2a13	0	0	0	0	0	0	0	1	0	0

Isolate No.	Ampicillin	Tetracycline	Erythromycin	Gentamycin	Rifampicin	Kanamycin	Nalidixic acid	Ciprofloxacin	Sulfadiazine	Vancomycin
					CULTURE COLLECTION					
2a14	1	0	0	1	1	0	0	1	1	1
2a15	0	0	0	0	0	0	0	1	1	1
2a16	0	0	0	0	1	0	0	0	0	0
2a17	0	0	0	0	0	0	0	0	0	0
2a18	1	0	0	0	1	0	0	1	0	0
2a19	1	0	0	0	0	0	0	1	1	1
2a20	0	0	0	0	1	0	0	0	1	1
2a21	0	0	0	0	0	0	0	1	1	1
2a22	0	1	1	0	1	1	0	0	1	0
2a23	0	0	1	1	0	1	0	1	0	1
2a24	0	0	0	0	0	0	1	0	1	1
2a25	0	0	0	0	0	0	0	0	0	0
2a26	0	0	0	0	1	0	0	1	0	0
2a27	0	0	0	0	0	0	0	0	0	0
2a28	0	0	0	0	0	0	0	0	0	0
2a29	0	0	0	0	1	0	0	1	0	0
2a30	1	0	0	0	1	0	0	1	1	0
2a31	1	0	0	0	0	0	0	1	0	1
2a32	1	0	0	1	0	1	1	1	1	1
2a33	0	0	0	0	0	0	0	1	0	1
2a34	1	0	0	0	1	0	0	1	0	0
2a35	0	0	0	0	0	0	0	1	1	0
2a36	0	0	1	0	0	0	0	0	0	1
2a37	0	0	0	0	1	0	0	0	0	0
2a38	1	0	0	0	0	0	0	0	1	1
2a39	0	0	1	1	1	0	0	0	1	1
2a40	1	0	1	1	1	0	0	0	0	0
2a41	0	0	0	0	0	0	0	0	0	0
2a42	0	0	0	0	1	0	0	1	1	1

Isolate No.	Ampicillin	Tetracycline	Erythromycin	Gentamycin	Rifampicin	Kanamycin	Nalidixic acid	Ciprofloxacin	Sulfadiazine	Vancomycin
CULTURE COLLECTION										
2a43	0	0	0	0	0	0	0	1	0	0
2a44	0	0	1	1	1	0	0	0	0	1
2a45	0	0	0	0	1	0	0	0	0	0
2a46	0	0	0	0	0	1	0	1	0	0
2a47	0	1	1	0	1	1	0	1	0	0
2a48	1	0	1	0	1	1	1	1	1	1
2a49	1	0	0	0	1	1	1	1	1	0
2a50	0	1	1	0	1	1	1	0	1	1
2b1	1	0	0	0	0	1	0	0	1	1
2b2	0	0	0	0	0	0	0	1	0	0
2b3	1	0	0	0	0	0	0	1	1	1
2b4	0	1	0	0	0	0	0	0	0	0
2b5	0	0	0	0	0	1	0	0	0	0
2b6	0	0	0	0	0	0	0	0	0	0
2b7	0	0	0	0	0	0	0	1	0	0
2b8	1	0	0	0	1	0	0	1	0	0
2b9	1	0	0	1	0	0	0	0	1	1
2b10	1	0	0	0	0	0	0	0	0	0
2b11	1	0	0	0	1	0	0	1	1	1
2b12	0	0	0	0	0	0	0	1	0	0
2b13	0	0	0	0	0	0	0	1	1	0
2b14	0	0	0	0	0	0	1	0	0	0
2b15	0	0	0	0	1	0	0	1	0	0
2b16	0	0	0	0	0	0	0	0	0	0
2b17	0	0	0	0	0	0	0	0	0	0
2b18	1	0	0	0	0	0	0	1	1	0
2b19	0	0	0	0	0	0	0	0	0	0
2b20	0	0	0	0	0	0	0	0	0	0
2b21	0	0	0	0	0	0	0	1	0	0

Isolate No.	Ampicillin	Tetracycline	Erythromycin	Gentamycin	Rifampicin	Kanamycin	Nalidixic acid	Ciprofloxacin	Sulfadiazine	Vancomycin
					CULTURE COLLECTION					
2b22	1	0	0	0	0	0	0	0	0	1
2b23	1	0	0	0	1	0	0	0	1	0
2b24	1	0	0	0	1	0	0	1	0	1
2b25	0	0	0	0	0	0	0	0	0	0
2b26	1	0	0	0	0	1	0	1	1	0
2b27	1	0	0	0	0	0	0	1	1	0
2b28	0	0	0	0	0	0	0	0	0	0
2b29	1	0	0	0	1	0	0	0	0	0
2b30	0	0	0	0	1	0	0	0	0	0
2b31	0	0	0	0	1	0	0	0	0	0
2b32	0	0	0	0	1	0	0	1	0	0
2b33	1	0	0	0	1	0	1	1	1	0
2b34	0	0	0	0	0	0	0	1	1	1
2b35	0	0	0	0	1	0	0	0	1	0
2b36	0	0	0	0	0	0	0	0	0	0
2b37	1	0	0	0	1	0	0	0	0	0
2b38	0	0	0	0	0	0	0	0	1	0
2b39	1	0	0	0	0	0	1	1	1	1
2b40	1	0	0	0	0	0	1	1	1	1
2b41	0	0	0	0	0	0	0	0	0	1
2b42	0	0	1	0	0	0	1	1	1	1
2b43	0	0	0	0	0	0	0	0	0	0
2b44	0	0	0	0	0	0	0	0	0	0
2b45	1	0	1	0	0	0	0	1	0	0
3a1	0	0	0	0	0	0	0	0	0	1
3a2	0	0	0	0	1	0	0	1	1	1
3a3	1	0	1	1	1	1	0	1	1	1
3a4	0	1	0	0	1	1	0	1	1	1
3a5	0	0	1	0	1	0	0	1	1	1

Isolate No.	Ampicillin	Tetracycline	Erythromycin	Gentamycin	Rifampicin	Kanamycin	Nalidixic acid	Ciprofloxacin	Sulfadiazine	Vancomycin
CULTURE COLLECTION										
3a6	1	0	1	0	1	0	0	1	1	1
3a7	1	0	0	0	0	0	0	1	1	1
3a8	0	0	0	0	1	0	0	0	0	0
3a9	0	0	0	0	0	0	0	0	0	0
3a10	0	0	0	0	1	0	0	0	0	0
3a11	1	0	1	0	0	0	1	1	1	1
3a12	1	0	0	0	0	0	1	1	1	1
3a13	0	0	0	0	0	0	0	1	1	1
3a14	0	0	1	0	0	0	1	1	1	1
3a15	1	0	1	0	1	0	1	1	1	1
3a16	1	0	0	0	1	0	0	0	0	0
3a17	0	0	0	0	0	0	0	0	0	0
3a18	0	0	0	0	1	0	0	0	0	0
3a19	1	0	0	0	1	0	0	1	1	1
3a20	1	0	1	0	1	0	0	1	1	1
3a21	0	0	1	0	0	0	0	1	0	1
3a22	0	0	1	0	0	0	1	1	1	1
3a23	0	0	0	0	1	0	0	0	0	0
3a24	1	0	0	0	1	0	0	1	1	1
3a25	1	0	1	1	1	1	0	1	1	0
3a26	0	0	0	0	0	0	0	0	0	0
3a27	1	0	1	0	1	0	0	0	0	1
3a28	0	0	0	0	0	0	0	0	0	0
3a29	0	0	0	0	0	0	0	1	1	1
3a30	0	0	0	0	0	0	0	1	1	1
3a31	1	0	1	0	0	0	0	1	1	1
3a32	0	0	0	0	0	0	0	0	0	0
3a33	1	0	0	0	0	0	0	0	0	0
3a34	1	0	0	0	1	0	0	1	1	1

Isolate No.	Ampicillin	Tetracycline	Erythromycin	Gentamycin	Rifampicin	Kanamycin	Nalidixic acid	Ciprofloxacin	Sulfadiazine	Vancomycin
CULTURE COLLECTION										
3a35	1	0	1	0	0	0	0	0	0	1
3a36	1	0	0	0	1	0	0	1	1	1
3a37	0	0	0	0	0	0	0	0	0	0
3a38	0	0	1	1	0	1	0	1	0	0
3a39	0	0	0	0	0	0	0	0	0	1
3a40	0	0	1	0	0	0	0	1	1	1
3a41	0	0	0	0	0	0	0	1	1	1
3a42	1	0	1	1	1	1	1	1	1	1
3a43	1	0	1	0	0	0	0	1	1	0
3a44	0	0	0	0	0	0	0	0	0	1
3a45	1	0	0	0	0	0	0	0	1	0
3a46	0	0	0	0	1	0	0	0	0	1
3a47	1	0	0	0	0	0	0	1	1	1
3a48	1	0	0	0	1	0	0	1	1	1
3a49	1	0	0	0	0	0	0	1	1	1
3a50	1	0	0	1	0	0	0	1	1	1
3b1	1	0	0	0	1	0	0	1	1	1
3b2	1	0	1	1	1	0	1	0	1	1
3b3	0	0	0	0	1	0	0	1	1	1
3b4	0	1	1	1	1	1	1	1	1	1
3b5	1	0	0	0	1	1	0	1	1	0
3b6	0	0	0	0	0	0	0	0	0	0
3b7	1	0	0	0	0	0	0	1	1	1
3b8	1	0	1	0	0	0	0	1	1	1
3b9	1	0	0	0	0	0	0	0	0	0
3b10	1	0	0	0	1	0	0	0	1	1
3b11	1	0	1	1	1	1	0	1	0	1
3b12	1	1	0	0	1	1	0	0	0	1
3b13	0	0	0	0	1	0	0	1	1	1

Isolate No.	Ampicillin	Tetracycline	Erythromycin	Gentamycin	Rifampicin	Kanamycin	Nalidixic acid	Ciprofloxacin	Sulfadiazine	Vancomycin
CULTURE COLLECTION										
3b14	0	0	0	0	0	0	0	0	1	0
3b15	1	0	0	0	0	0	0	1	1	0
3b16	0	0	0	0	1	0	0	1	1	0
3b17	0	0	0	0	0	0	0	0	0	0
3b18	0	0	1	1	1	1	0	1	0	0
3b19	0	0	0	0	0	0	0	0	0	0
3b20	1	0	0	0	1	0	0	1	1	1
3b21	1	0	0	0	1	0	0	1	1	1
3b22	0	0	1	1	1	0	0	1	0	1
3b23	0	0	0	0	1	0	0	1	1	1
3b24	0	0	0	0	0	0	0	0	0	0
3b25	0	0	0	0	1	0	0	1	1	0
3b26	0	0	0	0	0	0	0	1	1	0
3b27	1	0	0	0	0	0	1	0	1	1
3b28	0	0	0	0	0	0	1	1	0	1
3b29	0	0	0	0	0	0	0	0	1	0
3b30	1	0	0	0	0	0	0	0	0	1
3b31	1	0	0	0	1	0	1	1	1	1
3b32	0	0	0	0	1	0	0	1	1	1
3b33	1	0	0	0	0	0	0	1	1	0
3b34	1	0	0	0	0	0	0	1	1	0
3b35	0	0	0	0	1	1	1	0	1	1

Species	Ampicillin	Tetracycline	Erythromycin	Gentamycin	Rifampicin	Kanamycin	Nalidixic acid	Ciprofloxacin	Sulfadiazine	Vancomycin
					IDENTIFIED SPECIES					
Rouxiella chamberiensis	1	0	1	1	1	1	1	1	1	1
Pandoraea apista	1	0	1	1	1	1	1	1	1	1
Pandoraea terrae	0	1	1	1	1	1	1	1	1	1
Serratia marcescens	1	0	1	0	1	1	1	1	1	1
Rouxiella chamberiensis	1	1	1	0	1	0	1	1	1	1
Paraburkholderia phytofirmans	1	0	1	1	1	1	0	1	1	1

