This is the accepted version of the following article: Fewer, D.P. and Metsä‐Ketelä, M. (2020), A pharmaceutical model for the molecular evolution of microbial natural products. FEBS J. doi:10.1111/febs.15129, which has been published in final form at https://doi.org/10.1111/febs.15129.

Title

A Pharmaceutical Model for the Molecular Evolution of Microbial Natural Products

Authors

David P. Fewer[1] and Mikko Metsä-Ketelä[2,*]

Affiliations

[1]Department of Microbiology, University of Helsinki, FI-00014 Helsinki, Finland
[2]Department of Biochemistry, University of Turku, FI-20014 Turku, Finland

Corresponding Author

Mikko Metsä-Ketelä, Contact: E-mail: *mianme@utu.fi, Tel. +35823336847, Fax. +35823336860

Running Title (50 / 50 char)

Molecular Evolution of Microbial Natural Products

Keywords (up to 10)

molecular evolution; enzyme promiscuity; secondary metabolism; natural product; gene duplication; horizontal gene transfer; recombination; polyketide; non-ribosomal peptide; ribosomally synthesized and post translationally modified peptide; terpene.

Abbreviations

PKS, polyketide synthase, NRPS; non-ribosomal peptide synthetase; RiPP, ribosomally synthesized and post-translationally modified peptide; ACP, acyl carrier protein; AT, acyl transferase; KS, ketosynthase; DH, dehydratase; ER, enoyl reductase; KR, ketoreductase; C, condensation domain; A, adenylation domain; PCP, peptide carrier protein.

Conflict of interest

The authors declare no conflict of interest

ABSTRACT (250 / 250 words)
Microbes are talented chemists with the ability to generate tremendously complex and diverse natural products that harbor potent biological activities. Natural products are produced using sets of specialized biosynthetic enzymes encoded on secondary metabolism pathways. Here we present a two-step evolutionary model to explain the diversification of biosynthetic pathways that account for the proliferation of these molecules. We argue that the appearance of natural product families has been a slow and infrequent process. This first step led to the original emergence of bioactive molecules and different classes of natural products. However, much of the chemical diversity observed today has resulted from the endless modification of the ancestral biosynthetic pathways. This second step rapidly modulates the pre-existing biological activities to increase their potency and to adapt to changing environmental conditions. We highlight the importance of enzyme promiscuity in this process as it facilitates both the incorporation of horizontally transferred genes into secondary metabolic pathways and the functional differentiation of proteins to catalyze novel chemistry. We provide examples where single point mutations or recombination events have been sufficient for new enzymatic activities to emerge. A unique feature in the evolution of microbial secondary metabolism is that gene duplication is not essential, but offers opportunities to synthesize more complex metabolites. Microbial natural products are highly important for the pharmaceutical industry due to their unique bioactivities and understanding the natural mechanisms leading to the formation of diverse metabolic pathways is vital for future attempts to utilize synthetic biology for the generation of novel molecules.

INTRODUCTION

Microbial natural products are chemically complex metabolites that occupy an unusual chemical space [1–3]. A key feature of natural products is that they possess potent biological activities and have proven to be a rich source of pharmaceutical agents [1,2]. The potency of these molecules is reflected in their widespread use in the treatment of a variety of human diseases [1,2]. A large portion of drugs approved for clinical use are based on or inspired by natural products [1]. Fifty-eight percent of antibiotics and 32% of anticancer agents approved for medical use in the last 35 years are natural products or their semi-synthetic derivatives [2]. Natural products can target biological macromolecules in a highly selective fashion and are widely used as chemical probes [4]. In addition, natural products have traditionally been the pinnacle of organic synthesis due to their complexity and they have inspired the development of many advanced synthetic methods in organic chemistry [5].

The role of natural products in the biology of the producing organisms is not always clear. However many have important and diverse roles in the ecology and physiology of microbes [6,7]. Natural products with antimicrobial, antifungal and cytotoxic activities are widely believed to facilitate chemical warfare with other organisms [6] or as act as part of a chemical defense [8]. These compounds can inhibit key cellular processes such as DNA replication, transcription and translation [1,2], but may also act as highly selective and potent enzyme inhibitors or destroy cell membranes [4]. However, many secondary metabolites can also serve as iron-scavenging siderophores [9], function as natural sunscreens [10] or act as signaling molecules in microbial chemical communication [11]. The diverse biological activities are generally thought to improve fitness and confer a competitive advantage on the producing microbes [12].

Natural Product Biosynthesis and Comparative Genomics

The majority of microbial natural products share a common biosynthetic logic and can be assigned to just a handful of major classes, including terpenes, saccharides, polyketides, alkaloids, and peptides [13–16]. Natural products are synthesized through complex secondary metabolic pathways consisting of enzymes that can catalyze challenging sequential chemical transformations in a regioselective and stereoselective manner [13, 14, 17–19]. Microbial natural product biosynthetic pathways are typically organized in self-contained gene clusters that can be over 100 kb in length and encode over 40 biosynthetic enzymes [12, 20]. Natural product biosynthetic enzymes work in coordinated enzyme cascades to produce complex metabolites using a limited number of primary metabolites as starting material [12, 15, 21]. The chemical diversity of natural products can also be increased through enzymes that supply unusual dedicated precursor substrates or catalyze tailoring reactions [12, 20, 22]. Families of natural products that exhibit chemical variation typically have a set of conserved biosynthetic enzymes for the synthesis of a core that defines the family, but also a set of accessory tailoring enzymes that are not universally conserved [12, 17]. The gene clusters typically encode all of the enzymatic reactions needed to assemble a natural product using substrates recruited from primary metabolism, but also encode genes for regulation, resistance and transport of natural products through coordinated expression of the secondary metabolic pathways [12, 20].

Many microbes specialize in the production of natural products and have large genomes rich in natural product biosynthetic genes [12, 20, 23]. The organization of natural product biosynthetic enzymes into regulated gene clusters has facilitated the identification of biosynthetic gene clusters and enabled the detection of numerous cryptic pathways encoding novel natural products from rapidly accumulating public microbial genome sequence data [20,

23–25]. Many of these biosynthetic pathways appear silent and under strict regulation, and the metabolites are produced only in response to various environmental stimuli [26, 27]. Natural product biosynthetic gene clusters are much more wide-spread and more diverse in nature than anticipated and are frequently encountered in microbes not typically associated with natural products [23, 24]. For instance, numerous secondary metabolic pathways have been discovered by metagenomics sequencing of the uncultured microbes of the human microbiota [28, 29] and a variety of other microbial communities [30].

Enzyme Promiscuity Drives the Structural Diversification of Natural Products.

The biosynthesis of numerous chemical variants is a hallmark of secondary metabolism [1, 2]. A wealth of data indicates that enzyme promiscuity is an inherent property of natural product biosynthetic enzymes [31–35]. Enzyme promiscuity can be divided into substrate promiscuity, which describes the ability of the enzyme to utilize a wide range of substrates to catalyze similar chemical transformations, and catalytic promiscuity, where the enzyme catalyzes secondary reactions that differ from the canonical primary reaction [36]. Substrate and catalytic promiscuity can be considered key driving forces for the diversification of natural product biosynthetic enzymes [37].

Enzymes involved in natural product biosynthesis appear to harbor several important general features that promote promiscuity [38]. Natural product biosynthetic enzymes are "generalists", which typically display slow reaction rates and substrate promiscuity [31]. These enzymes have been widely shown to harbor weak secondary "moonlighting" activities [32, 39, 40]. This catalytic promiscuity promotes the ability to catalyze fortuitous chemical transformations and may form the seed for the emergence of new enzymatic activities [41]. Furthermore, natural product biosynthetic pathways often encode multi-functional enzymes that catalyze either several consecutive reactions with different regiospecificity [42, 43] or entirely different chemical transformations [39, 44].

In many ways, proteins involved in secondary metabolism are reminiscent of "generalist" primordial enzymes that were responsible for shaping the primary metabolisms of organisms. These ancestral proteins responsible for establishing the core primary metabolic pathways, which are shared by most organisms today, have subsequently gained specificity and efficiency in purifying selection [41]. In this respect, enzymes associated with biosynthesis of natural products could be considered "evolutionary intermediates" from the perspective of primary metabolism. However, it may be beneficial for enzymes associated with secondary metabolism to stay as "generalists" and to never reach the specialization state of their counterparts in primary metabolism. In secondary metabolism, negative selection may prevent the evolution of highly efficient and specific enzymes, since such a high specialization state could be detrimental for increasing chemical diversity [37]. In contrast, promiscuous "generalist" enzymes can play a major role in both the acquisition of novel genes by natural product biosynthetic pathways and the functional diversification of enzymes already encoded on those pathways

A Pharmaceutical Model for Pathway Evolution

The chemical complexity of secondary metabolites obfuscates both their evolutionary origins and the selective pressures driving the diversification of their biosynthetic pathways. The widely considered Screening Hypothesis theory by Firn and Jones originating in 1991 is based on the proposition that "potent biological activity is a rare property for any one molecule to possess" [45–47]. Consequently, evolution would favor organisms that can generate and retain chemical diversity at low cost, similarly to how compound libraries are

screened by pharmaceutical companies for desired bioactivities [45, 46]. However, the accumulating wealth of genomic data and the elucidation of the biosynthetic logic of all major classes of natural products [14, 15, 16] warrants a revision to the hypothesis.

We propose a two-step model that accounts for much of the chemical diversity of microbial natural products. The first step in this model is the slow appearance of ancestral biologically active metabolites. The first appearance of a new natural product family required the recruitment and cooperation of multi-gene combinations resulting in the formation of novel metabolic pathways. Spontaneous chemical transformations, side products produced through other enzymatic pathways and acquisition of new extracellular substrates may have contributed to the appearance of the first secondary metabolites in a manner similar to the evolution of early primary metabolism [48]. This first stage may have been a cyclical process that has required several rounds of metabolite-enzyme coevolution [48], which has finally resulted in the emergence of bioactive secondary metabolites. The relative low number of natural product classes suggests that the evolution and dissemination of new natural product classes were rare events (Fig. 1). This scenario is in agreement with current data that indicates that only a limited number of enzyme families are responsible for the bulk of the structural diversity found in microbial natural products [49].

Our model dictates that the majority of natural products have appeared in the second step through near endless modulation of pre-existing secondary metabolic pathways. The key difference is that the natural products emerging at this later stage have evolved from already existing bioactive secondary metabolites. The driving force for the chemical diversification has been the modulation of the biological activity due to selective pressures experienced under changing environmental conditions. Ancestral secondary metabolites may have displayed low target affinity or weak biological activity, but have served as templates to maximize potency through optimization of chemical structure. Formally, the second step follows the forward pathway evolution model by Granick [48, 50], which states that all pathway intermediates have been, at some stage, the end products of the given pathway, but have evolved further to carry out the same function more efficiently. In keeping with the pharmaceutical theme, this process would be analogous to Lead Optimization where the desired activity is fine-tuned. Even minor changes in the chemical structures of natural products may have significant consequences for the biological activity of the molecule [1, 2].

This two-part model implies that the formation of novel metabolite classes were rare events, but were followed by frequent and continuous structural diversification to form large natural product families (Fig. 1). In the next sections, we will focus specifically on the second step and provide examples how novel natural products can arise through simple evolutionary events. We highlight cases where single point mutations, gene duplication, recombination and horizontal gene transfer have allowed facile generation of chemical analogs.

