Science & Society

EMBO reports

The impact and prospect of natural product discovery in agriculture

New technologies to explore the diversity of secondary metabolites in plants and microorganisms for applications in agriculture

Yan Yan[1], Qikun Liu[2], Steven E Jacobsen[2] & Yi Tang[1,3]

Natural products (NPs) are small organic molecules of < 1,500 Da that are produced mostly by microorganisms and plants, and many of them have enormously benefited human health and agriculture. As NPs have co-evolved together with their molecular targets, they possess potent bioactivities and selectivities, which their producers use to kill or limit the growth of competitors by inhibiting vital metabolic enzymes. These chemicals are structurally complex and can include multiple stereocenters and intricate ring systems along with a plethora of functional groups that serve as warheads, hydrogen bond donors/acceptors, or target recognition moieties.

> *"Humans have been using the potent action of NPs for multiple purposes from medicinal to cosmetic and recreational use as well as in agriculture."*

Humans have been using the potent action of NPs for multiple purposes from medicinal to cosmetic and recreational use as well as in agriculture. During the golden age of NP discovery from the 1960s to the 1980s, scientists in academia and industry identified and characterized an impressive list of NPs that are still being used today: The antibiotic compounds penicillin or amphotericin, the cholesterol-lowering lovastatin, or the cancer drug taxol are just a few examples of how these microbial ecological weapons were repurposed for combatting diseases.

Global demand for new NPs

In agriculture, NPs have been applied as fungicides, insecticides, and herbicides that have contributed substantially to the increases of crop yield and quality worldwide. From 1997 to 2010, NPs and their derivatives made up about 36% of all new registered pesticide ingredients [1]. For example, spinosyn and avermectin, produced by soil-borne bacteria *Saccharopolyspora spinose* and *Streptomyces avermitilis*, can effectively paralyze insects through hyperexcitation of their nervous system (Fig 1). The discovery of avermectin by Satoshi Ōmura was awarded the 2015 Nobel Prize in Physiology or Medicine. Phosphinothricin, also known as glufosinate, produced by *Streptomyces*, has been commercialized by Bayer as an herbicide under the tradename of Finale® (Fig 1). By inhibiting glutamine synthetase, glufosinate kills plants via ammonia buildup in the thylakoid lumen, which leads to decoupling of photophosphorylation. Fenpicoxamid is a commercialized fungicide derived from the NP antimycin that inhibits cellular respiration (Fig 1). The sales of both glufosinate and fenpicoxamid exceed US$1 billion annually. There are many other NPs with unique modes of actions that have not been commercialized owing to the high cost of mass production. For example, potential herbicides thaxtomin and tentoxin are able to disrupt cellulose biosynthesis and energy transfer, respectively; cornexistin possesses broad-spectrum herbicidal activity via inactivation of aminotransferases but only low activity against maize (Fig 1).

> *"... there is a continuous demand to discover new insecticides, fungicides and herbicides with novel modes of action, accompanied by efforts to decrease their production cost."*

Nonetheless, new products are constantly needed: It is estimated that up to 50% of global crop yields are lost each year mainly due to pesticide resistance [2]. Hence, there is a continuous demand to discover new insecticides, fungicides and herbicides with novel modes of action, accompanied by efforts to decrease their production cost. Not surprisingly, NPs have remained important sources for such discovery efforts. Here, we describe how deeper understanding of NPs and their biosynthesis may lead us to new products for agricultural use.

Genome-driven discovery of NPs

NP discovery was traditionally performed by isolating organic molecules from an

1 Department of Chemical and Biomolecular Engineering, University of California Los Angeles, Los Angeles, CA, USA. E-mail:yitang@ucla.edu
2 Department of Molecular, Cell, and Developmental Biology, Howard Hughes Medical Institute, University of California Los Angeles, Los Angeles, CA, USA. E-mail: jacobsen@ucla.edu
3 Department of Chemistry and Biochemistry, University of California Los Angeles, Los Angeles, CA, USA
DOI 10.15252/embr.201846824 | EMBO Reports (2018) 19: e46824 | Published online 25 October 2018