Table S 2: CARD data.

gene	seq accession	ARO ID	ARAI (ppm)
	EFFLUX		
MuxB	NP_251217.1	ARO:3004074	0,09284924
mdtB	AAC75136.1	ARO:3000793	0,06455909
mdtC	AAC75137.1	ARO:3000794	0,06395501
MuxC	NP_251216.1	ARO:3004075	0,03980334
mexN	BAE06006.1	ARO:3003705	0,03865124
macB	AAV85982.1	ARO:3000535	0,04879058
msbA	NP_415434.1	ARO:3003950	0,05395834
novA	AAF67494.2	ARO:3002522	0,0497723
MexF	NP_251184.1	ARO:3000804	0,02825009
sav1866	YP_186749.1	ARO:3000489	0,04914918
tetA(48)	APB03214.1	ARO:3003980	0,06922454
ceoB	AAB58161.1	ARO:3003010	0,02032149
MexB	AAA74437.1	ARO:3000378	0,01955353
TaeA	APB03219.1	ARO:3003986	0,02970345
bcrA	AAA99504.1	ARO:3002987	0,06047777
MexW	AAG07763.1	ARO:3003031	0,01710895
oqxB	YP_001693238	ARO:3003923	0,01535147
adeG	YP_001706893.1	ARO:3000778	0,01514988
emrB	AAC75733.1	ARO:3000074	0,02935501
rosA	AAC60781.1	ARO:3003048	0,03619162
MuxA	NP_251218.1	ARO:3004073	0,03136404
mdsB	NP_459346.1	ARO:3000790	0,01256567
oleC	AAA26793	ARO:3003748	0,03993438
mexQ	BAE06008.1	ARO:3003699	0,01206131
efrB	CDO61516.1	ARO:3003949	0,03446019
srmB	CAA45050.1	ARO:3002828	0,02266
acrB	NP_414995.1	ARO:3000216	0,01174828
oleB	AAA50325.1	ARO:3003036	0,02114981
tetB(46)	AET10445.1	ARO:3004033	0,02065008
mdtA	AAC75135.2	ARO:3000792	0,02834198
AcrF	AAC76298.1	ARO:3000502	0,01122947
carA	AAC32027.1	ARO:3002817	0,02105207
lmrD	ABF66027.1	ARO:3002882	0,01615682
farA	NP_273367.1	ARO:3003961	0,02533388
smeR	AAD51348.1	ARO:3003066	0,04192463
MexI	NP_252896.1	ARO:3000808	0,0091894
tlrC	AAA26832.1	ARO:3002827	0,01686407
farB	NP_273368.1	ARO:3003962	0,01765589
tetB(60)	ANZ79241.1	ARO:3004036	0,01549083
patA	NP_417544.5	ARO:3000024	0,0195281
efrA	CDO61513.1	ARO:3003948	0,01840955
adeJ	AAX14802.1	ARO:3000781	0,00831868
acrD	YP_490697.1	ARO:3000491	0,00835304
OpmB	NP_251215.1	ARO:3004072	0,01673061
mtrA	CCP46065.1	ARO:3000816	0,03595869
lmrC	ABF66011.1	ARO:3002881	0,01410021
smeE	CAC14595.1	ARO:3003056	0,0076604
cmeB	ABS43151.1	ARO:3000784	0,00748212
arlR	YP_001332362.1	ARO:3000838	0,03436745
MexK	AAG07064.1	ARO:3003693	0,00725811
mdtD	AAC75138.1	ARO:3001330	0,01441747
emrY	BAA11237.1	ARO:3000254	0,01322899
arlS	YP_499945.1	ARO:3000839	0,01442731
emrA	BAA16547.1	ARO:3000027	0,01628277
evgS	AAC75429.1	ARO:3000833	0,00525192
MexD	AAB41957.1	ARO:3000801	0,00587738
adeH	YP_001713101.1	ARO:3000779	0,01267973

gene	seq accession	ARO ID	ARAI (ppm)
EFFLUX			
emrK	BAA11236.1	ARO:3000206	0,01693645
smeB	AAD51345.1	ARO:3003052	0,00498763
OprM	NP_249118.1	ARO:3000379	0,01077574
MexA	NP_249116.1	ARO:3000377	0,01319168
Pseudomonas aeruginosa CpxR	SIP52035.1	ARO:3004054	0,02183713
amrB	NP_250708.1	ARO:3002983	0,00450772
MexE	NP_251183.1	ARO:3000803	0,01119623
TriC	NP_248848.1	ARO:3003681	0,00451536
msrB	NP_416292.1	ARO:3002818	0,03336865
lsaC	AEA37904.1	ARO:3003112	0,00920924
baeR	YP_490321.1	ARO:3000828	0,01878238
adeF	YP_001706894.1	ARO:3000777	0,0106868
adeB	YP_002325611.1	ARO:3000775	0,00414261
OprN	NP_251185.1	ARO:3000805	0,00904707
mtrD	NP_274718.1	ARO:3000811	0,00396949
tetA(60)	ANZ79240.1	ARO:3004035	0,00678474
tetA(46)	AET10444.1	ARO:3004032	0,0065713
otrC	AAR96051.1	ARO:3002894	0,01071322
YojI	NP_416715.1	ARO:3003952	0,00671559
mdtN	BAE78084.1	ARO:3003548	0,01010158
opcM	AAC43969.1	ARO:3003037	0,00637119
adeL	ALH22601.1	ARO:3000620	0,00955932
baeS	BAA15934.1	ARO:3000829	0,00657569
msrC	AAK01167.1	ARO:3002819	0,00593537
mexM	BAE06005.1	ARO:3003704	0,00746454
cpxA	NP_312864.1	ARO:3000830	0,00613637
Enterobacter cloacae acrA	ABG77965.1	ARO:3004042	0,00680108
vgaE	CBY88983.1	ARO:3002833	0,00511955
vgaD	ACX92986.2	ARO:3002832	0,00497737

gene	seq accession	ARO ID	ARAI (ppm)
EFFLUX			
abcA	XP_753111.1	ARO:3003942	0,0017438
oqxA	YP_001693237.1	ARO:3003922	0,00643125
OprA	BAM10414.1	ARO:3003039	0,00529777
tcmA	AAA67509.1	ARO:3003554	0,00451247
tet(41)	AAP93922.1	ARO:3000569	0,00602994
AcrE	AAC76297.1	ARO:3000499	0,00597464
ceoA	AAB58160.1	ARO:3003009	0,00562237
smeA	AAD51344.1	ARO:3003051	0,0057067
MexV	AAG07762.1	ARO:3003030	0,00597897
efpA	NP_217362.1	ARO:3003955	0,00383719
smeS	AAD51347.1	ARO:3003067	0,00430522
lsaE	AFU35065.1	ARO:3003206	0,00362421
adeR	ADM92605.1	ARO:3000553	0,00717805
Klebsiella pneumoniae acrA	CAC41008.1	ARO:3004041	0,00442561
mel	YP_002743273.1	ARO:3000616	0,00361682
smeD	CAC14594.1	ARO:3003055	0,00432348
vgaALC	ABH10964.1	ARO:3002830	0,00316342
CRP	BAE77933.1	ARO:3000518	0,00783576
smeF	CAC14596.1	ARO:3003057	0,0034441
tet(30)	AAD09860.1	ARO:3000561	0,00398525
smeC	AAD51346.1	ARO:3003053	0,00327222
vgaB	AAB95639.1	ARO:3000118	0,00275007
rosB	AAC60780.1	ARO:3003049	0,00266547
MexC	AAB41956.1	ARO:3000800	0,00374293
Escherichia coli acrA	NP_414996.1	ARO:3004043	0,00357567
tcr3	BAA07390.1	ARO:3002893	0,0023991
mdsC	NP_459345.2	ARO:3000791	0,00243709
salA	AGN74946	ARO:3003749	0,00202417
tet(39)	AAW66497.1	ARO:3000566	0,00277234

gene	seq accession	ARO ID	ARAI (ppm)
EFFLUX			
MexL	NP_252368.1	ARO:3003710	0,00464617
msrE	YP_724476.1	ARO:3003109	0,00197069
tet(43)	ACS83748.1	ARO:3000573	0,00184204
mdsA	NP_459347.2	ARO:3000789	0,00236964
adeK	AAX14803.1	ARO:3000782	0,00185553
OprJ	AAB41958.1	ARO:3000802	0,00181442
opmE	BAE06009.1	ARO:3003700	0,00173467
mexP	BAE06007.1	ARO:3003698	0,00218217
MexY	BAA34300.1	ARO:3003033	0,00080319
adeI	YP_002320475.1	ARO:3000780	0,00188028
adeS	ADM92606.1	ARO:3000549	0,0021186
evgA	NP_311275.1	ARO:3000832	0,00369229
golS	NP_459349.1	ARO:3000504	0,00485346
MexH	NP_252895.1	ARO:3000807	0,00202009
otr(B)	AAD04032.1	ARO:3002892	0,00130701
tetB(48)	APB03215.1	ARO:3003981	0,00271529
OpmH	NP_253661.1	ARO:3003682	0,00141846
mtrE	CAA64891.1	ARO:3000812	0,00143921
mdtG	YP_489321.1	ARO:3001329	0,0012923
TriA	NP_248846.1	ARO:3003679	0,00131614
adeN	AGV28567.1	ARO:3000559	0,00226956
PmpM	NP_250052.1	ARO:3004077	0,0009839
tet(C)	AAO16462.1	ARO:3000167	0,00240391
tet(42)	ACD35503.1	ARO:3000572	0,00102885
qepA	AEZ36150.1	ARO:3000448	0,00082772
sdiA	NP_460903.1	ARO:3000826	0,00176236
lsaA	AAT46077.1	ARO:3000300	0,00083769
mdtP	BAE78082.1	ARO:3003550	0,00085486
lsaB	NP_899166.1	ARO:3003111	0,00082436

gene	seq accession	ARO ID	ARAI (ppm)
EFFLUX			
MexJ	NP_252367.1	ARO:3003692	0,00105777
pp-flo	BAA07072.1	ARO:3002812	0,00100699
tet(J)	AAD12753.1	ARO:3000177	0,00094626
lrfA	AAC43550.1	ARO:3003967	0,00070126
mtrC	NP_274719.1	ARO:3000810	0,00085786
tet(33)	CAD12227.1	ARO:3000478	0,00085416
OpmD	NP_252897.1	ARO:3000809	0,00070195
qacH	AAZ42322.1	ARO:3003836	0,00310772
adeC	ALX99516.1	ARO:3003811	0,00069778
msrA	CCQ20328.1	ARO:3000251	0,00279538
TolC	ACN32294.1	ARO:3000237	0,00064378
MexX	BAA34299.1	ARO:3003034	0,00072984
emrD	EFF04178.1	ARO:3000309	0,00061452
TriB	NP_248847.1	ARO:3003680	0,00065102
abeS	YP_002325052.1	ARO:3000768	0,0020731
mgrA	YP_003281576.1	ARO:3000815	0,0015372
tet(B)	BAC67143.1	ARO:3000166	0,00056351
tet(H)	CAA75663.1	ARO:3000175	0,00053595
macA	AAV85981.1	ARO:3000533	0,00051733
tet(Y)	AAC72341.1	ARO:3000182	0,00051865
cmeA	ABS43901.1	ARO:3000783	0,0005052
tet(E)	AAA71915.1	ARO:3000173	0,0004578
pmrA	NP_358469.1	ARO:3000822	0,00045016
tet(Z)	AAD25063.1	ARO:3000183	0,00046775
vgaA	AGN33258.1	ARO:3002829	0,00032189
lmrB	KIX81495.1	ARO:3002813	0,00030367
tap	CAA03986.1	ARO:3000343	0,00035416
abeM	BAD89844.2	ARO:3000753	0,00029746
adeA	YP_002325610.1	ARO:3000774	0,00032189

