ANNALS OF THE NEW YORK ACADEMY OF SCIENCES

Genomic and functional techniques to mine the microbiome for novel antimicrobials and antimicrobial resistance genes

Boahemaa Adu-Oppong,[1,a] Andrew J. Gasparrini,[1,a] and Gautam Dantas[1,2,3,4]

[1]Center for Genome Sciences and Systems Biology, Washington University School of Medicine, St. Louis, Missouri. [2]Department of Pathology & Immunology, Washington University School of Medicine, St. Louis, Missouri. [3]Department of Biomedical Engineering, Washington University in St. Louis, St. Louis, Missouri. [4]Department of Molecular Microbiology, Washington University School of Medicine, St. Louis, Missouri

Address for correspondence: Gautam Dantas, 4515 McKinley Ave, Room 5314, St. Louis, MO 63110. dantas@wustl.edu

Microbial communities contain diverse bacteria that play important roles in every environment. Advances in sequencing and computational methodologies over the past decades have illuminated the phylogenetic and functional diversity of microbial communities from diverse habitats. Among the activities encoded in microbiomes are the abilities to synthesize and resist small molecules, yielding antimicrobial activity. These functions are of particular interest when viewed in light of the public health emergency posed by the increase in clinical antimicrobial resistance and the dwindling antimicrobial discovery and approval pipeline, and given the intimate ecological and evolutionary relationship between antimicrobial biosynthesis and resistance. Here, we review genomic and functional methods that have been developed for accessing the antimicrobial biosynthesis and resistance capacity of microbiomes and highlight outstanding examples of their applications.

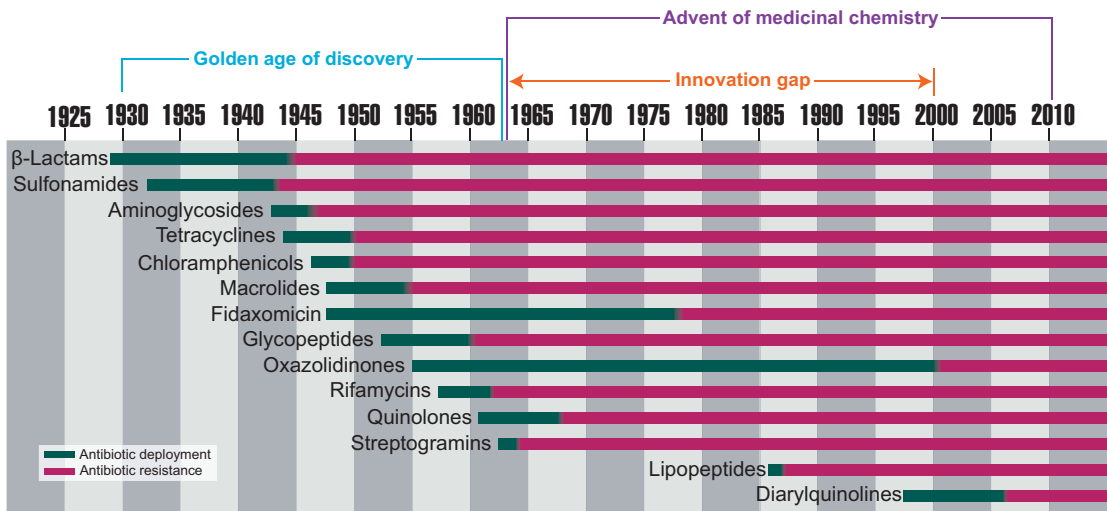**Keywords:** antimicrobials; antimicrobial resistance; microbiome; metagenomics

Introduction

Nobel laureate Joshua Lederberg was the first to propose the term "microbiome," which he defined as "the ecological community of commensal, symbiotic, and pathogenic microorganisms that literally share our body space."[1] While Lederberg coined the term in reference to the microorganisms that inhabit human body sites, microbial communities can be associated with any environmental, clinical, or engineered habitat. Concurrent with the expansion in sequencing capacity and the widespread adoption of "omics" technologies, the term microbiome was redefined as the collective genetic material of a microbial community, while "microbiota" is generally used to refer to the microorganisms themselves.[2] Initial studies of microbial communities attempted to characterize the phylogenetic and functional diversity by culturing the organisms

and performing targeted functional assays. However, culture-based assays are inherently biased, as a large fraction of microbes from most habitats are not readily cultured using standard methods.[3] Fortunately, metagenomic sequencing of microbial communities, facilitated by the revolution in genomics over the past several decades, has enabled increased investigation and appreciation of the breadth and depth of functions encoded in diverse microbiomes.

The study of antimicrobial resistance and biosynthesis in microbial communities is motivated by the ubiquity and diversity of microbes in wide-ranging habitats, the evolutionary and ecological origins of antimicrobial resistance, and the global health crisis posed by increasing antimicrobial resistance. The human commensal microbiota harbors approximately 10 trillion microbes,[4,5] while the diversity of the soil microbiota has been estimated to be as high as 10,000 unique species per gram.[6] Both culture-based and culture-independent analyses of these and other microbial communities have revealed

---

[a]Both authors contributed equally to this work.

Figure 1. Timeline illustrating the rapid emergence of clinical antimicrobial resistance. The introduction of an antimicrobial to the clinic (green bars) is quickly followed by the first observation and persistence of resistance of that antimicrobial (magenta). As a result, discovery and development of new antimicrobials have struggled to keep pace with the evolution of antimicrobial resistance.

biosynthetic pathways for small molecules with antimicrobial activity,[7,8] as well as abundant and diverse antimicrobial resistance genes.[9–11] This collection of resistance genes in a given microbiome, known as the antimicrobial *resistome*, consists of acquired, intrinsic, proto, and silent resistance genes.[12,13] It is important to characterize resistomes and biosynthetic pathways from varied microbial communities in order to better understand and combat the rise of clinical antimicrobial resistance.[10] This is ideally achieved by pairing functional and sequence-based genomic technologies with traditional culture-based analyses.

This review provides an overview of the functional and sequence-based methods that have been developed for discovery of antimicrobial biosynthesis pathways and resistance genes from microbial communities. We will critically evaluate these methods and highlight outstanding examples of their applications, with a specific focus on the advantages and disadvantages that each method provides.

Rise in antimicrobial resistance and decline in antimicrobial discovery

The World Health Organization has described antimicrobial resistance as the single greatest challenge in infectious disease today, posing a serious public health threat on a global scale.[14] Clinical and environmental antimicrobial resistance has been

increasing since the widespread anthropogenic deployment of antimicrobials. Indeed, clinical resistance is now frequently observed within a year of deployment of a new antimicrobial[15,16] (Fig. 1). Furthermore, polymerase chain reaction (PCR)–based surveys of archived soils have revealed that the prevalence of environmental antimicrobial resistance genes has been on the rise since 1940.[3,17] In the United States alone, two million people acquire serious infections that are resistant to the antimicrobials designed to treat those infections, resulting in an estimated 23,000 deaths.[18] Incidences of multidrug resistance in human pathogens are also increasing.[18] The proportion of clinical isolates that are resistant to antimicrobials, including deadly methicillin-resistant *Staphylococcus aureus* (MRSA), vancomycin-resistant *Enterococcus*, and fluoroquinolone-resistant *Pseudomonas aeruginosa* strains, has been steadily increasing since the 1980s.[19] The economic burden of antimicrobial resistance is also tremendous, with an estimated eight million additional hospital days resulting in $20 billion in direct healthcare costs and $35 billion in lost productivity in the United States alone.[18,20] If the current trajectory is not altered, it is estimated that worldwide deaths from antimicrobial resistance could climb to 10 million per year by 2050.[21]

Compounding the problem posed by antimicrobial resistance, the golden age of antimicrobial

discovery (ca. 1940–1960) was followed by a substantial innovation gap spanning 1962–2000 in which no new classes of antimicrobials were introduced (Fig. 1).[15,22] Only three new antimicrobials were approved for human use during the first decade of the 21st century.[23] The decrease in antimicrobial discovery can be partially attributed to low incentives for research and development of new drug classes.[14] While most antimicrobial products currently in the pipeline are active against at least some of the pathogens identified by the U.S. Food and Drug Administration as posing a "serious threat to public health," barely a third (16 drugs) show significant activity against multidrug-resistant Gram-negative species.[14] The issue is complicated by the collateral damage from excessive antimicrobial use: when an antimicrobial is administered clinically or agriculturally, it provides selective pressure for increased antimicrobial resistance in all microbes in those habitats, not simply the targeted pathogen(s), and thus has a negative societal impact.[24] To counter the withdrawal of pharmaceutical companies from the antimicrobial market, the U.S. Congress is considering legislation to create a streamlined antimicrobial approval pathway. This would approve drugs to treat life-threatening infections on the basis of smaller (and hence faster) clinical trials, provided that they meet the same phase I safety and efficacy standards as other drugs.[25] While these measures are being considered, it is imperative that we continue to better understand the ecological and evolutionary dimensions of antimicrobial production and resistance to support the improved stewardship of both our current and future arsenals of antimicrobial compounds.

