NATURAL PRODUCTS

Leveraging ecological theory to guide natural product discovery

Michael J. Smanski1,2 · Daniel C. Schlatter3 · Linda L. Kinkel2,3

Received: 17 June 2015 / Accepted: 29 August 2015 / Published online: 5 October 2015
© Society for Industrial Microbiology and Biotechnology 2015

Abstract Technological improvements have accelerated natural product (NP) discovery and engineering to the point that systematic genome mining for new molecules is on the horizon. NP biosynthetic potential is not equally distributed across organisms, environments, or microbial life histories, but instead is enriched in a number of prolific clades. Also, NPs are not equally abundant in nature; some are quite common and others markedly rare. Armed with this knowledge, random ‘fishing expeditions’ for new NPs are increasingly harder to justify. Understanding the ecological and evolutionary pressures that drive the non-uniform distribution of NP biosynthesis provides a rational framework for the targeted isolation of strains enriched in new NP potential. Additionally, ecological theory leads to testable hypotheses regarding the roles of NPs in shaping ecosystems. Here we review several recent strain prioritization practices and discuss the ecological and evolutionary underpinnings for each. Finally, we offer perspectives on leveraging microbial ecology and evolutionary biology for future NP discovery.

Keywords Biogeography · Coevolution · Arms race · Endosymbiont · Competition

Natural product (NP) discovery is poised to enter a second ‘golden age’ due to the abundance of recently identified biosynthetic potential present in sequenced genomes. Waning interest in NPs as drug leads near the turn of the century has been replaced with renewed optimism [33], and there are currently tens of thousands of uncharacterized gene cluster families suggesting that there are many unexplored scaffolds as well. Lastly, while the diversity of NPs is large, it is not without limit. Extrapolating data on BGC identification from a subset

Special Issue: Natural Product Discovery and Development in the Genomic Era. Dedicated to Professor Satoshi Ōmura for his numerous contributions to the field of natural products.

* Linda L. Kinkel
  kinkel@umn.edu

Michael J. Smanski
  smanski@umn.edu

Daniel C. Schlatter
  schl0453@umn.edu

1 Department of Biochemistry, Molecular Biology, and Biophysics, University of Minnesota-Twin Cities, Saint Paul, MN 55108, USA
2 BioTechnology Institute, University of Minnesota-Twin Cities, Saint Paul, MN 55108, USA
3 Department of Plant Pathology, University of Minnesota-Twin Cities, Saint Paul, MN 55108, USA
of the 830 available actinobacteria genomes, Doroghazi et al. estimate that 15,000 bacterial genomes would need to be sequenced to saturate the number of new BGC families [27]. While this extrapolation extends far beyond the empirical data and could be negatively affected by sampling biases in genomes that were analyzed, it is encouraging that these estimates are within reach for the coming decade given continued reduction in sequencing costs. However, filling in the gaps in our current knowledge base will take more than just random gathering of sequence information. Sequencing additional genomes from closely related isolates can provide insights to speciation and genome evolution, but may or may not increase the number of known BGC families [29]. The important question becomes, how can we prioritize strains for genome sequencing efforts to maximize the discovery of new NP biosynthetic potential?

The distribution of BGCs among microbial phylogenetic lineages, lifestyles, and habitats is non-random in nature and ecological theory can provide testable hypotheses to focus future genome sequencing and NP isolation programs [11]. Further, NP production plays a significant role in mediating ecological and coevolutionary interactions, and in population differentiation. Understanding the roles of NPs in the ecology and evolutionary biology of microbes in natural ecosystems will offer substantial insights for optimizing NP discovery efforts. In this mini-review, we highlight how information on microbial phylogeny, habitat characteristics, and species interactions can inform NP discovery. Further, we identify key knowledge gaps, and propose ecological and coevolutionary hypotheses that should guide future research, and increase the likelihood of novel NP discovery.

Microbial phylogeny and NP discovery

Tens of millions of microbial strains have been isolated worldwide and screened as part of drug-discovery campaigns, mostly by large pharmaceutical companies [4]. However, despite the geographic diversity of sampling locations, most standard isolation methods yield the same predominant taxa [34]. As a result, re-isolation of closely related genotypes has hampered the discovery rate of new NPs and made drug development inefficient. Early taxonomic characterization has been viewed as a reasonable means of de-replication, improving efficiency by focusing NP isolation and structure determination resources on strains that are likely to produce new molecules. In plants, correlations between taxonomy and NP production were described in the early 1800s [14]. However, the preponderance of horizontal gene transfer and the lack of a compelling species definition have blurred the relationship between taxonomy and NPs for microorganisms.