gene	seq accession	ARO ID	ARAI (ppm)
EFFLUX			
AcrS	AAC76296.1	ARO:3000656	0,00055307
cmeC	BAO79432.1	ARO:3000785	0,00024731
Pseudomonas aeruginosa emrE	NP_253677.1	ARO:3004038	0,00110614
Escherichia coli emrE	CAA77936.1	ARO:3004039	0,00105346
MdfA	AFH35853.1	ARO:3001328	0,0002685
tet(31)	CAC80727.1	ARO:3000476	0,00025437
tet(G)	AAD25538.1	ARO:3000174	0,00026673
amrA	NP_250709.1	ARO:3002982	0,00019021
emrR	NP_417169.1	ARO:3000516	0,00042797
ykkC	CAB13166.1	ARO:3003063	0,00051733
emeA	BAC11911.1	ARO:3003551	0,00013269
robA	AFK13827.1	ARO:3000825	0,00018044
tet(V)	AAB84282.1	ARO:3000181	0,00012445
efmA	BAG75524.1	ARO:3003954	9,4763E-05
mdtO	BAE78083.1	ARO:3003549	5,9383E-05
qacA	BAJ09383.1	ARO:3003046	7,8907E-05
ykkD	CAB13167.1	ARO:3003064	0,00038627
mdtM	AAC77293.1	ARO:3001214	8,4791E-05
mefC	BAL43360	ARO:3003745	8,5416E-05
H-NS	NP_309766.1	ARO:3000676	0,00021146
mdtH	AAC74149.2	ARO:3001216	7,2065E-05
patB	NP_358969.1	ARO:3000025	7,4666E-05
tet(D)	CAE51745.1	ARO:3000168	0,00020993
norB	CCQ22388.1	ARO:3000421	4,9734E-05
blt	AAC36944.1	ARO:3003006	4,3455E-05
bmr	AAA22277.1	ARO:3003007	4,4684E-05
cmeR	YP_002343805.1	ARO:3000526	8,2772E-05
hp1181	NP_207972.1	ARO:3003964	3,9237E-05
hmrM	WP_014550864.1	ARO:3003953	2,4974E-05

gene	seq accession	ARO ID	ARAI (ppm)
EFFLUX			
MdtK	AML99881.1	ARO:3001327	2,4447E-05
mdtL	AAC76733.1	ARO:3001215	2,9637E-05
mefB	ACJ63262.1	ARO:3003107	2,8333E-05
tet(35)	AAK37619.1	ARO:3000481	3,1404E-05
tetA(P)	AAA20116.1	ARO:3000180	2,7591E-05
floR	AAG16656.1	ARO:3002705	1,4342E-05
gadW	ANK04027.1	ARO:3003838	2,3942E-05
marA	YP_489794.1	ARO:3000263	4,5622E-05
MexG	NP_252894.1	ARO:3000806	3,9149E-05
qacB	AAQ10694.1	ARO:3003047	1,1272E-05
tet(A)	YP_007503840.1	ARO:3000165	1,4449E-05
vgaC	AMP35312.1	ARO:3002831	7,6238E-05

gene	seq accession	ARO ID	ARAI (ppm)
ANTIBIOTIC RESISTANCE GENES			
dfrE	AAD01867.1	ARO:3002875	0,04131574
parY	AAO47226.2	ARO:3003318	0,03759951
vanRF	AAR84672.1	ARO:3002925	0,01509965
vanRM	ACL82957.1	ARO:3002928	0,01371092
vanRO	AHA41505.1	ARO:3002930	0,01225951
tetB(P)	AAA20117.1	ARO:3000195	0,01174806
dfrA3	AAA25550.1	ARO:3003105	0,01015747
vanRI	WP_011461303	ARO:3003728	0,00974
arnA	NP_252244	ARO:3002985	0,00884032
vanRE	AAL27445.1	ARO:3002924	0,00807119
iri	AAB41059.1	ARO:3002884	0,00772943
FEZ-1 beta-lactamase	CAB96921.1	ARO:3000606	0,0077254
vanRC	AAF86641.1	ARO:3002922	0,00657161
vanRA	AAA65953.1	ARO:3002919	0,00601979
PmrF	AAC75314.1	ARO:3003578	0,00599198
arr-1	AAC05822.1	ARO:3002846	0,00551043
tetT	AAF01499.1	ARO:3000193	0,00533124
vanHD	AAM09850.1	ARO:3002944	0,00527384
vanRB	AAB05622.1	ARO:3002921	0,00521465
tet36	CAD55718.1	ARO:3000197	0,00487968
otr(A)	CAA37477.1	ARO:3002891	0,00465796
desR	AAC68679.1	ARO:3001293	0,00464097
tet34	BAB59035.1	ARO:3002870	0,00459009
VgbC	APB03225.1	ARO:3003990	0,00438469
vanSN	AEP40504.1	ARO:3002940	0,00415453
ileS	YP_003971446	ARO:3003730	0,0041087
vanRL	ABX54691.1	ARO:3002927	0,00401319
vanHO	AHA41499.1	ARO:3002948	0,00397306
sul2	AAL59753.1	ARO:3000412	0,00386983

gene	seq accession	ARO ID	ARAI (ppm)
ANTIBIOTIC RESISTANCE GENES			
vanSI	WP_011461302.1	ARO:3003726	0,00382821
rphB	APB03222.1	ARO:3003992	0,0037622
oleR	AAC12650.1	ARO:3001297	0,00361657
basS	AEX49906.1	ARO:3003583	0,00344971
vanSA	AAA65954.1	ARO:3002931	0,00324406
vanHF	AAF36802.1	ARO:3002945	0,00318493
cat	AAA22081.1	ARO:3002670	0,00317999
bacA	AAC76093.1	ARO:3002986	0,00301376
vatB	AAA86871.1	ARO:3002841	0,00300635
vanSO	AHA41504.1	ARO:3002941	0,00295581
vanRN	AEP40503.1	ARO:3002929	0,00292222
dfrC	AAO04716.1	ARO:3002865	0,00269909
vanSG	ABA71728.1	ARO:3002937	0,00266085
vatF	AAF63432	ARO:3003744	0,00246444
PmrE	AAC75089.1	ARO:3003577	0,00241917
vanO	AHA41500.1	ARO:3002913	0,0024114
vanHB	AAB05626.1	ARO:3002943	0,00240372
vanRG	ABA71727.1	ARO:3002926	0,00236693
tet44	CBH51823.1	ARO:3000556	0,00232667
LlmA 23S ribosomal RNA methyltransferase	APB03216.1	ARO:3003982	0,00226109
tetX	AAA27471.1	ARO:3000205	0,00210557
cmlv	AAB36568.1	ARO:3002700	0,00209968
rphA	AIA08936.1	ARO:3000444	0,00205639
sul3	ACJ63260.1	ARO:3000413	0,00200479
vatE	AAF86220.1	ARO:3002844	0,00197648
dfrA26	CAL48457.1	ARO:3002857	0,00196301
vanG	ABA71731.1	ARO:3002909	0,00195902
tetM	CAJ67339.1	ARO:3000186	0,00170467
vanD	AAM09849.1	ARO:3000005	0,00168923

gene	seq accession	ARO ID	ARAI (ppm)
ANTIBIOTIC RESISTANCE GENES			
vanL	ABX54687.1	ARO:3002910	0,00164358
tetQ	CAA79727.1	ARO:3000191	0,00164032
vanHM	ACL82960.1	ARO:3002947	0,00156404
dfrA20	CAE53424.1	ARO:3003016	0,00150851
vanC	AAA24786.1	ARO:3000368	0,00150341
RbpA	ADV91011.1	ARO:3000245	0,00147393
tet32	CAC41371.1	ARO:3000196	0,00145078
LRA-19	ACH59005.1	ARO:3002513	0,00137126
Brucella suis mprF	Q8FW76	ARO:3003772	0,0013695
VgbA	AAA98349.1	ARO:3001307	0,00135647
BJP-1 beta-lactamase	BAL75272.1	ARO:3000856	0,00134012
vanSL	ABX54692.1	ARO:3002938	0,00133709
apmA	CBL58181.1	ARO:3003918	0,00128992
Rm3 beta-lactamase	AGU01679.2	ARO:3003894	0,00126625
clbA	YP_001420189.1	ARO:3002814	0,00121194
APH(9)-Ia	AAB58447.1	ARO:3002662	0,00119032
clbB	YP_002773985.1	ARO:3002815	0,00118246
VgbB	AAC61670.1	ARO:3001308	0,00115881
vanSM	ACL82958.1	ARO:3002939	0,00114315
tsnr	CCP44409.1	ARO:3003060	0,00113653
dfrA12	AHW42429.1	ARO:3002858	0,00112369
vanRD	AAM09851.1	ARO:3002923	0,00109887
dfrA22	CAX16467.1	ARO:3003018	0,00108858
tetW	ACA23185.1	ARO:3000194	0,00108808
dfrG	BAE15963.1	ARO:3002868	0,00101835
vanSF	AAR84673.1	ARO:3002936	0,00096832
FosX	CWV56762.1	ARO:3000198	0,00095841
dfrF	AAD01868.1	ARO:3002867	0,0009539
vanB	AHH83938.1	ARO:3000013	0,00094873

gene	seq accession	ARO ID	ARAI (ppm)
ANTIBIOTIC RESISTANCE GENES			
y56 beta-lactamase	AAX55643.1	ARO:3003558	0,00094597
tetS	AAA25293.1	ARO:3000192	0,00094006
vatD	AAK84316.1	ARO:3002843	0,00093727
LRA-2	ACH58985.1	ARO:3002485	0,00092091
vanHA	AAA65955.1	ARO:3002942	0,00091769
vanSE	AAL27446.1	ARO:3002935	0,00090887
vanF	AAF36803.1	ARO:3002908	0,00089529
blaI	ABU39978.1	ARO:3000160	0,00087371
THIN-B beta-lactamase	CAC33832.1	ARO:3000851	0,00086177
vatH	ACX92987.1	ARO:3002845	0,00085838
TLE beta-lactamase	AAA19882.1	ARO:3003562	0,00082772
LRA-1	ACH58980.1	ARO:3002482	0,00082492
dfrK	CBL80435.1	ARO:3002869	0,00081757
sul1	AEJ33969.1	ARO:3000410	0,00080992
CAU-1 beta-lactamase	CAC87665.1	ARO:3000855	0,00080194
vanA	AAA65956.1	ARO:3000010	0,00079394
FosA	NP_249820.1	ARO:3000149	0,00077254
PEDO-1	AJP77059	ARO:3003670	0,00076984
vanE	AAL27442.1	ARO:3002907	0,00072426
catB6	CAA11473.1	ARO:3002678	0,00071736
SPG-1	AJP77080	ARO:3003720	0,00069122
SRT-1	BAA23130.1	ARO:3002493	0,00067444
LRA-17	ACH58994.1	ARO:3002512	0,00067006
VatI	APB03220.1	ARO:3003987	0,00066218
LRA-8	ACH58988.1	ARO:3002487	0,00064589
catB8	YP_009077553.1	ARO:3002680	0,00063459
mfpA	CCP46182.1	ARO:3003035	0,00063323
AER-1	AAC09015.1	ARO:3002481	0,00062896
APH(6)-Ic	CAA25854.1	ARO:3002659	0,0006099