Brief history of antimicrobials

Antimicrobials are defined as chemicals that either kill (bacteriocidal) or inhibit the growth (bacteriostatic) of bacteria at defined concentrations, and function by targeting systems critical to bacterial physiology. The five major modes of antimicrobial action are (1) inhibition of cell wall synthesis, (2) inhibition of protein synthesis via the bacterial ribosome, (3) inhibition of DNA or RNA synthesis, (4) inhibition of the folic acid pathway of nucleic acid synthesis, and (5) disruption of cell membrane integrity.[15,26–29] Antimicrobials have likely been naturally produced by environmental microbes as a means

of communication and defense for billions of years.[30–32] Residues of antimicrobials were discovered in human skeletal remains dating from 350 to 550 CE, suggesting that humans have been exposed to antimicrobials for thousands of years.[33–35] However, it was not until the 20th century that natural-product antimicrobials were deliberately repurposed for human use, revolutionizing our treatment of infectious diseases. Up until the 1920s, infectious diseases were the leading cause of death, with pneumonia, influenza, gastrointestinal infections, and tuberculosis accounting for a combined 540 per 100,000 deaths in the United States during the first decade of the 20th century.[36] Ehrlich's discovery of arsphenamine, an organoarsenic drug that could cure syphilis-infected rabbits[35,37] and Fleming's subsequent serendipitous discovery of penicillin, a β-lactam,[38] signaled the advent of the antimicrobial era. The golden age of antimicrobial discovery began in the 1940s and lasted through the 1960s, fueled by Selman Waksman's rapid antimicrobial discovery platform.[16,39] Waksman was awarded the Nobel Prize for the development of this method,[40] and the drugs developed during the golden age of antimicrobial discovery substantially improved the treatment of infectious disease.

Many natural-product antimicrobials were found to have pharmacological or toxicological drawbacks or eventually became obsolete owing to evolution and spread of resistance.[41] Synthetically altering natural-product antimicrobials to improve pharmacology or evade resistance launched a new era of medicinal chemistry. Many antimicrobials were synthesized by modifying molecular scaffolds of antimicrobials previously discovered during the golden era. As a result, the amphenicol, tetracycline, aminoglycoside, macrolide, glycopeptide, and quinolone classes were introduced or expanded, further improving our treatment of infectious diseases.[16,22,35,42,43]

Brief history of antimicrobial resistance

The introduction of antimicrobials to the clinic resulted in a precipitous drop in worldwide mortality from infectious disease.[44] However, the selective pressure placed on bacterial systems targeted by antimicrobials owing to the extensive use of antimicrobials in clinical and agricultural settings soon led to the steady evolution and increase of antimicrobial

resistance, systematically compromising each of the antimicrobials in our arsenal (Fig 1).[27,45–47]

The strategies that bacteria use to resist antimicrobials can be divided into four general categories: (1) drug efflux, (2) reducing the permeability of the cell wall or membrane, (3) target overexpression, modification, or protection, and (4) enzymatic inactivation of the drug.[27,48] In the first resistance mechanism of drug efflux, bacteria pump out the antimicrobial, keeping the intracellular concentration low and preventing the drug from reaching inhibitory concentrations. This mechanism encompasses efflux pumps that are either specific to single antimicrobials or classes, such as the tetracycline efflux pumps,[49] or nonspecific, multidrug-resistance efflux pumps.[50–53] The second resistance strategy, thematically related to the first, involves reducing the permeability of the cell wall or cell membrane to the antimicrobial, thereby preventing it from reaching its target. This can occur via reduction of porin expression or expression of a more selective porin variant.[54,55] The third resistance strategy is overexpression, modification, or protection of the drug target in the bacterium, enabling survival. This mechanism is used by MRSA, which expresses PBP2a, a redundant and methicillin-insensitive version of the native PBP2 penicillin-binding protein encoded by *mecA*.[56] The final (fourth) resistance strategy involves enzymatic inactivation of the drug, by either degrading or modifying the antimicrobial compound such that it no longer has activity. Examples of this mechanism include β-lactamases, which degrade β-lactam antimicrobials,[57,58] tetracycline destructases, such as Tet(X),[59,60] and chloramphenicol acetyltransferases.[61]

While we primarily evaluate antimicrobial resistance in terms of its effects on the clinical use of antimicrobials, the capacity to resist such compounds is a natural and ancient feature of all microbial communities.[62,63] It is important to consider that the genes that we refer to as "resistance genes" may have had functions unrelated to clinical antimicrobial resistance in their original context but could be repurposed for resisting antimicrobials when they are encountered in a new context. The original functions of such genes may include cell wall biosynthesis (e.g., β-lactamases); export of signaling molecules, metabolic intermediates, or plant-produced compounds (e.g., multidrug efflux pumps); or detoxifying the original host from the antimicrobial that it is producing (e.g., the tetracycline efflux pump induced just in time by its biosynthetic precursor anhydrotetracycline[64]).[65] The discovery of β-lactam, tetracycline, and vancomycin resistance genes in 30,000-year-old Beringian permafrost sediments confirms that genes that confer resistance to antimicrobials preceded the clinical use of antimicrobials by several thousand years.[62] Furthermore, while the first enzymatic degradation of penicillin was described in 1940, when researchers noted that extracts of certain bacterial cells were capable of inactivating penicillin,[66] it is estimated that environmental β-lactamases have existed for over 2 billion years.[67,68] Intensive human use of antimicrobials in clinical, agricultural, and industrial settings over the past 80 years has led to the fairly recent exaptation of such genes with resistance-conferring potential for surviving in the presence of antimicrobials.[69,70]

Since the introduction of antimicrobials to the clinic, the study of antimicrobial resistance has largely focused on isolates of human pathogens and has relied on information gathered from culturing these organisms.[71] However, it is thought that most clinical antimicrobial resistance has its evolutionary origins in the aforementioned genes, with resistance-conferring potential encoded by benign environmental and human-associated bacteria. Furthermore, a recent study identified numerous resistance genes from benign soil bacteria with 100% nucleotide identity to resistance genes that were present in diverse human pathogens, suggesting the relatively recent sharing of resistance genes via horizontal gene transfer.[72] This finding highlights the need to study antimicrobial resistance in benign as well as pathogenic bacteria and in the context of microbial communities as well as isolates, as horizontal gene transfer is the mechanism by which most pathogens become resistant to antimicrobials.[73]

Culture-based methods underestimate the functional diversity of microbiomes

Discovering novel antimicrobials and resistance from microbial communities through culture-based methods is challenging and limited to the cultivable minority of bacteria.[74] Genomic technologies offer the promise of substantial performance improvements for characterizing the functional capacity of microbiota, which has implications in both the basic
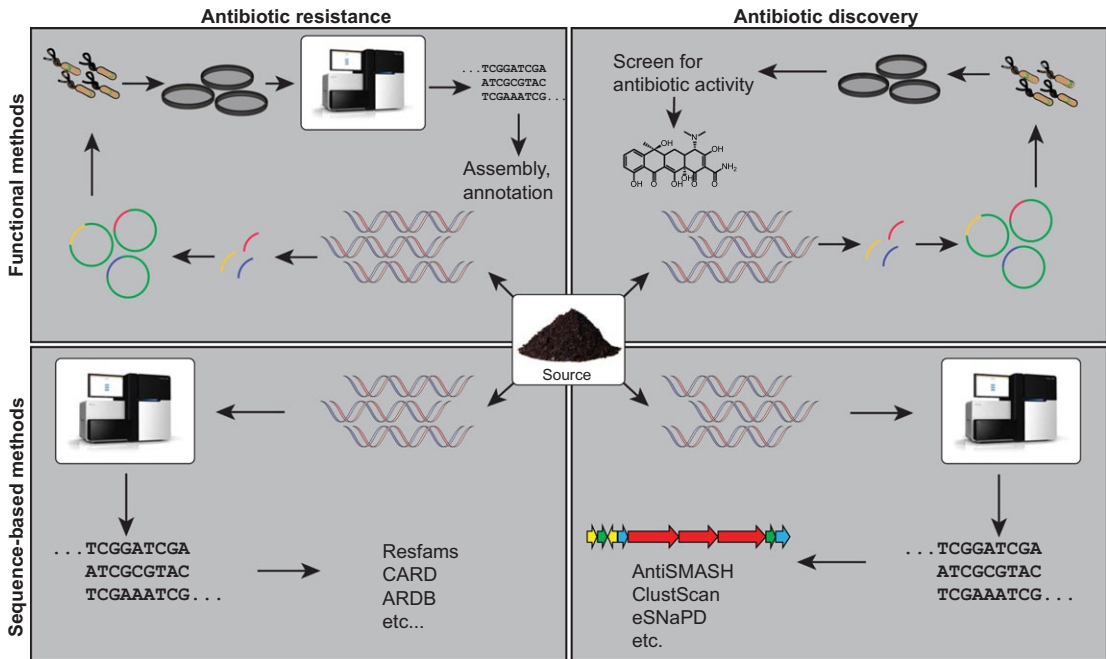
Figure 2. An overview of function- and sequence-based methods for the discovery of antimicrobials and antimicrobial resistance from microbiomes. DNA is isolated from source material (center), which can then be mined for antimicrobial resistance genes (left) or antimicrobial biosynthetic pathways (right). Functional metagenomic methods (top) typically entail shotgun cloning metagenomic DNA into an expression vector and selecting for a desired phenotype (e.g., antimicrobial resistance or antimicrobial activity). Sequence-based methods (bottom) customarily involve sequencing metagenomic DNA and annotating sequences using general or function-specific databases.

science and translational realms.[74,75] For example, characterizing antimicrobial resistance using current culture-based methods can take approximately 1–3 weeks, versus the proposed less than 12 h using genomic technologies.[74] Whole-genome sequencing has the potential to reduce exposure to ineffective drugs and the evolution of resistance by reducing diagnosis to days instead of weeks, as seen in the case of drug-resistant *Mycobacterium tuberculosis*.[75,76]

Depending on the habitat, traditional culturing efforts may capture as little as 1% of a microbial community.[3] The other 99% includes species that are recalcitrant to culture by standard methods.[77] This severe undersampling has motivated considerable recent advancements in culturing methods to enable improvements in the proportion of microbial communities that can be cultured. However, these methods, which may leverage microfluidic technologies,[78,79] clever simulations of natural growth conditions,[79,80] or exhaustive sampling of culture conditions,[81] can be technically challenging

and are not yet part of standard culturing workflows. Culture-based methods are further hindered by the possible unreliable identification of the cultured organisms, which is dependent on the expertise of the researcher. Culture-based methods are also riddled with irreproducibility on nonselective media.[82] It is often difficult to recapitulate the phenotype of resistance of a targeted bacterium from selective media to nonselective media.