Direct Modulation of Enzyme Function in Biosynthetic Gene Clusters

Most evolutionary theories highlight the importance of the gene duplication event in molecular evolution [41, 51, 52], since it provides a mechanism for generating novel chemistry while retaining the original function [51–53]. However, Ohno's law of *Evolution by Gene Duplication* [51] would appear not to apply to secondary metabolism. A growing body of evidence suggests that natural product gene functions may also evolve directly in the biosynthetic gene clusters without gene duplication [54], which is possible because genes involved in secondary metabolism are not essential for the survival of the host. This is most apparent in comparisons of related secondary metabolite biosynthetic gene clusters, where

highly conserved enzymes can nonetheless catalyze distinct chemistry. Perhaps the most illustrative examples are found from evolution-inspired protein engineering experiments, where the activities of functionally distinct enzymes can be interchanged with single point mutations [39, 43].

Anthracyclines are a family of cytotoxic tetracyclic polyphenolic compounds, which are synthesized by type II iterative polyketide synthases (PKS) and modified by a range of tailoring enzymes [17]. Many anthracycline biosynthetic gene clusters found in actinomycetes encode SAM-dependent methyltransferases (Fig. 2A). These promiscuous enzymes have diversified to catalyze 4-O-methylation, 10-decarboxylation or 10-hydroxylation, with additional changes in substrate specificity in regards to the glycosylation state [54]. The functional diversification appears to have occurred directly *in situ* within the gene clusters (Fig 2B), as demonstrated by the congruence of phylogenetic trees composed of the methyltransferases and core anthracycline biosynthetic genes [54]. The canonical member DnrK from the daunorubicin pathway is a 4-O-methyltransferase [55], but it harbors moonlighting 10-decarboxylation activity towards anthracyclines with a free carboxyl group at C10 [39]. In turn, this latter promiscuous activity has evolved to 10-hydroxylation functionality as observed in RdmB from the rhodomycin pathway [56]. The shift in catalysis could be pinpointed to insertion of a single serine residue S297 to DnrK, which resulted in gain of 10-hydroxylation activity [39]. Surprisingly, the inserted serine points away from the active site in DnrK-Ser, but since it is located in α16-helix, the insertion results in rotation of the preceding phenylalanine F296 towards the active site. The bulky F296 blocks a water channel to the surface of the protein in a manner similar to native RdmB (Fig. 2C). Closure of the active site and protection of a carbanion intermediate (Fig. 2A) from protonation by solvent molecules is likely to be crucial for the switch in activity [39]. The third enzyme type represented by EamK from the komodoquinone pathway catalyzes 10-decarboxylation solely (Fig. 2A). Even though the structural basis for 10-decarboxylation is still unresolved, the loss of methylation activity could result from misalignment of the substrate in relation to SAM, which would leave 10-decarboxylation as the sole functionality for EamK [54]. Interestingly, 10-hydroxylation may offer the producing organism significant selective advantage, since this functionality appears to have evolved independently on at least two occasions as exemplified by the CalMB-type proteins (Fig. 2B).

Angucyclines produced mainly by *Streptomyces* bacteria are another large group of aromatic polyketides that include gaudimycins, urdamycins, jadomycins and landomycins [57]. The various end products are formed via complex redox modification cascades and glycosylation from the last common intermediate prejadomycin (Fig. 3A) [58]. All pathways include FAD-dependent mono-oxygenases that are responsible for C12 hydroxylation and quinone formation [33]. However, in gaudimycin biosynthesis PgaE catalyses a secondary hydroxylation at C12b [42, 59], whereas the homologous flavoenzyme JadH from the jadomycin pathway performs 4a,12b-dehydration [42, 60]. The emergence of 12b-hydroxylation activity in PgaE could be linked to just two residues P78 and I79 distal from the active site [61]. These residues reside at the dimerization interface and, while the mutations allowed the ancestral protein to gain C12b hydroxylation functionality, lead simultaneously to substrate inhibition for the original C12 hydroxylation activity. Urdamycin biosynthesis continues from the double hydroxylated intermediate with C6 ketoreduction (Fig. 3B) by a short chain alcohol dehydrogenase/ reductase (SDR) UrdMred [62]. LanV from the related landomycin pathway has diverged from these canonical 6-ketoreductases to prefer the earlier 12-hydroxylated pathway intermediate as a substrate and catalyse the ketoreduction with opposite stereochemistry [63]. A single point mutation S192 to I192 in LanV, where the

longer side-chain pushes the substrate into a different angle above NADPH, was sufficient to convert the activity towards use of C12,C12b-hydroxylated substrates [62, 63]. The importance of enzyme promiscuity in these redox cascades was highlighted under artificial conditions where native JadH and LanV were used to catalyze ancestral non-native reactions leading to formation of the non-cognate urdamycin pathway product [33].

A key difference in landomycin and urdamycin lies in the glycosylation steps (Fig. 3C), where the first D-olivose unit is attached either *via* a canonical O-glycosidic [64] or an atypical C-C bond [65, 66], respectively. The LanGT2 has been engineered to catalyze C-glycosylation through exchange of just 12 residues in the α3-helix from the UrdGT2 template [67]. The modification led to reduction in the size of the aglycone binding pocket in LanGT2 particularly by I58 and I62 [68]. Docking studies indicated that subtle repositioning of the aglycone relative to the sugar nucleotide, with the C9 atom moving towards the co-substrate and the C8 hydroxyl group away from the co-substrate, leading to a switch in regiochemistry and even the catalytic mechanism of LanGT2 [68].

Natural enzyme promiscuity is also evident in the diversification of terpene natural products [69]. The squalene hopane synthase *Aac*SHC from *Alicyclobacillus acidocaldarius* is able to convert a wide range of functionalized long-chain linear polyprenoids [34, 70]. Catalysis is initiated by an aspartic acid D376 acting as an unusually strong Brønsted acid leading to C=C bond protonation [71]. The hydrophobic binding pocket of the enzyme chaperones the flexible substrates into product-like conformations that induce cyclization cascades that may be terminated either by releasing water molecules through specific channels or protonation of the terminal isoprene unit (Fig. 4A). However, the native enzyme generally does not accept relatively small substrates, in contrast to the related enzyme from *Zymomonas mobilis* [72]. Single point mutations in AscSHC were sufficient to the reshape the chiral hydrophobic environment in the active site to alter both the substrate and catalytic specificity of the enzyme to diverse compounds [73]. Similarly, six out of 11 SHCs could be converted to citronellal monoterpenoid cyclases with a single point mutation [72].

Enzymes of the 2-oxoglutarate (2OG) and non-heme iron dependent oxygenases are able to perform energetically demanding chemical transformations [18, 19] and are common components of secondary metabolic pathways [19]. Their functional divergence is particularly prevalent in the formation of β-lactams, where the enzymes catalyze various hydroxylation, cyclization, desaturation and epimerization reactions [18]. Studies on fungal meroterpenoids (Fig. 4B) have illustrated how highly divergent products may be obtained through dynamic skeletal rearrangement [43]. AusE [74] and PrhA [75] both utilize preaustinoid A1 as a substrate, but whereas the former catalyzes desaturation at C1–C2 and carbocyclic rearrangement to preaustinoid A3, the latter first desaturates at C5–C6, followed by rearrangement of the A/B-ring to form berkeleydione. The functions of the two proteins could be interchanged by manipulating two amino acids in the vicinity of the A-ring [43]. The importance of enzyme promiscuity was evident as a single point mutation AusE-S232A was sufficient for gaining a PrhA-type “B-ring expansion” activity, while still retaining its natural “spiro-lactone forming” activity. In case of a PhrA- V150L-A232S double mutant, the enzyme gained unnatural hydroxylation functionality as a third activity, while the addition of a third M241V mutation to this scaffold led to formation of even more novel reaction products [43].

Gene Duplication Allows Formation of More Complex Chemical Structures

Many natural product biosynthetic gene clusters encode two or more copies of homologous genes [12], which have presumably arisen through gene duplication. Gene duplication offers a means to build larger and chemically more complex natural products from ancestral pathways. This may result in the formation of novel compounds with enhanced or altered biological activity that confers a selective advantage on the producing organism. Enzyme promiscuity is likely to play a central role in the functional differentiation of duplicated gene products, and examples of both substrate and catalytic promiscuity are abundant in natural product biosynthesis [76–78]. The process is reminiscent to events in the early evolution of ancestral core metabolism, where Ycas [79] and Jensen [80] proposed independently that broad-specificity generalist enzymes originally participated in multiple pathways, but duplication and divergence to specialist enzymes allowed for the foundation of primary metabolic pathways [48].

Examples of catalytic promiscuity are found in many pathways where individual standalone genes have duplicated. Nogalamycin (Fig. 5A) represents a subgroup of anthracyclines that contain a dually linked aminosugar connected at C1 and C2 through both O- and C-glycosylation, respectively [81]. Gene duplication and catalytic promiscuity have been important in the evolution of the unique substructure. This chemistry originates from an unusual two-component mono-oxygenase system that is responsible for the installation of the C1 hydroxyl group [82, 83]. The system is composed of an SDR-enzyme, SnoaW, and a protein SnoaL2 homologous to canonical polyketide cyclases such as SnoaL [82]. In a manner similar to flavin chemistry, NADPH-dependent reduction of the anthracycline quinone by SnoaW allows the substrate to react with molecular oxygen to form a peroxide intermediate, which is converted to the hydroxylated reaction product by SnoaL2. The canonical reaction of the anthracycline fourth ring polyketide cyclases such as SnoaL [84] and AknH [85] is to catalyze aldol condensations that determine the stereochemical configuration at C9, but interestingly SnoaL2 has also been observed to harbor minor ancestral cyclization activity [76]. After attachment of L-rhodosamine by the glycosyltransferase SnogD [81, 86, 87], the subtype epitomizing C2–C5´´ bond is formed by the 2OG and non-heme iron dependent oxygenase SnoK [78]. However, a gene duplication event has led to the appearance of C4´´ epimerase SnoN from SnoK [78]. Despite drastic differences in the chemistries catalyzed by the two enzymes, both proteins are able to use the same substrate and the key difference in catalysis appears to be whether the C4´´ or C5´´ carbon is positioned in front of the high valent Fe=O center [78].

The biosynthesis of the fungal novofumigatonin monoterpenoid offers another fascinating example of gene duplication within a single biosynthetic pathway [77]. The novofumigatonin biosynthetic gene cluster encodes three copies of 2OG and non-heme iron-dependent enzymes responsible for endoperoxide formation and rearrangement (Fig. 5B). The endoperoxide forming NvfI utilizes asnovolin as a substrate and is responsible for the incorporation of three oxygen atoms [77]. The mechanistic proposal follows classical hydroxylation chemistry that is initiated by hydrogen abstraction, but with an unusual additional incorporation of molecular oxygen prior to completion of the reaction cascade through oxygen rebound. The second enzyme, NvfE, catalyzing orthoesterification has lost the ability to utilize 2OG as a co-substrate, but requires the non-heme iron for initial cleavage of the endoperoxide and subsequent rearrangements [77]. Finally, NvfF is engaged in two consecutive oxidative desaturation and hydroxylation reactions, which lead to further rearrangements and the formation of novofumigatonin [77].