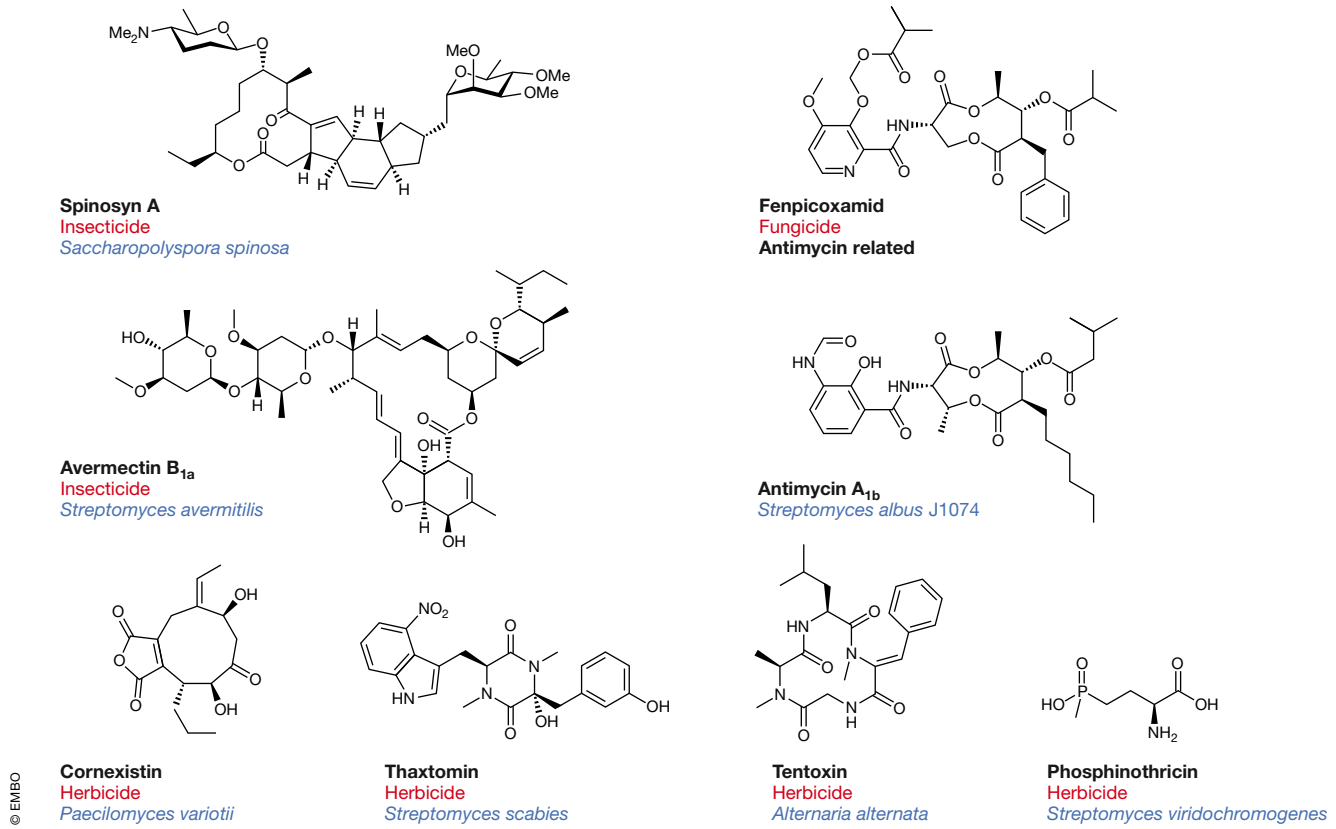

Figure 1. Microbial NPs that are used as pesticides.
The bioactivities (red) and sources (blue) of NP or NP-derived molecules.

organism of interests. The workflow involves the collection or growth of the organism, followed by extraction of organic molecules and fractionation of the extract. The isolation of a pure NP from complex extracts is typically guided by screens, either via direct biological assays of the target enzyme or through identification of novel structural features. These techniques have proven to be hugely successful during the golden age of antibiotic discovery. In recent years, however, reports of new structures and activities have slowed significantly, leading many pharmaceutical companies to abandon NP programs.

