Emerging evolutionary paradigms in antibiotic discovery

Marc G. Chevrette1,2 · Cameron R. Currie2

Received: 31 August 2018 / Accepted: 25 September 2018 © Society for Industrial Microbiology and Biotechnology 2018

Abstract
Antibiotics revolutionized medicine and remain its cornerstone. Despite their global importance and the continuous threat of resistant pathogens, few antibiotics have been discovered in recent years. Natural products, especially the secondary metabolites of Actinobacteria, have been the traditional discovery source of antibiotics. In nature, the chemistry of antibiotic natural products is shaped by the unique evolution and ecology of their producing organisms, yet these influences remain largely unknown. Here, we highlight the ecology of antibiotics employed by microbes in defensive symbioses and review the evolutionary processes underlying the chemical diversity and activity of microbe-derived antibiotics, including the dynamics of vertical and lateral transmission of biosynthetic pathways and the evolution of efficacy, targeting specificity, and toxicity. We argue that a deeper understanding of the ecology and evolution of microbial interactions and the metabolites that mediate them will allow for an alternative, rational approach to discover new antibiotics.

Keywords Natural products · Secondary metabolism · Evolution · Ecology · Antibiotics

Introduction
Antibiotics are essential in modern medicine, both as therapeutics for life-threatening infections and as enabling prophylactics in vital surgeries and cancer therapies. The discovery of streptomycin by Albert Schatz and Selman Waksman in 1943 was the first effective treatment for tuberculosis, decreasing its mortality rate by over 85% from 1940 to 1960 [37]. The subsequent discovery of many other molecules to treat infectious diseases ushered in the “Golden Age” of antibiotics in the mid-20th century. By 2015, 42.3 billion defined daily doses of antibiotics were used as medicine globally, a number expected to increase 202% to 128 billion in 2030 [69]. By 2050, deaths attributable to antimicrobial resistance are expected to grow from 700,000 to 10 million, surpassing deaths from cancer (8.2 million) and diabetes (1.5 million) [98]. Further, the widespread use of antibiotics in animal husbandry, plant agriculture, and biotechnology accentuates the critical importance of effective antibiotics across society.
Clinical antibiotic use places selective pressure on pathogens to evolve increasing levels of resistance, rapidly compromising the efficacy of antibiotic molecules [46]. The emergence of resistance mechanisms in many virulent human pathogens represents a burgeoning global public health crisis. Recently, the World Health Organization (WHO) has focused on the so-called ESKAPE (or more ominously, “No ESKAPE”) pathogens [12, 103], yet resistance continues to emerge across many more bacterial and fungal pathogenic lineages [48, 62, 124]. Gram-negative bacterial infections present a particularly difficult challenge, due in large part to the unique physiologies of their cell envelope [13]. Fungal infections are also exceedingly difficult to treat because many of the cellular processes targeted by antifungal molecules are conserved in humans, leading to off-target effects, high toxicity, and small therapeutic windows [72, 94, 105]. The threat of resistant infections is often exacerbated in hospital settings where many patients are already immunocompromised or otherwise stressed. An increasing number of pathogens now exhibit resistance mechanisms against multiple, or even all, known antibiotics making these “superbugs” a paramount clinical concern [48, 62, 124]. Deaths attributed to antibiotic resistant infections are expected to rise from the current estimate of 700,000 annually to over
10 million by 2050 at a total societal cost of US$100 trillion [98]. Emerging resistance to the few therapeutics available for these pathogens underscores the need to discover novel antibiotics to combat infectious disease.

To combat resistance, novel therapeutics to control and treat infectious diseases are required [12, 13]. Natural products have been the primary source of clinically approved antibiotics, as well as anticancer, immunosuppressant, and other therapeutically relevant molecules. Actinobacteria, especially Streptomyces from the soil, are prolific producers of antibiotic secondary metabolites [59, 66]. Exploring and harnessing the chemical potential of Streptomyces have changed the course of modern medicine and biotechnology, providing an invaluable source of molecules to treat infectious disease [11, 24, 108, 130]. As early as 1876, Louis Pasteur noted that the ability to understand and intervene in antagonism between microbes would offer “perhaps the greatest hopes for therapeutics” [68] and the molecules that mediate these antagonisms have become critical pillars of modern medicine. Following the serendipitous discovery of penicillin in 1928, Waksman and others pioneered the traditional antibiotic discovery pipeline: culturing Actinobacteria from the soil, screening them for in vitro inhibitory activity, and isolating the molecules responsible. These early efforts led to the discovery of the aminoglycosides (e.g., the antituberculosis molecule streptomycin), glycopeptides (e.g., the Gram-positive active vancomycin), macrolides (e.g., erythromycin), ansamycins (e.g., rifamycin), cephalosporins, tetracyclines, and many other examples from the early 1940s to the 1960s [66]. Despite these successes, discovery rates have recently stagnated [59, 136]. The antibiotic discovery drought coincided with a shift towards combinatorial synthesis and target-based drug discovery [7, 136], which further divested resources away from natural product screening efforts. In part, this was due to the assumption that the antibiotic discovery “well was dry” and early soil screens had exhausted nature’s medicine cabinet. A notable exception is the lipopeptide daptomycin (discovered in the early 1980s and approved for clinical use in 2003; [3]); however, approximately 10^7 soil Actinomycetes were screened to attain this discovery [4]. More recently, a 2015 discovery described the novel antibiotic teixobactin, a promising lead targeting Gram-positive peptidoglycan biosynthesis [75]. Further, the teixobactin scaffold is amenable to derivatization and analogs have shown activity against a various target organisms [52, 96, 101]. However, its discovery was not the result of a simple screen, but rather involved the engineering of new isolation techniques for phylogenetically divergent bacteria. These few exceptions fail to keep stride with the ever-growing threat of resistance [102, 108]. Many of the Key Access Antibiotics on the WHO List of Essential Medicines (20th edition, amended August 2017) have been in commercial use for over 50 years (Fig. 1). Further, there is a dire need for new classes of antifungal molecules, as there are few drugs currently available to treat life-threatening fungal infections [94]. To reinvigorate the antibiotic discovery pipeline, alternative approaches and new paradigms are needed.