gene	seq accession	ARO ID	ARAI (ppm)
ANTIBIOTIC RESISTANCE GENES			
catB10	CAI47810.1	ARO:3003110	0,000607
AIM-1	CAQ53840.1	ARO:3000853	0,00059279
TLA-2	CAG27800.1	ARO:3003203	0,00059084
vanXO	AHA41501.1	ARO:3002954	0,00057367
LRA-10	ACH58999.1	ARO:3002489	0,00057168
MSI-1	AJP77057	ARO:3003718	0,00056071
LRA-13	ACH58991.1	ARO:3002484	0,0005423
arr-4	ABV26705.1	ARO:3002849	0,00054078
AAC(2')-Ic	CCP42991.1	ARO:3002525	0,00051218
vanM	ACL82961.1	ARO:3002911	0,00050677
EXO beta-lactamase	AAA26775.1	ARO:3003564	0,00049821
GOB-1 beta-lactamase	AAF04458.1	ARO:3000850	0,00047951
CTX-M-151	BAP34782.1	ARO:3002008	0,00047648
cipA	WP_015735625.1	ARO:3003907	0,00046888
rgt1438	AFO53532.1	ARO:3002883	0,00045112
clbC	YP_174574.1	ARO:3002816	0,00044697
vanI	AEP96393.1	ARO:3003723	0,00044307
LRA-12	ACH58990.1	ARO:3002511	0,00043505
vanSC	AAF86642.1	ARO:3002933	0,00043335
BEL-1	AAZ04368.1	ARO:3002385	0,00042995
arr-5	ABV26707.1	ARO:3002850	0,0004249
vanTG	ABA71733.1	ARO:3002972	0,00042316
Erm(37)	CCP44758.1	ARO:3000392	0,0004208
ESP-1	AJP77085	ARO:3003717	0,00041957
TLA-1	AAD37403.1	ARO:3003202	0,00040595
vanN	AEP40500.1	ARO:3002912	0,00040541
CTX-M-38	AAV70602.1	ARO:3001900	0,00039822
vatA	AAA26683.1	ARO:3002840	0,00039685
CARB-5	AAF61417.1	ARO:3002244	0,00038886

gene	seq accession	ARO ID	ARAI (ppm)
ANTIBIOTIC RESISTANCE GENES			
oleI	ABA42118.2	ARO:3000866	0,00038263
OXA-29	CAC35728.1	ARO:3001424	0,0003703
SHV-137	AEI83430.1	ARO:3001181	0,00036466
r39 beta-lactamase	CAA37699.1	ARO:3003565	0,00036213
catB2	NP_848167.1	ARO:3002675	0,00035868
catB7	NP_249397.1	ARO:3002679	0,0003553
dfrA24	CAI99385.1	ARO:3002856	0,00034451
OXA-243	AFQ90085.1	ARO:3001610	0,00033711
catB3	YP_006965431.1	ARO:3002676	0,00033109
vatC	AAC61671.1	ARO:3002842	0,00032797
PEDO-2	AJP77071	ARO:3003714	0,00032528
vanSB	AAB05623.1	ARO:3002932	0,00032405
dfrA13	CAA90683.1	ARO:3003012	0,00031604
arr-3	ACD56151.1	ARO:3002848	0,00030902
FosC2	BAJ10053.1	ARO:3002874	0,00030726
SRT-2	AAS07017.3	ARO:3002494	0,00030656
murA	CCE36834	ARO:3003784	0,00030495
KPC-7	ACE62798.1	ARO:3002317	0,00029662
PEDO-3	AJP77076	ARO:3003715	0,00029337
SMB-1 beta-lactamase	BAL14456.1	ARO:3000854	0,0002897
vanXI	WP_015943580.1	ARO:3003725	0,00028683
OXA-129	CAP69660.1	ARO:3001811	0,00028211
OXA-18	AAB58555.1	ARO:3001413	0,0002739
LRA-18	ACH58997.1	ARO:3002492	0,00027019
BahA	APB03218.1	ARO:3003984	0,00026967
AAC(6')-Iak	BAO21229.1	ARO:3003199	0,00026509
ACC-3	AAF86697.1	ARO:3001817	0,00026127
vanXA	AAA65957.1	ARO:3002949	0,00025815
vanXB	AAB05628.1	ARO:3002950	0,00025815

gene	seq accession	ARO ID	ARAI (ppm)
ANTIBIOTIC RESISTANCE GENES			
JOHN-1 beta-lactamase	AAK38324.1	ARO:3000840	0,00025699
catB9	AAL68645.1	ARO:3002681	0,0002495
AAC(6')-Iy	AAF03531.1	ARO:3002569	0,00023975
AAC(6')-Ic	AAA26549.1	ARO:3002549	0,00023811
Sed1 beta-lactamase	AAK63223.1	ARO:3003561	0,00023569
VEB-3	AAS48620.1	ARO:3002372	0,00023254
vanXM	ACL82962.1	ARO:3002953	0,00022947
Bcl	CAA29819.1	ARO:3002877	0,00022722
TEM-141	AAX56615.1	ARO:3001004	0,00022285
PmrC	BAE78116.1	ARO:3003576	0,00022244
AAC(6')-32	ABR10839.1	ARO:3002586	0,00022043
OXA-12	AAA83417.1	ARO:3001407	0,00021947
KPC-2	AAK70220.1	ARO:3002312	0,00021752
L1 beta-lactamase	CAB75346.1	ARO:3000582	0,00021605
vanZA	AAA65959.1	ARO:3002962	0,00021593
AAC(2')-Ib	AAC44793.1	ARO:3002524	0,00020799
myrA	BAA03674.1	ARO:3001300	0,00019641
AAC(6')-IIb	AAA25680.1	ARO:3002595	0,00019314
arr-8	AGC29882.1	ARO:3002853	0,00019314
OXA-209	AEM66528.1	ARO:3001809	0,00019032
OXA-31	AAK52604.1	ARO:3001425	0,00018894
PER-2	CAA63714.1	ARO:3002364	0,00018812
AAC(3)-Ic	CAD53575.1	ARO:3002531	0,00018571
Erm(34)	AAP74657.1	ARO:3000600	0,00018557
cfrA	CAL64019.1	ARO:3003441	0,00018262
dfrA14	ACI32877.1	ARO:3002859	0,00018106
ACC-4	ABP49606.1	ARO:3001818	0,00018013
NPS beta-lactamase	CAA33795.1	ARO:3003563	0,00017797
FosK	BAO79518.1	ARO:3003207	0,00017558

gene	seq accession	ARO ID	ARAI (ppm)
ANTIBIOTIC RESISTANCE GENES			
OXA-43	CAD32565.1	ARO:3001770	0,00017231
FosA5	AJE60855.1	ARO:3003209	0,00016674
FosA2	ACC85616.1	ARO:3002804	0,00016437
VCC-1	ALU64000	ARO:3003713	0,00016321
Bacillus subtilis mprF	CAX52582.1	ARO:3003324	0,00016245
SHV-41	AAN04883.1	ARO:3001099	0,00016207
SHV-70	AAY42633.1	ARO:3001124	0,00016207
QnrA5	AAZ04784.1	ARO:3002711	0,00015947
CTX-M-122	AFA51700.1	ARO:3001981	0,00015929
vanWI	WP_005813024.1	ARO:3003724	0,00015534
vanKI	WP_011461306	ARO:3003727	0,00015451
tetO	AAA23033.1	ARO:3000190	0,00015415
ROB-1	CAA37052.1	ARO:3002995	0,00015198
OXA-5	CAA41211.1	ARO:3001400	0,0001519
PER-3	AAU89132.1	ARO:3002365	0,00015049
TLA-3	WP_059512353.1	ARO:3003204	0,00015001
TEM-43	AAC32889.2	ARO:3000912	0,00014181
TEM-57	ACJ43254.1	ARO:3000926	0,00014181
ErmD	AAA22597.1	ARO:3000495	0,00014132
CTX-M-64	BAF63422.1	ARO:3001925	0,00013938
cphA7	CAA40386.1	ARO:3003103	0,00013687
MOX-9	WP_042649345	ARO:3002191	0,00013615
AAC(6')-31	CAK55557.1	ARO:3002585	0,00013475
LRA-3	ACH58987.1	ARO:3002510	0,0001343
OXA-198	ADT70779.1	ARO:3001805	0,00013269
Clostridium perfringens mprF	Q0SSM7	ARO:3003773	0,00013238
PER-4	ACE77058.1	ARO:3002366	0,00013168
EF-Tu	CAA67349.1	ARO:3003359	0,00013135
vanXF	AAF36804.1	ARO:3002952	0,00012991

gene	seq accession	ARO ID	ARAI (ppm)
ANTIBIOTIC RESISTANCE GENES			
OXA-60	AAQ08905.1	ARO:3001808	0,00012828
AAC(6')-Iih	CAE50926.1	ARO:3002590	0,00012665
vanTrL	ABX54690.1	ARO:3002974	0,00012665
LRA-5	ACH59002.1	ARO:3002483	0,00012441
chrB	AAS79458.1	ARO:3001302	0,00012416
AQU-1	BAM76830.1	ARO:3002993	0,00012198
OKP-A-9	CAJ19607.1	ARO:3002426	0,00012155
TEM-214	AJO16044.1	ARO:3001391	0,00012155
OXY-2-9	ACV44455.1	ARO:3002404	0,00012029
CPS-1	AJP77054	ARO:3003716	0,00011988
CTX-M-131	AEW46676.2	ARO:3001990	0,00011988
CTX-M-87	ACB41777.1	ARO:3001947	0,00011946
AAC(6')-Iu	AAD03493.1	ARO:3002565	0,00011906
APH(4)-Ia	CAA24743.1	ARO:3002655	0,00011894
vanTC	AAD22403.1	ARO:3002970	0,00011621
arr-2	AAC64366.1	ARO:3002847	0,00011588
arr-7	CAZ48628.1	ARO:3002852	0,00011588
catP	AAB51421.1	ARO:3002686	0,00011196
mecB	BAI83385.1	ARO:3003440	0,00011175
CARB-23	AHJ02283.1	ARO:3003186	0,00011142
vanTE	AAL27444.1	ARO:3002971	0,0001073
dfrD	AAA85213.1	ARO:3002866	0,0001073
MSI-OXA	AJP77058	ARO:3003719	0,00010651
vanWB	AAB05625.1	ARO:3002964	0,00010535
dfrA21	CAP69659.1	ARO:3003017	0,00010535
OCH-3	CAC17623.1	ARO:3002516	0,000104
PDC-7	ACQ82812.1	ARO:3002506	0,00010216
OKP-A-6	CAJ19603.1	ARO:3002423	0,00010129
SHV-140	AEK80394.1	ARO:3001182	0,00010129

gene	seq accession	ARO ID	ARAI (ppm)
ANTIBIOTIC RESISTANCE GENES			
SHV-42	AAN04884.1	ARO:3001100	0,00010129
TEM-105	AAM61953.1	ARO:3000968	0,00010129
TEM-90	AAK30619.1	ARO:3000957	0,00010129
OXA-55	AAR03105.1	ARO:3001813	0,00010024
CTX-M-148	AHX39589.1	ARO:3002006	9,9554E-05
CTX-M-42	AAY84742.1	ARO:3001904	9,9554E-05
CTX-M-58	ABM97538.1	ARO:3001919	9,9554E-05
CTX-M-78	CAQ42481.2	ARO:3001939	9,9554E-05
vanSD	ACM47284	ARO:3002934	9,9469E-05
AAC(6')-Ib-SK	BAD11815.1	ARO:3002593	9,8762E-05
SHV-53	AAT01223.1	ARO:3001110	9,5769E-05
determinant of bleomycin resistance	YP_007652797.1	ARO:3001205	9,5769E-05
AAC(6')-33	AEZ05106.1	ARO:3002587	9,4468E-05
AAC(6')-Ia	AAA98298.1	ARO:3002545	9,3958E-05
BLA1	AAR20595.1	ARO:3000090	9,3755E-05
BlaB beta-lactamase	WP_024565805.1	ARO:3000579	9,3452E-05
DHA-14	AIT76107.1	ARO:3002145	9,1726E-05
OXA-16	AAB97924.1	ARO:3001411	9,018E-05
OXA-61	AAT01092.1	ARO:3001773	9,018E-05
ACC-1	CAB46491.1	ARO:3001815	9,0063E-05
ACC-5	CCK86740.1	ARO:3001819	9,0063E-05
AAC(3)-IV	ABB43029.1	ARO:3002539	8,983E-05
OCH-5	CAC17625.1	ARO:3002518	8,9139E-05
tmrB	CAB12108.2	ARO:3003059	8,8234E-05
AAC(3)-IIb	AAA26548.1	ARO:3002534	8,6157E-05
OXA-22	AAD12233.1	ARO:3001417	8,4585E-05
FosA3	AEG78825.1	ARO:3002872	8,3972E-05
Erm(41)	ABW06859.1	ARO:3000603	8,3972E-05
CARB-22	BAC61820.1	ARO:3003151	8,1895E-05