The advancement of new techniques in genomics has brought a renaissance to NP discovery. Thanks to the rapid development of DNA sequencing technologies, an increasing number of whole genome sequences are now available for research. Extensive studies into the biosynthesis of NPs and the genes encoding the enzymes involved have shown that the genes for one NP are typically clustered, which presumably facilitates co-regulation during transcription, and horizontal cluster transfer between species. A biosynthetic gene cluster (BGC) can be readily identified using powerful software packages through an anchoring biosynthetic enzyme that produces the core of a NP. Such anchoring enzymes include polyketide synthases, nonribosomal peptide synthetases, or terpene synthases.

> *"It is therefore tantalizing to speculate how many new NPs could be discovered if we can efficiently tap into these silent BGCs."*

The number of BGCs in a microorganisms identified *in silico* is therefore a reasonable estimation of the total number of NPs an organism can potentially produce. Given that only a small fraction of BGCs are associated with known compounds, the true biosynthetic potential of microbes is much larger than the number of known NPs. Indeed, most BGCs remain silent owing to their complex regulation and our inability to reproduce the natural environmental cues that are needed to turn them on: It is estimated that more than 90% of BGCs remain as genomic "dark matter" encoding secondary metabolites that have eluded traditional NP discovery.

It is therefore tantalizing to speculate how many new NPs could be discovered if we can efficiently tap into these silent BGCs. Different approaches have been applied to awaken these gene clusters, including constitutively expressing pathway-specific transcription factors, epigenetic modifications to alter chromatin structure and transcriptional activities, and heterologous expression of desired pathways in model hosts [3]. While these approaches are successful in inducing BGCs to produce new NPs, their true biological activities are typically unknown: Compared to more traditional NP discovery, the genomic approaches are not activity-guided. Given

the large number of BGCs available, it is essential to prioritize genome-driven discovery of NPs by biological activity. How can we predict the activity of a NP based on genomic sequence? The answer to this question can unlock the true untapped potential of the tens of thousands of BGCs.

Resistance-directed approaches to discover new bioactive NPs

To find agriculturally useful NPs with new modes of action from tens of thousands predicted BGCs, we developed a resistance gene-guided approach [4]. The rationale is that host organisms producing NPs that target housekeeping enzymes must have a method of protecting themselves. Several mechanisms of self-resistance are known: efflux pumps that actively transport the metabolite to the extracellular space; proteins that stoichiometrically bind to the NPs; and enzymes that modify the housekeeping target to evade NPs. Nature also evolved the clever strategy to encode a mutated copy of the sensitive housekeeping gene in the NP BGC (Fig 2). This self-resistance enzyme can carry out the same function as the housekeeping enzyme, but is sufficiently mutated to be insensitive to the NP. Because the self-resistance gene is required for survival during NP production, it is frequently co-localized in the same BGC. An example is the lovastatin BGC: A second copy of 3-hydroxy-3-methylglutaryl-coenzyme A reductase (HMGR), which is the target of lovastatin, is encoded in the lovastatin BGC in *Aspergillus terreus* (Fig 2). This co-localization has been exploited to link BGCs to compounds with known targets [5,6].

We propose that using the self-resistance gene as a predictive marker, one can mine NPs from collections of BGCs with desired bioactivity. A workflow for such guided genome mining looks as follows: After identifying a desired (plant) target enzyme that is also present in microorganisms, one can search through genome databases for BGC carrying duplicate copies of the target gene that is located close to a biosynthetic anchoring enzyme; different synthetic biology approaches can be applied to produce the NP encoded in the cluster; the NP is isolated and the structure is elucidated using NMR spectroscopy; and inhibition of the housekeeping enzyme and insensitivity toward the self-resistance enzyme are validated biochemically or genetically.