Given the somewhat controversial discussions of their function as antimicrobials in nature, we present the case that in some natural systems, especially defensive symbioses, microbial natural products do indeed act as antibiotics. Next, we outline the processes underlying the evolution of the genes responsible for the biosynthesis of antibiotics. Building on these genetic and genomic principles, we provide examples of the phenotypic consequences of evolution on antibiotic chemistry. We then contrast the evolutionary constraints of these natural systems to the clinical settings in which antibiotics are used to combat human disease and
highlight nature’s influence on both target specificity and toxicity as dictated by evolutionary selection. Finally, we summarize the implications of these evolutionary insights on antibiotic discovery, argue their importance, and provide an outlook for the future of the field. Herein, we propose that this evolutionary framework provides valuable insights into how our field can: (1) discover antibiotic-producing microbes; (2) influence antibiotic production; (3) minimize toxicity; (4) maximize target efficacy; and (5) increase efficiency in the engineering and synthetic biology of antibiotic production.

**Antibiotics as antibiotics in nature**

Since the beginnings of the antibiotic era, thousands of natural products from microbes have been discovered [108], yet the ecological functions of the vast majority remain unknown. It has been widely assumed that natural products with in vitro antimicrobial activity function as antibiotics in nature; a microbe produces these chemical “weapons” to engage in interference competition within their ecological niche. This view is congruent with both antibiotics being energetically costly to produce and the occurrence of antibiotic-producing microbes in competitive niches (e.g., soil). However, early on, Waksman recognized that the evidence surrounding their natural functions was unclear [133, 134] and evidence for natural products acting as antibiotics in nature remains largely elusive. Further, Julian Davies and others have shown that the response to antibiotics at sub-lethal concentrations adds another layer of complexity to the contextual nature of how they mediate specific interactions, especially in those that target the ribosome and other core cellular machinery [34, 35, 54, 77, 113, 128]. Some antibiotics modulate gene expression in susceptible target organisms depending on antibiotic concentration [54], while others can act as species–species signals [44, 74, 78, 109, 138], calling into question whether antibiotics actually act as antibiotics in their natural contexts [35].

The microbes that produce antibiotics do not occur as pure, isolated cultures in nature. Instead they are individual players in a complex web of interactions, where they may also be the target of a neighbor’s chemical weaponry. However, the concentrations of antibiotics used in the clinic likely do not reflect the in situ concentrations of these molecules in their natural settings and the ecology of producing microbes is divorced from the laboratory conditions in which they are studied. Thus, the ecological roles of these molecules cannot be assumed to be bacterio- or fungicidal. Although others championed the idea that they may play a protective role in multi-species systems [70], the ability of a microbe to produce a molecule with antimicrobial activity does not indicate it has that ecological role. Synergistic and contingent response against competitors is thought to drive the evolution of antibiotics [21]; however, the large distribution of intrinsic resistance seen in microbes that themselves cannot produce a given antibiotic further highlights the incomplete understanding of the in situ function of these molecules [24, 25]. Further, antibiotic resistance mechanisms are both ancient and widespread [33, 127], even in isolated environments [8], suggesting that antibiotics themselves have a similarly long natural history shaped by evolution. The chemistry-mediated interaction networks between antibiotic producers, their competitors, and other sympatric organisms, all with varying response and resistance, represent a complicated fitness landscape. While these interactions can be difficult to parse in complex communities like soil, pairwise co-culture of sympatric microbes has been used to elicit antibiotic chemistry and discover new molecules [2, 36].