Recombination and Substrate Promiscuity Allow Diversification of Modular Biosynthetic Systems

Polyketides and non-ribosomal peptides belong to two of the most important, structurally diverse and widespread natural product classes in nature. Type I polyketides and non-ribosomal peptides are synthesized on modular enzyme complexes that act as assembly lines [14,15]. Nonribosomal peptide synthetase (NRPS) and PKS biosynthetic pathway have been discovered from all three domains of life [24], and the thiotemplate based biosynthetic logic underpinning their biosynthesis has been elucidated in detail [14, 15]. Briefly, these module systems are comprised of loading and elongation modules responsible for the initiation of biosynthesis and the growth of the polyketide/peptide intermediate [14, 15]. The elongation modules of type I polyketides requires three essential domains (Fig. 6A); the polyketide chain is tethered to the acyl carrier protein (ACP), while the acyltransferase (AT) selects the extended units to be incorporated by the ketosynthase (KS) domain [14]. The β-keto functional group may be reduced in a sequential manner by auxiliary ketoreductase- (KR), dehydratase- (DH) and enoylreductase- (ER) domains to generate structural diversity [14]. The essential domains in NRPS elongation modules (Fig. 6A) include the adenylation (A) domain for selection and activation of a range of proteinogenic and non-proteinogenic amino acids, the condensation (C) domain for peptide bond formation and the peptidyl carrier protein (PCP) domain, to which the nascent peptide intermediate is bound as a thioester [15]. In both of these systems, elongation modules are responsible for the growth of the polyketide or peptide intermediates, which are passed from one ACP/PCP domain to another until the biosynthesis is typically terminated in the final module by thioesterase (TE) domain [14,15]. A range of additional optional domains increase the structural diversity of the final polyketide and non-ribosomal peptide end-product [14,15]. The modular architecture of NRPS and PKS enzymes coupled with high sequence similarity promotes recombination events between modules. This recombination process can result in the loss, gain or replacement of modules, as well as specific domains within a module, resulting in reprogramming of the chemical diversity encoded by the pathway.

Microcystins (Fig. 6A) are potent inhibitors of eukaryotic protein phosphatases 1 and 2 A that are responsible for infrequent animal toxicosis around the world [8, 88]. These hepatotoxins are one of the most structurally diverse families of toxins with almost 250 reported variants [88, 89]. Microcystins are the end-products of mixed NRPS-PKS enzyme complexes encoded in a 55-kb microcystin (*mcy*) gene cluster [8]. Recurrent homologous recombination between McyB1 and McyC modules resulted in the replacement of the substrate conferring portions of the A domains leading to the incorporation of a range of amino acids at the variable X and Z positions of the toxin [8, 90]. In some cases, the substrate conferring portions of the A domains in the resultant McyB1 and McyC are identical, but nevertheless incorporate different amino acids [90]. This phenomenon was hypothesized to be the result of a conflict between the substrate specificities of the gatekeeping C domain and the recombinant A domain [90]. Experimental work to recapitulate entire McyB1 and McyC modules using artificial combinations of domains confirmed this hypothesis and demonstrated that the structural diversity observed could be best explained by a series of point mutations and homologous recombination events [91]. A larger recombination event resulting in the fusion of the McyA2 and McyB2 modules concomitant with deletion of McyB1 module and a change in the substrate specificity of the McyA1 module from L-Ser to L-Tyr is thought to have resulted in the creation of nodularin biosynthetic pathway from an ancestral microcystin synthetase pathway [8].

Few studies report the redesign of modular PKS natural product biosynthetic pathways using natural recombination. Pikromycin and erythromycin are related macrolide antibiotics produced by *Streptomyces* bacteria (Fig. 6B). Homologous recombination in yeast was utilized to generate hybrid pikromycin genes either with artificial combinations of two pikromycin modules or recombination with a module from the erythromycin pathway [92]. Recombination occurred within the KS or AT domains in the majority of cases with in-frame recombination frequencies of 90% [92]. Many of the chimeras led to formation of soluble proteins and expected modified macrolactones were synthesized with high efficiency [92]. A similar approach resulted in the successful engineering of the rapamycin (Fig. 6C), a type I polyketide macrolide with immunosuppressive and anti-proliferative properties, biosynthetic pathway [93]. Serendipitous natural recombination events were used directly to generate a library of novel analogs in a rapamycin-producing *Streptomyces* strain [93]. A temperature sensitive plasmid was first integrated into the gene cluster using homologous recombination. Surprisingly, activation of the temperature sensitive replicon, while still integrated into the genome, led to the production of twenty chemical variants of rapamycin [93]. Genome sequencing revealed that the progeny harbored new PKS architectures, including both deletion and duplication of modules, which were generated at positions of high sequence similarity [93]. The biological activity of the novel truncated analogs were modulated, which is in agreement with our two-step model. Rapamycin harbors peptidyl-prolyl isomerase (PPIase) inhibitory activity and is able to modulate the activity of the mTORC1 pathway [94, 95]. One of the rapamycin chemical variants (Fig. 6C) retained high PPIase inhibitory activity, but displayed 100-fold reduced activity against p-S6 formation in the mTORC1 pathway [93].

Recent evolution-guided protein engineering approaches that mimic natural recombination events have been utilized in rewiring natural product biosynthetic pathways. Aureothin and neoaureothin (Fig. 6D) are antifungal metabolites produced on related modular PKS enzyme complexes by *Streptomyces* spp. [96]. The neoaureothin (*nor)* assembly line contains two additional modules in comparison to the aureothin (*aur*) gene cluster suggesting the two pathways are related through the gain or loss of modules due to recombination events [97]. Engineering experiments to insert modules into the *aur* pathway or deleting the modules from the *nor* PKS resulted in the expected end-products of the hybrid pathways [97]. The antifungal activity is dependent on the presence of the tetrahydrofuran ring system, which is formed during tailoring steps of the biosynthesis by the oxygenases AurH and NorH [97]. Only AurH was able to utilize both aureothin and neoaureothin as substrates, suggesting that the aureothin pathway may be the ancestral pathway [97]. Antimycins (Fig. 6E) are a family of di-lactone depsipeptides with anti-fungal, anti-cancer, and anti-inflammatory bioactivities [98]. Homologous recombination inspired protein engineering experiments allowed the successful reprogramming of the antimycin biosynthetic pathway by expanding and contracting the modular architecture using three related hybrid NRPS-PKS pathways encoding antimycin, JBIR-06 and neoantimycin (Fig. 6E) as templates [98].

Horizontal Gene Transfer Facilitates the Distribution of Natural Product Biosynthesis Enzymes

Horizontal gene transfer is thought to play a central role in the diversification of natural product biosynthetic pathways [12, 99]. It is well established that entire biosynthetic gene clusters can spread through horizontal gene transfer where they confer a selective advantage upon the recipient organism [12, 20, 99]. The spread of natural product pathways is facilitated by the organization of biosynthetic enzymes into gene clusters, which may be further encoded

in genomic islands or on plasmids [20, 99, 100]. The limited host range of genetic elements like conjugal plasmids or pathogenicity islands as well as codon usage, the availability of metabolic precursors, or differences in transcription and regulation place restrictions on horizontal gene cluster transfer making the spread of natural product biosynthetic pathways more common amongst closely related organisms [12, 20, 99]. However, horizontal gene transfers are necessary to explain the production of identical secondary metabolites by microbes belonging to different phyla, for example the lyngbyatoxin biosynthetic pathway is found in both the actinobacteria and cyanobacteria phyla [101], while the macrolide scytophycin is found in both the proteobacteria and cyanobacteria phyla [102]. The same natural product can also be shared between organisms from different domains, for example the UV-sunscreen mycosporine is produced by a range of photosynthetic algae and bacteria [7], the beta-lactam antibiotic penicillins and cephalosporins are produced by fungi and bacteria [103] and potent neurotoxic saxitoxins [104] are produced by both cyanobacteria and dinoflagellates.

In addition to transfer of entire gene clusters, natural product biosynthetic pathways also evolve through the acquisition of new individual enzymes by horizontal gene transfer, which leads to the diversification of natural product chemical structures [12, 99]. The *O*-acetylation of the cyanobacterial microcystin toxin is the result of a gain of an acetyltransferase by strains of the cosmopolitan genus *Nostoc* [105] concomitant with the loss of a SAM dependent O-methyltransferase found in all other microcystin biosynthetic pathways [8]. This enzyme belongs to the XAT class of hexapeptide acyltransferases that catalyze the CoA-dependent acetylation of a variety of hydroxyl-bearing acceptors including chloramphenicol and streptogramin antibiotics [105, 106]. Since these enzymes confer resistance to antibiotics through covalent modifications, they are widely found in microbes due to the spread of antibiotic resistance genes [106, 107]. Members of the xenobiotic acetyltransferase family exhibit some degree of substrate promiscuity [105–109], which may promote the incorporation of this enzyme into non-cognate natural product biosynthetic pathways. The gain of natural product biosynthetic enzymes in this manner would obviate the need for gene duplication or point mutations affecting the activity of pre-existing enzymes.

The ABBA superfamily of prenyltransferases (Fig. 7) catalyze the aromatic prenylation of an unusually wide variety of natural products in bacteria and fungi [110–113]. The enzymes share low primary sequence similarity and are instead defined by a conserved tertiary architecture consisting of a ring of solvent-exposed α-helices surrounding a central barrel of ten antiparallel β-strands [111, 112, 114]. They catalyze the prenylation of aromatic substrates by electrophilic aromatic substitution similar to Friedel-Crafts alkylation [113]. ABBA prenyltransferases catalyze the transfer of C5, C10, or C15 prenyl groups derived from the corresponding isoprenyl diphosphate metabolites onto a variety of electron-rich aromatic acceptors [112]. Characterized ABBA prenyltransferases typically exhibit strict specificity for their prenyl donors, but can accept diverse aromatic acceptors [115]. ABBA prenyltransferases commonly exhibit substrate promiscuity and typically work on a variety aromatic acceptors [111, 113, 115, 116]. For example, representative members of the indole alkaloid ABBA prenyltransferase family catalyze the regioselective normal or reverse prenylation of a variety of natural products [114, 117–120]. The KgpF, SirD, an PagF prenyltransferases from the kawaguchipeptin, sirodesmin PL, and prenylagaramide biosynthetic pathways, respectively, have all demonstrated substrate promiscuity *in vitro* [121–123]. Such broad substrate specificity could be expected to facilitate the spread and integration of these enzymes into new natural product biosynthetic pathways. Interestingly, the permissive prenyltransferases PriB, FgaPT2 and CdpNPT were all shown to catalyze the

prenylation of daptomycin [121], an unrelated macrocyclic lipopeptide antibiotic in clinical use (Fig. 7). Prenylation of daptomycin using these enzymes improved the antibacterial activity of this bactericidal antibiotic [121]. ABBA prenyltransferases are found in a wide variety of natural product classes including indole alkaloids, peptides, flavonoids, coumarins, terpenoids, and phenazines [110–114, 124]. Indeed phylogenetic analysis suggests the active exchange of ABBA prenyltransferases between different natural product classes results in increased structural diversity of the natural products they encode [125].

Novel natural product enzymes are often thought to be recruited from primary metabolism [126]. However, the origin of natural product biosynthetic enzymes has not been widely studied and many could originate from other natural product pathways in addition to primary metabolism. Substrate promiscuity and the ability to act on non-cognate substrates should facilitate the dissemination and integration of natural product biosynthetic enzymes into novel natural product biosynthetic pathways through horizontal gene transfer. Natural product biosynthesis is traditionally viewed as a constrained self-contained process with all of the enzymes necessary for the biosynthesis of the natural product encoded in a single gene cluster [127]. However, the biosynthesis of erythrochelin requires crosstalk between two separate non-ribosomal peptide synthetase biosynthetic gene clusters [128]. Here the authors demonstrate that a delta-N-acetyltransferase enzyme located in separate non-functional non-ribosomal peptide synthetase biosynthetic gene cluster was necessary for the biosynthesis of erythrochelin [128]. Similarly, hybrid peptides produced through the use of non-proteinogenic amino acids encoded in unrelated biosynthetic pathways are commonly reported in cyanobacteria [129, 130]. The formation of natural products that are intermediates between two pathways could be seen as a step towards full integration of novel biosynthetic genes into natural product pathways. Overall, it is clear that substrate and catalytic promiscuity is commonly exhibited by enzymes encoded in many natural product pathways [32–36]. However, the implications of this promiscuity for the spread and integration of natural product enzymes into new natural product pathways are not yet widely considered.