Advancements in genomic and computational technologies and the concomitant dramatic reductions in sequencing costs over the past decades have allowed for an explosion of microbial genome and metagenome sequencing, illuminating the diversity of microbial communities and the breadth and depth of functions encoded therein.[74,83–89] Genomic technologies can overcome culture-based limitations and facilitate the discovery of novel antimicrobials and resistance determinants. The remaining review details the functional and sequencing-based methods that can be employed to quantitatively interrogate microbiomes and resistomes (Fig. 2).

Function-based methods for resistance discovery

Functional metagenomics

Functional metagenomic methods for the discovery of antimicrobial resistance genes sit at the intersection of culture-based and genomic techniques.[90] The general functional metagenomic pipeline consists of extracting total metagenomic DNA from a microbial community, shearing the metagenomic DNA to a target size distribution (e.g., 1–5 kb), and shotgun cloning these fragments into an expression vector. The expression library is then transformed into a suitable heterologous host, such as *Escherichia coli*. A phenotype, such as antimicrobial resistance, is selected by plating a metagenomic library on media supplemented with antimicrobials.[71,91] The combination of functional metagenomics with next-generation sequencing and improved computational sequence-assembly algorithms has dramatically increased the throughput of functional metagenomics while also decreasing the cost.[72] This method is particularly well suited for discovering novel genes with antimicrobial resistance function, including those that are less than 65% identical at the amino acid level to known resistance genes.[81] Functional metagenomics vastly outperforms methods that rely solely on alignment to reference databases to predict resistance functionality because it is independent of prior knowledge of resistance function.[72,90,92–96] However, a limitation of this method is that the identification of novel resistance genes depends on functionality of genes in the heterologous host.[90] It is thus best to use a range of heterologous hosts in order to truly capture all phenotypic resistance.[90,97] Additionally, some combinatorial antimicrobial resistance mechanisms require the presence of multiple genes, which may not be captured within a functional metagenomic library depending on the insert size.[98]

One of the primary strengths of functional metagenomics is that it enables the identification of novel resistance-conferring genes from diverse microbiomes.[71,72,96] For example, a recent study used functional metagenomics to identify 1100 genes encoding resistance functions, 121 of which were unique, from 96 fecal and environmental samples from rural areas in El Salvador and a peri-urban shanty town in Lima, Peru.[96] By complementing this functional metagenomic profiling with

deep metagenomic shotgun sequencing of 285 fecal and environmental samples from these same habitats (representing 344 Gb), researchers identified chicken coops and wastewater treatment systems as hotspots of resistance gene exchange and provided a new framework for risk assessment of resistance gene exchange across interconnected habitats. In addition to clearly identifying novel antimicrobial-resistance determinants, functional metagenomics can also be used to predict horizontal gene transfer of antimicrobial resistance genes by identifying colocalized mobility elements.[99] For instance, the aforementioned cross-habitat resistome study in El Salvador and Peru identified a single β-lactamase (TEM-1) in over 25 different genomic contexts, highlighting its high genomic mobility.[96]

Functional metagenomic selections in phylogenetically diverse expression hosts are powerful tools for expanding our knowledge of novel, functional resistance determinants in the face of rising clinical and environmental resistance. Going forward, it is important that functional methods be used to identify novel resistance genes for existing antimicrobials and to quantify the prevalence of different resistance mechanisms across habitats. Furthermore, new antimicrobial lead compounds should be screened for potentially cryptic resistance genes from diverse environments using functional metagenomic assays. This would allow us to anticipate the specific mechanism of inevitable evolution and dissemination of antimicrobial resistance to next-generation antimicrobials before the widespread emergence of these resistance genes in clinical and agricultural settings. Such information would enable design of preemptive molecular surveillance and diagnostics of these resistance determinants, as well as potential targets for design of resistance gene inhibitors.

Hybridization/PCR-based methods

Another method that integrates culturing and sequencing for resistance characterization is a new technology that utilizes molecular padlock probes.[100] In this method, which shows particular clinical promise in predicting the antimicrobial susceptibility profile of patient samples, the microbial community is exposed to the antimicrobial of interest for a short period of time, after which metagenomic DNA is extracted.[100] The metagenomic DNA is then mixed with complementary

biotinylated oligonucleotides of targeted rRNA gene sequences and streptavidin-coated magnetic beads. This is to ensure the capture of the relevant rRNA gene sequences. The mixture is then hybridized and ligated to the designed padlock probes, usually designed to detect certain positions in the 16S rRNA gene.[100] The probes are then amplified through rolling-circle amplification.[100] After amplification, the final product is fluorescently labeled, and DNA copy number is quantified using a high-performance fluorescence detector.[100] An antimicrobial susceptibility profile can be determined using the DNA copy number of the sample without antimicrobials and comparing it to the DNA copy number of the sample after exposure to antimicrobials.[100] If there is an increase in DNA copy number after exposure to antimicrobials, this indicates resistance. If there is no significant growth after exposure to antimicrobials, this indicates susceptibility.[100] This method has been used to characterize the antimicrobial susceptibility profile of uropathogens from complex urinary tract infection samples.[100] Researchers accurately predicted the antimicrobial susceptibility profile for 55 of 56 samples, highlighting the clinical utility of this method.[100] An advantage that the padlock probe method presents is short turnaround time, leveraging high throughput to give results in a matter of hours and help clinicians prescribe the correct antimicrobials to patients with bacterial infections. However, the probes that are used in the assays have to be designed for individual bacterial species and antimicrobials that are targeted to be used in the assay and cannot characterize novel resistance genes.

Sequence-based methods for resistance discovery

While function-based methods have been valuable in identifying novel resistance determinants from diverse microbiomes, they are limited in throughput and are ideally complemented by sequence-based methods, which we will review here. The ability to analyze microbial communities through culture-independent shotgun metagenomic sequencing has revolutionized our ability to characterize resistomes.[96,101–103] These analyses have benefited from databasing projects, including the 35,000+ public metagenomes that have been uploaded to the Metagenomics Rapid Annotations using Subsystems Technology server.[104] Shotgun metagenomic

sequencing involves extracting and sequencing total metagenomic DNA from a microbial community. Once sequenced, a specific subset of informative marker genes from the metagenome can be used to infer phylogeny and functions that are present in the microbial community of interest.[105] To predict antimicrobial resistance from shotgun sequencing, most studies have relied on pairwise comparisons (e.g., by Basic Local Alignment Search Tool (BLAST)[106]) of sequences to databases of known or predicted resistance genes.[107–111] Many antimicrobial resistance databases are highly biased toward human-associated organisms; therefore, environmental antimicrobial resistance is comparatively ignored.[112] Prediction algorithms based on hidden Markov models (HMMs) trained on functionally validated resistance genes from diverse environments have proven to be superior to traditional pairwise annotators in both precision and accuracy of resistance gene annotation, since HMM profiles allow for the discovery of highly diverse and understudied antimicrobial resistance genes.[112]

Antimicrobial resistance gene databases

A number of databases that store nucleotide and protein sequence information for antimicrobial resistance genes have been developed. Query sequences are aligned to sequences of known antimicrobial resistance determinants in the databases by BLAST,[106] which performs a strict pairwise DNA or protein sequence alignment, to catalog the antimicrobial resistance gene content of a genome or metagenome. These databases have dramatically advanced the field of genomic and metagenomic analyses of antimicrobial resistance by allowing the sequence-based identification of known resistance determinants.[113]

The Antibiotic Resistance Gene Database (ARDB) (http://ardb.cbcb.umd.edu/) is a manually curated antimicrobial resistance database.[114] The ARDB provides a comprehensive ontology that creates a resistance profile by matching specific genes and mechanisms of action. It uses BLAST to identify and annotate antimicrobial resistance genes. Many studies have used the ARDB as their sole method to identify antimicrobial resistance genes[115–117] and have revealed that a majority of antimicrobial resistance genes cluster based on ecology.[107,116,118] Although this database was the first of its kind, it has not been updated since 2009.

The Comprehensive Antibiotic Resistance Database (CARD) (https://card.mcmaster.ca/) is a collection of known resistance determinants and associated antimicrobials.[119] It is designed to predict known antimicrobial resistance from genome sequence data using a BLAST-based pairwise sequence alignment to known antimicrobial resistance genes in the database. The database is organized on the basis of ontologies, linking genes to their function. This is an ideal database to use if the antimicrobial resistance mechanism of interest has been thoroughly researched. However, it is not suited for the detection of point mutations in chromosomal target genes known to be associated with antimicrobial resistance genes.[120] Many studies have annotated antimicrobial resistance using the CARD alone or in combination with the ARDB.[110,111,121,122] Unlike the ARDB, the CARD is actively managed and periodically updated.