Discussion of the ambiguities surrounding the ecological function of antibiotics by Waksman, Julian Davies, and others argues that an understanding of the ecological role of these molecules in nature can help inform the field. Kraklinov [70] suggested that over evolutionary timescales, microbes have come to produce antibiotics as weapons against sympatric competitors, yet evidence suggests that their natural functions are rarely this simple [126]. Nonetheless, there are natural systems in which secondary metabolites do function as antibiotics. It is not our intent to comprehensively describe all such examples herein. Instead, we restrict our attention to a few notable and well-established examples below to highlight that microbially produced antibiotics can and do act as antibiotics in nature, specifically within symbiotic systems.

Arguably, the ecological niches with the most compelling evidence of natural products acting as antibiotics are within symbiotic relationships in which microbes help defend a multi-species system. These defensive symbioses offer ecological and evolutionary insight into microbial molecular interactions, particularly how microbial secondary metabolites operate in defense. One example is the fungus-farming ant system, where ants cultivate a symbiotic fungus as their sole food source [22]. Like many agricultural crops, the fungal cultivar is of limited genetic diversity and thus susceptible to pathogens. Actinobacteria of the genus *Pseudonocardia* reside in coevolved structures on the ant exoskeleton [31] and provide in situ protection for the cultivar by suppressing the growth and morbidity of the pathogenic fungus *Escovopsis* [28, 106]. This defense is mediated through *Pseudonocardia*’s production of metabolites with antibiotic activity [30, 32, 73]. Reciprocal selection has apparently generated an arms-race dynamic between the *Pseudonocardia* and *Escovopsis*, the study of which has resulted in the discovery of antibiotic secondary metabolites with remarkable structural and functional diversity, including dentigerumycin,
pseudonocardone, selvamycin, and others [20, 29, 99, 131, 132]. Inhibition of Escovopsis by Pseudonocardia is specific and is stronger than inhibition of related, non-Escovopsis fungi [17]. Further, free-living Pseudonocardia from outside the fungus-growing ant system (e.g., soil) have significantly less inhibition compared to Pseudonocardia from ant symbioses [17]. Importantly, Pseudonocardia does not deter the mutualistic fungal cultivar [29]. Escovopsis can elicit specific metabolic responses from near-identical strains of Pseudonocardia [50], suggesting that the production of antibiotic metabolites is strongly influenced by ant host, responses to infection, and other biological and environmental contexts. This variation in metabolite production parallels results from in vitro assays between Pseudonocardia and Escovopsis that show variation in inhibitory phenotype across strains [17] and environmental stimuli [76].

As in the fungus-growing ant system, other insects have formed defensive mutualisms with antibiotic-producing bacteria. One example is the southern pine beetle, Dendroctonus frontalis, that associates with a Streptomyces sp. that densely coats the walls of their larval galleries and produces antibiotics to protect the developing larvae [117]. Similar to the microbes in ant system, these Streptomyces produce antifungal compounds, including mycangimycin and the frontalamides, that inhibit fungal pathogens while the beneficial fungal food source remains unperturbed [10, 117]. Beewolves also associate with Streptomyces spp. that are critical for successful beewolf development [64, 71]. Streptomyces grow within specialized glands on the antennae of female beewolves [53] and are inoculated to the ceiling of the brood cell before eggs are laid [71]. These Streptomyces are incorporated into the beewolf cocoon where they produce a white film [64]. Limiting or removing access to this white substance dramatically reduces survival probability in developing larvae and renders them susceptible to pathogenic infection [64, 71]. Streptomyces protection is accomplished through a cocktail of antibiotics, including streptochlorin and multiple piericidin derivatives [71].

Hoopoes (genus Upupa) are colorful birds with crested head feathers found throughout Africa, Europe, and Asia. The Eurasian hoopoe, Upupa epops, secretes antibiotic-producing Enterococci from its uropygial glands that it inoculates on its eggs for protection against opportunistic microbes [80, 115, 123]. The bacteriocins and other antibiotics produced by these Enterococci protect the eggs and increase successful hatch rates [80, 115]. The eggs themselves have specialized structures that encourage colonization by the Enterococci [81], suggesting that this mutualism has been a longstanding association.

The cutaneous microbial community of eastern red-backed salamanders (Plethodon cinereus) protects with antibiotics that inhibit the virulent amphibian pathogen Batrachochytrium dendrobatidis [14, 79]. Janthinobacterium from the salamander skin microbiome produces the antibiotic molecules indole-3-carboxaldehyde and violacein in vitro at concentrations lethal to B. dendrobatidis [14]. Further, synergistic antibiotic activity is seen against B. dendrobatidis when cutaneous microbiome isolates are co-cultured, leading to the production of tryptophol and other cryptic antibiotics active against the pathogen [79].