Conclusions and future perspectives

Here we present a two-step evolutionary model to explain how much of the chemical diversity of microbial natural products has arisen. The fact that only a handful of enzymatic machineries are responsible for much of the observed chemical diversity reinforces the challenges in the *de novo* evolution of new natural product metabolic pathways. We propose that the rate-limiting step has been the original appearance of bioactive molecules, which requires sets of enzymes from primary metabolism to act in a concerted manner to build novel natural product biosynthetic pathways. It is interesting to note that the distribution of natural product biosynthetic pathways in living organisms is typically patchy with polyketide, terpene and peptidic natural product biosynthetic gene clusters being widely distributed and common in diverse bacterial phyla [13–16, 23]. However, other secondary metabolites are found from a much narrower range of organisms. For example, recent work indicates that the relatively rare C-nucleoside metabolites are the result of convergent evolution that are synthesized using three distinct biosynthetic logic in actinobacteria [131–133]. An interesting question from an evolutionary perspective is whether these rare metabolites may have come into existence only at a later stage of evolution and have not had the time for diversification and dissemination, or whether the enzymes residing in these pathways are not as suitable for functional diversification as those found on the more commonly distributed pathways. Nonetheless, this provides reassurances that new natural product families with novel biosynthetic logic and a limited phylogenetic distribution are still awaiting discovery.

Our model predicts that the explosion in the numbers of natural products has occurred at a later stage in a process where pre-existing scaffolds and gene clusters are modified to alter the biological activities. We highlight with examples that in many cases single evolutionary events, such as point mutations, recombination or horizontal gene transfer, have been sufficient to convert catalytic activities towards novel chemistry. In particular, recombination and horizontal gene transfer events have been shown to lead to the production of novel secondary metabolite libraries that harbor modulated biological activities [93, 121. It has also become apparent that gene duplication events are not essential to the diversification of natural product biosynthetic pathways. Instead gene duplication facilitates the functional diversification of gene products *in situ*, and that the major role of gene duplication is to allow formation of more complex chemical structures.

Enzyme promiscuity is important in understanding how natural product biosynthetic gene clusters diversify and gain new functions. Both substrate and catalytic promiscuity play key roles in the diversification of natural products by promoting the incorporation of horizontally transferred genes into secondary metabolic pathways, but also for the functional differentiation of natural product enzymes to catalyze novel chemistry. This intrinsic property of natural product enzymes can be exploited in biotechnology, but offers challenges for the rational design of new pathways. Nonetheless, the pervasive substrate promiscuous activities of natural product enzymes provide a major advantage for synthetic biology efforts to construct novel natural product pathways. A better understanding of how natural product biosynthetic pathways evolve should provide a source of inspiration for the rational design and reprogramming of biosynthetic pathways in order to produce new natural product analogs.

Acknowledgements
The authors would like to than financial support from the Academy of Finland (grant no. 285971 to M. M.-K.), the Novo Nordisk Foundation (grant no. 21684 to M.M.-K. and grant no. 34836 to D.P.F.), the Juselius Foundation to M. M.-K., the Jane and Aatos Erkko Foundation to M. M.-K. and from Nordforsk (Nordic Center of Excellency NordAqua project no. 82845 to D. P. F).

References

1. Newman DJ & Cragg GM (2012) Natural products as sources of new drugs over the 30 years from 1981 to 2010. *J Nat Prod* **75**, 311–335.
2. Newman DJ & Cragg GM (2016) Natural products as sources of new drugs from 1981 to 2014. *J Nat Prod* **79**, 629–661.
3. Pye CR, Bertin MJ, Lokey RS, Gerwick WH & Linington RG (2017) Retrospective analysis of natural products provides insights for future discovery trends. *Proc Natl Acad Sci U S A* **114**, 5601–5606.
4. Carlson EE (2010) Natural products as chemical probes. *ACS Chem Biol* **5**, 639–653.
5. Nicolaou KC (2014) Organic synthesis: the art and science of replicating the molecules of living nature and creating others like them in the laboratory. *Proc Math Phys Eng Sci* **470**, 20130690.
6. Barka EA, Vatsa P, Sanchez L, Gaveau-Vaillant N, Jacquard C, Meier-Kolthoff JP, Klenk HP, Clément C, Ouhdouch Y & van Wezel GP (2016) Taxonomy, physiology, and natural products of Actinobacteria. *Microbiol Mol Biol Rev* **80**, 1–43.

7. Colabella F, Moline M & Libkind D (2014) UV Sunscreens of microbial origin: Mycosporines and mycosporine- like amino acids. *Recent Pat Biotechnol* **8**, 179–193.
8. Dittmann E, Fewer DP & Neilan BA (2013) Cyanobacterial toxins: biosynthetic routes and evolutionary roots. *FEMS Microbiol Rev* **37**, 23–43.
9. Carroll CS & Moore MM (2018) Ironing out siderophore biosynthesis: a review of non-ribosomal peptide synthetase (NRPS)-independent siderophore synthetases. *Crit Rev Biochem Mol Biol* **53**, 356–381.
10. Gao Q & Garcia-Pichel F (2011) Microbial ultraviolet sunscreens *Nature Rev Microbiol* **9**, 791–802.
11. Kenny DJ & Balskus EP (2018) Engineering chemical interactions in microbial communities. *Chem Soc Rev* **47**, 1705–1729.
12. Fischbach MA, Walsh CT & Clardy J (2008) The evolution of gene collectives: How natural selection drives chemical innovation. *Proc Natl Acad Sci U S A* **105**, 4601–4608.
13. Malik VS (1980) Microbial secondary metabolism *Trends Biochem Sci* **5**, 68–72.
14. Robbins T, Liu YC, Cane DE, & Khosla C (2016) Structure and mechanism of assembly line polyketide synthases. *Curr Opin Struct Biol* **41**, 10–18.
15. Payne JA, Schoppet M, Hansen MH & Cryle MJ (2016) Diversity of nature's assembly lines - recent discoveries in non-ribosomal peptide synthesis. *Mol Biosyst* **13**, 9–22.
16. Hudson GA & Mitchell DA. (2018) RiPP antibiotics: biosynthesis and engineering potential. *Curr Opin Microbiol* **45**, 61–69.
17. Metsä-Ketelä M, Niemi J, Mäntsälä P & Schneider G (2008) Anthracycline Biosynthesis: Genes, Enzymes and Mechanisms In Anthracycline Chemistry and Biology I: Biological Occurrence and Biosynthesis, Synthesis and Chemistry, (Krohn, K., Ed.) pp. 101–140, Springer-Verlag, Berlin/Heidelberg
18. Rabe P, Kamps JJAG, Schofield CJ & Lohans CT (2018) Roles of 2-oxoglutarate oxygenases and isopenicillin N synthase in β-lactam biosynthesis. *Nat Prod Rep* **35**, 735–756.
19. Herr CQ & Hausinger RP (2018) Amazing diversity in biochemical roles of Fe(II)/2-oxoglutarate oxygenases. *Trends Biochem Sci* **43**, 517–532.
20. Jensen PR (2016) Natural products and the gene cluster revolution. *Trends Microbiol* **24**, 968–977.
21. Walsh CT & Moore BS (2019) Enzymatic cascade reactions in biosynthesis. *Angew Chem Int Ed Engl* **58**, 6846–6879.
22. Walsh CT, Chen H, Keating TA, Hubbard BK, Losey HC, Luo L, Marshall CG, Miller DA, & Patel HM (2001) Tailoring enzymes that modify nonribosomal peptides during and after chain elongation on NRPS assembly lines. *Curr Opin Chem Biol* **5**, 525–534.
23. Cimermancic P, Medema MH, Claesen J, Kurita K, Wieland Brown LC, Mavrommatis K, Pati A, Godfrey PA, Koehrsen M, Clardy J, Birren BW, Takano E, Sali A, Linington RG & Fischbach MA. (2014) Insights into secondary metabolism from a global analysis of prokaryotic biosynthetic gene clusters. *Cell* **158**, 412–421.
24. Wang H, Fewer DP, Holm L, Rouhiainen L & Sivonen K. (2014) Atlas of nonribosomal peptide and polyketide biosynthetic pathways reveals common occurrence of nonmodular enzymes. *Proc Natl Acad Sci U S A* **111**, 9259–9264.
25. Harvey AL, Edrada-Ebel R & Quinn RJ (2015). The re-emergence of natural products for drug discovery in the genomics era. *Nat Rev Drug Discovery* **14**, 111–129.
26. Rutledge PJ & Challis GL (2015) Discovery of microbial natural products by activation of silent biosynthetic gene clusters. *Nat Rev Microbiol* **13**, 509–523.

27. Baral B, Akhgari A & Metsä-Ketelä M (2018) Activation of microbial secondary metabolic pathways: Avenues and challenges. *Synth Syst Biotechnol* **3**, 163–178.

28. Donia MS, Cimermancic P, Schulze CJ, Wieland Brown LC, Martin J, Mitreva M, Clardy J, Linington RG & Fischbach MA (2014). A systematic analysis of biosynthetic gene clusters in the human microbiome reveals a common family of antibiotics. *Cell* **158**, 1402–1414.

29. Mousa WK, Athar B, Merwin NJ & Magarvey NA (2017) Antibiotics and specialized metabolites from the human microbiota. *Nat Prod Rep* **34**, 1302–1331.

30. Milshteyn A, Schneider JS & Brady SF (2014) Mining the metabiome: identifying novel natural products from microbial communities. *Chem Biol* **21**, 1211–1223.

31. Bar-Even A, Noor E, Savir Y, Liebermeister W, Davidi D, Tawfik DS & Milo R (2011) The moderately efficient enzyme: evolutionary and physicochemical trends shaping enzyme parameters. *Biochemistry* **50**, 4402–4410.

32. Li B, Sher D, Kelly L, Shi Y, Huang K, Knerr PJ, Joewono I, Rusch D, Chisholm SW & van der Donk WA (2010) Catalytic promiscuity in the biosynthesis of cyclic peptide secondary metabolites in planktonic marine cyanobacteria. *Proc Natl Acad Sci U S A* **107**, 10430–10435.

33. Patrikainen P, Kallio P, Fan K, Klika KD, Shaaban KA, Mäntsälä P, Rohr J, Yang K, Niemi J & Metsä-Ketelä M (2012) Tailoring enzymes involved in the biosynthesis of angucyclines contain latent context-dependent catalytic activities. *Chem Biol* **19**, 647–655.

34. Seitz M, Syrén PO, Steiner L, Klebensberger J, Nestl BM & Hauer B (2013) Synthesis of heterocyclic terpenoids by promiscuous squalene-hopene cyclases. *ChemBioChem* **14**, 436–439.

35. Zhang Q, Yang X, Wang H & van der Donk WA (2014) High divergence of the precursor peptides in combinatorial lanthipeptide biosynthesis. *ACS Chem Biol* **9**, 2686–2694.

36. Copley SD (2015) An evolutionary biochemist's perspective on promiscuity. *Trends Biochem Sci* **40**, 72-8.

37. Metsä-Ketelä M (2017) Evolution inspired engineering of antibiotic biosynthesis enzymes *Org Biomol Chem* **15,** 4036–4041.

38. Bar-Even A & Tawfik DS (2013) Engineering specialized metabolic pathways - is there a room for enzyme improvements? *Curr Opin Biotechnol* **24**, 310–319.

39. Grocholski T, Dinis P, Niiranen L, Niemi J & Metsä-Ketelä M (2015) Divergent evolution of an atypical S-adenosyl-l-methionine-dependent monooxygenase involved in anthracycline biosynthesis. *Proc Natl Acad Sci U S A* **112**, 9866–9871.

40. Henderson B, Fares, MA & Martin AC (2016) Protein moonlighting in biology and medicine. John Wiley & Sons.

41. Conant GC & Wolfe KH (2008) Turning a hobby into a job: how duplicated genes find new functions. *Nat Rev Genet* **9**, 938–950.