Historically, bacterial species were defined primarily by morphology and phenotypic characteristics. When using phenotype and morphology, members of the same species sometimes had distinct NP production profiles [105], while unique species could produce identical suites of NPs [51]. However, more recent genome-based analyses demonstrate that NP gene clusters can be highly conserved within some species [29, 42] and that there is a clear correlation between phylogeny and NP biosynthetic repertoires [27, 29, 42]. These correlations explain the successful efforts to find new NPs by looking to undersampled genera, for example the discovery of gentamicin by researchers at Schering Corp who turned to members of Micromonospora [113]. What remains uncertain are the phylogenetic scales (e.g., within or across species, genera, or phyla) at which there is significant conservation (or divergence) of NP production among microbial populations. Such information is critical to determining the value of sampling novel strains or additional strains of known taxa for yielding novel NPs. For example, although two actinomycetes with 99 % conservation in rpoB sequence are expected to share 80 % of NP gene clusters, this correlation drops off rapidly with increasing evolutionary distance [27]. This suggests that even though there is phylogenetic conservation of many NPs, perhaps representing a core biosynthetic capacity that reflects the life-history strategy of a specific phylogenetic lineage [24], there is still considerable intra-specific variation in 20 % of specific NP production capacity. The NPs encoded in this flexible genome are likely to be highly relevant to interactions with other species and especially responsive to selection pressures that generate unique biological diversity [20, 24]. This leads to a testable hypothesis that combining ecological data, including those pertaining to sympatrically co-evolved community members, biogeography, and selective pressures with molecular phylogeny data will improve our ability to identify closely related strains with unique NP potential. In other words, a more holistic approach that considers more than sequence identity of phylogenetic markers will be important for strain prioritization as we examine increasingly fine-scale operational taxonomic units.

From an evolutionary perspective, correlations between phylogeny and NP production in bacteria are related to the microbial speciation process. On one hand, widespread intraspecific recombination among populations may homogenize the gene pool of a species when barriers to gene flow are absent, resulting in a strong linkage between phylogeny and NP production across a species [28, 29]. However, when a strain variant exploits a new ecological niche, within-niche recombination is likely to be greater than between-niche recombination and distinct populations are likely to become genetically and ecologically coherent units [18, 87, 88]. This process may be especially crucial for generating intraspecific NP diversity among ecotypes.
of a species that has accessed a new niche by employing NPs \[13, 42, 84\]. NPs themselves may subsequently represent a barrier to recombination and promote the speciation process if they exhibit anti-microbial activity against sympatric populations or closely related populations found in jointly colonized niches \[21\]. Thus, although encompassing a broad phylogenetic scope in microbial strain collections for NP discovery programs is valuable, additional consideration of microbial populations that have undergone divergent selection, such as ecotypes from unique habitats, geographic locations, or biological niches, is also warranted.

Encouragingly, we have still only scratched the surface in terms of exploiting diverse organisms for NP discovery (Fig. 1). Most actinobacteria in strain collections are from the genus Streptomyces, as are a majority of described microbial NPs. Even within these well studied taxa, new compounds are still being discovered. Platensimycin and platencin are quite common, present in 0.3% of Streptomyces isolates, yet they evaded detection for decades \[34, 40, 106, 107\]. Rare actinobacteria, and other phyla are still underrepresented and are attractive targets for future NP discovery efforts \[97, 98\]. Further exploration of new branches in the tree of life, as well as more in-depth studies of the genomic content of microbial populations belonging to clades with especially high potential to produce NPs, are both likely bear fruit.

**Biogeography of microbial NPs: habitat variation and spatial distributions of NP diversity**

In efforts to enrich for under-represented taxa and capture strains with unique evolutionary and ecological histories, researchers have often focused on sampling from geographically isolated locations or environments with novel physical or chemical characteristics. Implicit in this approach is the assumption that distinct habitats and/or microbial community composition will impose divergent selection pressures that will yield distinct NP capacities. However, other than harboring unique or dissimilar microbial communities, are there distinct characteristics that make particular habitats attractive targets for NP discovery? Specific habitats that have received substantial focus in NP discovery include the following:

**Soil environments** Since the discovery of microbial NPs in the early 1900’s, soil microbes have received intense scrutiny as producers of useful NPs. The term ‘soil’ is used as a blanket term to describe what in reality are scores of spatially structured micro-habitats, each with unique chemical and physical properties. As such, soils harbor an enormous diversity of microorganisms \[99\]. In soil, NPs are believed to be particularly important to microbial fitness as mediators of resource competition and other species interactions \[39\]. In general, resource availability in soils is thought to be generally low and highly aggregated, though areas directly around plant roots (rhizosphere soil) are more resource-rich. In general, the low nutrient availability in soil may select for potent antibiotic producers able to defend their resource pool from competitors \[48, 115\]. Soil physical environments are also quite variable through time, with frequent drying/rewetting and temperature fluctuations, which may generate nutrient flushes. Strong competition for limited resources, diverse species interactions, and the highly variable range of conditions that microbes experience in soil likely play key roles in generating the high functional and structural diversity of NPs found in soil communities.