Host-associated microbes can protect from infection by producing antimicrobial compounds in humans as well. Human microbiota (and thus their antibiotic metabolites) typically localize to distinct body sites and can provide antimicrobial defenses specific to their niche [92]. One example is the use of the cyclic peptide antibiotic lugdunin by a nasal commensal Staphylococcus lugdunensis [141]. Lugdunin exhibits potent inhibitory activity against methicillin- and vancomycin-resistant S. aureus, vancomycin-resistant Enterococci, Listeria monocytogenes, and Streptococcus pneumoniae [141]. Due to lugdunin’s high inhibition of S. aureus and the ubiquity of S. aureus throughout the nasal passage of both healthy and unhealthy individuals, the prevalent interaction mediated by lugdunin is likely to control S. aureus carriage. The human vaginal microbial community can also provide antibiotic defense to pathogenic invasion. The lactocilllin family of antibiotics produced by the vaginal isolate Lactobacillus gasseri is active against invasive pathogens but not vaginal commensals [38] suggesting a protective role that does not disturb the microbiome.

Natural products drive species–species interactions, some antagonistic, and as such are shaped by fitness landscapes and other evolutionary pressures [88]. In Darwin’s study of Galapagos finches, the beak morphology of certain finches corresponded to the interactions they had with their food source. For example, specialization of those with large, short beaks was able to break the shells of tough nuts that finches with more narrow beaks were not able to crack. The evolutionary forces that act on microbial chemistry can be thought of in the same way. There are structural and biosynthetic similarities within families of molecules, but (often subtle) differences may significantly influence function. Further, it is important to remember that antibiotics are one of many mechanisms that modulate competitive interactions and influence the complex microbial fitness landscape [126]. Just as the complex phenotype of beak size is driven by the finches’ interactions with their food, microbial chemistry is driven by their interactions with the world around them. The examples above of natural products acting as antibiotics within defensive symbioses demonstrate that these antibiotics are important for mediating host–pathogen interactions and that they represent a promising source of new therapeutics for human use. Further, close associations with antibiotic-producing defensive symbions have evolved independently multiple times. Understanding the unique evolutionary pressures and distinct ecologies of these systems
can inform on taxa and/or environments to look for novel antibiotics and help shape stewardship policies to mitigate clinical resistance.

**Evolutionary dynamics of antibiotic biosynthesis**

An understanding of the evolutionary forces that shape the chemical diversity of antibiotics can provide valuable insight on where to look for chemical novelty and how to rationally engineer antibiotic-producing pathways. Bacteria, specifically *Streptomyces*, have been the primary source of antimicrobial compounds used clinically [66, 95]. The suites of genes required to assemble, transport, and regulate these molecules are typically genomically adjacent, either on the chromosome or on plasmids, forming functionally cohesive units called biosynthetic gene clusters (BGCs), often over 100 kb in length. Recent sequencing campaigns have revealed biosynthetic potential of bacteria to be much greater than initially estimated [23, 40]. Genetic signatures for the substrate specificity of certain biosynthetic enzymes can enable in silico predictions of metabolite structure and genome mining of BGCs has become a powerful discovery tool [82, 85, 88, 129]. BGCs are transmitted through both vertical and lateral gene transfer of selfish genetic elements [47, 83], such as genomic islands [104] and plasmids [47]. The rate and nature of evolutionary change appear to correlate with chemical class [84]. Furthermore, selective pressures within individual pathways can vary at different biosynthetic loci [142] suggesting selection towards a metabolite’s in situ function. Enzymes and functional domains found in BGCs are often products of concerted evolution with origins from within the BGC itself, other BGCs, or central metabolism [16, 27, 84]. The resulting functional redundancy allows for expansion of substrate specificities and/or metabolic functions while reaction mechanisms are maintained [27, 84]. BGCs can arise from concatenations of functionally independent subunits [84] and inter-module recombination has been observed in polyketide [61, 112] and nonribosomal peptide [26, 110] machineries that confer new chemistry through shifts of substrate specificities, stereochemistry, and/or processing. Further understanding of the underlying evolutionary mechanisms of biosynthesis would allow for the rational engineering of novel compounds.

The progressive and modular logic of polyketide synthase (PKS) and nonribosomal synthetase (NRPS) BGCs allows for genomic predictions of chemistry [9, 23]. Similarly, chemistry can be predicted from biosynthetic signatures of other BGCs, including terpenes, bacteriocins, aminoglycosides, saccharides, and ribosomally synthesized and post-translationally modified peptides (RiPPs), among others [9]. In large part, the identification of BGCs and predictions of their chemistry rely on their genes clustering into discrete units, yet the evolutionary rationale for why this might happen remains unclear. A strong hypothesis for the clustering of secondary metabolic genes into BGCs stems from their interdependence in producing a single phenotype that conveys a fitness benefit, in this case generating a metabolite that endows a benefit upon the producing species and/or any of its symbiotic partners. This cohesion and interconnectivity between gene function can cause clustering over evolutionary timeframes [111]. Adjacent clustering would allow for single lateral transfer events to be beneficial [111]. If a full BGC is necessary for a molecule’s function (e.g., assembly, regulation, resistance, transport, etc.), then transfer of one or a few of these functionally coupled genes would not otherwise be beneficial. Nonetheless, the evolutionary pressures shaping the fitness landscape of microbial secondary metabolism are both complex and difficult to assay, and as such most of the ecological constraints driving selection remain unknown.