gene	seq accession	ARO ID	ARAI (ppm)
ANTIBIOTIC RESISTANCE GENES			
SHV-9	AAB37395.2	ARO:3001068	8,132E-05
SHV-89	ABA60809.1	ARO:3001141	8,1036E-05
TEM-157	ABI81768.1	ARO:3001023	8,1036E-05
AAC(6')-If	CAA39038.1	ARO:3002553	8,0473E-05
CTX-M-40	AAV28215.1	ARO:3001902	7,9918E-05
CTX-M-137	BAO37256.1	ARO:3001994	7,9643E-05
CTX-M-160	AJU57235.1	ARO:3003168	7,9643E-05
CTX-M-93	ADN26580.1	ARO:3001953	7,9643E-05
AAC(6')-Ir	AAD03490.1	ARO:3002561	7,9371E-05
KPC-16	AGJ01153.1	ARO:3002326	7,91E-05
KPC-22	AIX87991.1	ARO:3003180	7,91E-05
mecC	WP_000725529.1	ARO:3001209	7,8416E-05
ErmH	AAC32026.1	ARO:3002823	7,7512E-05
LRA-7	ACH58998.1	ARO:3002486	7,6742E-05
ACT-22	AHM76774.1	ARO:3001843	7,6037E-05
FOX-2	CAA71325.1	ARO:3002156	7,5838E-05
AAC(6')-Iz	AAD52985.1	ARO:3002570	7,5739E-05
PER-6	ADD80743.1	ARO:3002368	7,5247E-05
PER-7	AEI54993.1	ARO:3002369	7,5247E-05
MCR-1	AKF16168	ARO:3003689	7,4969E-05
ACC-2	AAF86691.1	ARO:3001816	7,4283E-05
OCH-4	CAC17624.1	ARO:3002517	7,4283E-05
dfrB2	FAA00064.1	ARO:3003021	7,4283E-05
AAC(6')-Isa	BAD10948.2	ARO:3002563	7,381E-05
AAC(3)-Id	AAR21614.1	ARO:3002529	7,3342E-05
PDC-4	ACQ82809.1	ARO:3002501	7,2973E-05
PDC-5	ACQ82810.1	ARO:3002502	7,2973E-05
PDC-8	ACQ82813.1	ARO:3002507	7,2973E-05
lnuD	ABR14060.1	ARO:3002838	7,0659E-05

gene	seq accession	ARO ID	ARAI (ppm)
ANTIBIOTIC RESISTANCE GENES			
TUS-1 beta-lactamase	AAN63648.1	ARO:3000844	7,0089E-05
ErmB	AAF86219.1	ARO:3000375	7,0089E-05
dfrA8	AHV80711.1	ARO:3002863	6,8569E-05
OXA-251	AER57903.1	ARO:3001500	6,5346E-05
OXA-59	CAG15145.1	ARO:3001772	6,4618E-05
AAC(6')-IIa	AAM92464.1	ARO:3002594	6,2979E-05
tlrB conferring tylosin resistance	AAD12162.1	ARO:3001299	6,2079E-05
BEL-2	ACV69996.1	ARO:3002386	6,1421E-05
DHA-12	CDL68900.1	ARO:3002143	6,1151E-05
OXA-62	AAR32134.1	ARO:3001792	6,099E-05
ACT-37	AJG06172.1	ARO:3003172	6,083E-05
CMY-73	ACU00152.1	ARO:3002086	6,083E-05
LEN-22	CAP12350.2	ARO:3002469	6,0777E-05
OKP-B-10	CAJ19619.1	ARO:3002443	6,0777E-05
SHV-25	AAF37209.2	ARO:3001083	6,0777E-05
SHV-37	AAL82593.1	ARO:3001095	6,0777E-05
SHV-48	AAP03063.1	ARO:3001105	6,0777E-05
TEM-146	AAZ14084.2	ARO:3001013	6,0777E-05
CepS beta-lactamase	CAA56561.1	ARO:3003553	6,0671E-05
CMY-9	BAB72158.1	ARO:3002020	6,0671E-05
FOX-3	CAA71947.1	ARO:3002157	6,0671E-05
MOX-2	CAB82578.1	ARO:3002183	6,0671E-05
MOX-3	ACA30419.1	ARO:3002186	6,0671E-05
CTX-M-152	AHY20039.1	ARO:3002009	5,9938E-05
OXY-6-2	CAI43423.1	ARO:3002414	5,9938E-05
CTX-M-126	BAL72196.1	ARO:3001985	5,9732E-05
CTX-M-134	AFX60298.1	ARO:3001992	5,9732E-05
CTX-M-60	CAL80726.1	ARO:3001921	5,9732E-05
CTX-M-7	CAA06312.1	ARO:3001870	5,9732E-05

gene	seq accession	ARO ID	ARAI (ppm)
ANTIBIOTIC RESISTANCE GENES			
CTX-M-74	ACS32293.1	ARO:3001935	5,9732E-05
ErmN	CAA66307.1	ARO:3000592	5,9732E-05
OCH-7	CAC17627.1	ARO:3002520	5,9426E-05
KPC-3	AAL05630.1	ARO:3002313	5,9325E-05
VEB-1b	AAK14294.1	ARO:3003712	5,8134E-05
VEB-5	ABN80430.1	ARO:3002375	5,8134E-05
Erm(42)	CBY77552.1	ARO:3003106	5,7748E-05
vanXD	AAM09852.1	ARO:3003070	5,7367E-05
bcrC	AAA99503.1	ARO:3003250	5,7084E-05
APH(6)-Ia	CAA68516.1	ARO:3002657	5,6619E-05
APH(6)-Ib	CAA29136.1	ARO:3002658	5,6619E-05
gimA	CAA11707.1	ARO:3000463	5,5578E-05
catIII	CAB75601.1	ARO:3002685	5,4404E-05
mphE	WP_010550189.1	ARO:3003071	5,3319E-05
QnrS6	AEG74318.1	ARO:3002795	5,3156E-05
catQ	AAA23215.1	ARO:3002687	5,2914E-05
SFH-1	AAF09244.1	ARO:3000849	4,9522E-05
ANT(9)-Ia	AFU35063.1	ARO:3002630	4,7885E-05
IMP-16	CAE48334.1	ARO:3002207	4,7106E-05
MUS-1 beta-lactamase	AAN63647.1	ARO:3000843	4,7106E-05
mecI	NP_373280.1	ARO:3000124	4,7106E-05
mprF	NP_465220	ARO:3003770	4,6888E-05
DHA-15	AIT76106.1	ARO:3002146	4,5863E-05
DHA-2	AAG36927.1	ARO:3002133	4,5863E-05
ACT-19	AHM76779.1	ARO:3001840	4,5622E-05
ACT-27	AHL39340.1	ARO:3001847	4,5622E-05
OXA-85	AAP69916.1	ARO:3001780	4,5622E-05
BcII	AAA22562.1	ARO:3002878	4,5266E-05
spd	AGW81558.1	ARO:3002631	4,509E-05

gene	seq accession	ARO ID	ARAI (ppm)
ANTIBIOTIC RESISTANCE GENES			
ErmV	AAB51440.1	ARO:3002824	4,4742E-05
OCH-8	ABF50909.1	ARO:3002521	4,457E-05
LCR-1	CAA40146.1	ARO:3002997	4,457E-05
SHV-22	AAF34336.1	ARO:3001080	4,457E-05
ErmO	CAA11706.1	ARO:3001303	4,457E-05
PDC-6	ACQ82811.1	ARO:3002505	4,3784E-05
PDC-9	ACQ82814.1	ARO:3002508	4,3784E-05
OXA-232	AGD91915.1	ARO:3001778	4,3729E-05
OXA-11	CAA80304.1	ARO:3001406	4,3564E-05
APH(3')-IIb	CAA62365.1	ARO:3002645	4,3239E-05
AAC(3)-IIIa	CAA39184.1	ARO:3002536	4,2761E-05
APH(3')-Ib	AAA26412.1	ARO:3002642	4,2761E-05
OXA-258	CCE73593.2	ARO:3001503	4,2292E-05
OXA-347	AET35493.1	ARO:3001777	4,2292E-05
OXA-9	AAA98406.1	ARO:3001404	4,2292E-05
oleD	ABA42119.1	ARO:3000865	4,1885E-05
LEN-15	AAL50725.1	ARO:3002478	4,1534E-05
Erm(36)	AAL68827.1	ARO:3000605	4,1239E-05
BEL-3	ACT09140.1	ARO:3002387	4,0947E-05
AAC(3)-Xa	BAA78619.1	ARO:3002544	4,0803E-05
LEN-2	AAK69830.1	ARO:3002455	4,0518E-05
LEN-21	CAP12349.2	ARO:3002468	4,0518E-05
LEN-5	AAU25807.1	ARO:3002456	4,0518E-05
OKP-A-14	ACL68095.1	ARO:3002431	4,0518E-05
OKP-A-5	CAJ19602.1	ARO:3002422	4,0518E-05
OKP-B-2	CAJ19610.1	ARO:3002435	4,0518E-05
OKP-B-20	CAP12360.2	ARO:3002453	4,0518E-05
OKP-B-9	CAJ19618.1	ARO:3002442	4,0518E-05
SHV-121	AEI83429.1	ARO:3001168	4,0518E-05

gene	seq accession	ARO ID	ARAI (ppm)
ANTIBIOTIC RESISTANCE GENES			
TEM-166	ACI25375.1	ARO:3001032	4,0518E-05
TEM-208	AGL39384.1	ARO:3001385	4,0518E-05
TEM-220	AIW68620.1	ARO:3003158	4,0518E-05
TEM-55	ABB97007.1	ARO:3000924	4,0518E-05
TEM-80	AAM15527.1	ARO:3000947	4,0518E-05
GES-18	AEX59144.1	ARO:3002347	4,0377E-05
GES-21	AFK80745.1	ARO:3002350	4,0377E-05
CARB-1	AAK96394.1	ARO:3002240	4,0236E-05
CARB-6	AAD19217.1	ARO:3002245	4,0236E-05
CARB-9	AAP22374.1	ARO:3002248	4,0236E-05
CTX-M-109	AEM44654.1	ARO:3001969	4,0236E-05
OXY-3-1	AAN28732.1	ARO:3002409	4,0236E-05
AAC(6')-Iad	BAD12078.1	ARO:3002572	4,0236E-05
ANT(6)-Ib	CBH51824.1	ARO:3002629	4,0236E-05
OXY-2-4	CAB42614.1	ARO:3002399	4,0097E-05
OXY-5-1	CAI43417.1	ARO:3002411	3,9959E-05
AAC(6')-Iaa	NP_460578.1	ARO:3002571	3,9959E-05
AAC(6')-Ik	AAA87229.1	ARO:3002558	3,9959E-05
CTX-M-102	ADY02546.1	ARO:3001961	3,9822E-05
CTX-M-13	AAF72531.1	ARO:3001876	3,9822E-05
CTX-M-141	AGN11769.1	ARO:3001999	3,9822E-05
CTX-M-158	AIT97310.1	ARO:3003166	3,9822E-05
CTX-M-159	AJO16046.1	ARO:3003167	3,9822E-05
CTX-M-20	CAC95175.1	ARO:3001882	3,9822E-05
CTX-M-24	AAN38836.1	ARO:3001886	3,9822E-05
CTX-M-39	AAX54694.1	ARO:3001901	3,9822E-05
CTX-M-4	CAA74573.1	ARO:3001867	3,9822E-05
CTX-M-41	AAY43008.1	ARO:3001903	3,9822E-05
CTX-M-5	AAC32890.1	ARO:3001868	3,9822E-05