We applied this approach to search for herbicide leads with novel modes of action to target dihydroxy acid dehydratase (DHAD) within the branched-chain amino-acid biosynthetic pathway (Fig 3) [4]. We first scanned fungal genomes in publicly available databases for a BGC that encodes a possible resistant copy of DHAD. We eventually found a conserved four-gene cluster in *Aspergillus terreus*, which encodes a terpene synthase, two cytochrome P450s and a duplicate copy of DHAD that is about 60% identical to the well-conserved housekeeping DHAD. The

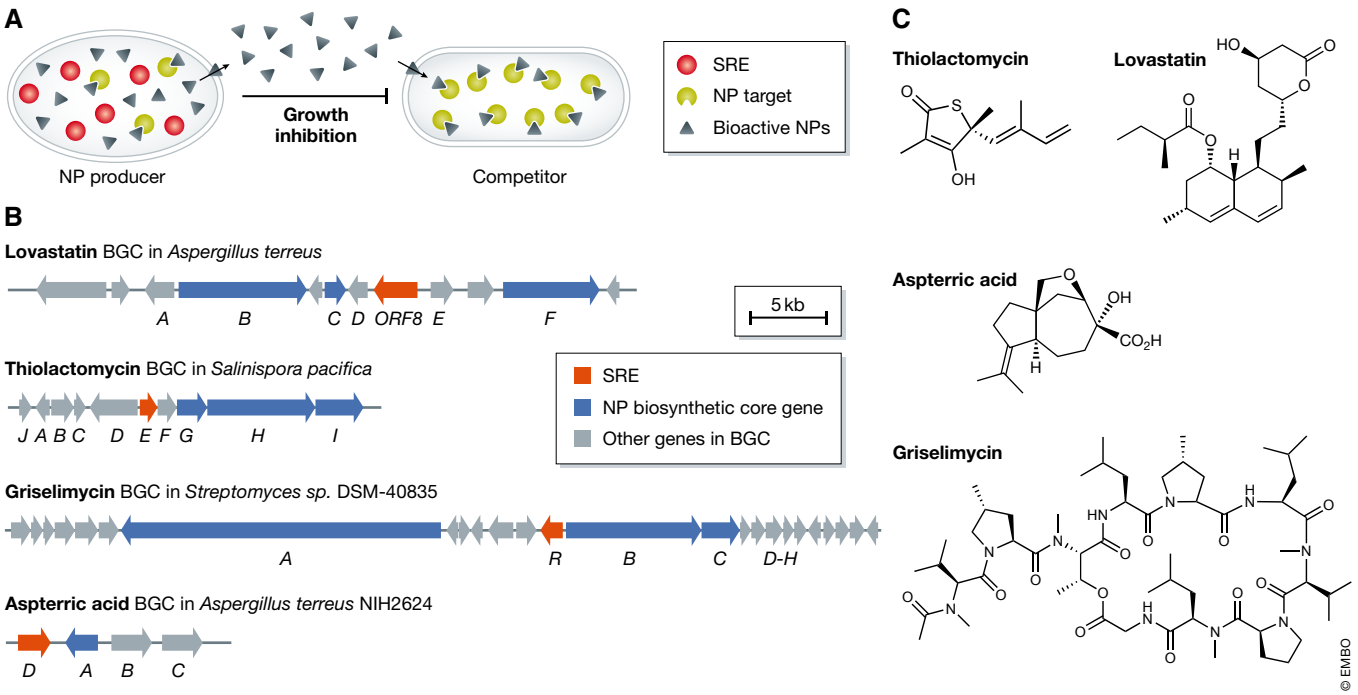

Figure 2. Examples of co-localization of biosynthetic gene clusters (BGCs) and target enzymes.
(A) Microorganisms produce NPs (black triangles) to kill competitors by inhibiting essential metabolic enzymes (yellow spheres). The organism produces self-resistance enzymes (SRE, red spheres) that are able to complement the function of the targeted metabolic enzymes for survival. (B) Representative BGCs show co-localization of biosynthetic core enzymes for lovastatin (blue) and self-resistance genes of HMG-CoA reductase in eukaryotes (red). The fungus *A. terreus* that produces lovastatin has a second copy of HMGR encoded by *ORF8* in the gene cluster; BGC of the fatty acid synthase inhibitor thiolactomycin from *Salinispora pacifica* was located using a second copy of fatty acid synthase, which was further proved to be a SRE. The antituberculosis agent griselimycin was found to target DNA polymerase, hinted by the presence of a self-resistance gene that encodes DNA polymerase in its BGC. A natural herbicide aspterric acid was discovered using resistance gene-directed approach. (C) The structure of NPs relevant to (B).