The extreme repetition and high GC content within the modular biosynthetic logic of PKS and NRPS BGCs have evolutionary implications on the domain level changes that impact metabolite chemistry [23, 61, 84, 110, 114]. For example, in *Streptomyces* spp. that have the avermectin type 1 PKS (T1PKS) BGC, variation in cluster content suggests gains, losses and swaps of reductive domains within the PKS itself [61]. T1PKS modules can sometimes have over 85% local GC content and form direct repeats of many kilobases. This is thought to allow for recombination at homologous sequences at interdomain regions [61]. The sequence of adenylation (A) domains within NRPSs and acyltransferase (AT) domains within T1PKSs reflects phylogenetic signatures of their elongation substrate specificities [23, 45, 51, 67]. Though these domain-level substrate–sequence relationships are informative, small changes within A-domains have led to changes from one substrate specificity to another [97], changing the overall structure and potentially the activity of the produced molecule. Similarly, point mutations in dehydratase or ketoeductase T1PKS reductive domains can render them enzymatically inactive and change the resulting chemistry [82]. Other domains within modular biosynthesis show contrasting evolutionary patterns. T1PKS ketosynthase (KS) domains group phylogenetically based on their BGC family rather than by phylogeny of the producing organism [51, 61, 84, 87, 140]. This phenomenon suggests that within-cluster duplication of these domains or the modules containing them are important means of evolution in T1PKS biosynthesis. Further, there are recombination “hot-spots” flanking the acyltransferase domains of T1PKSs [112], which provide a means for both intra- and inter-BGC chemical diversification. In contrast to T1PKSs, the KS domains of trans-acting acyltransferase (transAT) PKSs group phylogenetically based on elongation substrate specificity, not...
based on pathway, suggesting lateral transfer may be a more prevalent means of chemical diversification in these systems [55]. The condensation (C) domains of NRPSs group based on their upstream and downstream A-domain substrate specificities [110], highlighting evolutionary constraints between domains. Within modular BGCs, the evolutionary relationships of domains and modules of their assembly line biosynthesis appear to be cluster dependent. Despite our knowledge of modular biosynthesis, engineering efforts towards designer molecules are still nascent, with relatively few successful examples of reprogramming and domain swaps that can produce desired molecules in sufficient quantities, if at all [18, 56, 135, 139]. This limited success is an opportunity for evolutionary genomics to expand our knowledge of modular biosynthesis, the rules of protein–protein interactions within these molecular assembly lines, and the “engineering” events that happen in the natural world.

Gains, losses, and variable genetic contexts of enzymes involved in precursor biosynthesis, tailoring of a core chemical scaffold, and regulation are also often observed. For example, in the rapamycin family of BGCs (including FK520, FK506, and meridamycin) the presence or absence of oxygen-installing cytochrome p450s impacts the metabolite’s final structure. Also within this BGC family, variable placement of the T1PKS, chorismatase, large ATP-binding LuxR (LAL) regulator, and lysine cyclodeaminase between clusters suggests gene shuffling over evolutionary timescales. More subtle changes can have large impacts on chemistry as well. Point mutations within the specificity-conferring active sites of certain enzymes (e.g., A-domains, AT-domains, glycosyltransferases, etc.) can shift or expand substrate specificity [88]. Even more extreme, point mutations within the precursor peptides of RiPPs can have large and adaptive effects on chemistry yet are not bound by the same limits of opportunity as they can arise spontaneously.

In a 2014 systematic assessment of known BGCs, Medema et al. [84] found exceptionally high rates of within-cluster evolutionary events, including indels, duplications, and rearrangements compared to primary metabolic operons. They also note a “high frequency” of putative lateral gene transfer (LGT) of BGCs, as noted by discrepancies between 16S rDNA similarity, primary metabolic pathway divergence, and domain overlap between BGCs [84]. It is important to note that the description of BGC LGT frequency as high is relative to other genomic transfers and does not represent a commonly occurring event. Detailed genomic investigations have revealed that the Streptomyces, a lineage that originated approximately 382 million years ago, has an observed LGT rate approximately 10 genes per million years [83]. However, many other genes undergo LGT and the rate described is an aggregate of all possibilities. While secondary metabolism is an overrepresented ontology group for genes that have undergone LGT [83], these events are (1) very rare [83], (2) typically happen at the gene or sub-cluster level [83, 84]. In the MIBiG repository v1.3 [86], there are 24,085 genes in 1393 BGCs for an average of 17.3 genes per cluster. Thus, the transfer of a full BGC is an unlikely event over millions of years. LGT itself is a function of phylogenetic divergence and the similarity of donor and recipient organisms; those that are more distantly related have dramatically reduced rates of transfer [83]. Rampant LGT of BGCs is a prevailing hypothesis in the field, but this view often overlooks the large evolutionary timescales in which an LGT event is observed. Vertical inheritance of BGCs represents an alternative mechanism. Many clusters appear to be retained within specific lineages over millions of years through vertical inheritance (Figs. 2, 3), suggesting that this is a major, albeit often underappreciated, means of BGC evolution within nature. Likely, the biological reality is a combination of the two, heavily influenced by how the
changes affect chemistry, changing environmental pressures, and opportunities for genetic exchange bounded by both phylogenetic relatedness and opportunity for gene flow.