The Bush–Jacoby β-lactamase list is a curated database that matches β-lactamase gene sequences to resistance to β-lactam antimicrobials. This database will soon be hosted on the National Center for Biotechnology Information (NCBI) website. This database was originally focused heavily on TEM, SHV, and OXA type β-lactamases, only a subset of β-lactam antimicrobial resistance genes. Other β-lactamases that are now included in the database are CTX-M, CMY, AmpC, CARB, IMP, VIM, KPC, GES, PER, and VEB.[113]

The Lactamase Engineering Database (LacED) (http://www.laced.uni-stuttgart.de/)[123] is an extension of the Bush–Jacoby β-lactamase list. This database includes a way to predict TEM, SHV, and class B enzymes by merging the information from both the NCBI peptide database and the TEM mutation table. The LacED builds protein families from these databases and integrates protein sequence and structure information using DWARF.[124]

The Repository of Antibiotic Resistance Cassettes (http://rac.aihi.mq.edu.au/rac/) takes advantage of the fact that resistance genes are often syntenic with mobile genetic elements, forming resistance cassettes. This database was the first to simultaneously automate the identification of antimicrobial resistance genes and place the genes in their broader genetic context.[125]

ResFinder (https://cge.cbs.dtu.dk//services/ResFinder/)[126] combines the Bush–Jacoby β-lactamase list with the ARDB and other published antimicrobial resistance gene sequences. Similar to other resistance databases, query sequences are matched to the sequences in the database via BLAST. The difference between ResFinder and most other resistance gene databases is the ability for the user to tailor the identity and length coverage thresholds. This allows the user to specify settings that are suited for the depth and quality of sequencing for each project. The current challenge in using ResFinder and the previously described databases is the inability to classify chromosomal mutations that can lead to antimicrobial resistance or to classify novel resistance genes, and it is thus limited to resistance genes commonly acquired through horizontal gene transfer.[120]

The Antibiotic Resistance Gene-Annotation (ARG-ANNOT) (http://www.mediterranee-infection.com/article.php?laref=282&titer=arg-annot)[120] uses a BLAST-enabled search on a curated antimicrobial resistance gene database compiled from the Bush–Jacoby β-lactamase list, Resfinder,[126] the ARDB,[114] and others. A main difference between this program and others is the ability to additionally predict resistance function on the basis of chromosomal point mutations. ARG-ANNOT contains a curated database focused on mutations found in many chromosomal genes (i.e., *rpoB, gyrA1, gyrA2, parC*) that can confer resistance to antimicrobials.

It is difficult to identify novel resistance genes when using pairwise comparisons (e.g., by BLAST) to resistance databases, which are largely based on resistance genes from cultured bacterial isolates and are especially biased toward human-associated pathogens. Resfams (http://www.dantaslab.org/resfams/) implements an alternative method in order to identify known and novel resistance genes with high precision and accuracy.[112] Resfams is a curated database of profile HMMs built on resistance proteins compiled from the CARD,[119] the LacED,[123] and Jacoby and Bush's collection of curated β-lactamases, based on the gene ontology from the CARD database. Additionally, the Resfams profile HMMs incorporate antimicrobial resistance genes from environmental, human commensal, and pathogenic bacteria that have been discovered through culture-independent functional metagenomic selections. The HMM approach is superior to pairwise annotation approaches in both precision and accuracy. Pairwise BLAST searches against

resistance gene databases were unable to identify over 60% of resistance genes identified by Resfams from a functionally validated resistance gene test set.[112] Perhaps more importantly, differences in HMM-based versus BLAST-based pairwise resistome analyses of the same data can lead to significantly different (and sometimes completely contradictory) biological conclusions. For instance, comparison of resistome predictions from over 6000 bacterial genomes by Resfams versus predictions by pairwise alignment of resistance gene databases revealed that pairwise methods are inaccurately biased in both their phylogenetic and ecological distributions toward the relatively oversampled human-associated environments that they represent.[112]

These myriad resistance gene databases are clearly important tools for interrogating microbiomes for antimicrobial resistance genes, but some challenges in their use and opportunities for their improvement exist. An important limitation to both HMM- and BLAST-based searches is the implicit requirement for users to determine the accuracy of the predicted resistance gene function, especially when the exact sequence being annotated has not previously been functionally assayed. Additionally, many of these databases do not incorporate information on the host-specific functionality of some antimicrobial resistance genes.[127] For example, most Gram-negative bacteria are intrinsically resistant to important classes of Gram-positive–specific drugs (e.g., vancomycin, linezolid, and macrolides) because they cannot cross the Gram-negative outer membrane.[48] Accordingly, it is likely inappropriate to annotate homologs of genes that provide resistance to these antimicrobials in Gram-positive bacteria when they are identified in Gram-negative bacteria. An important future goal in these databases is a species-dependent risk estimator for horizontal gene transfer to quantify the likelihood of resistance gene dissemination between species and habitats. Additionally, it is important that environmental or nonclinical resistance genes, which are currently underrepresented in these databases, be included.[112] Because most clinical resistance originated in the environment, this will facilitate the surveillance of emerging resistance mechanisms as they disseminate beyond benign organisms. Despite the inherent limitations in ascribing resistance function based on sequence alone, antimicrobial

resistance databases play a key role in curating sequences of known resistance determinants. It is crucial that these databases are regularly updated and simultaneously cross-validated to maintain accuracy and relevance. Continued maintenance and improvement of these databases, including the ongoing incorporation of novel, functionally validated resistance determinants and associated metadata from diverse habitats, will ensure that they remain valuable resources for the antimicrobial resistance community for years to come.

It is likely that the selective pressure that drove the evolution of most antimicrobial resistance genes is the incredible biosynthetic potential of microbial communities to synthesize natural products with antimicrobial activity. This motivates the need to understand antimicrobial resistance in microbial communities in the context of antimicrobial biosynthesis and vice versa. Paralleling recent improvements in methods to characterize resistomes, there have also been significant advances in functional and sequence-based methods for discovering novel antimicrobial biosynthesis machinery, which we will review below.

Function-based methods for antimicrobial discovery from microbiomes

Function-based methods for the discovery of antimicrobials from microbiomes present a clear advantage over sequence-based methods in that no a priori knowledge of the sequence of a natural product biosynthetic pathway is required. This facilitates the discovery of novel natural products that are produced by biosynthetic pathways that are divergent from previously characterized pathways and provides access to novel chemical diversity. This unbiased approach has been successfully used for the discovery of natural products with antimicrobial activity.[128,129] The method depends on shotgun cloning of metagenomic DNA from a given environment into an expression vector, heterologous expression of pathways encoded in the metagenomic DNA in an appropriate host, and screening for production of bioactive natural products.[130]

Despite the advantages that this unbiased approach provides, there are several challenges that have hindered the discovery of natural-product antimicrobials using functional metagenomics. The main disadvantage is that construction and heterologous expression of large insert libraries are

technically challenging.[131] Strategies to improve the heterologous expression of biosynthetic pathways, including overexpression of alternative σ factors in *E. coli*, have proven successful for heterologous production of polyketides and hold promise for the development of improved metagenomic screening hosts.[132] Additionally, nontraditional heterologous hosts, such as *Streptomyces lividans*, have demonstrated utility in the discovery of bioactive natural products from metagenomic libraries.[133] A final challenge inherent in functional metagenomic antimicrobial discovery lies in the fact that antimicrobial production is a more complicated phenotype to select for than antimicrobial resistance, for reasons that we will detail below. This necessitates the use of clever screening strategies to identify and isolate clones of interest.[131]

One screening strategy that successfully enables the discovery of antimicrobial biosynthesis from metagenomic libraries is to directly assay for antimicrobial activity against an indicator strain using a double–agar layer method.[134] In this method, the heterologous host of the metagenomic library is allowed to grow on plates, after which top agar containing the indicator strain is overlaid. This method was used to discover two clones from a 113,500 member fosmid library representing 4 Gb constructed from a soil metagenome that were able to inhibit the growth of *Bacillus subtilis*. The clones encoded a biosynthetic pathway that produced indigo and indirubicin, pigments that were confirmed to be responsible for the observed antibacterial activity.[128] Another application of this method resulted in the identification of 65 clones with antibacterial activity from a 700,000-member soil library. A small molecule from the clone with the highest apparent bioactivity was purified and identified as a long-chain *N*-acyl amino acid. This represented the first report of the antibacterial activity of long-chain *N*-acyl amino acids, as well as the first identification of genes involved in long-chain *N*-acyl amino acid biosynthesis, highlighting two key advantages of the functional metagenomic approach to antimicrobial discovery.[129]

An alternative approach that extends traditional culturing methods uses the innovative iChip, a multichannel, microfluidic diffusion growth chamber.[79] This device was recently used to discover a novel nonribosomal peptide with antimicrobial activity from a soil microbial community.[135] The iChip requires that input microbial community samples are diluted such that each channel receives approximately one bacterial cell. The device is sealed with a semipermeable membrane and returned to the source so that nutrients and growth factors from the natural environment of the source organisms can diffuse into the chamber, increasing the likelihood of culturing fastidious organisms.[79] This high-throughput method was used to culture 10,000 diverse isolates from a soil sample, which were predicted to be recalcitrant to culture using traditional methods. Screening of extracts from these isolates for antimicrobial activity revealed the presence of a nonribosomal peptide named teixobactin, synthesized by a previously uncultured betaproteobacterial species. Teixobactin was isolated and characterized and found to have activity against multidrug-resistant Gram-positive pathogens *in vitro*, as well as in animal models for MRSA and *Streptococcus pneumoniae*.[135] Although this method relied on classical antimicrobial screening platforms, innovative culture methods facilitated the discovery of a novel antimicrobial and revealed the biosynthetic potential of the uncultivable fraction of microbial communities.