**Marine environments** In addition to being prolific producers of NPs \[25, 30, 66\], marine microorganisms belong to phylogenetically distinct clades from their terrestrial cousins and are specially adapted to marine environments. In these habitats there is expected to be high rates of diffusion, which may limit the benefit of producing extracellular
NPs. Thus, NP production in these habitats may be concentrated in endosymbionts, or microbes that inhabit spatially structured components of the marine environment (sediments or particles) where benefits of locally produced compounds may accrue to the producer [42, 55, 56]. Marine environments are typically more stable than most terrestrial environments in respect to physical perturbations like temperature or desiccation. Microbial communities are spatially stratified in both pelagic (open water) and benthic (seafloor) environments, suggesting the need for widespread and systematic sampling [121]. Despite oceans covering ~70% of the Earth’s surface, difficulty in sampling the ocean floor has limited sampling to relatively shallow depths that are accessible to divers. However, 95% of the ocean floor is located at depths greater than 1000 meters [90], and the deep ocean environments are arguably less well understood than the surface of the moon. New instruments that enable economical sampling of deep sea sediments are uncovering novel chemistry and these technologies continue to provide access to new and diverse ecosystems.

Many marine microbes require salt for growth, are physiologically distinct from their terrestrial cousins. The relative enrichment of halide elements in marine environments compared to terrestrial soils is reflected in their abundance in marine NPs [30]. The enrichment of halide elements can impact the chemical and biological properties of marine NPs. For example, the antineoplastic agent salinosporamide A (marizomib), contains a chlorinated ethyl group that is important for the mechanism of proteasome inhibition, increasing its potency by three orders of magnitude [35, 58]. With this combination of unique chemical environments, specially evolved microbial lineages, and dramatic undersampling, marine habitats promise to provide continued NP discovery opportunities for years.

Deep biosphere Caves, mines, and other deep underground environments contain unique edaphic qualities and are often characterized by low nutrient availability, limited nutrient diversity, stable physical environments, low microbial biomass, low microbial diversity, high humidity, and high metal concentrations, [5, 32]. They display increased stability to physical perturbations compared to terrestrial environments and less mixing than marine environments. Although few troglodytic NPs have been structurally elucidated, high frequencies of antagonistic phenotypes are seen in cave isolates [45], again perhaps reflecting the significance of competition for limited resources to microbial fitness. Similarly, the abundance of antibiotic resistance phenotypes in subterranean microbes suggests that antibiotic NPs are important in shaping the ecology of cave microorganisms [8]. Since some caves have been isolated from surface populations for millions of years [8] there may be higher potential for sustained coevolutionary interactions, including arms race dynamics, among microbial populations than in open surface environments. Arms race dynamics are hypothesized to be especially critical for generating biological novelty, including in NPs [96].

**Extreme environments** Extremophiles, or organisms that thrive in environments with extreme pH, temperature, pressure, salinity, or radiation, have had a large impact on industrial biotechnology as sources of enzymes with unique capabilities, and several groups have started looking to these microbes as sources of novel NPs. For example, a halophilic strain of the genus Actinopolyspora was isolated from a salt field in the Xinjiang Province of China and has yielded new analogs of erythromycin as well as new linear polyketides, actinopolysporins [120]. NPs isolated from mine waste pits with pH extremes include the novel anticancer spiroketal, berkelic acid, from an acidic (pH 2.7) abandoned copper mine [92], as well as a cytotoxic compound with a unique ring system, naphthospirinone A, from an alkaline (pH 10.0) tin mine [26]. Deep sea hydrothermal vents exist at temperature and pressure extremes and are home to many specially adapted animals and microorganisms. NP discovery efforts from vents have yielded novel cyclopesipeptides, clavatustides A and B [43], with anticancer activity, as well as the apoptosis-inducing chromene derivatives, ammonificin C and D [2]. Lastly, slow-burning underground coal fires are proving to be a fruitful habitat for isolating unique NP producing bacteria, including those producing new antifungal macrolides, venturicidin C and D [86], cyclopeptides, nulliamide A and B [110], and tetracyclic polyketide, ruthmycin [109].