gene	seq accession	ARO ID	ARAI (ppm)
ANTIBIOTIC RESISTANCE GENES			
CTX-M-6	CAA06311.1	ARO:3001869	3,9822E-05
CTX-M-77	CAQ42480.1	ARO:3001938	3,9822E-05
CTX-M-8	AAF04388.1	ARO:3001871	3,9822E-05
CTX-M-82	ABB59946.1	ARO:3001943	3,9822E-05
CTX-M-83	ACI29345.1	ARO:3001944	3,9822E-05
CTX-M-9	AAF05311.2	ARO:3001872	3,9822E-05
IMI-7	AIS19858.1	ARO:3003177	3,9685E-05
KPC-4	AAU06362.1	ARO:3002314	3,9685E-05
KPC-6	ACB71165.1	ARO:3002316	3,955E-05
blaR1	ABU39979.1	ARO:3000217	3,9415E-05
mphD	ANP63073.1	ARO:3003741	3,9415E-05
SME-2	AAG29813.1	ARO:3002380	3,9415E-05
VEB-7	ACO56763.1	ARO:3002376	3,8756E-05
LEN-14	AAP93847.1	ARO:3002477	3,8119E-05
APH(2")-IIIa	AAB49832.1	ARO:3002636	3,787E-05
lnuC	AAY32951.1	ARO:3002837	3,533E-05
APH(9)-Ib	AAB66655.1	ARO:3002663	3,5115E-05
FomB	BAA32494.1	ARO:3000449	3,5115E-05
vanJ	NP_627787	ARO:3002914	3,5115E-05
Erm(30)	AAC69328.1	ARO:3001265	3,4488E-05
mprF	Q8DWT2	ARO:3003774	3,4203E-05
ErmR	ALX06067.1	ARO:3000594	3,4083E-05
AAC(6')-Iid	CAE50925.1	ARO:3002589	3,1835E-05
AAC(6')-Ib	AFJ11384.1	ARO:3002546	3,1489E-05
AAC(6')-Ib'	AAT74613.1	ARO:3003676	3,1489E-05
dfrA10	AHG97174.1	ARO:3003011	3,0984E-05
CMY-98	AGH70380.1	ARO:3002111	3,0819E-05
DHA-16	AIT76105.1	ARO:3002147	3,0575E-05
ACT-1	AAC45086.2	ARO:3001821	3,0415E-05

gene	seq accession	ARO ID	ARAI (ppm)
ANTIBIOTIC RESISTANCE GENES			
ACT-28	AHL39333.1	ARO:3001848	3,0415E-05
CMY-105	AHL39330.1	ARO:3002117	3,0415E-05
ErmE	CAB60001.1	ARO:3000326	3,0415E-05
CMY-8	AAD50818.2	ARO:3002019	3,0335E-05
OCH-1	CAC04522.1	ARO:3002514	2,9713E-05
OCH-2	CAC17622.1	ARO:3002515	2,9713E-05
OCH-6	CAC17626.1	ARO:3002519	2,9713E-05
PDC-3	ACQ82808.1	ARO:3002500	2,9189E-05
mgtA	ABA28305.2	ARO:3000462	2,7723E-05
QnrB32	AEL00450.1	ARO:3002747	2,7075E-05
QnrB65	AGL43626.1	ARO:3002777	2,7075E-05
CatU	APB03217.1	ARO:3003983	2,6701E-05
catB	AAA73865.1	ARO:3002674	2,6457E-05
catI	CAA23899.1	ARO:3002683	2,6457E-05
QnrC	ACK75961.1	ARO:3002787	2,6217E-05
AAC(6')-Ib8	YP_009062819.1	ARO:3002579	2,5751E-05
QnrB45	AFA52644.1	ARO:3002760	2,5637E-05
IND-7	BAJ05825.1	ARO:3002263	2,4243E-05
CGB-1 beta-lactamase	ABS29619.1	ARO:3000841	2,3942E-05
IND-4	AAG29765.2	ARO:3002260	2,3942E-05
IND-11	ADK25050.1	ARO:3002267	2,3844E-05
ErmC	CAA70208.1	ARO:3000250	2,3746E-05
IMP-27	AEH41427.1	ARO:3002218	2,3553E-05
IMP-51	BAQ56016.1	ARO:3003659	2,3553E-05
Erm(47)	ANZ79476.1	ARO:3003908	2,3553E-05
SLB-1	AAT90846.1	ARO:3003556	2,3269E-05
GIM-1	CAF05908.1	ARO:3000845	2,3176E-05
GIM-2	AIY26289.1	ARO:3003194	2,3176E-05
CEPH-A3	AAM63403.1	ARO:3003093	2,2811E-05
cphA5	AAP69913.1	ARO:3003101	2,2811E-05
cphA8	AAP97129.1	ARO:3003104	2,2811E-05
imiS	CAA71441.1	ARO:3003095	2,2722E-05
OXA-14	AAA93528.1	ARO:3001409	2,2545E-05
SHV-23	AAF34337.1	ARO:3001081	2,2285E-05
aadA25	AET15272.1	ARO:3003197	2,2031E-05
LEN-12	CAG25817.1	ARO:3002462	2,2031E-05
LEN-8	CAG25835.1	ARO:3002458	2,2031E-05
LEN-9	CAG25834.1	ARO:3002459	2,2031E-05
APH(3')-IIa	WP_000572405.1	ARO:3002644	2,1947E-05
APH(3')-IIIa	AGV10830.1	ARO:3002647	2,1947E-05
OXA-45	CAD58780.1	ARO:3001794	2,1947E-05
OXA-145	ACN85419.1	ARO:3001804	2,1864E-05
OXA-48	AAP70012.1	ARO:3001782	2,1864E-05
OXA-240	AFN20670.1	ARO:3001499	2,1782E-05
OXA-246	AHC31001.1	ARO:3001807	2,1782E-05
OXA-256	CCE94500.1	ARO:3001502	2,1782E-05
VIM-10	AAS13761.1	ARO:3002280	2,1782E-05
Erm(35)	AAK07612.1	ARO:3000604	2,1782E-05
ErmF	AAA88675.1	ARO:3000498	2,1782E-05
OXA-42	CAD32564.1	ARO:3001769	2,1539E-05
APH(3')-Ia	CAE51638.1	ARO:3002641	2,138E-05
APH(3")-Ia	CAA37605.1	ARO:3002638	2,1302E-05
OXA-165	ADK35869.1	ARO:3001465	2,1224E-05
LEN-3	AAN05030.1	ARO:3002475	2,1146E-05
LEN-4	AAN05031.1	ARO:3002472	2,1146E-05
OXA-215	AEV91554.1	ARO:3001714	2,1069E-05
OXA-415	AIG94927.1	ARO:3001606	2,1069E-05
aad(6)	AAU10334.1	ARO:3002628	2,0993E-05

gene	seq accession	ARO ID	ARAI (ppm)
ANTIBIOTIC RESISTANCE GENES			
OXA-356	AGW83454.1	ARO:3001543	2,0993E-05
ErmU	CAA44667.1	ARO:3001305	2,0842E-05
AAC(3)-IIIc	AAA25683.1	ARO:3002538	2,0767E-05
OXA-397	AIT76118.1	ARO:3001583	2,0693E-05
AAC(3)-IXa	AAA25334.1	ARO:3002543	2,0619E-05
OKP-A-15	ACN65419.1	ARO:3002432	2,0619E-05
SHV-167	BAM28879.1	ARO:3001204	2,0546E-05
CARB-18	AIL92327.1	ARO:3003174	2,0474E-05
CARB-19	AIL92328.1	ARO:3003175	2,0474E-05
aadK	CAB14620.1	ARO:3002627	2,0402E-05
TEM-178	CAA65888.1	ARO:3001043	2,033E-05
LEN-16	AAU89706.1	ARO:3002464	2,0259E-05
LEN-18	CAP12346.2	ARO:3002465	2,0259E-05
LEN-19	CAP12347.2	ARO:3002466	2,0259E-05
LEN-26	WP_025712239	ARO:3002474	2,0259E-05
OKP-A-1	CAG25812.2	ARO:3002418	2,0259E-05
OKP-A-2	CAG25815.2	ARO:3002419	2,0259E-05
OKP-A-7	CAJ19605.1	ARO:3002424	2,0259E-05
OKP-B-13	AAV80715.1	ARO:3002446	2,0259E-05
OKP-B-18	CAP12358.1	ARO:3002451	2,0259E-05
OKP-B-8	CAJ19616.1	ARO:3002441	2,0259E-05
SHV-109	ACM04459.1	ARO:3001158	2,0259E-05
SHV-128	ADE58494.1	ARO:3001175	2,0259E-05
SHV-149	AFQ23955.1	ARO:3001188	2,0259E-05
SHV-157	AFQ23963.1	ARO:3001196	2,0259E-05
SHV-18	AAF64386.1	ARO:3001076	2,0259E-05
SHV-32	AAK69828.1	ARO:3001090	2,0259E-05
SHV-55	CAI10727.2	ARO:3001111	2,0259E-05
SHV-62	CAI30651.2	ARO:3001117	2,0259E-05

gene	seq accession	ARO ID	ARAI (ppm)
ANTIBIOTIC RESISTANCE GENES			
SHV-7	AAA87176.1	ARO:3001066	2,0259E-05
SHV-79	CAJ47134.2	ARO:3001133	2,0259E-05
SHV-80	CAJ47135.2	ARO:3001134	2,0259E-05
SHV-99	CAQ03504.1	ARO:3001337	2,0259E-05
TEM-102	AAK82652.1	ARO:3000965	2,0259E-05
TEM-112	AAS89982.1	ARO:3000975	2,0259E-05
TEM-156	CAQ00120.1	ARO:3001022	2,0259E-05
TEM-16	CAA46346.1	ARO:3000887	2,0259E-05
TEM-162	ABO64442.1	ARO:3001028	2,0259E-05
TEM-216	AHJ78622.1	ARO:3001393	2,0259E-05
TEM-34	AGE11905.1	ARO:3000904	2,0259E-05
TEM-47	CAA71322.1	ARO:3000916	2,0259E-05
TEM-52	CAA73933.1	ARO:3000921	2,0259E-05
TEM-53	AAD22538.1	ARO:3000922	2,0259E-05
TEM-67	AAD33116.2	ARO:3000934	2,0259E-05
TEM-87	AAG44570.1	ARO:3000954	2,0259E-05
TEM-91	BAB16308.1	ARO:3000958	2,0259E-05
TEM-92	AAF66653.1	ARO:3000959	2,0259E-05
TEM-94	CAC85661.1	ARO:3000961	2,0259E-05
ANT(6)-Ia	YP_008997281.1	ARO:3002626	2,0188E-05
GES-12	CBG22732.1	ARO:3002341	2,0188E-05
GES-13	ACZ54536.1	ARO:3002342	2,0188E-05
GES-2	AAK58421.1	ARO:3002331	2,0188E-05
GES-20	AEZ05108.1	ARO:3002349	2,0188E-05
GES-3	BAD06399.1	ARO:3002332	2,0188E-05
GES-8	AAK18183.1	ARO:3002337	2,0188E-05
OXY-1-4	AAL78276.1	ARO:3002392	1,9979E-05
OXY-2-1	CAI43414.1	ARO:3002396	1,9979E-05
OXY-2-6	AAL78278.2	ARO:3002401	1,9979E-05