Sequence-based methods for antimicrobial discovery from microbiomes

Sequence-based methods for the discovery of antimicrobial biosynthetic pathways from microbiomes depend on homology between the query sequence and known biosynthetic pathways. A clear advantage of this method is throughput. Advancements in annotation and prediction algorithms coupled with the abundance of microbial genome and metagenome sequences have made feasible the analysis of sequence data for the identification of biosynthetic gene clusters that may encode novel bioactive natural products. A disadvantage of this strategy is that prediction of the structure of a small molecule from the sequence of a biosynthetic gene cluster is nontrivial, as is prediction of a small molecule's activity from its structure.

One of the first attempts to identify biosynthetic pathways from an uncultured community relied on PCR amplification of ketosynthase genes from *Streptomyces* isolates, as well as from metagenomic DNA isolated from soil.[136] The amplicons were purified, cloned into an expression vector, and sequenced. Of 20 randomly selected transformants

that were sequenced, only two had similarity to any known ketosynthase gene products. Four of the novel ketosynthases were shown to be functional when expressed in a *S. lividans* or *Streptomyces glaucescens* host as part of a hybrid polyketide synthase pathway (i.e., the native ketosynthase was replaced with the newly discovered ketosynthase). The polyketide products of the hybrid biosynthetic pathways were purified and characterized to be novel octaketide and decaketide natural products. This highlights the utility of culture-free methods for identifying novel, sequence-divergent genes involved in biosynthetic pathways to expand access to novel natural products with diverse bioactivities, including potential antimicrobial activity.

Tools for identifying genes involved in natural-product biosynthesis from sequence data are a significant advancement over targeted PCR-based studies. One of the resources developed for mining genomic data for biosynthetic gene clusters is antiSMASH.[137] AntiSMASH makes use of the ClusterFinder algorithm,[138] which is an HMM-based algorithm for the discovery of biosynthetic gene clusters from genomic data. The ClusterFinder algorithm comprises a four-step prediction pipeline that can identify biosynthetic gene clusters in genomic data. The use of an HMM-based, probabilistic algorithm allows identification of biosynthetic gene clusters that lack high similarity in domain structure to any known biosynthetic gene clusters, circumventing the problem that had previously been posed in identification of biosynthetic gene clusters from sequence data. AntiSMASH also includes an active site finder module that identifies conserved amino acid motifs in the active sites of key biosynthetic enzymes, including those involved in polyketide biosynthesis. A final, key feature of antiSMASH is an algorithm that predicts the structure of a natural product synthesized by the annotated biosynthetic gene cluster. Furthermore, the development of a web-based server (http://antismash. secondarymetabolites.org/) has democratized the process of identifying biosynthetic gene clusters, because it is available to those without access to high-throughput computing facilities.[137]

The ClusterFinder/antiSMASH workflow was successfully used to identify a natural product with antimicrobial activity from metagenomic sequencing data generated as part of the National Institutes of Health Human Microbiome Project. Genomes

of 2430 members of the human microbiota were queried using ClusterFinder to identify over 14,000 putative biosynthetic gene clusters. Metagenomic reads from 752 samples representing five body sites of healthy subjects were mapped to the 14,000 predicted biosynthetic gene clusters. Among the natural products predicted was a thiopeptide from the vaginal commensal *Lactobacillus gasseri*. The thiopeptide, named lactocillin, was purified and characterized. Lactocillin was found to have antimicrobial activity against Gram-positive bacteria, including *Enterococus faecalis*, *Gardnerella vaginalis*, *S. aureus*, and *Corynebacterium aurimucosum*. The discovery and characterization of the structure and bioactivity of lactocillin from human microbiome metagenomic sequencing data demonstrate the potential that diverse metagenomes hold for encoding natural products with antimicrobial activity.[8]

A current shortcoming of the software available to identify biosynthetic gene clusters from metagenomic sequencing data is that the assembly of short-read data into genomes or large contigs is required. Although there have been recent advances in metagenomic assemblies of this type,[139–141] it remains a challenge. One way to circumvent the requirement for metagenomic assembly is the use of ShortBRED,[142] which finds short, unique amino acid markers from reference proteins of interest and maps metagenomic sequencing reads to these markers. This tool has been successfully implemented for the identification of antimicrobial resistance genes[101] and has potential for the identification of genes involved in natural product biosynthesis directly from short-read metagenomic sequencing data.

An alternative method that can be used to access the biochemical diversity of microbiomes without the necessity for metagenomic assembly uses short DNA sequences that are conserved between biosynthetic gene clusters. PCR amplification of these conserved biosynthetic motifs from a metagenomic source using degenerate primers followed by amplicon sequencing can provide an overview of the biosynthetic capabilities of a given microbiome and highlight promising leads for heterologous expression and further characterization.[143] A bioinformatics platform for mapping short DNA sequence tags to known biosynthetic gene clusters, termed Environmental Surveyor of Natural Product Diversity, has been developed to facilitate analysis of

amplicon sequencing data in the context of natural-product biosynthesis.[144] This strategy has been successfully used in the identification of novel bioactive small molecules, including some with antimicrobial activity.[143,145]

A final, promising method for sequence-based identification of biosynthetic gene clusters from microbiomes involves assembly of complete or near-complete genomes using single-molecule, real-time sequencing (SMRT),[146] followed by annotation with antiSMASH or a similar annotation pipeline. This approach was successfully applied in tandem with short-read data reference–independent assembly of a high-quality *Corynebacterium simulans* genome *de novo* from a skin microbiota sample. This previously uncharacterized genome was annotated with antiSMASH, revealing predicted nonribosomal peptide synthetases, type I polyketide synthetases, terpene synthases, and bacteriocin production genes.[147] Interestingly, the bacteriocin production locus was homologous to the locus for biosynthesis of lactococcin, a bacteriocin that is known to be bactericidal.[148] Although the natural products whose biosynthetic gene clusters were identified in this study were not isolated or confirmed for bioactivity, the authors demonstrate the power of metagenomic assembly in the identification of biosynthetic gene clusters from microbiomes.

Future directions in antimicrobial and antimicrobial resistance discovery

The increasing threat to human health from multidrug-resistant pathogens underscores the need to develop tools to interrogate microbiomes for novel antimicrobials and to understand and mitigate the evolution and transmission of resistance genes. A more complete understanding of antimicrobial resistance in the context of microbial communities will advance our treatment of antimicrobial resistant infections and help us choose therapies that will reduce selection for resistance. Additionally, novel antimicrobials discovered within microbial communities will lay the groundwork for narrow-spectrum, next-generation antimicrobials that reduce collateral microbiome damage. Given that most clinical antimicrobial resistance genes originated in environmental microbes, and most antimicrobials are natural products or derivatives thereof, it is critical that these two microbial functions be studied in tandem and in the context

of microbial communities in order to understand the dynamics governing the evolution and transmission of antimicrobial resistance and biosynthetic capacities.

Advances in sequencing technology will facilitate the discovery of complex functions from metagenomes on the basis of sequence alone. For example, long-read sequencing technologies, such as the NanoPore Minion and the PacBio SMRT, will help to overcome the challenges currently posed by metagenomic assembly from short-read data, providing greater information about functions encoded in large operons, such as biosynthetic and multidrug-resistance gene clusters.[140,149,150] Furthermore, cheaper sequencing will allow for deeper sequencing, improving our ability to assemble metagenomes and profile antimicrobial resistance and biosynthetic capacities.[83] Single-molecule sequencing, particularly when combined with short-read shotgun metagenomic data, may also improve metagenomic assembly and allow for better identification of biosynthetic gene clusters.[147] In addition to progress resulting from improvements to DNA reading, advancements in DNA writing will likely benefit the field. Dramatic reductions in the cost and efficiency of DNA synthesis may soon make synthesis and recoding of entire biosynthetic gene clusters feasible, simplifying heterologous expression in genetically tractable hosts and downstream characterization of bioactive natural products.[151,152] An integration of functional and genomic techniques, aided by these and other inevitable technical and computational developments, will help us better understand the delicate interplay between antimicrobial biosynthesis and resistance in microbiomes.

Acknowledgments

Conflicts of interest

The authors declare no conflicts of interest.

References

1. Lederberg, J. & A.T. McCray. 2001. 'Ome sweet' omics—a genealogical treasury of words. *The Scientist.* **15:** 8.
2. Turnbaugh, P.J., R.E. Ley, M. Hamady, *et al.* 2007. The human microbiome project. *Nature* **449:** 804–810.
3. Allen, H.K., J. Donato, H.H. Wang, *et al.* 2010. Call of the wild: antibiotic resistance genes in natural environments. *Nat. Rev. Microbiol.* **8:** 251–259.
4. Sender, R., S. Fuchs & R. Milo. 2016. Are we really vastly outnumbered? Revisiting the ratio of bacterial to host cells in humans. *Cell* **164:** 337–340.
5. Ursell, L.K., J.L. Metcalf, L.W. Parfrey, *et al.* 2012. Defining the human microbiome. *Nutr. Rev.* **70**(Suppl. 1): S38–S44.
6. Roesch, L.F.W., R.R. Fulthorpe, A. Riva, *et al.* 2007. Pyrosequencing enumerates and contrasts soil microbial diversity. *ISME J.* **1:** 283–290.
7. Waldor, M.K., G. Tyson, E. Borenstein, *et al.* 2015. Where next for microbiome research? *PLoS Biol.* **13:** e1002050.
8. Donia, M.S., P. Cimermancic, C.J. Schulze, *et al.* 2014. A systematic analysis of biosynthetic gene clusters in the human microbiome reveals a common family of antibiotics. *Cell* **158:** 1402–1414.
9. Sommer, M.O., G.M. Church & G. Dantas. 2010. The human microbiome harbors a diverse reservoir of antibiotic resistance genes. *Virulence* **1:** 299–303.
10. Perry, J.A., E.L. Westman & G.D. Wright. 2014. The antibiotic resistome: what's new? *Curr. Opin. Microbiol.* **21:** 45–50.
11. Sommer, M.O. & G. Dantas. 2011. Antibiotics and the resistant microbiome. *Curr. Opin. Microbiol.* **14:** 556–563.
12. D'Costa, V.M., K.M. McGrann, D.W. Hughes, *et al.* 2006. Sampling the antibiotic resistome. *Science* **311:** 374–377.
13. Wright, G.D. 2007. The antibiotic resistome: the nexus of chemical and genetic diversity. *Nat. Rev. Microbiol.* **5:** 175–186.
14. O'Neill, J. 2016. Securing new drugs for future generations: the pipeline of antibiotics. Accesssed August 22, 2016. http://amr-review.org/sites/default/files/SECURING%20 NEW%20DRUGS%20FOR%20FUTURE%20 GENERATIONS%20FINAL%20WEB_0.pdf.
15. Walsh, C.T. & T.A. Wencewicz. 2014. Prospects for new antibiotics: a molecule-centered perspective. *J. Antibiot. (Tokyo)* **67:** 7–22.
16. Lewis, K. 2013. Platforms for antibiotic discovery. *Nat. Rev. Drug Discov.* **12:** 371–387.