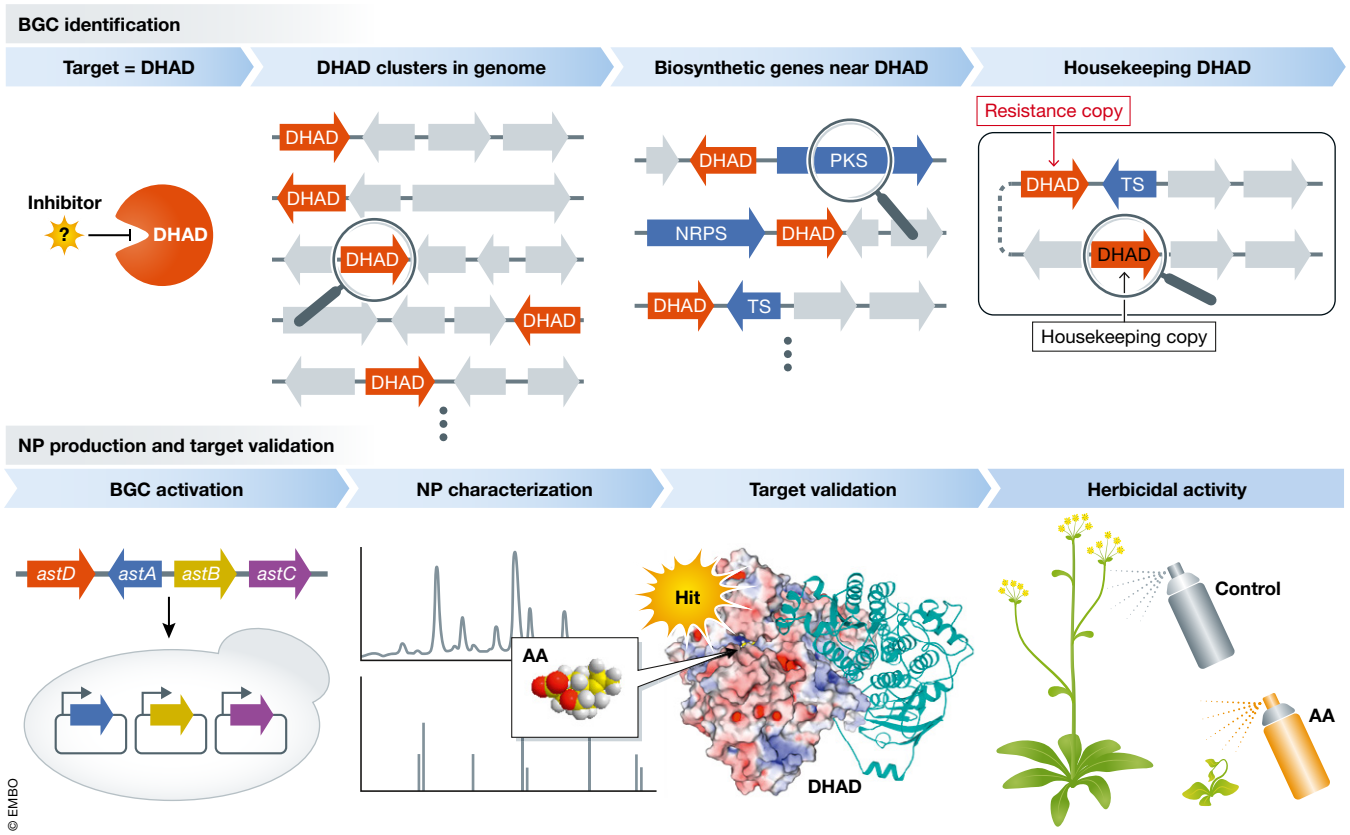

Figure 3. Workflow of resistance gene-guided discovery of NP herbicide aspterric acid (AA).

cluster was introduced into *Saccharomyces cerevisiae*, which produced aspterric acid at 20 mg/l. Consistent with our hypothesis, aspterric acid was verified as a potent competitive inhibitor of the housekeeping DHAD enzyme from *A. terreus* and *Arabidopsis thaliana*. In contrast, the *Aspergillus* self-resistance DHAD (AstD) was insensitive to aspterric acid.