42. Kallio P, Patrikainen P, Suomela JP, Mäntsälä P, Metsä-Ketelä M & Niemi J (2011) Flavoprotein hydroxylase PgaE catalyzes two consecutive oxygen-dependent tailoring reactions in angucycline biosynthesis. *Biochemistry* **50**, 5535–5543.

43. Nakashima Y, Mori T, Nakamura H, Awakawa T, Hoshino S, Senda M, Senda T & Abe I (2018) Structure function and engineering of multifunctional non-heme iron dependent oxygenases in fungal meroterpenoid biosynthesis. *Nat Commun* **9**, 104.

44. Chen JM, Hou, C, Wang G, Tsodikov OV & Rohr J (2015) Structural insight into MtmC, a bifunctional ketoreductase-methyltransferase involved in the assembly of the mithramycin trisaccharide chain. *Biochemistry* **54**, 2481–2489.

45. Firn RD & Jones CG (2003) Natural products – a simple model to explain chemical diversity. *Nat Prod Rep* **20**, 382–391.
46. Firn RD & Jones CG (2000) The evolution of secondary metabolism - a unifying model *Mol Microbiol* **37**, 989–994.
47. Jones CG & Firn RD (1991) On the evolution of plant secondary chemical diversity. *Phil Trans R Soc Lond B* **333**, 273–280.
48. Noda-Garcia L, Liebermeister W & Tawfik DS (2018) Metabolite-Enzyme Coevolution: From Single Enzymes to Metabolic Pathways and Networks. *Annu Rev Biochem* **87**,187–216.
49. Walsh CT & Tang Y (2017) Natural Product Biosynthesis: Chemical Logic and Enzymatic Machinery. Royal Society of Chemistry. ISBN 978-1-78801-076-4.
50. Granick S (1965) The evolution of heme and chlorophyll. In Evolving Genes and Proteins, (Bryson, V., Vogel, H.J. Eds.), pp. 67–88: New York: Academic
51. Ohno S (1970) Evolution by gene duplication, Springer, NY, 1970.
52. Khersonsky O & Tawfik DS (2010) Enzyme promiscuity: a mechanistic and evolutionary perspective. *Annu Rev Biochem* **79**, 471–505.
53. Moghe GD & Last RL (2015) Something old, something new: conserved enzymes and the evolution of novelty in plant specialized metabolism. *Plant Physiol* **169**, 1512–1523.
54. Grocholski T, Yamada K, Sinkkonen J, Tirkkonen H, Niemi J & Metsä-Ketelä M (2019) Evolutionary trajectories for the functional diversification of anthracycline methyltransferases. *ACS Chem Biol,* **14**, 850–856.
55. Jansson A, Koskiniemi H, Mäntsälä P, Niemi J & Schneider G (2004) Crystal structure of a ternary complex of DnrK, a methyltransferase in daunorubicin biosynthesis, with bound products. *J Biol Chem* **24**, 41149–41156.
56. Jansson A, Koskiniemi H, Erola A, Wang J, Mäntsälä P, Schneider G & Niemi J (2005) Aclacinomycin 10-hydroxylase is a novel substrate-assisted hydroxylase requiring S-adenosyl-L-methionine as cofactor. *J Biol Chem* **280**, 3636–3644.
57. Kharel MK, Pahari P, Shepherd MD, Tibrewal N, Nybo SE, Shaaban KA & Rohr J (2012) Angucyclines: Biosynthesis, mode-of-action, new natural products, and synthesis. *Nat Prod Rep* **29**, 264–325.
58. Fan K & Zhang Q (2018) The functional differentiation of the post-PKS tailoring oxygenases contributed to the chemical diversities of atypical angucyclines. *Synth Syst Biotechnol* **3**, 275–282.
59. Koskiniemi H, Metsä-Ketelä M, Dobritzsch D, Kallio P, Korhonen H, Mäntsälä P, Schneider G & Niemi J (2007) Crystal structures of two aromatic hydroxylases involved in the early tailoring steps of angucycline biosynthesis. *J Mol Biol* **372**, 633–648.
60. Chen Y, Fan K, He Y, Xu X, Peng Y, Yu T, Jia C & Yang K (2010) Characterization of JadH as an FAD- and NAD(P)H-dependent bifunctional hydroxylase/dehydrase in jadomycin biosynthesis. *ChemBioChem* **11**, 1055–1060.
61. Kallio P, Patrikainen P, Belogurov GA, Mäntsälä P, Yang K, Niemi J & Metsä-Ketelä M (2013) Tracing the evolution of angucyclinone monooxygenases: structural determinants for C-12b hydroxylation and substrate inhibition in PgaE. *Biochemistry* **52**, 4507–4516.
62. Patrikainen P, Niiranen L, Thapa K, Paananen P, Tähtinen P, Mäntsälä P, Niemi J & Metsä-Ketelä M (2014) Structure-based engineering of angucyclinone 6-ketoreductases. *Chem Biol* **21**, 1381–1391.
63. Paananen P, Patrikainen P, Kallio P, Mäntsälä P, Niemi J, Niiranen L & Metsä-Ketelä M (2013) Structural and functional analysis of angucycline C-6 ketoreductase LanV involved in landomycin biosynthesis. *Biochemistry* **52**, 5304–5314.

64. Luzhetskyy A, Taguchi T, Fedoryshyn M, Dürr C, Wohlert SE, Novikov V & Bechthold A (2005) LanGT2 catalyzes the first glycosylation step during landomycin A biosynthesis. *Chembiochem* **6**, 1406–1410.
65. Dürr C, Hoffmeister D, Wohlert SE, Ichinose K, Weber M, Von Mulert U, Thorson JS & Bechthold A (2004) The glycosyltransferase UrdGT2 catalyzes both C- and O-glycosidic sugar transfers. *Angew Chem Int Ed Engl* **43**, 2962–2965.
66. Mittler M, Bechthold A & Schulz GE (2007) Structure and action of the C-C bond-forming glycosyltransferase UrdGT2 involved in the biosynthesis of the antibiotic urdamycin. *J Mol Biol* **372**, 67–76.
67. Härle J, Günther S, Lauinger B, Weber M, Kammerer B, Zechel DL, Luzhetskyy A & Bechthold A (2011) Rational design of an aryl-C-glycoside catalyst from a natural product O-glycosyltransferase. *Chem Biol* **18**, 520–530.
68. Tam HK, Härle J, Gerhardt S, Rohr J, Wang G, Thorson JS, Bigot A, Lutterbeck M, Seiche W, Breit B, Bechthold A & Einsle O (2015) Structural characterization of O- and C-glycosylating variants of the landomycin glycosyltransferase LanGT2. *Angew Chem Int Ed Engl* **54**, 2811–2815.
69. Dickschat JS (2015) Bacterial terpene cyclases. *Nat Prod Rep* **33**, 87–110.
70. Yonemura Y, Ohyama T & Hoshino T (2012) Chemo-enzymatic syntheses of drimane-type sesquiterpenes and the fundamental core of hongoquercin meroterpenoid by recombinant squalene-hopene cyclase. *Org Biomol Chem* **10**, 440–446.
71. Syrén PO, Hammer SC, Claasen B & Hauer B (2014) Entropy is key to the formation of pentacyclic terpenoids by enzyme-catalyzed polycyclization. *Angew Chem Int Ed Engl* **53**, 4845–4849.
72. Siedenburg G, Breuer M & Jendrossek D (2013) Prokaryotic squalene-hopene cyclases can be converted to citronellal cyclases by single amino acid exchange. *Appl Microbiol Biotechnol* **97**, 1571–1580.
73. Hammer SC, Marjanovic A, Dominicus JM, Nestl BM & Hauer B (2015) Squalene hopene cyclases are protonases for stereoselective Brønsted acid catalysis. *Nat Chem Biol* **11**, 121–126.
74. Matsuda Y, Awakawa T, Wakimoto T & Abe I (2013) Spiro-ring formation is catalyzed by a multifunctional dioxygenase in austinol biosynthesis. *J Am Chem Soc* **135**, 10962–10965.
75. Matsuda Y, Iwabuchi T, Fujimoto T, Awakawa T, Nakashima Y, Mori T, Zhang H, Hayashi F & Abe I (2016) Discovery of key dioxygenases that diverged the paraherquonin and acetoxydehydroaustin pathways in *Penicillium brasilianum*. *J Am Chem Soc* **138**, 12671–12677.
76. Beinker P, Lohkamp B, Peltonen T, Niemi J, Mäntsälä P & Schneider G (2006) Crystal structures of SnoaL2 and AclR: two putative hydroxylases in the biosynthesis of aromatic polyketide antibiotics. *J Mol Biol* **359**, 728–740.
77. Matsuda Y, Bai T, Phippen CB, Nødvig CS, Kjærbølling I, Vesth TC, Andersen MR, Mortensen UH, Gotfredsen CH, Abe I & Larsen TO (2018) Novofumigatonin biosynthesis involves a non-heme iron-dependent endoperoxide isomerase for orthoester formation. *Nat commun* **9**, 2587.
78. Siitonen V, Selvaraj B, Niiranen L, Lindqvist Y, Schneider G & Metsä-Ketelä M (2016) Divergent non-heme iron enzymes in the nogalamycin biosynthetic pathway. *Proc Natl Acad Sci U S A* **113**, 5251–5256.
79. Ycas M (1974) On earlier states of the biochemical system. *J Theor Biol* **44**, 145–160.
80. Jensen RA (1976) Enzyme recruitment in evolution of new function. *Annu Rev Microbiol* **30**, 409–425

81. Siitonen V, Claesson M, Patrikainen P, Aromaa M, Mäntsälä P, Schneider G & Metsä-Ketelä M (2012) Identification of late-stage glycosylation steps in the biosynthetic pathway of the anthracycline nogalamycin. *ChemBioChem* **13**, 120–128.
82. Siitonen V, Blauenburg B, Kallio P, Mäntsälä P & Metsä-Ketelä M (2012) Discovery of a two-component monooxygenase SnoaW/SnoaL2 involved in nogalamycin biosynthesis. *Chem Biol* **19**, 638–646.
83. Zhang Z, Gong Y-K, Zhou Q, Hu Y, Ma H-M, Chen Y-S, Igarashi Y, Pan l & Tang G-L (2017) Hydroxyl regioisomerization of anthracycline catalyzed by a four-enzyme cascade. *Proc Natl Acad Sci U S A* **114**, 1554–1559.
84. Sultana A, Kallio P, Jansson A, Wang JS, Niemi J, Mäntsälä P & Schneider G (2004) Structure of the polyketide cyclase SnoaL reveals a novel mechanism for enzymatic aldol condensation. *EMBO J* **23**, 1911–1921
85. Kallio P, Sultana A, Niemi J, Mäntsälä P & Schneider G (2006) Crystal structure of the polyketide cyclase AknH with bound substrate and product analogue: implications for catalytic mechanism and product stereoselectivity. *J Mol Biol* **357**, 210–220.
86. Claesson M, Siitonen V, Dobritzsch D, Metsä-Ketelä M & Schneider G (2012) Crystal structure of the glycosyltransferase SnogD from the biosynthetic pathway of nogalamycin in *Streptomyces nogalater*. *FEBS J* **279**, 3251–3263.
87. Siitonen V, Nji Wandi B, Törmänen AP & Metsä-Ketelä M (2018) Enzymatic synthesis of the C- glycosidic moiety of nogalamycin R. *ACS Chem Biol* **13**, 2433–2437.
88. Meriluoto J, Spoof L & Codd G (2017) Handbook of cyanobacterial monitoring and cyanotoxin analysis; Publisher: John Wiley & Sons, Ltd, Chichester ISBN: 978-1-119-06868-6.
89. Dittmann E, Gugger M, Sivonen K & Fewer DP (2015) Natural product biosynthetic diversity and comparative genomics of the cyanobacteria. *Trends Microbiol* **23**, 642–652.
90. Fewer DP, Rouhiainen L, Jokela J, Wahlsten M, Laakso K, Wang H & Sivonen K (2007) Recurrent adenylation domain replacement in the microcystin synthetase gene cluster. *BMC Evol Biol* **7**, 183.
91. Meyer S, Kehr JC, Mainz A, Dehm D, Petras D, Süssmuth RD & Dittmann E (2016) Biochemical dissection of the natural diversification of microcystin provides lessons for synthetic biology of NRPS. *Cell Chem Biol* **23**, 462–471.
92. Chemler JA, Tripathi A, Hansen DA, O'Neil-Johnson M, Williams RB, Starks C, Park SR, Sherman DH (2015) Evolution of efficient modular polyketide synthases by homologous recombination. *J Am Chem Soc* **137**, 10603–10609.
93. Wlodek A, Kendrew SG, Coates NJ, Hold A, Pogwizd J, Rudder S, Sheehan LS, Higginbotham SJ, Stanley-Smith AE, Warneck T, Nur-E-Alam M, Radzom M, Martin CJ, Overvoorde L, Samborskyy M, Alt S, Heine D, Carter GT, Graziani EI, Koehn FE, McDonald L, Alanine A, Rodríguez Sarmiento RM, Chao SK, Ratni H, Steward L, Norville IH, Sarkar-Tyson M, Moss SJ, Leadlay PF, Wilkinson B & Gregory MA (2017) Diversity oriented biosynthesis via accelerated evolution of modular gene clusters. *Nat Commun* **8**, 1206.
94. Park ST, Aldape RA, Futer O, DeCenzo MT & Livingston DJ (1992) PPIase catalysis by human FK506-binding protein proceeds through a conformational twist mechanism. *Proc Natl Acad Sci USA* **267**, 3316–3324.
95. Laplante M & Sabatini DM (2012) mTOR signalling in growth control and disease. *Cell* **149**, 274–293.
96. Traitcheva N, Jenke-Kodama H, He J, Dittmann E & Hertweck C (2007) Non-colinear polyketide biosynthesis in the aureothin and neoaureothin pathways: an evolutionary perspective. *Chembiochem* **8**, 1841–1849.