17. Knapp, C.W., J. Dolfing, P.A. Ehlert, *et al.* 2010. Evidence of increasing antibiotic resistance gene abundances in archived soils since 1940. *Environ. Sci. Technol.* **44:** 580–587.
18. Centers for Disease Control and Prevention, 2013. Antibiotic resistance threats in the United States. Accessed August 22, 2016. http://www.cdc.gov/drugresistance/pdf/ ar-threats-2013-508.pdf.
19. Cooper, M.A. & D. Shlaes. 2011. Fix the antibiotics pipeline. *Nature* **472:** 32.
20. Roberts, R.R., B. Hota, I. Ahmad, *et al.* 2009. Hospital and societal costs of antimicrobial-resistant infections in a Chicago teaching hospital: implications for antibiotic stewardship. *Clin. Infect. Dis.* **49:** 1175–1184.
21. O'Neill, J. 2016. Tackling drug-resistant infections globally: final report and recommendations. London: Review on Antimicrobial Resistance.
22. Fischbach, M.A. & C.T. Walsh. 2009. Antibiotics for emerging pathogens. *Science* **325:** 1089–1093.
23. Kinch, M.S., E. Patridge, M. Plummer, *et al.* 2014. An analysis of FDA-approved drugs for infectious disease: antibacterial agents. *Drug Discov. Today* **19:** 1283–1287.
24. Howard, D.H. & R.D. Scott. 2005. The economic burden of drug resistance. *Clin. Infect. Dis.* **41:** S283–S286.
25. Miller, A. 2016. Antibacterial development: a changing landscape. *Microbe Magazine.* Accessed August 22, 2016. http://www.asmscience.org/content/journal/microbe/ 10.1128/microbe.11.111.1.
26. Tenover, F.C. 2006. Mechanisms of antimicrobial resistance in bacteria. *Am. J. Infect. Control.* **34:** S3–S10; discussion S64–S73.
27. Walsh, C. 2000. Molecular mechanisms that confer antibacterial drug resistance. *Nature* **406:** 775–781.
28. Alekshun, M.N. & S.B. Levy. 2007. Molecular mechanisms of antibacterial multidrug resistance. *Cell* **128:** 1037–1050.
29. MacLean, R.C., A.R. Hall, G.G. Perron, *et al.* 2010. The population genetics of antibiotic resistance: integrating molecular mechanisms and treatment contexts. *Nat. Rev. Genet.* **11:** 405–414.
30. Baltz, R.H. 2010. Genomics and the ancient origins of the daptomycin biosynthetic gene cluster. *J. Antibiot. (Tokyo)* **63:** 506–511.
31. Fajardo, A. & J.L. Martinez. 2008. Antibiotics as signals that trigger specific bacterial responses. *Curr. Opin. Microbiol.* **11:** 161–167.
32. Clardy, J., M.A. Fischbach & C.R. Currie. 2009. The natural history of antibiotics. *Curr. Biol.* **19:** R437–R441.
33. Nelson, M.L., A. Dinardo, J. Hochberg, *et al.* 2010. Brief communication: mass spectroscopic characterization of tetracycline in the skeletal remains of an ancient population from Sudanese Nubia 350–550 CE. *Am. J. Phys. Anthropol.* **143:** 151–154.
34. Bassett, E.J., M.S. Keith, G.J. Armelagos, *et al.* 1980. Tetracycline-labeled human bone from ancient Sudanese Nubia (A.D. 350). *Science* **209:** 1532–1534.
35. Aminov, R.I. 2010. A brief history of the antibiotic era: lessons learned and challenges for the future. *Front. Microbiol.* **1:** 134.

36. Jones, D.S., S.H. Podolsky & J.A. Greene. 2012. The burden of disease and the changing task of medicine. *N. Engl. J. Med.* **366:** 2333–2338.

37. Elsner, H.L. 1910. The new treatment of syphilis (ehrlichhata): observations and results. *J. Am. Med. Assoc.* **55:** 2052–2057.

38. Fleming, A. 2001. On the antibacterial action of cultures of a penicillium, with special reference to their use in the isolation of *B. influenzae*. 1929. *Bull. World Health Organ.* **79:** 780–790.

39. Schatz, A., E. Bugie & S.A. Waksman. 1944. Streptomycin, a substance exhibiting antibiotic activity against gram-positive and gram-negative bacteria. *Clin. Orthop. Relat. Res.* **55:** 66–69.

40. Zetterstrom, R. 2007. Selman A. Waksman (1888–1973) Nobel Prize in 1952 for the discovery of streptomycin, the first antibiotic effective against tuberculosis. *Acta Paediatr.* **96:** 317–319.

41. Brown, E.D. & G.D. Wright. 2016. Antibacterial drug discovery in the resistance era. *Nature* **529:** 336–343.

42. Pelaez, F. 2006. The historical delivery of antibiotics from microbial natural products—can history repeat? *Biochem. Pharmacol.* **71:** 981–990.

43. Fernandes, P. 2006. Antibacterial discovery and development—the failure of success? *Nat. Biotechnol.* **24:** 1497–1503.

44. Zaffiri, L., J. Gardner & L.H. Toledo-Pereyra. 2012. History of antibiotics. From salvarsan to cephalosporins. *J. Invest. Surg.* **25:** 67–77.

45. Walsh, C. 2003. *Antibiotics: Actions, Origins, Resistance.* Washington, DC: ASM Press.

46. Davies, J. & D. Davies. 2010. Origins and evolution of antibiotic resistance. *Microbiol. Mol. Biol. Rev.* **74:** 417–433.

47. de la Cruz, F. & J. Davies. 2000. Horizontal gene transfer and the origin of species: lessons from bacteria. *Trends Microbiol.* **8:** 128–133.

48. Blair, J.M., M.A. Webber, A.J. Baylay, *et al.* 2015. Molecular mechanisms of antibiotic resistance. *Nat. Rev. Microbiol.* **13:** 42–51.

49. Roberts, M.C. 1996. Tetracycline resistance determinants: mechanisms of action, regulation of expression, genetic mobility, and distribution. *FEMS Microbiol. Rev.* **19:** 1–24.

50. Ogawa, W., M. Onishi, R. Ni, *et al.* 2012. Functional study of the novel multidrug efflux pump KexD from *Klebsiella pneumoniae*. *Gene* **498:** 177–182.

51. Ruggerone, P., S. Murakami, K.M. Pos, *et al.* 2013. RND efflux pumps: structural information translated into function and inhibition mechanisms. *Curr. Top. Med. Chem.* **13:** 3079–3100.

52. Hinchliffe, P., M.F. Symmons, C. Hughes, *et al.* 2013. Structure and operation of bacterial tripartite pumps. *Annu. Rev. Microbiol.* **67:** 221–242.

53. Piddock, L.J. 2006. Clinically relevant chromosomally encoded multidrug resistance efflux pumps in bacteria. *Clin. Microbiol. Rev.* **19:** 382–402.

54. Tamber, S. & R.E. Hancock. 2003. On the mechanism of solute uptake in *Pseudomonas*. *Front. Biosci.* **8:** s472–s483.

55. Poulou, A., E. Voulgari, G. Vrioni, *et al.* 2013. Outbreak caused by an ertapenem-resistant, CTX-M-15-producing *Klebsiella pneumoniae* sequence type 101 clone carrying an OmpK36 porin variant. *J. Clin. Microbiol.* **51:** 3176–3182.

56. Katayama, Y., T. Ito & K. Hiramatsu. 2000. A new class of genetic element, staphylococcus cassette chromosome *mec*, encodes methicillin resistance in *Staphylococcus aureus*. *Antimicrob. Agents Chemother.* **44:** 1549–1555.

57. Neu, H.C. 1969. Effect of β-lactamase location in *Escherichia coli* on penicillin synergy. *Appl. Microbiol.* **17:** 783–786.

58. Bush, K. 2010. Bench-to-bedside review: the role of β-lactamases in antibiotic-resistant Gram-negative infections. *Crit. Care* **14:** 224.

59. Yang, W., I.F. Moore, K.P. Koteva, *et al.* 2004. TetX is a flavin-dependent monooxygenase conferring resistance to tetracycline antibiotics. *J. Biol. Chem.* **279:** 52346–52352.