To illustrate the evolutionary dynamics of BGCs, we constructed a core-genomic phylogeny of *Streptomyces* based on a concatenated 93-gene multilocus sequence of single copy genes (http://www.jcvi.org/cgi-bin/genome-properties/GenomePropDefinition.cgi?prop_acc=GenProp0799) with *Gordonia* sp. SID5947 as an outgroup (Fig. 2a). Prodigal v2.6.046 [60] was used to identify open reading frames

Fig. 2 a Genomic phylogeny of *Streptomyces* and outgroups. b Presence of the BGCs encoding the antibiotics kanamycin, streptomycin, valinomycin, and candicidin are indicated by filled boxes. c Number of full-length BGCs (BGCs on contigs that extend 500+ bp on either side of the BGC) and BGC fragments are shown to the right. BGCs were grouped into families that occur in either multiple (green) or single (purple) strains within the dataset
and HMMER v3.1b247 [43] was used to identify protein sequences of each GenProp0799 protein family. These were then aligned with MAFFT v7.24539 [65], converted to codon alignments, and concatenated. A phylogeny was generated with RAxML v8.1.2448 [125] under the GTRgamma substitution model (100 rapid bootstraps). DIAMOND [15] hits of protein identity > 75% to proteins within the kanamycin (MIBiG [86] accessions BGC0000702 to BGC0000706),

Fig. 3  a Genomic tree overlaid with an ancestral reconstruction of the streptomycin BGC. Pies indicate the probability that an ancestral node had the streptomycin BGC. Distinct genomic lineages are marked in orange and blue. b ASTRAL2 BGC tree built from genes in the streptomycin BGC only. Clades are colored based on B. c BGC overlap between strains is a function of time. Error bars, standard error of the mean; blue line, local regression; gray, 95% confidence. d Percent of the genome dedicated to secondary metabolism differs by genus. Center line, median; box, upper and lower quantiles; whiskers, ×1.5 interquartile range
variability in the amount of their genomes involved in BGCs, a portion (~15%) of their genomes to secondary metabolism in GenBank. Typically dedicate a large proportion (~15%) of their genomes to secondary metabolism, antiSMASH4 was used to identify strains with at least 50% BGC coverage, marked as the presence of a BGC family to the right of the tree. Kanamycin is widespread across the *Streptomyces* phylogeny, while streptomycin, valinomycin, and candidin are mostly confined to individual clades, suggesting vertical inheritance (Fig. 2b). BGCs from antiSMASH4 with BiG-SCAPE (https://git.wageningenur.nl/medema-group/big-scape/) distances of less than 0.5 were used to identify BGC families, many of which are found in more than one strain in this sampling (Fig. 2c).

A maximum likelihood ancestral character estimation [100] of the presence of the streptomycin gene cluster (as identified by methods above) is shown in Fig. 3a. *Streptomyces* split from *Kitasatospora* approximately 380 million years ago [83], which suggests that the likely origin of the streptomycin BGC within *Streptomyces* is approximately 40MYA ± 10MY. Outside of the major streptomycin lineage is SID8375, which likely represents a lateral transfer event of the BGC. Individual gene trees of streptomycin BGC proteins were created with Fasttree2 [107] and summarized into a BGC tree with ASTRAL2 [90]. The topology of the streptomycin BGC ASTRAL2 tree is concordant with the genomic phylogeny of these strains (Fig. 3b) further supporting that streptomycin has been inherited vertically. Placement of the SID8375 BGC between the major lineages (orange and blue in Fig. 3a, b) of streptomycin in our unrooted tree suggests that this transfer was prior to the divergence of the two clades.

The number of BGCs shared between *Streptomyces* strains is a function of divergence time (Fig. 3c). Branch lengths from the genomic phylogeny were scaled to millions of years (based on the *Streptomyces–Kitasatospora* split [83]) and BGC family overlap was calculated between all strains. As before, BGCs of less than 0.5 BiG-SCAPE distance were grouped into families. Strains with longer divergence times have less BGC overlap than those that have recently split, consistent with vertical inheritance playing a major role in BGC evolution.

In Actinobacteria, millions of base pairs and upwards of 15% of genomic content (Fig. 3d) can be dedicated to BGCs, most of which are silent when grown in isolation in the laboratory [1]. To assess the percentage of the genome that bacteria dedicate to secondary metabolism, antiSMASH4 was used to identify BGCs in all *Bacillus* (n = 557), *Burkholderia* (n = 354), *Mycobacterium* (n = 1837), *Pseudomonas* (n = 492), *Streptomyces* (n = 167), and *Vibrio* (n = 189) good quality genomes (eight or less contigs) publically available in GenBank. *Streptomyces* typically dedicate a large proportion (~15%) of their genomes to secondary metabolism (Fig. 3d). *Vibrio* and *Mycobacterium* have relatively low variability in the amount of their genomes involved in BGCs, while *Burkholderia* are the most variable of these genera. Together, these distributions suggest that the specific and situational context in which a BGC is produced influences organismal fitness in a given environment.