Despite the demonstrated value of sampling unique environments to uncover rare taxa and novel NPs, many sampling efforts have largely lacked a strong theoretical rationale for targeting specific locations or habitats. Uncovering the central ecological principles that govern the distribution of BGCs across environments (biogeography), such as at the correlates of BGC composition and diversity within and among habitats, is needed for identifying specific locations or habitats most likely to be rich in novel NPs [11, 60]. As a first step, describing the variation in NP composition and diversity among similar habitats from different geographic locations offers insight into the spatial distribution of NPs and speaks to the value of sampling a wide geographic range [16, 112]. Further, focus on environmental characteristics (physical, chemical, or biotic) that are correlated with NP production potential among habitats, will provide criteria for targeted discovery efforts [67].

Geographic and environmental correlates of BGC composition and diversity may be used to develop predictive models of BGC composition for particular environments, as has been done for microbial community composition [50]. Fortunately, advances in microbial ecology enabled by powerful high-throughput sequencing technologies and
Improved analytical pipelines have vastly improved our understanding of microbial biogeography and paved the way for investigations seeking to describe biogeographic patterns of BGCs. Recent metagenomic work targeting PKS and NRPS domains found substantial differences in the identity and abundances (i.e., composition) of BGCs in soil types from various sites in the United States (Fig. 2); erythromycin-, bleomycin-, and nystatin-like biosynthetic gene clusters were enriched in soils from the arid Southwest, whereas enduracin-, oxazolamycin-, and arthrobactin-like biosynthetic gene clusters were enriched in soils collected from temperate soils in New England [16]. Moreover, correlations between soil characteristics (pH, K, and Ca) and estimated PKS and NRPS domain richness among these soils suggest that investigating soils with specific chemical properties may increase the likelihood of discovering new compounds, though targeting soil communities with more even distributions of BGCs rather than those with greater overall BGC richness will increase recovery of rare molecules.

Developing more explicit information on distance-decay (how NP composition within communities or populations varies with increasing spatial distance), and NP richness-area (how NP richness increases with broader geographic sampling) relationships for BGCs in different environments will provide useful tools for informing sampling strategies in searches for novel NPs. Importantly, relationships between the BGC composition in habitats, geographic distance, and environmental characteristics may depend on the spatial scale investigated [61, 67] (Fig. 3). In particular, the distance scale of sampling efforts needs to be determined by the hypothesis being addressed. For instance, at small spatial scales (meters to kilometers) physicochemical or abiotic habitat characteristics may be important correlates of BGC composition, yet across continental or global scales (kilometers to megameters) geographic distance among samples is likely to play a significant role in predicting diversity in BGC composition among samples. Such biogeographic patterns in BGC composition are likely to vary among habitats of interest, and ecological trends established for one habitat type may not apply to others. Because of this, the distance scale of sampling also needs to be tailored to the specific environment. For example, distance-decay and taxa-area relationships for bacteria in marine environments are likely to differ for communities from coastal sediments, surface waters, and...
deep sea waters, due to differences in spatial heterogeneity and/or spatial isolation among these habitats [121].

Though environmental characteristics likely play a significant role in the ecology of NPs, it seems improbable that specific environments impose direct selection for the production of a particular chemical family. While stochastic and biogeographically-structured dispersal dynamics will generate substantial variation in microbial community composition over space and time, the environment selects for which microbial species are present and provides a biological context for species interactions. In this way, variation in dispersal, habitat, and environmental characteristics across the landscape produces a ‘geographic mosaic’ within which species interactions occur [96]. This variation in community composition, and thus in the nature and dynamics of within-community species interactions, is a critical covariate to selection for novel antibiotics. This suggests that enhanced understanding of the spatial and temporal scales of accumulation of distinct community assemblages (community biogeography) rather than accumulation of individual taxa will provide important insight into guiding NP discovery (Fig. 3). More broadly, integrating biotic interactions into biogeographic studies of NPs will be crucial for identifying competitive and coevolutionary hotspots that are likely to be sources of new chemical diversity.