97. Peng H, Ishida K, Sugimoto Y, Jenke-Kodama H & Hertweck C (2019) Emulating evolutionary processes to morph aureothin-type modular polyketide synthases and associated oxygenases. *Nat Commun* **10**, 3918.
98. Awakawa T, Fujioka T, Zhang L, Hoshino S, Hu Z, Hashimoto J, Kozone I, Ikeda H, Shin-Ya K, Liu W & Abe I (2018) Reprogramming of the antimycin NRPS-PKS assembly lines inspired by gene evolution. *Nat Commun* **9**, 3534.
99. Ruzzini AC & Clardy J (2016) Gene flow and molecular innovation in bacteria. *Curr Biol* **26**, R859–R864.
100. Kinashi H, Shimaji M & Sakai A (1987) Giant linear plasmids in *Streptomyces* which code for antibiotic biosynthesis genes. *Nature* **328**, 454–456.
101. Zhang L, Hoshino S, Awakawa T, Wakimoto T & Abe I (2016) Structural diversification of lyngbyatoxin A by host-dependent heterologous expression of the *tleABC* biosynthetic gene cluster. *Chembiochem* **17**, 1407–1411.
102. Humisto A, Jokela J, Liu L, Wahlsten M, Wang H, Permi P, Machado JP, Antunes A, Fewer DP & Sivonen K (2017). The swinholide biosynthetic gene cluster from a terrestrial cyanobacterium *Nostoc* sp. UHCC 0450. *Appl Environ Microbiol* **84**, e02321-17.2017
103. Liras P, Rodríguez-García A & Martín JF (1998) Evolution of the clusters of genes for beta-lactam antibiotics: a model for evolutive combinatorial assembly of new beta-lactams. *Int Microbiol* **1**, 271–278.
104. Orr RJ, Stüken A, Murray SA & Jakobsen KS (2013) Evolution and distribution of saxitoxin biosynthesis in dinoflagellates. *Mar Drugs* **11**, 2814–2828.
105. Fewer DP, Wahlsten M, Österholm J, Jokela J, Rouhiainen L, Kaasalainen U, Rikkinen J & Sivonen K (2013) The genetic basis for O-acetylation of the microcystin toxin in cyanobacteria. *Chem Biol* **20**, 861–869.
106. Murray IA & Shaw WV (1997) O-acetyltransferases for chloramphenicol and other natural products. *Antimicrob Agents Chemother* **41**, 1–6.
107. Stogios PJ, Kuhn ML, Evdokimova E, Courvalin P, Anderson WF & Savchenko A (2014) Potential for reduction of streptogramin A resistance revealed by structural analysis of acetyltransferase VatA. *Antimicrob Agents Chemother* **58**, 7083–7092.
108. Biswas T, Houghton JL, Garneau-Tsodikova S & Tsodikov OV (2012) The structural basis for substrate versatility of chloramphenicol acetyltransferase CATI. *Protein Sci* **21**, 520–530.
109. Mosa A, Hutter MC, Zapp J, Bernhardt R & Hannemann F (2015) Regioselective acetylation of C21 hydroxysteroids by the bacterial chloramphenicol acetyltransferase I. *Chembiochem* **16**, 1670–1679.
110. Tello M, Kuzuyama T, Heide L, Noel JP & Richard SB (2008) The ABBA family of aromatic prenyltransferases: broadening natural product diversity. *Cell Mol Life Sci* **65**, 1459–1463.
111. Kuzuyama T, Noel JP & Richard SB (2005) Structural basis for the promiscuous biosynthetic prenylation of aromatic natural products. *Nature* **435**, 983–987.
112. Bonitz T, Alva V, Saleh O, Lupas AN & Heide L (2011) Evolutionary relationships of microbial aromatic prenyltransferases. *PLoS One* **6**, e27336.
113. Tanner ME (2015) Mechanistic studies on the indole prenyltransferases. *Nat Prod Rep* **32**, 88–101.
114. Metzger U, Schall C, Zocher G, Unsöld I, Stec E, Li SM, Heide L & Stehle T (2009) The structure of dimethylallyl tryptophan synthase reveals a common architecture of aromatic prenyltransferases in fungi and bacteria. *Proc Natl Acad Sci U S A* **106**, 14309–14314.

115. Chen R, Gao B, Liu X, Ruan F, Zhang Y, Lou J, Feng K, Wunsch C, Li SM, Dai J & Sun F (2017) Molecular insights into the enzyme promiscuity of an aromatic prenyltransferase. *Nat Chem Biol* **13**, 226–234.
116. Mori T, Zhang L, Awakawa T, Hoshino S, Okada M, Morita H & Abe I. (2016) Manipulation of prenylation reactions by structure-based engineering of bacterial indolactam prenyltransferases. *Nat Commun* **7**, 10849.
117. Luk LY & Tanner ME (2009) Mechanism of dimethylallyltryptophan synthase: Evidence for a dimethylallyl cation intermediate in an aromatic prenyltransferase reaction. *J Am Chem Soc* **131**, 13932–13933.
118. Grundmann A, Kuznetsova T, Afiyatullov S Sh & Li SM (2008) FtmPT2, an N-prenyltransferase from *Aspergillus fumigatus*, catalyses the last step in the biosynthesis of fumitremorgin B. *ChemBioChem* **9**, 2059–2063.
119. Takahashi S, Takagi H, Toyoda A, Uramoto M, Nogawa T, Ueki M, Sakaki Y & Osada H (2010) Biochemical characterization of a novel indole prenyltransferase from *Streptomyces* sp. SN-593. *J Bacteriol* **192**, 2839–2851.
120. Rudolf JD, Wang H & Poulter CD (2013) Multisite prenylation of 4-substituted tryptophans by dimethylallyltryptophan synthase. *J Am Chem Soc* **135**, 1895–1902.
121. Elshahawi SI, Cao H, Shaaban KA, Ponomareva LV, Subramanian T, Farman ML, Spielmann HP, Phillips GN Jr, Thorson JS & Singh S (2017) Structure and specificity of a permissive bacterial C-prenyltransferase. *Nat Chem Biol* **13**, 366–368.
122. Dalponte L, Parajuli A, Younger E, Mattila A, Jokela J, Wahlsten M, Leikoski N, Sivonen K, Jarmusch SA, Houssen WE & Fewer DP (2018) N-Prenylation of tryptophan by an aromatic prenyltransferase from the cyanobactin biosynthetic pathway. *Biochemistry* **57**, 6860–6867.
123. Parajuli A, Kwak DH, Dalponte L, Leikoski N, Galica T, Umeobika U, Trembleau L, Bent A, Sivonen K, Wahlsten M, Wang H, Rizzi E, De Bellis G, Naismith J, Jaspars M, Liu X, Houssen W & Fewer DP (2016) A unique tryptophan C-prenyltransferase from the kawaguchipeptin biosynthetic pathway. *Angew Chem Int Ed Engl* **55**, 3596–3599.
124. Heide L (2009) Prenyl transfer to aromatic substrates: genetics and enzymology. *Curr Opin Chem Biol* **13**, 171–179.
125. Gallagher KA & Jensen PR (2015) Genomic insights into the evolution of hybrid isoprenoid biosynthetic gene clusters in the MAR4 marine streptomycete clade. *BMC Genomics* **16**, 960.
126. Qi X, Bakht S, Leggett M, Maxwell C, Melton R & Osbourn A (2004) A gene cluster for secondary metabolism in oat: implications for the evolution of metabolic diversity in plants. *Proc Natl Acad Sci U S A* **101**, 8233–8238.
127. Sieber SA & Marahiel MA (2005) Molecular mechanisms underlying nonribosomal peptide synthesis: approaches to new antibiotics. *Chem Rev* **105**, 715–738.
128. Lazos O, Tosin M, Slusarczyk AL, Boakes S, Cortés J, Sidebottom PJ & Leadlay PF (2010) Biosynthesis of the putative siderophore erythrochelin requires unprecedented crosstalk between separate nonribosomal peptide gene clusters. *Chem Biol* **17**, 160–173.
129. Liu L, Budnjo A, Jokela J, Haug BE, Fewer DP, Wahlsten M, Rouhiainen L, Permi P, Fossen T, & Sivonen K (2015). Pseudoaeruginosins, nonribosomal peptides in *Nodularia spumigena*. *ACS Chem Biol* **10**, 725–733.
130. Shishido TK, Jokela J, Fewer DP, Wahlsten M, Fiore MF & Sivonen K (2017) Simultaneous production of anabaenopeptins and namalides by the cyanobacterium *Nostoc* sp. CENA543. *ACS Chem Biol* **12**, 2746–2755.
131. Palmu K, Rosenqvist P, Thapa K, Ilina Y, Siitonen V, Baral B, Mäkinen J, Belogurov G, Virta P, Niemi J & Metsä-Ketelä M (2017) Discovery of the showdomycin gene cluster

from *Streptomyces showdoensis* ATCC 15227 yields insight into the biosynthetic logic of C- nucleoside antibiotics. *ACS Chem Biol* **12**, 1472–1477.
132. Sosio M, Gaspari E, Iorio M, Pessina S, Medema MH, Bernasconi A, Simone M, Maffioli SI, Ebright RH & Donadio S (2018) Analysis of the pseudouridimycin biosynthetic pathway provides insights into the formation of C-nucleoside antibiotics. *Cell Chem Biol* **25**, 540–549.
133. Wang SA, Ko Y, Zeng J, Geng Y, Ren D, Ogasawara Y, Irani S, Zhang YJ & Liu HW (2019) Identification of the formycin A biosynthetic gene cluster from *Streptomyces kaniharaensis* illustrates the interplay between biological pyrazolopyrimidine formation and de novo purine biosynthesis. *J Am Chem Soc* **141**, 6127–6131.