60. Forsberg, K.J., S. Patel, T.A. Wencewicz, *et al.* 2015. The tetracycline destructases: a novel family of tetracycline-inactivating enzymes. *Chem. Biol.* **22:** 888–897.

61. Shaw, W.V., L.C. Packman, B.D. Burleigh, *et al.* 1979. Primary structure of a chloramphenicol acetyltransferase specified by R plasmids. *Nature* **282:** 870–872.

62. D'Costa, V.M., C.E. King, L. Kalan, *et al.* 2011. Antibiotic resistance is ancient. *Nature* **477:** 457–461.

63. Bhullar, K., N. Waglechner, A. Pawlowski, *et al.* 2012. Antibiotic resistance is prevalent in an isolated cave microbiome. *PLoS One* **7:** e34953.

64. Palmer, A.C., E. Angelino & R. Kishony. 2010. Chemical decay of an antibiotic inverts selection for resistance. *Nat. Chem. Biol.* **6:** 105–107.

65. Martinez, J.L. 2012. Natural antibiotic resistance and contamination by antibiotic resistance determinants: the two ages in the evolution of resistance to antimicrobials. *Front. Microbiol.* **3:** 1.

66. Abraham, E.P. & E. Chain. 1988. An enzyme from bacteria able to destroy penicillin. 1940. *Rev. Infect. Dis.* **10:** 677–678.

67. Hall, B.G., S.J. Salipante & M. Barlow. 2004. Independent origins of subgroup Bl + B2 and subgroup B3 metallo-beta-lactamases. *J. Mol. Evol.* **59:** 133–141.

68. Hall, B.G. & M. Barlow. 2004. Evolution of the serine beta-lactamases: past, present and future. *Drug Resist. Updat.* **7:** 111–123.

69. Stokes, H.W. & M.R. Gillings. 2011. Gene flow, mobile genetic elements and the recruitment of antibiotic resistance genes into Gram-negative pathogens. *FEMS Microbiol. Rev.* **35:** 790–819.

70. Gould, S.J. & E.S. Vrba. 1982. Exaptation—a missing term in the science of form. *Paleobiology* **8:** 4–15.

71. Sommer, M.O., G. Dantas & G.M. Church. 2009. Functional characterization of the antibiotic resistance reservoir in the human microflora. *Science* **325:** 1128–1131.

72. Forsberg, K.J., A. Reyes, B. Wang, *et al.* 2012. The shared antibiotic resistome of soil bacteria and human pathogens. *Science* **337:** 1107–1111.

73. D'Costa, V.M., E. Griffiths & G.D. Wright. 2007. Expanding the soil antibiotic resistome: exploring environmental diversity. *Curr. Opin. Microbiol.* **10:** 481–489.

74. Didelot, X., R. Bowden, D.J. Wilson, *et al.* 2012. Transforming clinical microbiology with bacterial genome sequencing. *Nat. Rev. Genet.* **13:** 601–612.

75. Kwong, J.C., N. McCallum, V. Sintchenko, *et al.* 2015. Whole genome sequencing in clinical and public health microbiology. *Pathology* **47:** 199–210.

76. Koser, C.U., J.M. Bryant, J. Becq, *et al.* 2013. Whole-genome sequencing for rapid susceptibility testing of *M. tuberculosis.* *N. Engl. J. Med.* **369:** 290–292.

77. Li, L., N. Mendis, H. Trigui, *et al.* 2014. The importance of the viable but non-culturable state in human bacterial pathogens. *Front. Microbiol.* **5:** 258.

78. Ma, L., J. Kim, R. Hatzenpichler, *et al.* 2014. Gene-targeted microfluidic cultivation validated by isolation of a gut bacterium listed in Human Microbiome Project's Most Wanted taxa. *Proc. Natl. Acad. Sci. U.S.A.* **111:** 9768–9773.

79. Nichols, D., N. Cahoon, E.M. Trakhtenberg, *et al.* 2010. Use of ichip for high-throughput *in situ* cultivation of "uncultivable" microbial species. *Appl. Environ. Microbiol.* **76:** 2445–2450.

80. Kaeberlein, T., K. Lewis & S.S. Epstein. 2002. Isolating "uncultivable" microorganisms in pure culture in a simulated natural environment. *Science* **296:** 1127–1129.

81. Lagier, J.C., F. Armougom, M. Million, *et al.* 2012. Microbial culturomics: paradigm shift in the human gut microbiome study. *Clin. Microbiol. Infect.* **18:** 1185–1193.

82. McLain, J.E., E. Cytryn, L.M. Durso, *et al.* 2016. Culture-based methods for detection of antibiotic resistance in agroecosystems: advantages, challenges, and gaps in knowledge. *J. Environ. Qual.* **45:** 432–440.

83. Wetterstrand, K.A. DNA sequencing costs: data from the NHGRI Genome Sequencing Program (GSP). Accessed August 22, 2016. www.genome.gov/sequencingcostsdata.

84. Pettersson, E., J. Lundeberg & A. Ahmadian. 2009. Generations of sequencing technologies. *Genomics* **93:** 105–111.

85. Bonetta, L. 2010. Whole-genome sequencing breaks the cost barrier. *Cell* **141:** 917–919.

86. Niedringhaus, T.P., D. Milanova, M.B. Kerby, *et al.* 2011. Landscape of next-generation sequencing technologies. *Anal. Chem.* **83:** 4327–4341.

87. Christensen, K.D., D. Dukhovny, U. Siebert, *et al.* 2015. Assessing the costs and cost-effectiveness of genomic sequencing. *J. Pers. Med.* **5:** 470–486.

88. Mardis, E.R. 2008. Next-generation DNA sequencing methods. *Annu. Rev. Genomics Hum. Genet.* **9:** 387–402.

89. Koboldt, D.C., K.M. Steinberg, D.E. Larson, *et al.* 2013. The next-generation sequencing revolution and its impact on genomics. *Cell* **155:** 27–38.

90. Pehrsson, E.C., K.J. Forsberg, M.K. Gibson, *et al.* 2013. Novel resistance functions uncovered using functional metagenomic investigations of resistance reservoirs. *Front. Microbiol.* **4:** 145.

91. Allen, H.K., L.A. Moe, J. Rodbumrer, *et al.* 2009. Functional metagenomics reveals diverse β-lactamases in a remote Alaskan soil. *ISME J.* **3:** 243–251.

92. Moore, A.M., S. Patel, K.J. Forsberg, *et al.* 2013. Pediatric fecal microbiota harbor diverse and novel antibiotic resistance genes. *PLoS One* **8:** e78822.

93. Moore, A.M., C. Munck, M.O. Sommer, *et al.* 2011. Functional metagenomic investigations of the human intestinal microbiota. *Front. Microbiol.* **2:** 188.

94. Perron, G.G., L. Whyte, P.J. Turnbaugh, *et al.* 2015. Functional characterization of bacteria isolated from ancient arctic soil exposes diverse resistance mechanisms to modern antibiotics. *PLoS One* **10:** e0069533.

95. Clemente, J.C., E.C. Pehrsson, M.J. Blaser, *et al.* 2015. The microbiome of uncontacted Amerindians. *Sci. Adv.* **1.** pii: e1500183.

96. Pehrsson, E.C., P. Tsukayama, S. Patel, *et al.* 2016. Interconnected microbiomes and resistomes in low-income human habitats. *Nature* **533:** 212–216.

97. Lam, K.N., J. Cheng, K. Engel, *et al.* 2015. Current and future resources for functional metagenomics. *Front. Microbiol.* **6:** 1196.

98. Schmieder, R. & R. Edwards. 2012. Insights into antibiotic resistance through metagenomic approaches. *Future Microbiol.* **7:** 73–89.

99. Forsberg, K.J., S. Patel, M.K. Gibson, *et al.* 2014. Bacterial phylogeny structures soil resistomes across habitats. *Nature* **509:** 612–616.

100. Mezger, A., E. Gullberg, J. Goransson, *et al.* 2015. A general method for rapid determination of antibiotic susceptibility and species in bacterial infections. *J. Clin. Microbiol.* **53:** 425–432.

101. Gibson, M.K., B. Wang, S. Ahmadi, *et al.* 2016. Developmental dynamics of the preterm infant gut microbiota and antibiotic resistome. *Nat. Microbiol.* **1:** 16024.

102. Buelow, E., T.B. Gonzalez, D. Versluis, *et al.* 2014. Effects of selective digestive decontamination (SDD) on the gut resistome. *J. Antimicrob. Chemother.* **69:** 2215–2223.

103. Bengtsson-Palme, J., F. Boulund, J. Fick, *et al.* 2014. Shotgun metagenomics reveals a wide array of antibiotic resistance genes and mobile elements in a polluted lake in India. *Front. Microbiol.* **5:** 648.

104. Meyer, F., D. Paarmann, M. D'Souza, *et al.* 2008. The metagenomics RAST server—a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinformatics* **9:** 386.

105. Segata, N., L. Waldron, A. Ballarini, *et al.* 2012. Metagenomic microbial community profiling using unique clade-specific marker genes. *Nat. Methods* **9:** 811–814.

106. Altschul, S.F., W. Gish, W. Miller, *et al.* 1990. Basic local alignment search tool. *J. Mol. Biol.* **215:** 403–410.

107. Hu, Y., X. Yang, J. Qin, *et al.* 2013. Metagenome-wide analysis of antibiotic resistance genes in a large cohort of human gut microbiota. *Nat. Commun.* **4:** 2151.

108. Forslund, K., S. Sunagawa, J.R. Kultima, *et al.* 2013. Country-specific antibiotic use practices impact the human gut resistome. *Genome Res.* **23:** 1163–1169.