**Roles of NPs in biotic interactions**

Microbial NPs mediate diverse species interactions in natural environments and are important in microbial antagonism, cooperation, and exploitation. Best known is the role that antibiotic NPs play in microbial warfare, where they are well-documented to influence antagonistic competition in terrestrial soils [31, 48, 102, 116]. For example, *Streptomyces* populations from the same location are significantly more effective at inhibiting one another than *Streptomyces* from different soil locations, confirming the significance of NP-based antagonistic interactions to local fitness [48]. Beyond their role in mediating antagonistic competitive interactions, NPs, including subinhibitory concentrations of compounds traditionally recognized as antibiotics, can also act as microbial signaling compounds [102, 108, 118]. NPs employed as signaling compounds may mediate microbial coexistence, including microbial mutualistic and exploitative interactions. For example, shifts in nutrient utilization among microbial populations in response to subinhibitory concentrations of antibiotics can alter the likelihood of microbial coexistence in complex environments [102]. Similarly, up-regulation of NP production by one strain in response to subinhibitory NP concentrations in another [65, 103] provides opportunities for mutualistic associations via biochemical synergy among NPs produced by different strains [15, 29]. An alternative consequence of up-regulation of NP production by one strain in response to another strain may be exploitation of the NP producer by the signaling strain [6, 102]. Among *Streptomyces* from soil environments, shifts in NP production in response to coexisting strains are significantly more frequent among sympatric (coexisting) than among allopatric (from different locations) populations, suggesting that these interactions are significant to microbial fitness [103]. Overall, the roles of NPs in mediating antagonistic, mutualistic, or exploitative associations among microbial populations has significant implications for their distribution in the environment, and suggests testable hypotheses with significant potential to inform NP discovery efforts. For example, the prevalence of sympatric inhibitory and signaling interactions suggests that efforts to co-culture isolates to activate silent BGCs [7, 59, 100] should focus on sympatric isolates that share a co-evolutionary history. Moreover, in cases where cooperative interactions occur, molecules from sympatric isolates may be hypothesized to be more likely to have synergistic activities.

NPs are also hypothesized to play a significant role in mediating cross-kingdom interactions, including plant–microbe, insect-microbe, and animal-microbe interactions [85] (Fig. 4). Microbial symbionts of plants and animals have garnered much attention for their NP-production capabilities over the past few decades, and for good reason. Symbiotic bacteria from marine animals, plants, and insects have been a tremendous source of new chemical diversity (Fig. 4), and rational NP discovery efforts based targeting these niches have paid off. In many cases, microbial NPs are produced by endosymbiotic microbes that form intimate, highly coevolved associations with the host organism, though more diffuse (non-obligate, not physically intimate) coevolutionary associations between hosts and exosymbiotic or free-living NP-producers are also documented. Microbial NP production is perceived to be critical to the symbiotic relationship from the perspective of the host, while the host organism provides food and protection from abiotic and/or biotic environments to the microbial symbiont. Endosymbiotic or host-associated NP-producers are often suggested to mediate interactions of the plant, insect, or animal host with other organisms (especially pathogenic microbes), or with the physical environment, by contributing novel chemical phenotypes to the host organism [53, 82, 111]. However, though a handful of well-studied protective mutualisms have been documented among hosts and NP-producing microbes, the specific functional effects of the vast majority of NP-producing endosymbionts or host-associated microbes on host fitness remain poorly understood. Similarly, the specific benefits of the host association to the microbes are often unclear. Further consideration of the correlates of enhanced NP production
in populations, the distinct life history strategies of NP-producing microbes, and the effects of NPs on microbial or host or symbiont fitness, are critical for informing the design of efficient NP discovery strategies. Systems of particular interest to understanding the dynamics of NP-producing phenotypes are summarized below.

**NPs from endosymbionts** More bioactive NPs have been isolated from sponges than any other group of marine organisms [9]. Marine sponges are densely colonized by bacteria, with up to a third of their mass coming from bacterial symbionts [38]. Marine sponges contain highly stable bacterial communities that are markedly different from the surrounding communities in seawater [95]. Microbial communities of sponges are very diverse, with up to thousands of bacterial species suggested to occur in an individual sponge. However, NP production within sponge microbiomes seems to come from one or a few of these members. For example, each of the bioactive metabolites from the sponge *Theonella swinhoei*, including diverse polyketides, non-ribosomal peptides, and ribosomally-synthesized and post-translationally modified peptides, are all produced by *Entotheonella*, a single NP-producing taxon [117]. This sponge endosymbiont has not yet been cultivated and accessing its NP biosynthetic potential relies upon production in heterologous hosts. In other sponges, culturable marine actinomycetes are responsible for NP biosynthesis. Several hundred NPs have been isolated from marine sponge-associated actinomycetes, according to a recent review [1]. While NPs from sponges have been found to have diverse activities, including antibacterial, antifungal, anti-inflammatory, anti-oxidant, and anticancer, the importance of these NPs in the biology of the sponge host remains largely speculative. NP-producing bacterial endosymbionts have been hypothesized to act as protective mutualists, to enhance nutrient acquisition, and to process metabolic waste [1], but the role of specific NPs in these functions remains unexplored.