> *"Microbial fermentation using genetically modified organisms has therefore great potential to produce a given compound at high titer."*

When applied *in planta*, aspterric acid showed strong growth inhibition of representative monocots and dicots. When applied at lower concentration, aspterric acid could specifically inhibit the formation of pollen without harming pistil development. Hence, aspterric acid may be used as a chemical hybridization agent in the field to facilitate out cross and hybrid seed production. Motivated by the success of using a combination of glyphosate-based herbicide and glyphosate-tolerant crops in weed control, we also demonstrated the possibility of creating transgenic plants that can tolerate aspterric acid treatment through expression of the resistance gene *astD*.

Metabolic engineering to improve yield

Discovery and proof of action is just one part of the process though; for using NPs in agriculture also requires large-scale and cost-effective production. Microbial fermentation using genetically modified organisms has therefore great potential to produce a given compound at high titer. Furthermore, the use of a generally regarded as safe (GRAS) organisms such as *S. cerevisiae* as a production host can alleviate public concerns. In recent years, *S. cerevisiae* has been intensely pursued as a host for production of biofuels, chemicals, and pharmaceuticals, the last of which mostly consist of NPs. For example, a 23-step biosynthesis of opioids was recently achieved by combining enzymes from different organisms into a biosynthetic pathway in yeast [7]. Similarly, strictosidine, the common precursor to thousands of plant monoterpene indole alkaloid NPs, can now be produced from yeast after introducing more than twenty genetic changes [8].

NPs are secondary metabolites that are synthesized from primary metabolites as building blocks. For example, aspterric acid is a sesquiterpene, which is synthesized from five-carbon isoprenoid building blocks that are universally used in terpene biosynthesis. The isoprenoid building blocks, IPP and DMAPP, are in turn synthesized from acetyl-CoA, a central metabolite in aerobic respiration, fatty acid biosynthesis, and protein acetylation. Thus, the yield of terpene-derived NPs can be increased through engineering of the host primary metabolism to elevate acetyl-CoA concentrations.

"A diversified library of NPs can help to illuminate the structure-activity relationships of the compound and allow screening of analogs that are more potent, or that can overcome evolved resistance mechanisms."

In recent years, metabolic engineering and synthetic biology have turned Baker's yeast into efficient microbial factories. Metabolic engineering approaches, such as overexpression of pathway genes to produce NPs, or increasing flux of building blocks, such as acetyl-CoA, can lead to dramatic increases in target compound titers. One milestone example is the production of the plant metabolite artemisinic acid at titers of 25 g/l from yeast, as a sustainable source of the antimalarial compound artemisinin [9]. Artemisinic acid, like aspterric acid, is an oxidized sesquiterpene synthesized by the collaborative action of terpene synthase and a P450 monooxygenase. It is therefore reasonable to expect there is ample room to further increase the titers of aspterric acid from the current levels of ~20 mg/l.

Moreover, gene-editing tools have further revolutionized our capability to genetically engineer microorganisms. A suite of recently developed CRISPR-based strain evolution strategies are promising multiplex tools for strain engineering. Other synthetic biology toolkits for yeast, including product compartmentalization and enzyme prospecting, as well as directed evolution of bottleneck enzymes, will help to further increase NP titers.

Diversity-oriented biosynthesis

Although some NPs can be directly used in agriculture or medicine, most compounds require further modification to improve their biological activity. A diversified library of NPs can help to illuminate the structure–activity relationships of the compound and allow screening of analogs that are more potent, or that can overcome evolved resistance mechanisms. Chemical synthesis is a widely adopted approach to generate such analogs of existing NPs. However, many NP structures are difficult to manipulate chemically and often degrade quickly if chemists try to modify them. Moreover, chemically modified compounds are no longer considered "natural" and can face significant regulatory hurdles. Therefore, engineering of the biosynthetic pathway to create structural analogs is an attractive alternative to chemical modification or synthesis. This requires a thorough understanding of the microbial NPs biosynthetic machinery, including the sequence of enzymatic transformations, the mechanisms and substrate flexibilities of individual enzymes.