109. Durso, L.M., G.P. Harhay, J.L. Bono, *et al.* 2011. Virulence-associated and antibiotic resistance genes of microbial populations in cattle feces analyzed using a metagenomic approach. *J. Microbiol. Methods* **84:** 278–282.

110. Ma, L., B. Li & T. Zhang. 2014. Abundant rifampin resistance genes and significant correlations of antibiotic resistance genes and plasmids in various environments revealed by metagenomic analysis. *Appl. Microbiol. Biotechnol.* **98:** 5195–5204.

111. Abeles, S.R., M. Ly, T.M. Santiago-Rodriguez, *et al.* 2015. Effects of long term antibiotic therapy on human oral and fecal viromes. *PLoS One* **10:** e0134941.

112. Gibson, M.K., K.J. Forsberg & G. Dantas. 2015. Improved annotation of antibiotic resistance determinants reveals microbial resistomes cluster by ecology. *ISME J.* **9:** 207–216.

113. Xavier, B.B., A.J. Das, G. Cochrane, *et al.* 2016. Consolidating and exploring antibiotic resistance gene data resources. *J. Clin. Microbiol.* **54:** 851–859.

114. Liu, B. & M. Pop. 2009. ARDB—antibiotic resistance genes database. *Nucleic Acids Res.* **37:** D443–D447.

115. Looft, T., T.A. Johnson, H.K. Allen, *et al.* 2012. In-feed antibiotic effects on the swine intestinal microbiome. *Proc. Natl. Acad. Sci. U.S.A.* **109:** 1691–1696.

116. Nesme, J., S. Cecillon, T.O. Delmont, *et al.* 2014. Large-scale metagenomic-based study of antibiotic resistance in the environment. *Curr. Biol.* **24:** 1096–1100.

117. Modi, S.R., H.H. Lee, C.S. Spina, *et al.* 2013. Antibiotic treatment expands the resistance reservoir and ecological network of the phage metagenome. *Nature* **499:** 219–222.

118. Smillie, C.S., M.B. Smith, J. Friedman, *et al.* 2011. Ecology drives a global network of gene exchange connecting the human microbiome. *Nature* **480:** 241–244.

119. McArthur, A.G., N. Waglechner, F. Nizam, *et al.* 2013. The comprehensive antibiotic resistance database. *Antimicrob. Agents Chemother.* **57:** 3348–3357.

120. Gupta, S.K., B.R. Padmanabhan, S.M. Diene, *et al.* 2014. ARG-ANNOT, a new bioinformatic tool to discover antibiotic resistance genes in bacterial genomes. *Antimicrob. Agents Chemother.* **58:** 212–220.

121. Yamashita, A., T. Sekizuka & M. Kuroda. 2014. Characterization of antimicrobial resistance dissemination across plasmid communities classified by network analysis. *Pathogens* **3:** 356–376.

122. McCall, C.A., E. Bent, T.S. Jorgensen, *et al.* 2016. Metagenomic comparison of antibiotic resistance genes associated with liquid and dewatered biosolids. *J. Environ. Qual.* **45:** 463–470.

123. Thai, Q.K., F. Bos & J. Pleiss. 2009. The Lactamase Engineering Database: a critical survey of TEM sequences in public databases. *BMC Genomics* **10:** 390.

124. Fischer, M., Q.K. Thai, M. Grieb, *et al.* 2006. DWARF—a data warehouse system for analyzing protein families. *BMC Bioinformatics* **7:** 495.

125. Tsafnat, G., J. Copty & S.R. Partridge. 2011. RAC: repository of antibiotic resistance cassettes. *Database (Oxford)* **2011:** bar054.

126. Zankari, E., H. Hasman, S. Cosentino, *et al.* 2012. Identification of acquired antimicrobial resistance genes. *J. Antimicrob. Chemother.* **67:** 2640–2644.

127. McArthur, A.G. & G.D. Wright. 2015. Bioinformatics of antimicrobial resistance in the age of molecular epidemiology. *Curr. Opin. Microbiol.* **27:** 45–50.

128. Lim, H.K., E.J. Chung, J.C. Kim, *et al.* 2005. Characterization of a forest soil metagenome clone that confers indirubin and indigo production on *Escherichia coli*. *Appl. Environ. Microbiol.* **71:** 7768–7777.

129. Brady, S.F. & J. Clardy. 2000. Long-chain *N*-acyl amino acid antibiotics isolated from heterologously expressed environmental DNA. *J. Am. Chem. Soc.* **122:** 12903–12904.

130. Banik, J.J. & S.F. Brady. 2010. Recent application of metagenomic approaches toward the discovery of antimicrobials and other bioactive small molecules. *Curr. Opin. Microbiol.* **13:** 603–609.

131. Katz, M., B.M. Hover & S.F. Brady. 2016. Culture-independent discovery of natural products from soil metagenomes. *J. Ind. Microbiol. Biotechnol.* **43:** 129–141.

132. Stevens, D.C., K.R. Conway, N. Pearce, *et al.* 2013. Alternative sigma factor over-expression enables heterologous expression of a type II polyketide biosynthetic pathway in *Escherichia coli*. *PLoS One* **8:** e64858.

133. McMahon, M.D., C. Guan, J. Handelsman, *et al.* 2012. Metagenomic analysis of *Streptomyces lividans* reveals host-dependent functional expression. *Appl. Environ. Microbiol.* **78:** 3622–3629.

134. Lee, M.H. & S.W. Lee. 2013. Bioprospecting potential of the soil metagenome: novel enzymes and bioactivities. *Genomics Inform.* **11:** 114–120.

135. Ling, L.L., T. Schneider, A.J. Peoples, *et al.* 2015. A new antibiotic kills pathogens without detectable resistance. *Nature* **517:** 455–459.

136. Seow, K.T., G. Meurer, M. Gerlitz, *et al.* 1997. A study of iterative type II polyketide synthases, using bacterial genes cloned from soil DNA: a means to access and use genes from uncultured microorganisms. *J. Bacteriol.* **179:** 7360–7368.

137. Weber, T., K. Blin, S. Duddela, *et al.* 2015. AntiSMASH 3.0—a comprehensive resource for the genome mining of biosynthetic gene clusters. *Nucleic Acids Res.* **43:** W237–W243.

138. Cimermancic, P., M.H. Medema, J. Claesen, *et al.* 2014. Insights into secondary metabolism from a global analysis of prokaryotic biosynthetic gene clusters. *Cell* **158:** 412–421.

139. Raveh-Sadka, T., B.C. Thomas, A. Singh, *et al.* 2015. Gut bacteria are rarely shared by co-hospitalized premature infants, regardless of necrotizing enterocolitis development. *Elife* **4**. doi: 10.7554/eLife.05477.

140. Frank, J.A., Y. Pan, A. Tooming-Klunderud, *et al.* 2016. Improved metagenome assemblies and taxonomic binning using long-read circular consensus sequence data. *Sci. Rep.* **6:** 25373.

141. Sangwan, N., F. Xia & J.A. Gilbert. 2016. Recovering complete and draft population genomes from metagenome datasets. *Microbiome* **4:** 8.

142. Kaminski, J., M.K. Gibson, E.A. Franzosa, *et al.* 2015. High-specificity targeted functional profiling in microbial communities with shortBRED. *PLoS Comput. Biol.* **11:** e1004557.

143. Owen, J.G., Z. Charlop-Powers, A.G. Smith, *et al.* 2015. Multiplexed metagenome mining using short DNA sequence tags facilitates targeted discovery of epoxyketone proteasome inhibitors. *Proc. Natl. Acad. Sci. U.S.A.* **112:** 4221–4226.

144. Reddy, B.V., A. Milshteyn, Z. Charlop-Powers, *et al.* 2014. eSNaPD: a versatile, web-based bioinformatics platform for surveying and mining natural product biosynthetic diversity from metagenomes. *Chem. Biol.* **21:** 1023–1033.

145. Reddy, B.V., D. Kallifidas, J.H. Kim, *et al.* 2012. Natural product biosynthetic gene diversity in geographically distinct soil microbiomes. *Appl. Environ. Microbiol.* **78:** 3744–3752.

146. Eid, J., A. Fehr, J. Gray, *et al.* 2009. Real-time DNA sequencing from single polymerase molecules. *Science* **323:** 133–138.

147. Tsai, Y.C., S. Conlan, C. Deming, *et al.* 2016. Resolving the complexity of human skin metagenomes using single-molecule sequencing. *MBio* **7:** e01948–e01915.

148. Oppegard, C., P. Rogne, L. Emanuelsen, *et al.* 2007. The two-peptide class II bacteriocins: structure, production, and mode of action. *J. Mol. Microbiol. Biotechnol.* **13:** 210–219.

149. Koren, S., G.P. Harhay, T.P. Smith, *et al.* 2013. Reducing assembly complexity of microbial genomes with single-molecule sequencing. *Genome Biol.* **14:** R101.

150. Gomez-Escribano, J.P., S. Alt & M.J. Bibb. 2016. Next generation sequencing of actinobacteria for the discovery of novel natural products. *Mar. Drugs* **14.** doi: 10.3390/md14040078.

151. Meyer, Q.C., S.G. Burton & D.A. Cowan. 2007. Subtractive hybridization magnetic bead capture: a new technique for the recovery of full-length ORFs from the metagenome. *Biotechnol. J.* **2:** 36–40.

152. Kosuri, S. & G.M. Church. 2014. Large-scale *de novo* DNA synthesis: technologies and applications. *Nat. Methods* **11:** 499–507.