Endosymbiotic NP production is common in terrestrial environments as well. While the 1993 report that a fungal endosymbiont of the Pacific yew tree (*Taxus brevifolia*) can produce paclitaxel in axenic cultures has been contested, it jump-started the search for medicinal NPs in plant endosymbionts [37, 91]. Since then, microbial endophytes have been a recognized source for NP discovery [54, 80, 93, 104]. All plants harbor diverse endophytic microbiomes, yet the majority of NP research effort on endophytes in recent years has focused predominantly on plants used in traditional medicines [75]. For example, the antitubercular pluramycin analogs were recently isolated from *Streptomyces* purified from within the stalks of the traditional Chinese medicine plant *Heracluem souliei* [54]. However, although metagenomic analyses continue to expand our understanding of the incredible diversity of endophytic populations in both medicinal and non-medicinal plants, the roles of microbial NPs in the plant–microbe symbiosis remain poorly understood. Endophytic microbes have been reported to produce diverse NPs including polyketides, non-ribosomal peptides, alkaloids, terpenoids, and other compounds [68]. These compounds are hypothesized to influence plant fitness via diverse mechanisms, including in
plant growth promotion, pathogen or herbivore protection, or protection for physical or osmotic stress, but in most cases little direct evidence for plant fitness benefits exists. In return, the host plant provides both physical protection (or entrapment?) and food to the microbial symbiont.

The current focus on endosymbiotic populations as a source for NPs raises the question as to whether endosymbiotic populations are uniquely enriched in NP production capacities. That is, considering closely related microbial lineages, are endosymbiotic lineages more likely to produce diverse or novel NPs than free-living populations? One hypothesis is that endosymbionts may be especially enriched in NP production due to active selection by hosts for microbial populations with host-beneficial traits, such as the production of a NP. Endosymbiotic communities differ among species of plants and marine sponges [81, 111], suggesting that each species can select for specific microbial inhabitants. If hosts control endophytic colonization, fitness benefits conferred to the host as a result of microbial NP production may, through long periods of host-symbiont coevolution, lead to enriched plant colonization by NP-producing symbionts. Although both vertical (from parents to offspring) and horizontal (inoculation from environmental populations of bacteria) transmission are possible in sponges and plants [89, 101], vertically-transmitted microbes, which are likely to be especially co-adapted to their hosts, may be most fruitful as targets for novel NPs.

An alternative possibility is that endosymbiotic microbes are enhanced in NP capacities relative to free-living populations not because of the benefits to the host, but because the endosymbiotic lifestyle amplifies the potential benefits of NP production to the microbe. For example, given the physical constraints of plant colonization, and the limited capacities to disperse or escape plant tissues, NP production may be an important means for mediating growth and survival. Furthermore, because cell walls, vascular membranes, and plant structures are likely to significantly limit diffusion of NPs, NPs may be more likely to be present at biologically active concentrations within plants than in an open soil environment. This could magnify the benefits of NP biosynthetic capacities to endophytic microbial populations. Finally, NPs may be especially important in endosymbiotic lifestyles because of their capacities to alter their niche (e.g., plant host or sponge) in ways that benefit the microbial NP producer. For example, although a free-living NP-producing microbe in soil may be able to access a greater fraction of a limited nutrient than a non-producing microbe, an endosymbiotic NP-producer might create a larger, healthier, and more vigorous host, and thus a larger and potentially more-long-lasting pool of resources. A larger, stable host food source should result in greater microbial fitness than increased access to a modest or finite resource pool among free-living NP-producers. Future efforts aimed at estimating the rates of evolution for NP biosynthetic capabilities in endosymbionts and their free-living cousins might shed some light on these questions.

NPs from exosymbionts Several species of insects cultivate fungal gardens as a food source. Similar to agricultural systems of humans, these fungal gardens are often susceptible to invasion or parasitism by other fungi, significantly reducing the capacity of the fungal garden to support the insect population. One group of fungus-farming leaf-cutter ants has evolved symbioses with actinomycete bacteria, which produce the cyclic depsipeptide, dentigerumycin. This NP has selective capacity to inhibit the pathogenic Escovopsis fungus, but not the fungal crop [72]. Morphological adaptations allow the ants to house the bacterial exosymbionts in specialized crypts on the surface of their bodies, providing both food and protection to the microbial symbiont [53]. There is evidence for host specificity, with different genera of ants housing Pseudonocardia strains with unique morphotypes [12, 78]. Other fungus-farming insects follow a similar paradigm, but with distinct symbionts and NPs. For example, the Southern pine beetle harbors a Streptomyces symbiont for production of mycangimycin, a polyene peroxide, with selective activity against a pathogenic fungus [73]. Non-farming insects, including wasps, likewise harbor symbiotic actinomyces that produce diverse NPs, possibly to protect their eggs from pathogenic microbes [44, 79]. These systems suggest that for exosymbionts, NP capacities are likely to be associated with specific vulnerabilities in the host population life cycle, including food production or acquisition, and offspring survival. Identifying these vulnerabilities, and specifically when or where (perhaps in physically stressful habitats or in the presence of competitors) host populations are especially vulnerable directly or indirectly to microbial pathogens may be especially fruitful in defining targets for NP discovery. Further studies should also explore the specific means by which hosts select for and support NP-producing exosymbionts; such information may provide insights into strategies for manipulating selection to maximize the likelihood for capturing novel NP-production capacities.