NP biosynthesis usually follows a "linear" sequence of enzymatically catalyzed reactions, reflecting nature's biosynthetic logic in constructing a complex molecule. Some of these enzymes are promiscuous and can function out of sequence or use alternative substrates. This allows synthetic biologists to exploit the biosynthetic machineries for the synthesis of many NPs to produce "unnatural" NPs. One effective method to expand NP structural diversity is through precursor-directed biosynthesis and mutasynthesis (Fig 4A). In contrast, *de novo* biosynthesis of diversified NP molecules can be achieved by mixing and matching enzymes from different biosynthetic pathways (Fig 4B). Combinatorial biosynthesis, which parallels the concept of combinatorial synthesis, is particular successful in biosynthetic pathways that utilize modular enzymes such as polyketide synthases (PKSs) and nonribosomal peptide synthetases (NRPSs). Individual domains of these "assembly-line" enzymatic machineries can

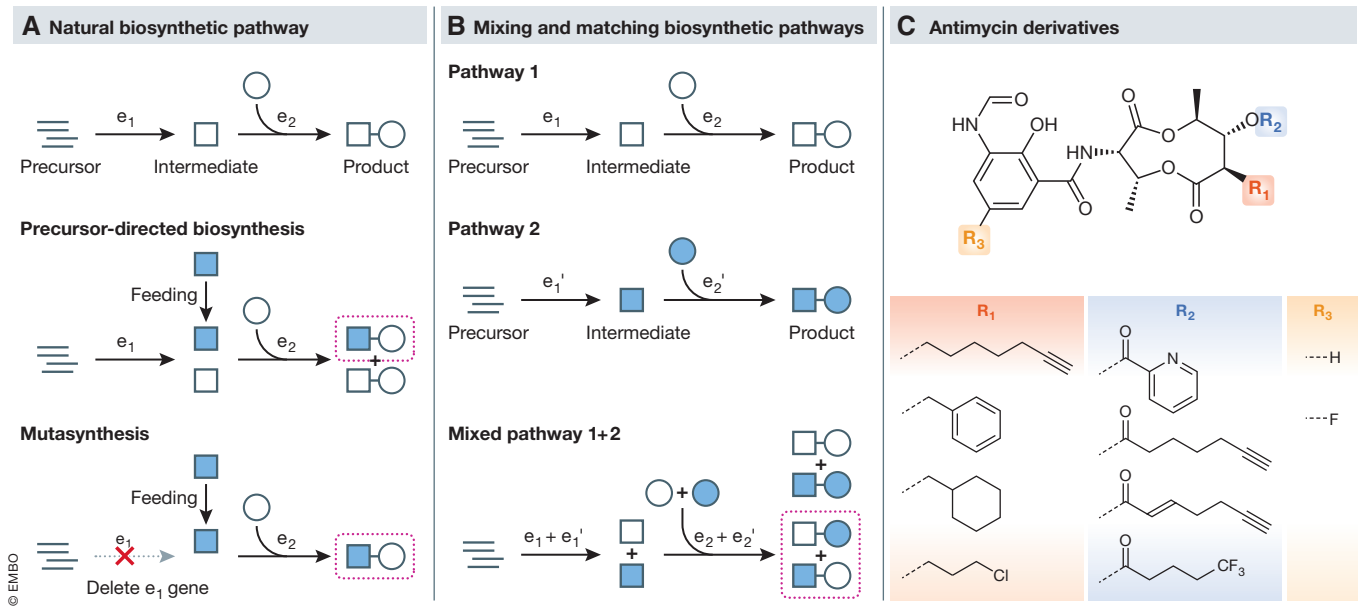

Figure 4. Strategies for diversity-oriented biosynthesis.
(A) Precursor-directed biosynthesis and mutasynthesis to obtain "unnatural" NPs. (B) Mixing and matching biosynthetic pathways to produce NP derivatives. (C) Representative antimycin derivatives produced using combinatorial strategies of diversity-oriented biosynthesis.

be inactivated, inserted, or swapped to precisely introduce modifications to the final NP structure.