gene	seq accession	ARO ID	ARAI (ppm)
ANTIBIOTIC RESISTANCE GENES			
CTX-M-10	AAT68658.1	ARO:3001873	1,9911E-05
CTX-M-100	CBW46935.1	ARO:3001959	1,9911E-05
CTX-M-105	ADY02554.1	ARO:3001965	1,9911E-05
CTX-M-112	AEM44652.1	ARO:3001972	1,9911E-05
CTX-M-114	ACU00153.1	ARO:3001974	1,9911E-05
CTX-M-116	AEI70324.1	ARO:3001976	1,9911E-05
CTX-M-117	AET99223.1	ARO:3001977	1,9911E-05
CTX-M-121	AFA51699.1	ARO:3001980	1,9911E-05
CTX-M-124	AFH88134.1	ARO:3001983	1,9911E-05
CTX-M-125	AFO69261.1	ARO:3001984	1,9911E-05
CTX-M-139	AFY98865.1	ARO:3001997	1,9911E-05
CTX-M-147	AHA80961.1	ARO:3002005	1,9911E-05
CTX-M-15	AAL02127.1	ARO:3001878	1,9911E-05
CTX-M-22	AAL86924.1	ARO:3001884	1,9911E-05
CTX-M-23	AAL99990.1	ARO:3001885	1,9911E-05
CTX-M-27	AAO61597.1	ARO:3001889	1,9911E-05
CTX-M-30	AAP43508.1	ARO:3001892	1,9911E-05
CTX-M-31	CAD99181.1	ARO:3001893	1,9911E-05
CTX-M-36	BAD16611.1	ARO:3001898	1,9911E-05
CTX-M-43	AAZ14955.1	ARO:3001905	1,9911E-05
CTX-M-44	BAA07082.1	ARO:3001906	1,9911E-05
CTX-M-52	ABB17185.1	ARO:3001914	1,9911E-05
CTX-M-53	ABB72225.1	ARO:3001915	1,9911E-05
CTX-M-62	ABP04245.1	ARO:3001923	1,9911E-05
CTX-M-69	ABY91281.1	ARO:3001930	1,9911E-05
CTX-M-75	ACS32294.1	ARO:3001936	1,9911E-05
CTX-M-91	ACX34099.1	ARO:3001951	1,9911E-05
CTX-M-94	ADK11041.1	ARO:3001954	1,9911E-05
CTX-M-95	CBL80615.1	ARO:3001955	1,9911E-05
CTX-M-98	ADO17948.1	ARO:3001957	1,9911E-05
KPC-14	AFV48348.1	ARO:3002324	1,9911E-05
OXY-1-6	CAB42615.1	ARO:3002394	1,9911E-05
OXY-6-1	CAI43422.1	ARO:3002413	1,9911E-05
OXY-6-3	CAI43424.1	ARO:3002415	1,9911E-05
OXY-6-4	CAI43425.1	ARO:3002416	1,9911E-05
SHV-16	AAC98092.2	ARO:3001075	1,9911E-05
IMI-1	AAA93461.1	ARO:3001858	1,9843E-05
IMI-2	ABA00479.1	ARO:3001859	1,9843E-05
NmcA beta-lactamase	CAA79967.1	ARO:3003589	1,9843E-05
mecR1	YP_001245420.1	ARO:3000215	1,9809E-05
KPC-10	ACS35345.1	ARO:3002320	1,9775E-05
KPC-13	AEA73284.1	ARO:3002323	1,9775E-05
SME-3	AAS92558.1	ARO:3002381	1,9708E-05
SME-5	AHV85514.1	ARO:3002383	1,9708E-05
CARB-10	ACJ61335.1	ARO:3002249	1,9443E-05
CARB-16	CCW43444.1	ARO:3002255	1,9443E-05
CARB-8	AAO59455.1	ARO:3002247	1,9443E-05
mph	EOO80837.1	ARO:3003767	1,9378E-05
VEB-1a	AAK14293.1	ARO:3003711	1,9378E-05
cepA beta-lactamase	AAA21532.1	ARO:3003559	1,9314E-05
APH(2")-IVa	AAC14693.1	ARO:3002637	1,9249E-05
CARB-3	AAB19430.2	ARO:3002242	1,9059E-05
ErmW	BAA03402.1	ARO:3001306	1,863E-05
ErmS	AAA26742.1	ARO:3001304	1,8163E-05
CfxA2	AAD23513.1	ARO:3003002	1,805E-05
AAC(6')-30/AAC(6')-Ib' fusion protein	CAE48335.2	ARO:3002599	1,7665E-05
DHA-17	AIT76103.1	ARO:3002148	1,5288E-05
DHA-22	AIT76109.1	ARO:3002153	1,5288E-05

gene	seq accession	ARO ID	ARAI (ppm)
	ANTIBIOTIC RESISTANCE GENES		
DHA-3	AAR87489.1	ARO:3002134	1,5288E-05
ACT-12	AFU25650.1	ARO:3001834	1,5207E-05
ACT-17	AHM76771.1	ARO:3001838	1,5207E-05
ACT-20	AHA80105.1	ARO:3001841	1,5207E-05
ACT-35	BAP68758.1	ARO:3001855	1,5207E-05
ACT-36	AJG06170.1	ARO:3003171	1,5207E-05
ACT-5	ACJ05689.1	ARO:3001824	1,5207E-05
CMY-100	AHA80101.1	ARO:3002112	1,5207E-05
CMY-29	ABS12248.1	ARO:3002040	1,5207E-05
CMY-42	ADM21467.1	ARO:3002053	1,5207E-05
CMY-65	AEI52842.1	ARO:3002078	1,5207E-05
CMY-87	BAL63057.1	ARO:3002100	1,5207E-05
CMY-90	CCK86743.1	ARO:3002103	1,5207E-05
MIR-15	AIT76104.1	ARO:3002180	1,5207E-05
MIR-6	AFJ79785.1	ARO:3002171	1,5207E-05
MIR-8	ACJ05687.1	ARO:3002173	1,5207E-05
CMY-1	CAA63264.1	ARO:3002012	1,5168E-05
CMY-10	AAK31368.1	ARO:3002021	1,5168E-05
FOX-1	CAA54602.1	ARO:3002155	1,5168E-05
FOX-10	AGE45503.1	ARO:3002162	1,5168E-05
FOX-9	AEK78851.1	ARO:3002163	1,5168E-05
MOX-4	ACI89425.1	ARO:3002184	1,5168E-05
ADC-2	WP_004746565.1	ARO:3003848	1,5128E-05
MOX-6	ACS44784.1	ARO:3002185	1,5128E-05
MOX-7	ACS44785.1	ARO:3002189	1,5128E-05
MOX-8	AGH56079.1	ARO:3002190	1,5128E-05
PDC-2	ACQ82806.1	ARO:3002498	1,4595E-05
mecA	AGC51118.1	ARO:3000617	8,6737E-06
vanTN	AEP40502.2	ARO:3002975	8,2772E-06

Table S 3: List of manually curated resistance determinants. Abundance is given in number of assigned reads before and after curation and after normalisation of the curated abundance as ARAI*. a) Genes which represent antibiotic targets, for which point mutations are known to confer resistance. Assigned reads which align to the mutation area were filtered and only those containing the point mutation were retained in the dataset. b) Genes and their assigned Antibiotic Resistance Ontologies (AROs) tow which two or more genes sequences are attributed to. To avoid multiple representation of such AROs, hits of homolog genes were merges. c) Genes sourced out due to high similarity to common and widely spread genes, which are not directly related to antibiotic resistance.

	Gene (ARO)	Abundance before curation	A) Mutation (No. of reads aligning to mutation area) B) Accession numbers of homologous genes	Abundance after curation	ARAI* (ppm) after normalisation of curated reads
A) Curated for mutations	*Chlamydia trachomatis murA* (ARO:3003785)	194	C119D (0)	0	0
	Mycobacterium tuberculosis murA (ARO:3003784)	1058	C117D (309)	22	0.00030495
	Streptomyces cinnamoneus EF-Tu (ARO:3003359)	13978	A379T (1039)	9	0.00013135
B) Curated for gene redundancy	*arnA* (ARO:3002985)	361	NP_252244	1007	0.00884032
		646	AAC75315.1		
	cat (ARO:3002670)	114	AAA22081.1	118	0.00317999
		2	AAA23018.1		
		1	BAC11901.1		
		1	AAB23649.1		
	catIII (ARO:3002685)	1	CAB75601.1	2	5.4404E-05
		1	CAA30695.1		
	ANT(6)-Ib (ARO:3002629)	1	CBH51824.1	2	4.0236E-05
		1	AIJ27543.1		
C) Other	*mfd* (ARO:3003844)	8097	**	0	0
	NmcR (ARO:3003665)	442	***	0	0

*ARAI (antibiotic resistance abundance index): number of reads assigned to an antibiotic resistance gene per total number of reads and respective gene length in ppm (≙reads per million reads) (Elbehery, Aziz and Siam 2016)
** *Mfd* influences the spontaneous mutation rate that can give rise to ciprofloxacin resistances in *Campylobacter jejuni* (Han *et al.* 2008). It is, however, a wide-spread protein involved in DNA repair and by itself not directly related in antibiotic resistance (Savery 2007).
*** *NmcR* regulates the *NmcA* β-lactamase, to which one metagenomic read was assigned resulting in an abundance of 1,98E-05 ppm (Table S2). However, *NmcR* is a homolog of the widely conserved *lysR* regulators (Naas and Nordmann 1994).

Table S 4. Minimal inhibitory concentrations of β-lactam resistant metagenomic clones. Minimal inhibitory concentration (µg ml^{-1}) of penam and cephalosporin antibiotics for the metagenomic clones *E. coli* EPI300 pCC2FOS-Mm1, Mm2 and Mm3 and the empty vector control *E. coli* EPI300 pCC2FOS (X). Mean values, n=3.

Clone	Ampicillin	Carbenicillin	Cefotaxime	Cefalothin	Cephalexin
Mm1	64	16	8	8	4
Mm2	64	32	8	8	8
Mm3	>512	>1024	<0.5	64	8
X	32	8	<0.5	8	4

Table S 5: Antibiotic concentrations for resistance screenings used.