**Evolution of chemistry**

The evolutionary forces that shape and maintain BGCs can have meaningful consequences on the structure, and thus the activity, of their produced molecules. Convergence of antibiotic chemistry has been observed within the fungus-growing ant system. Dentigerumycin has activity against the fungal pathogen *Escovopsis* [99]. *Pseudonocardia* from Panamanian fungus-growing ants can produce the molecules dentigerumycin, gerumycin [120] (or both, or neither) and these differences can be seen over fine geographic scales, with strains from colonies showing variation in (dent)igerumycin biosynthetic potential over only a few kilometers [19]. Although their BGCs are distinct and unrelated (no BiG-SCAPE similarity), they produce structurally similar molecules with a Tanimoto similarity coefficient of 0.528 between gerumycin A and the core scaffold of dentigerumycin A (excluding the acyl side chain). This similarity suggests that the shared chemistry of gerumycin and dentigerumycin is important in mediating *Escovopsis* infection, yet important differences in their structures may contribute to differential activity that is more active in certain ecological and/or molecular contexts.

Minor changes in chemical structure can have large effects on the activity and function of antibiotics. Some BGCs can produce many molecules and this “combinatorial” chemistry can serve as diversity upon which evolution can act. For example, the cyanobactins that are produced by a single RiPP BGC show extreme variability in the precursor peptide, deploying many chemical variants that have different activities [39, 116]. In this example, there is high fidelity between the leader peptide and its protease and also low substrate specificity in the BGC’s tailoring enzymes (relative to other pathways) that together allow for variable precursor architectures [39, 116]. As leader peptide is discarded post-cleavage, fidelity to its protease ensures proper cleavage but has no influence on the final chemical structure. Further, this frees the core peptide to be hyper-variable (as it is not needed for cleavage recognition). This variability is accommodated by promiscuity in downstream tailoring enzymes. Similarly, a single T1PKS in *Streptomyces* sp. Mg1 produces over 50 structural variants of linearmycin [58]. In these examples, combinatorial biosynthesis of many variants may result in activity against a broader spectrum of target alleles and thus can act as a form of natural structure–activity relationship exploration to identify new variants to produce at higher frequencies. Balancing selection to
lower fidelity of biosynthetic machinery and to produce high numbers of products could be a response to recent antibiotic resistance against a specific molecule, where exploration of the chemical space may lead to recovered or even improved activity. This may also be a mechanism of niche partitioning whereby the producing organisms and/or systems can explore the ecological landscape.

**Antibiotics in nature vs. the clinic**

Broad-spectrum antibiotic use in the clinic encourages the development of resistance without the reciprocal selection for changes to the antibiotic molecule itself. In nature, the evolution of molecule–target interactions is dependent on both population and opportunity. For these dynamics to manifest in an evolutionary arms race, repeated interactions (here between molecule and target) are necessary at high enough frequency to drive selection. That is, selection of a pathogen to evade or circumvent an antibiotic (1) needs to have enough pressure to drive evolution in the first place, and (2) will be met with reciprocal selection on the genes assembling, transporting, or otherwise influencing the activity of the antibiotic itself. As these targets change, those microbes that produce more effective antibiotics will survive over those with less effective, representing the selection to remain effective. Here, the antibiotic activity is a selected phenotype and the genes that influence this are themselves under selection to maintain activity in the face of resistance, resulting in an evolutionary arms race. Sometimes referred to as the Red Queen hypothesis, this phenomenon is a type of antagonistic coevolution where both sides are evolving just to maintain the same equilibrium, much like the Red Queen in Lewis Carroll’s *Through the Looking Glass* who runs faster and faster just to stay in the same place. This dynamic contrasts with antimicrobial resistance in the clinic where a “pure molecule” treatment remains static and is itself not beheld to selection for activity. Further, if a host population dies from infection in nature, so too does the pathogen population, equilibrating to weaker selective forces compared to clinic (where there are increased rates of transmission and allowing people to die is not an ethical option). Few antibiotics and a limited range of mechanisms are in use in the clinic compared to the vast chemical diversity of nature, further encouraging resistance mechanisms to propagate, especially in hospital settings [13].