Figure legends

Figure 1. A Pharmaceutical model for evolution of secondary metabolites. Inner ring, the majority of natural products are derived from primary metabolites such as malonyl-CoA, amino acids and glucose-1-phosphate. Middle ring, the Screening Hypothesis represents the emergence of the various classes of biologically active natural products. This requires the organization of gene collectives to form novel biosynthetic pathways and has therefore occurred slowly. Outer ring, the Lead Optimization Hypothesis dictates that once a bioactive compound has arisen, chemical derivatives to produce compound families can evolve quickly. Novel metabolites have appeared in some instances through single evolutionary events such as point mutations, recombination or horizontal gene transfer.

Figure 2. Evolution of enzyme function without gene duplication in the biosynthesis of anthracycline SAM-dependent methyltransferases. (A) The 4-O-methyltransferase DnrK harbors hidden 10-carboxylation moonlighting activity. Formation of a carbanion intermediate in the 10-carboxylation reaction has allowed development of novel chemistry through substrate-activated reaction with molecular oxygen leading to 10-hydroxylation by RdmB. The loss of 4-O-methylation functionality has left 10-decarboxylation the sole enzymatic activity of EamK. Legend: R_1= -OH or l-rhodosamine or l-rhodosamine-2-deoxyfucose-cinerulose. R_2 = thiol reducing agent such as glutathione. (B) Phylogenetic trees demonstrating the co-evolution of the methyltransferases (*left*) with proteins involved in formation of the common core anthracycline scaffold (*right*). The evolution of anthracyclines is represented by concatenated sequences of ketosynthase α-subunits, 9-ketoreductases, first ring aromatases, and 15-methyl esterases. (C) Close-up views of the active sites of DnrK, DnrK-Ser and RdmB. DnrK contains a channel to the surface of the protein lined by Q295 that enables water molecules to protonate the carbanion intermediate. In contrast, this channel is blocked by F296 in DnrK-Ser and by F300 in RdmB.

Figure 3. Enzyme promiscuity and the diversification of angucycline biosynthetic pathways. (A) The multifunctional flavoenzymes JadH, PgaE and LanE all use prejadomycin as a substrate, but catalyze distinct chemistry. In PgaE, the ability to catalyze 12b-hydroxylation has emerged through mutations at the dimerization interface at P78 and I79, which has resulted in appearance of substrate inhibition in the 12-hydroxylation reaction. (B) A single point mutation in the 6-ketoreductases LanV was sufficient to convert 30% of the enzymatic

activity towards UrdMred-type chemistry. (C) Modification of the α3-helix in the glycosyltransferase LanG according to the UrdGT2 template repositions I58 and I62 towards the active site, leading to the conversion of an O-gycosyltransferase into a C-glycosyltransferase.

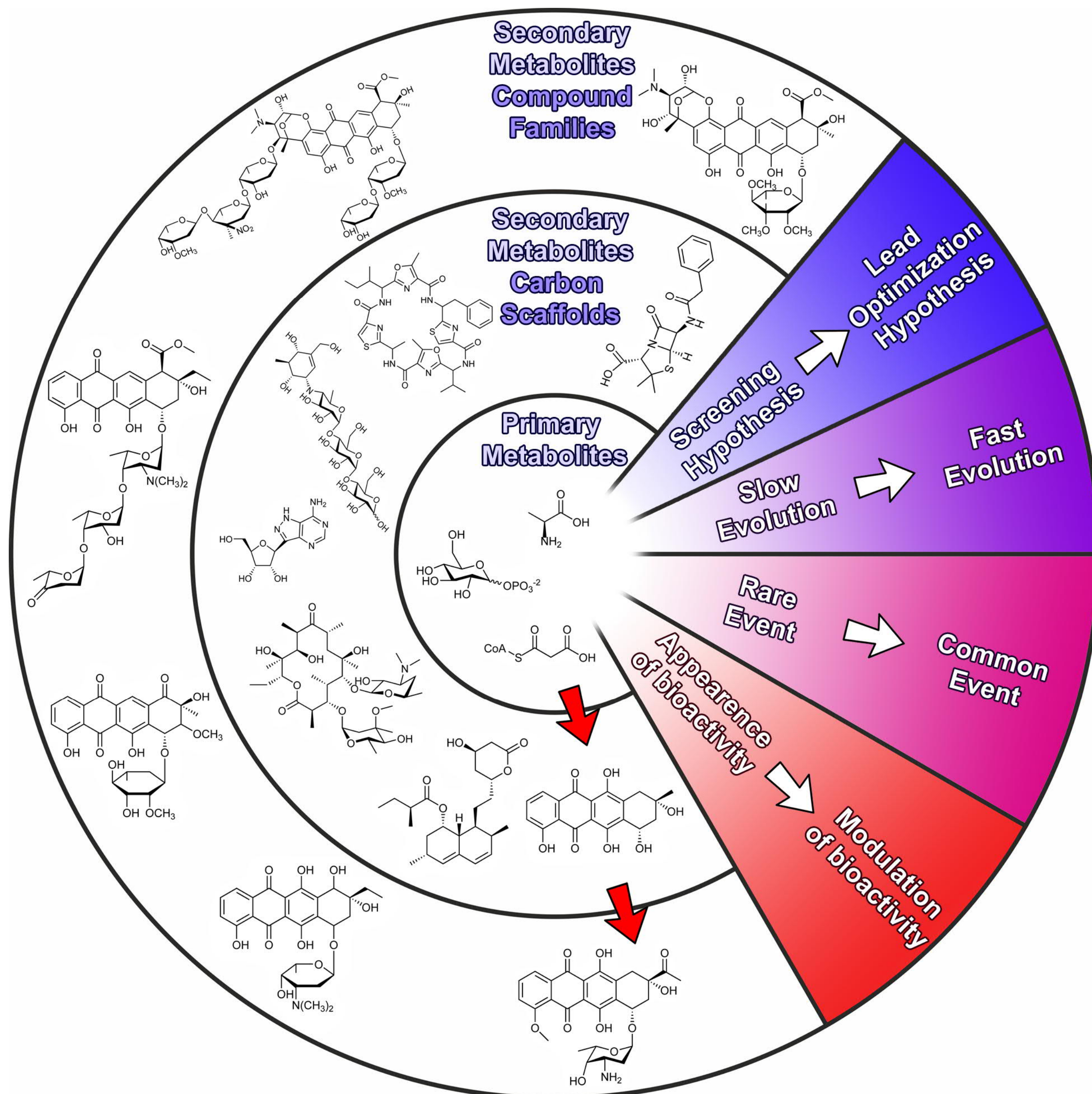
Figure 4. Evolution and rearrangement of terpenoid scaffolds. (A) The hydrophobic binding pockets of squalene hopane cyclases such as *Aac*SHC are able to fold polyprenoids resulting in distinct outcomes. Single point mutations (e.g. F385C, G600F or Y420W) have been shown to be sufficient to alter substrate specificities and cyclization patterns. (B) The non-heme iron and 2OG dependent AusE and PrhA catalyze distinct skeletal rearrangement reactions of meroterpenoids. A few point mutations, influencing positioning of the A-ring in front of the high-valent iron-oxo center, are sufficient to both interconvert the enzymatic activities and for novel chemistry to emerge. In these examples, formation of novel products is facilitated by (A) the highly reactive nature of the substrates and (B) the potency of the catalytic machineries of the enzymes.

Figure 5. Evolution of complex chemical structures through gene duplication and functional diversification. (A) The biosynthetic pathways of anthracycline antibiotics nogalamycin and aclacinomycin contain conserved fourth ring cyclases SnoaL and AknH, respectively. An atypical two-component mono-oxygenase system has emerged after gene duplication and recruitment of the cyclase-like SnoaL2 and AclR to function together with the quinone reductases SnoaW and AclQ, respectively. In nogalamycin biosynthesis, another gene duplication event has occurred at a later stage, which has led to conversion of the non-heme iron and 2OG dependent carbocyclase SnoK into the epimerase SnoN. (B) Three functionally distinct proteins belonging to the non-heme iron and 2OG dependent enzyme family, NvFI, NvfE and NVfF, are involved in complex oxidative rearrangement processes in the biosynthesis of novofumigatonin. In contrast to other members of the family, NvfE does not utilize 2OG as a cosubstrate and relies solely on the non-heme iron for the orthoesterification reaction.

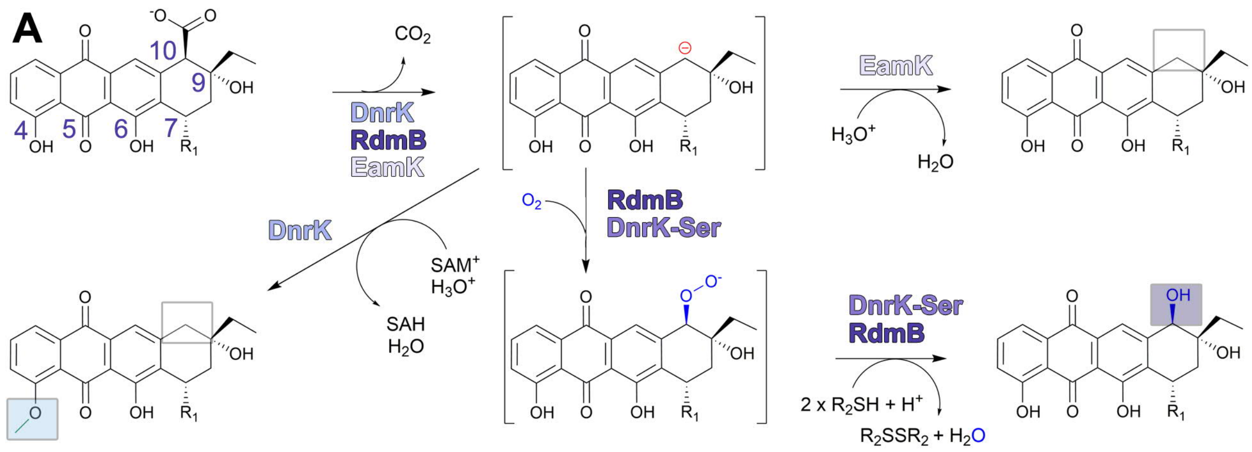
Figure 6. Biosynthesis of natural products through modular assembly lines allows the emergence of novel metabolites by homologous recombination. (A) The biosynthetic logic for microcystin formation. A series of recombination events leading to replacement of the A domains of the McyB1 and McyC modules in different strains of cyanobacteria is responsible for the bulk of the observed chemical variation in the microcystin family of hepatotoxins. Chemical structures of closely related secondary metabolites (B) pikromycin and erythromycin, (C) rapamycin and a novel analog with altered biological activity, (D) aureothin and neoaureothin, and (E) antimycin and JBIR-06. R_1 and R_2 various alkyl groups. Abbreviations: acyl carrier protein (●), ketoreductase domain (KR), C-methyltransferase domain (CM), dehydratase domain (DH), acyltransferase domain (AT), ketosynthase domain (KS), aminotransferase domain (AMT), adenylation domain (A), N-methyltransferase domain (M), codensation domain (C) peptidyl carrier protein (T), thioesterase domain (TE).