NPs among free-living microbial populations: diffuse symbioses? Among free-living and rhizosphere soil populations, the frequencies of NP production and the diversity of NP phenotypes vary as a function of both abiotic and biotic habitat characteristics, including soil type, plant host identity, and plant community characteristics [3, 16, 83]. Among the most well-studied systems associated with enhanced NP production are naturally-occurring plant disease-suppressive soils. Specifically, in multiple cropping systems long-term monoculture of crop plants has led to soils where free-living microbes prevent the development of soil-borne plant diseases. While all soils support
indigenous microbial populations with potential to suppress plant pathogens, disease-suppressive soils are uniquely enriched with high frequencies and densities of NP-producing, pathogen-inhibitory microbial populations [3, 41, 46, 62, 63, 114]. Both Streptomyces and Pseudomonas species are often associated with naturally occurring plant disease-suppressive soils [46, 63]. However, most of the effort to discover the NPs responsible for plant protection have focused on Pseudomonas spp., and several bioactive metabolites have been found [36, 52]. For example, non-ribosomal peptides nunamycin and nunapeptin were isolated from Pseudomonas fluorescens in a disease-suppressive potato field in Greenland [64]. The cyclic lipodepsipeptide vicosinamide from Pseudomonas fluorescens DR54 is a potent antifungal agent and biosurfactant [71]. Similarly, the 16-membered polyketide macrolide rhizoxin is produced by numerous plant-associated pseudomonads [10]. This antimitotic compound contributes to antifungal biocontrol by pseudomonads, but has also been exploited by the fungal plant pathogen Rhizopus chinensis to cause disease, including rice seedling blight [94]. Numerous other small aromatic and halogenated compounds have been described from disease-suppressing Pseudomonas, perhaps most notably phenazine and 2,4-diacetylphloroglucinol, which are significant in the suppression of take-all of wheat.

The dependence for both Streptomyces and Pseudomonas-based suppressive soils on long-term monoculture has led to a model describing the development of disease-suppressive soils, or soils with high NP activities and/or diversities [46]. Specifically, the association of long-term monoculture with NP-producing populations suggests that stable and simple (monoculture) plant communities that provide consistent but low-diversity resource inputs to soil maximize selection for inhibitory phenotypes among microbial populations (Fig. 5). In such a soil environment, it is hypothesized that NP are significant in protecting access to a limited diversity of nutrient sources (antagonistic resource competition), and that an ongoing coevolutionary arms race among microbial populations is likely to generate high NP frequencies and diversity. In contrast, high-diversity plant communities are expected to contribute a broad array of resources to the soil habitat, so that microbes are able to co-evolve to diversify in their nutrient preferences, resulting in niche differentiation as a stable coexistence strategy (Fig. 5). In niche differentiated communities it is hypothesized that because there is less competition for resources, antibiotic NPs are less beneficial to producers and are selected against. Experimental field studies support these predictions, showing that long-term prairie monocultures produce more antagonistic soil communities than polyculture plant communities [3]. In total, these data suggest that plant community diversity, by influencing the diversity of soil nutrients, plays a critical role in determining the coevolutionary trajectories among NP-producing microbes in soil, and thus in the NP phenotypes of soil microbes (Fig. 5).
Distinct from plant community effects on NP phenotypes, individual plant hosts can also vary in the rhizosphere densities of inhibitory microbial populations [3], suggesting that plant hosts have capacity to influence selection for NP phenotypes in soil populations. Because high densities of inhibitory populations can reduce the incidence and severity of diverse plant diseases, this plant or plant community-based selection can feedback to influence host plant fitness in both natural and agricultural systems [3]. It is unclear whether the differences in antagonistic populations associated with plant species is a function of direct plant selection for microbial inhibitory phenotypes, or reflects the impact of plant host on soil resources that mediate microbial competitive outcomes (Fig. 5).

In total, these studies suggest that NP discovery among free-living micro-organisms should focus on highly competitive communities with limited resource diversity in order to maximize the likelihood of detection of novel NPs. This framework identifies long-term monoculture plant ecosystems as particularly likely to impose selection for highly inhibitory microbial populations in soil. For example, silver beech forests in New Zealand, long-established dense perennial grass stands, or long-term experimental monocultures may be especially suitable sites for NP exploration. Focusing on habitats that are most likely to support arms race microbial coevolutionary dynamics may be important not only for capturing high densities of NP-producing microbes, but also for generating high diversities of NPs [22].