"Resistance gene-directed NP discovery has been demonstrated to be an effective strategy to uncover novel NPs with desired modes of action as lead candidates for new insecticides, fungicides or herbicides…"

Combinatorial biosynthesis has been successfully applied to generate a library of fungicidal antimycin analogs, which are cytochrome C reductase inhibitors; fenpicoxamid for instance has been developed by Dow AgroSciences to control the wheat pathogen *Zymoseptoria tritici*. Based on detailed understanding of the biosynthetic pathway of antimycin, diversity-oriented biosynthesis of about 400 analogs was achieved by altering the chemical identities of priming, extending, and tailoring building blocks. Several of these analogs exhibited stronger biological activities than the original NPs, while a few introduced orthogonal reactive handles in the molecules that enabled further chemical derivatization [10] (Fig 4C).

Conclusion

The application of insecticides, herbicides, and fungicides with potent bioactivities and good safety profiles has played an indispensable role in improving the yield and quality of agricultural products. However, their continuous and excessive use has led to the emergence of resistance among plants and plant pathogens. Resistance gene-directed NP discovery has been demonstrated to be an effective strategy to uncover novel NPs with desired modes of action as lead candidates for new insecticides, fungicides, or herbicides to address the problem of growing resistance. Metabolic and biosynthetic engineering of NP synthetic pathways in yeast can further improve titers for microbial production and biological activities for commercial applications. The increasing sophistication of these tools means that we are entering a renaissance of NP discovery for both pharmaceutical and agricultural applications.

Conflict of interest

The authors declare that they have no conflict of interest.

References

1. Cantrell CL, Dayan FE, Duke SO (2012) Natural products as sources for new pesticides. *J Nat Prod* 75: 1231 – 1242
2. Gould F, Brown ZS, Kuzma J (2018) Wicked evolution: can we address the sociobiological dilemma of pesticide resistance? *Science* 360: 728 – 732
3. Rutledge PJ, Challis GL (2015) Discovery of microbial natural products by activation of silent biosynthetic gene clusters. *Nat Rev Microbiol* 13: 509 – 523
4. Yan Y, Liu Q, Zang X, Yuan S, Bat-Erdene U, Nguyen C, Gan J, Zhou J, Jacobsen SE, Tang Y (2018) Resistance-gene-directed discovery of a natural-product herbicide with a new mode of action. *Nature* 559: 415 – 418
5. Tang X, Li J, Millán-Aguiñaga N, Zhang JJ, O'Neill EC, Ugalde JA, Jensen PR, Mantovani SM, Moore BS (2015) Identification of thiotetronic acid antibiotic biosynthetic pathways by target-directed genome mining. *ACS Chem Biol* 10: 2841 – 2849
6. Kling A, Lukat P, Almeida DV, Bauer A, Fontaine E, Sordello S, Zaburannyi N, Herrmann J, Wenzel SC, König C *et al* (2015) Targeting DnaN for tuberculosis therapy using novel griselimycins. *Science* 348: 1106 – 1112
7. Galanie S, Thodey K, Trenchard IJ, Filsinger Interrante M, Smolke CD (2015) Complete biosynthesis of opioids in yeast. *Science* 349: 1095 – 1100
8. Brown S, Clastre M, Courdavault V, O'Connor SE (2015) *De novo* production of the plant-derived alkaloid strictosidine in yeast. *Proc Natl Acad Sci USA* 112: 3205 – 3210
9. Paddon CJ, Westfall PJ, Pitera DJ, Benjamin K, Fisher K, McPhee D, Leavell MD, Tai A, Main A, Eng D *et al* (2013) High-level semi-synthetic production of the potent antimalarial artemisinin. *Nature* 496: 528 – 532
10. Yan Y, Jing C, Lihan Z, Qingfei Z, Ying H, Hua Z, Daozhong Z, Takayoshi A, Ikuro A, Wen L (2013) Multiplexing of combinatorial chemistry in antimycin biosynthesis: expansion of molecular diversity and utility. *Angew Chem Int Ed Engl* 52: 12308 – 12312