Soil is a tremendously complex natural microbial community and as such many soil-centered screens for antibiotics have failed to investigate the underlying interactions driving antibiotic chemistry. Microbiomes with less complexity, while more simplistic, can help dissect the mechanisms through which microbes use secondary metabolism to interact with their environments. For example, microbiomes of marine invertebrates have been a valuable discovery resource for diverse antibiotic chemistry [57, 89]. Unpacking interactions between isolates from the sea squirt microbiome resulted in the discovery of a new anthracycline antibiotic, keyicin [2]. Here, expression of the keyicin BGC by the producing Micromonospora is induced only in coculture with a sympatric *Rhodococcus* isolated from the same microbiome, indicating that complex species–species interactions can influence the “cryptic” nature of BGCs. There is no observed bactericidal or bacteriostatic effect against the *Rhodococcus* itself to keyicin, suggesting that these bacteria work together to sense and respond to Gram-positive invaders to their system, but not kill the native Gram-positive community. More examples of antibiotics crucial to the success of a multi-organismal systems are listed above (section titled *Antibiotics as antibiotics*).

**Targeting specificity and toxicity in the face of selection**

Many examples of chemically mediated defensive symbioses have been reviewed elsewhere [49, 91, 131]. We rather draw the reader’s attention to the high targeting specificity observed in natural systems and posit this is influenced by evolutionary or ecological constraints. Waksman noted that the most important biological properties of antibiotics are antibiotic spectrum and their selective toxicity [134]. Spectrum is closely tied to resistance and its prevalence in nature. The population of resistance genes within a gene pool can facilitate rapid emergence when targeting core processes that are under strict selection, leading to the evolution of both pathogens and commensals with resistance. One alternative to broad-spectrum treatments is to move toward targeted therapy for individual pathogens as resistance continues. More specific targeting of pathogens through species-specific, non-core processes likely has less standing resistance in the gene pool, which would lead to effective antibiotics and suggests a more sustainable way to think about antibiotic use against resistant pathogens.

Clinical antibiotic use draws many parallels to the herbicidal use of glyphosate. Marketed as RoundUp and many other tradenames, glyphosate interferes with the synthesis of the essential amino acids tryptophan, phenylalanine, and tyrosine by disrupting 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS), a conserved enzyme in the shikimate pathway. Crops engineered with resistant EPSPS allow for the broad-spectrum killing of weeds while the crop itself is resistant. Glyphosate-resistant crops have seen widespread adoption throughout agriculture and glyphosate has become the most heavily applied herbicide in the world [93]. Human artificial selection has driven the widespread emergence of resistance in many weeds, through de novo EPSPS point
mutations, extreme EPSPS copy-number variation, and other genetic changes [41, 93]. One strategy to address this issue has been to move towards more narrow-spectrum herbicides, either in combination with glyphosate or alone, to target the specific genotypes of individual weeds. Highly selective treatments that perturb a narrow-spectrum of organisms would likely have less probability for environmental effects compared to broad-spectrum treatments that directly impact many different organisms. Further, a targeted approach may slow the emergence of resistance while mitigating off-target effects.

Selection can also act to maintain and/or avoid destruction of the complex interaction networks of multi-organism systems. In the late 1950s, Krasilnikov observed that microbe-produced antibiotics in the rhizosphere show little if any toxicity to their associated plant and proposed that this low toxicity might be due to evolutionary constraint on the system [70]. Since, there have been a few other examples of low toxicity, notably within defensive symbioses. In the fungus-growing ant system, dentigerumycin produced by symbiotic *Pseudonocardia* has potent activity against the pathogen *Escovopsis*, but negligible toxicity against the ant host and the fungal cultivar [99]. In Southern pine beetle larval galleries, mycangimycin similarly selectively kills invading beetle pathogens and does not harm the beneficial fungal food source [117]. *Lactobacillus* and *Staphylococcus* from the human vaginal and nasal microbiomes, respectively, produce protective antibiotics that are not active against the healthy microbial community or the human host [38, 141]. Together, these examples argue that the ecological and evolutionary pressures on antibiotic-producing microbes, especially those in close associations with eukaryotic hosts, are an exciting source of new antibiotics, as selection acts to exhibit targeting specificity with limited host toxicity.

**Outlook**

Taking an evolutionary perspective on natural products offers a rational framework for revitalizing the antibiotic discovery pipeline. While negative findings in blind “grind and find” bioprospecting efforts lead to no actionable information, an ecological and evolutionary framework can offer insights to drive forward discovery and engineering even without the discovery of a new molecule. The evolutionary histories and trajectories of antibiotics, their BGCs, and their producing organisms can begin to inform: (1) where we might look for antibiotic-producing microbes, (2) how environmental cues influence the production of antibiotics, (3) where we might find increased efficacy towards certain targets, (4) how to limit off-target toxicity, and (5) how evolutionary genomic information can inform engineering and synthetic biology. More studies investigating the ecology and evolution of antibiotics are necessary to glean insights into the generation and maintenance of natural chemical diversity and biological activity in the face of antibiotic resistance.

**Acknowledgements** The authors would like to thank Heidi Horn, Reed Stubben Dieck, Lily Khadempour, Don Huang, Jennifer Bratburd, Alex Cheong, Daniel May, and Ian Miller for meaningful discussion and critical appraisal of the manuscript. This project was supported through National Institutes of Health (NIH) U19 AI109673. Additional support was provided to MGC through NIH National Research Service Award T32 GM008505.

**References**