**Conclusion**

Ongoing exploration of rare taxa and unusual habitats, coupled with creative culturing techniques, has continued to yield new and useful NPs. However, given the immense global environmental and biological diversity, it is difficult to justify random ‘fishing expeditions’ as a way to explore NP diversity. Applying ecological concepts and theory to NP discovery has the potential to inform our understanding of NP production across taxa, environments, symbiotic organisms, and geographic scales, and can be especially powerful in identifying optimal settings for discovery of new NPs.

A strain’s NP biosynthetic potential can be influenced by its life history as well as the long-term selective pressures placed on it by the surrounding abiotic and biotic environments. The potent biological activities that make NPs attractive lead compounds for drug discovery [69] likely makes them strong drivers of the species–species interactions. The abiotic and biotic environment generates selective pressures that drive NP evolution. For example, it is the combination of leaf-cutter ant, fungal food crop, and fungal parasite that creates a niche for a symbiotic organism that produces a NP to kills only the parasite, but not the fungal food crop [72]. In this way, the ecology drives chemical innovation by selecting for a unique biological activity. Also, the ecological environment can define and constrain the genetic building blocks available to create new biosynthetic gene clusters through horizontal gene transfer and recombination, which is the evolutionary model for some, but not all classes of NPs [70].

We propose that incorporating different types of data to strain prioritization can focus genome sequencing efforts to uncover new BGCs that yield novel structures. The readily-analyzed phylogenetic markers may not provide fine-enough resolution to predict the full complement of BGCs, which varies much faster than housekeeping genes [84]. NP diversity is known to vary over geographic spatial scales, habitat physicochemical characteristics, and biotic community composition [16, 17, 23, 48, 67]. While these correlations are poorly understood, they lead to readily testable hypotheses that can lead to more systematic characterization of NP potential in the future. For example, we hypothesize that niche overlap within a sympatric population of soil microbes should be positively correlated with the intensity of antibiotic inhibition. If this is true, then
future efforts could focus on microenvironments known to give rise to high niche overlap. Some of the salient biotic and abiotic metadata that are important to report with each new strain isolate include microbial community characteristics (density, diversity, composition); lifestyle (endosymbiotic, exosymbiotic, free-living); source habitat characteristics (nutrient availability or content and physicochemical characteristics, e.g., carbon, nitrogen, pH); ecosystem (e.g., forest, grassland, temperate, tropical); and global position (GPS coordinates). Individual isolate characteristics of interest include nutrient use preferences, temperature optima for growth, growth rates, antibiotic resistance profile, and signal production and response capacities. Systematic characterization of habitat and community characteristics as well as isolate phenotypic characteristics associated with novel or diverse BGCs will provide important insights for guiding future BGC discovery efforts.

Natural product discovery stands at an important crossroads. New technologies for rapid and cost effective DNA sequencing and synthesis are poised to contribute to a second golden age in NP discovery. However, NP chemical diversity is not without limit, and current sequencing efforts allow for global estimates on the limits of NP diversity [16, 19, 27]. In light of the emerging problem of antibiotic resistant infectious disease, it is prudent to view NPs as a valuable natural resource that need to be systematically characterized and carefully managed for long-term societal benefits. If the technological advancements in NP identification and characterization are mirrored by a greater understanding of the technological advancements in NP discovery. However, NP chemical diversity is not without limit, and current sequencing efforts allow for global estimates on the limits of NP diversity [16, 19, 27]. In light of the emerging problem of antibiotic resistant infectious disease, it is prudent to view NPs as a valuable natural resource that need to be systematically characterized and carefully managed for long-term societal benefits. If the technological advancements in NP identification and characterization are mirrored by a greater understanding of the ecological and evolution of NP biosynthesis, it will be possible to sustain discovery efforts into the future by focusing resources on the specific physical and biological habitats that are enriched in new NP production capabilities.

Acknowledgements This work has been supported by the National Institute of Food and Agriculture, U.S. Department of Agriculture, under agreement No. 2011-67019-30530 and the University of Minnesota Agricultural Experiment Station Project (#MIN 22-018). Resources from the University of Minnesota Supercomputing Institute are gratefully acknowledged.

Reference

populations of bacteria act as socially cohesive units of antibiotic production and resistance. Science 337:1228–1231. doi:10.1126/science.1219385


89. Sipkema D, de Caralt S, Ja Morillo, Al-Soud WA, Sørensen SJ, Smidt H, Uriz MJ (2015) Similar sponge-associated bacteria...