Antibiotic	Antibiotic class	Manufacturer	Spectrum	Concentration [µg ml^{-1}] used for	
				Isolates	Metagenomic clones
Ampicillin	β-Lactam	Roth, Germany	Gram +/-	10	50
Ciprofloxacin	Fluoroquinolone	Sigma-Aldrich, Missouri, USA	Gram +/-	5	1
Erythromycin	Macrolide	Roth, Germany	Gram +/-	15	150
Gentamycin	Aminoglycoside	Roth, Germany	Gram +/-	10	10
Kanamycin sulfate		Roth, Germany	Gram +/-	30	20
Nalidixic acid	Quinolone	Merck, Germany	Gram +/-	30	15
Tetracycline	Tetracycline	Merck, Germany	Gram +/-	30	4
Rifampicin	Ansamycin	Duchefa Biochemie, Netherlands	Gram +/-	5	20
Sulfadiazine	Sulfonamide	Sigma-Aldrich, Missouri, USA	Gram +/-	300	≤ 2250
Vancomycin	Glycopeptide	Sigma-Aldrich, Missouri, USA	Gram +	30	1000

Supplementary figures

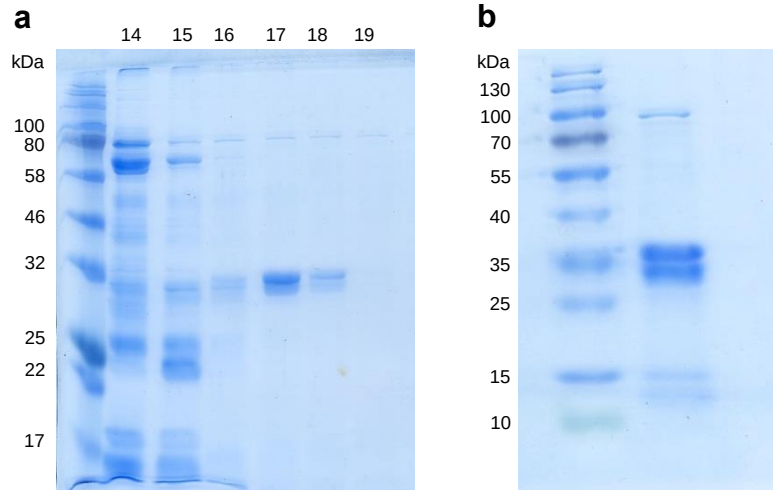


Figure S 1: Purified β-lactamase Mm3. The *blaMm3* gene encoding a 304 amino acid long β-lactamase was cloned under an N-terminal His Tag and the tagged protein purified by affinity chromatography. a) Elution fractions no. 14 to 19; no. 17 and 18 were selected for further use. b) SDS-PAGE of the pure enzyme shows two bands, one with the estimated molecular weight of the his-tagged protein of about 35 kDA and a smaller protein band around 32 kDA. As identified by LC-MS/MS analysis the higher molecular weight band contains the β-lactamase still adjunct to the His-Tag, while the lower band stems from a smaller version of the purified β-lactamase which lost the His-Tag. Lower molecular weight proteolytic products of approximately 15 to 17 kDA are visible as well, corresponding to the His-tagged termini of the protein. Protein Ladder: a) Color prestained Protein Standard, Broad Range (New England Biolabs); b) PageRuler Prestained Protein Ladder (Thermo Scientific).

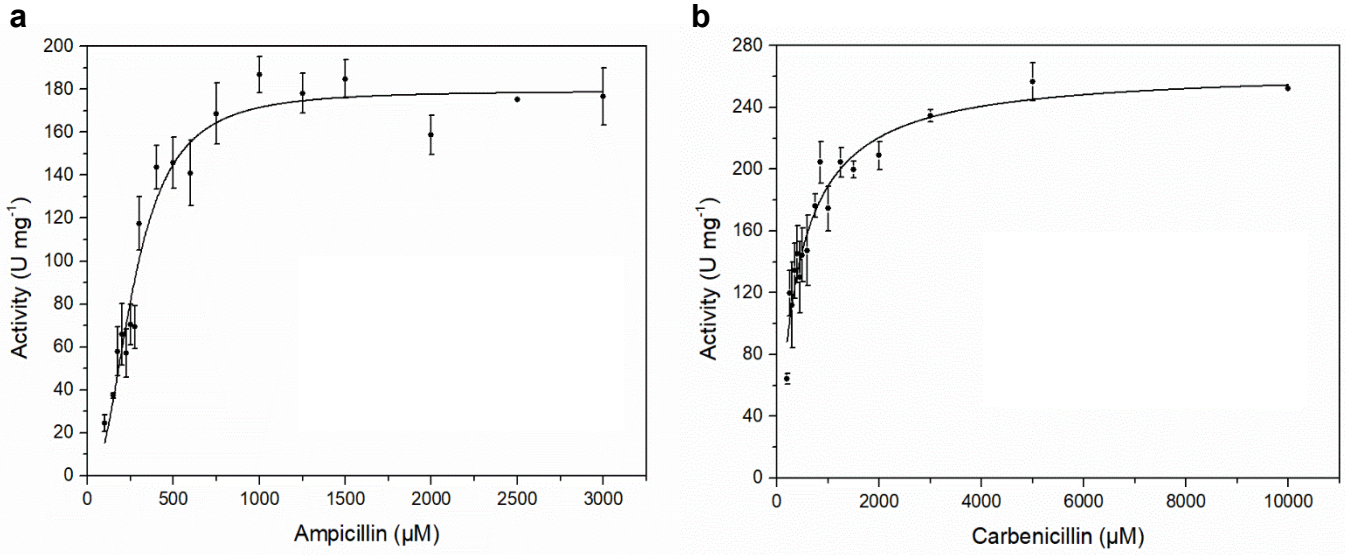


Figure S 2: Kinetic characterisation of the β-lactamase Mm3. The initial hydrolysis of the substrate was followed spectrophotometrically at 235 nm. Points, mean values (n= 2 to 8) for all measurements, except for 2.5 mM carbenicillin (n01); error bars, standard deviations. Data for ampicillin (a) and carbenicillin (b) were fitted according to the Hill or Michaelis-Menten equation, respectively.

Additional methods

Screening for antimicrobial VOCS. Using the htTCVA as initial step metagenomic libraries were screened against several target organisms. Generally, two 96-well plates were assembled such that the top sides of two plates, separated by perforated silicon (0.2 mm width, 0.5 mm hole diameter), faced one another and fixed with clams on either side. One plate, thereby, contained the library clones, the other the target organism. For filamentous fungi a spore suspension was prepared by adding 10 ml of sterile dH_2O onto fungal mycelium grown for four days on the respective culture medium. In case of *Verticillium spp.* 10 ml of Czapek Dox broth were used. Spores were then mobilised into the liquid by scraping the mycelium with a sterile drigalski spatula and subsequently separated from hyphae by passing the suspension through sterile mull. The spore concentration was adjusted to 100 spores μl^{-1} and 20 000 spores μl^{-1} for *F. culmorum* and *V. longisporum*, respectively, and stored at 4 °C for up to 48 h. The wells of 96-well plates were filled with 70 µl of agar onto which 10 µl of the spore suspension was pipetted. For bacteria and yeast, cell material from grown colonies was used to prepare a cell suspension in 0.85% sodium chloride solution (w/v) to match the turbidity of McFarland 0.5. The suspension was prepared freshly and directly before usage and diluted 1:20 into soft agar (0.7% w/v) of which 70 µl were then pipetted into the wells of 96-well plates. The metagenomic clones were transferred onto 70 µl of LB agar amended with 12.5 µg μl^{-1} chloramphenicol and 0.01 % arabinose (w/v) by pipetting 10 µl of glycerol stock onto the agar. An empty vector strain served as negative control. After incubation at 37 °C for 24 h, 30 µl bacillol were added into dedicated wells as positive control and plates assembled as described above. All plates were incubated 20 °C. The turbidity of the agar served as indicator to determine whether the test organism was inhibited or not.

In a second step, positive tested clones were re-evaluated with the TCVA using 6- and 12-well plates. The wells were filled with 1 ml or 3 ml agar, respectively. Of the library stocks 30 µl or 70 µl, of the spore suspensions 20 µl or 35 µl, and 150 µl or 500 µl of bacillol were used, respectively. Otherwise the assay was performed as described above.

Inhibition efficacy against filamentous fungi was determined using a petri dish based VOC assay. For this purpose, overnight cultures of the metagenomic clones and the empty vector strain in LB amended with 12.5 µg μl^{-1} chloramphenicol and 0.01 % arabinose (w/v) or alternatively to arabinose CopyControl induction solution (epicentre) were diluted to an OD_{600} of 0.8. Of the adjusted cultures, 100 µl were plated onto selective culture medium and incubated for 24 h at 37 °C. *F. culmorum* was transferred onto agar plates by: 1) placing a plaque of mycelium (Ø 8 mm) in the centre of the petri dish, or 2) plating 5 µl of the spore suspension of the agar plate. Plates were again assembled face to face and sealed with parafilm. As positive control, three pieces of paper towel soaked with 500 µl bacillol were placed on the agar plate. The assembled plates were incubated at 20 °C for 5 days or until the mycelium reached the rim of the petri dish.

The pure, purchased substances were tested against filamentous fungi in a standard petri dish. A PTFE-lined silicon rubber septum (LaPhaPack) holding 150 µl of the purchased substance was placed into the centre of the petri dish lid. The bottom plate, in its centre containing a plaque of mycelium on respective culture medium, was placed upside down onto the lid and the petri dish sealed with parafilm. Plates were incubated at 20 °C until the mycelia had reached the rim of the petri dish. Bacillol soaked paper towels served as positive control and the sole PTFE-lined silicon rubber septum as negative control.

VOCs profiling. Headspace solid phase microextraction gas chromatrography-mass spectrometry was used to measure the volatilome of active library clones.

Prior to the measurement library clones were streaked from glycerol stocks onto an 7 ml LB slope agar amended with 12.5 μg μl^{-1} chloramphenicol and 0.01 % arabinose (w/v) or alternatively to arabinose CopyControl induction solution inside a 20 ml headspace vial (75.5 × 22.5 mm, Chromtech, Germany). Closed with a cotton plug and covered with tin foil, the vials were incubated for 24 h at 37°C and then sealed with magnetic crimp caps with a PTFE-lined silicon rubber septum (LaPhaPack, Langerwehe, Germany) followed by another two hours of incubation at 20 °C.

The purchased substances were dissolved in dimethyl sulfoxide to a concentration of 0.1 mol l^{-1}. For the GC-MS measurement, 10 μl were placed into a sterile 20 ml headspace vial (75.5 × 22.5 mm) which were sealed with magnetic crimp caps with a PTFE-lined silicon rubber septum and incubated for two hours.

The profiling was performed with an automated sampler and a 50/30 μm Divinylbenzen/CarboxenTM/Polydimethylsiloxane 2 cm Stableflex/SS fiber (Supelco, Bellefonte, PA, USA), a GC 7890A combined with a quadrupole MS 5974C (Agilent Technologies, Waldbronn, Germany). After enrichment of volatile compounds for 30 min at 30 °C, samples were run through a (5% phenyl)-methylpolysiloxane column, 60 m × 0.25 mm i.d., 0.25 μm film thickness (DB-5MS; Agilent Technologies, Waldbronn, Germany) followed by electron ionisation (EI; 70 eV) and compound detection (mass range 25 to 350 u). The helium flow rate was set to 1.2 ml min^{-1} and the inlet temperature set to 270 °C. The applied temperature gradient was as following: 40 °C for 2 min, 40 °C – 110 °C with 5 °C min^{-1}, 110 °C – 280 °C with 10 °C min^{-1}, 280 °C for 3 min. Spectra were compared with NIST Mass Spectral Database 08 entries. Compounds were identified by their retention indices and comparison to reference substances (Sigma-Aldrich, St. Louis, MO, USA).

List of abbreviations

ACC	Aminoclycopropane 1-carboxylic acid
AHL	N-acylhomoserine-lactone
AMR	Antimicrobial resistance
ARAI	Antibiotic resistance abundance index
ARG	Antibiotic resistance gene
BCA	Biological control agent
CARD	Comprehensive antibiotic resistance database
CC	*E. coli* EPI300 pCC2FOS-CC, clone against *C. albicans*
ESLB	Extended spectrum β-lactamase
FACS	Fluorescence-activated cell sorting
GHases	Glycoside hydrolases
HTS	High-throughput screening
htTCVA	high-throughput Two Clamps VOCs Assay
LB	Luria Bertani
LC-MS/MS	Liquid Chromatography-Mass Spectrometry/ Mass Spectrometry
LSBL	Limited-spectrum β-lactamase
METREX	Metabolic regulated expression screening
MIC	Minimal inhibitory concentration
NB	Nutrient broth
Ni	Nickel
NP	Natural product
NRPS	Non-ribosomal peptide synthetase
PCR	Polymerase chain reaction
PKS	Polyketide synthase
PPT	4´phosphopantetheinyl transferase
SIGEX	Substrate induced gene expression screening
TB	Terrific broth
TCVA	Two Clamps VOCs Assay
VOC	Volatile organic compound
XANT	*Xanthomonas* group of β-lactamases

List of tables

List of figures