Figure 7. Enzyme promiscuity and gain of new functionality in natural product biosynthetic gene clusters. (A-B) ABBA prenyltransferases catalyze the isoprenylation of a variety of aromatic natural products and can be divided into a number of families widely distributed in natural product biosynthetic gene clusters of a variety of bacteria and fungi (C) ABBA prenyltransferases display a distinctively barrel shape fold (e.g. PriB and FgaPT2), where the

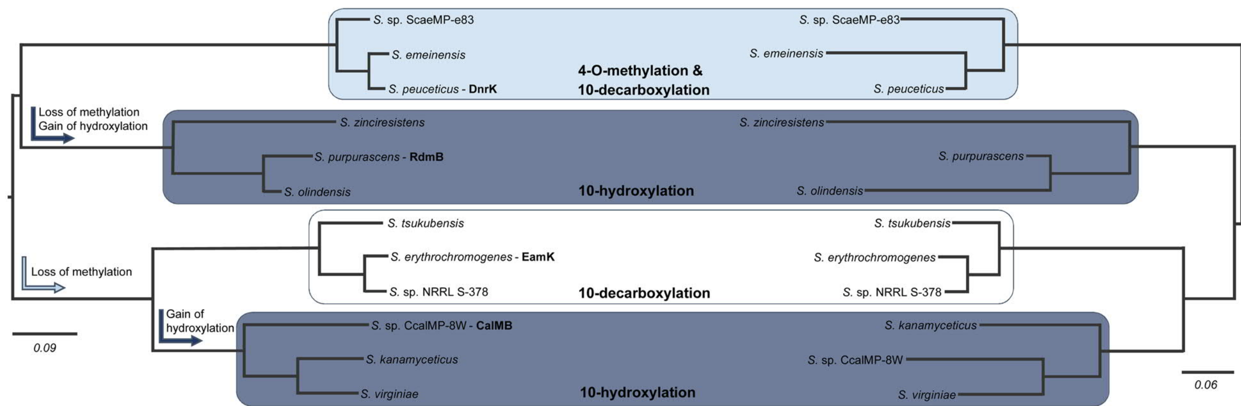
substrate binding site is lined with β-sheets. Daptomycin is a non-ribosomal peptide antibiotic that can be prenylated by the unrelated permissive PriB and FgaPT2 enzymes *in vitro*.



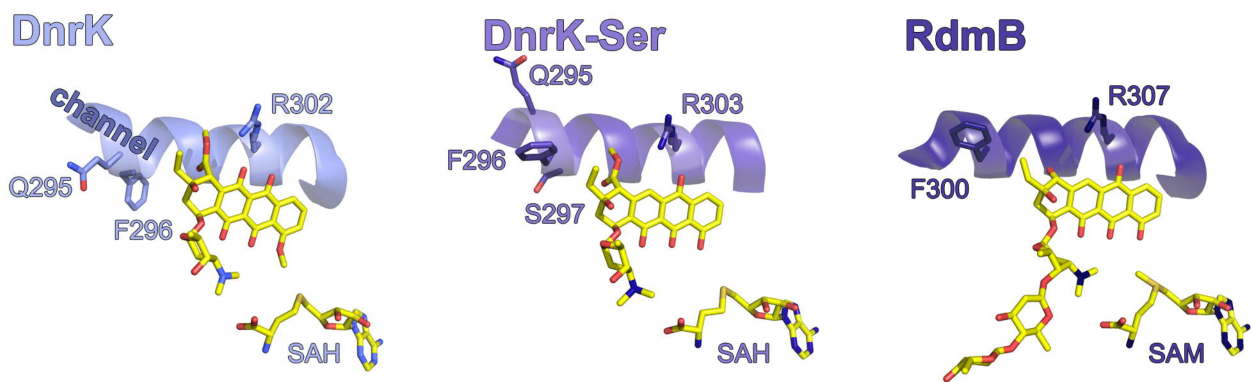
Secondary Metabolites Compound Families
Secondary Metabolites Carbon Scaffolds
Primary Metabolites
Screening Hypothesis
Lead Optimization Hypothesis
Slow Evolution
Fast Evolution
Rare Event
Common Event
Appearence of bioactivity
Modulation of bioactivity



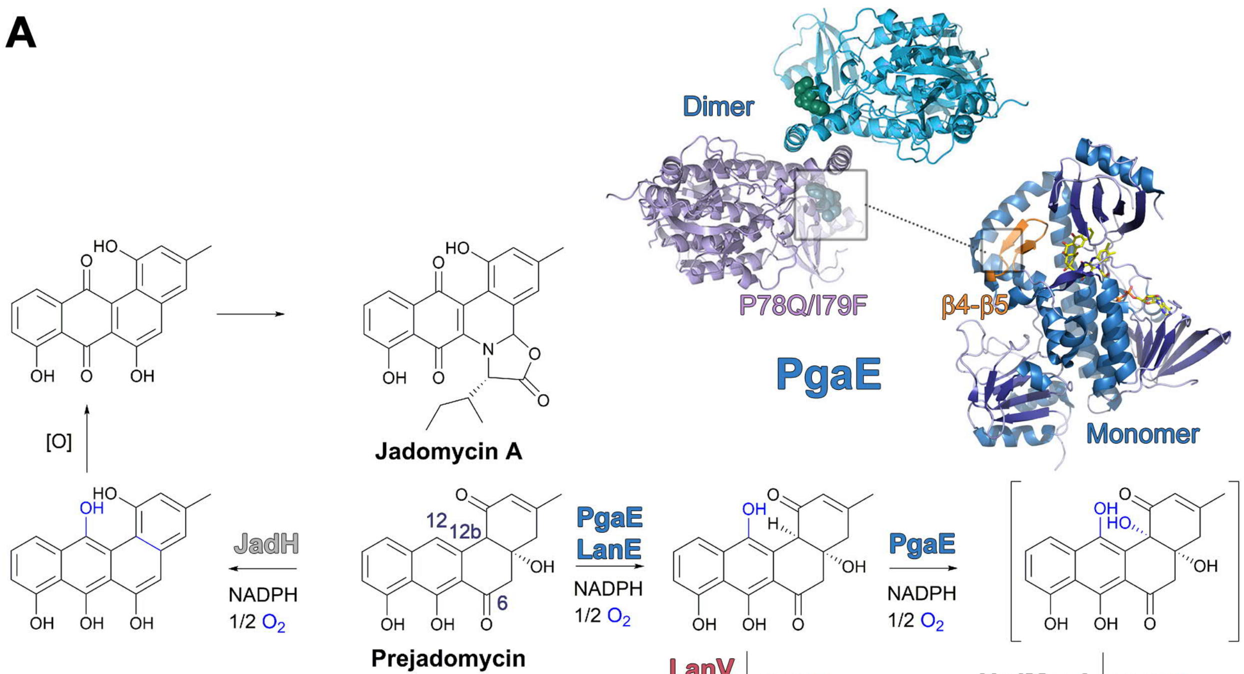
A
10
9
4
5
6
7
OH
O
R_1
CO_2
DnrK
RdmB
EamK
EamK
H_3O^+
H_2O
O_2
RdmB
DnrK-Ser
DnrK
SAM^+
H_3O^+
SAH
H_2O
DnrK-Ser
RdmB
2 x R_2SH + H^+
R_2SSR_2 + H_2O



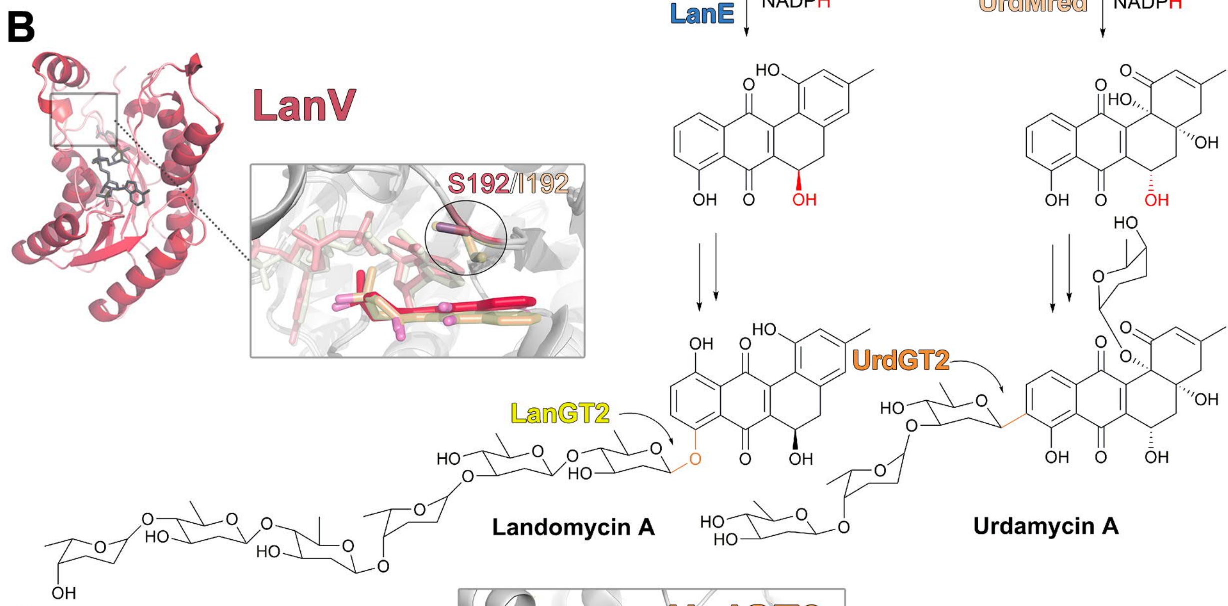
B
S. sp. ScaeMP-e83
S. emeinensis
S. peuceticus - DnrK
4-O-methylation &
10-decarboxylation
S. peuceticus
Loss of methylation
Gain of hydroxylation
S. zincìresistens
S. purpurascens - RdmB
S. olindensis
10-hydroxylation
S. purpurascens
S. tsukubensis
S. erythrochromogenes - EamK
S. sp. NRRL S-378
10-decarboxylation
S. erythrochromogenes
Loss of methylation
Gain of hydroxylation
S. sp. CcalMP-8W - CalMB
S. kanamyceticus
S. virginiae
10-hydroxylation
S. sp. CcalMP-8W
0.09
0.06



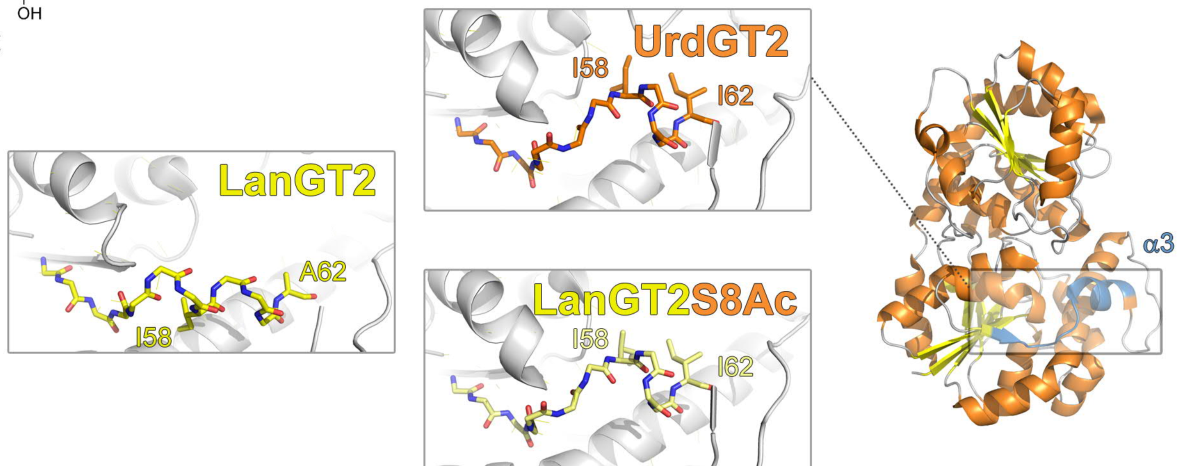
C
DnrK
channel
R302
Q295
F296
SAH
DnrK-Ser
Q295
R303
F296
S297
SAH
RdmB
R307
F300
SAM



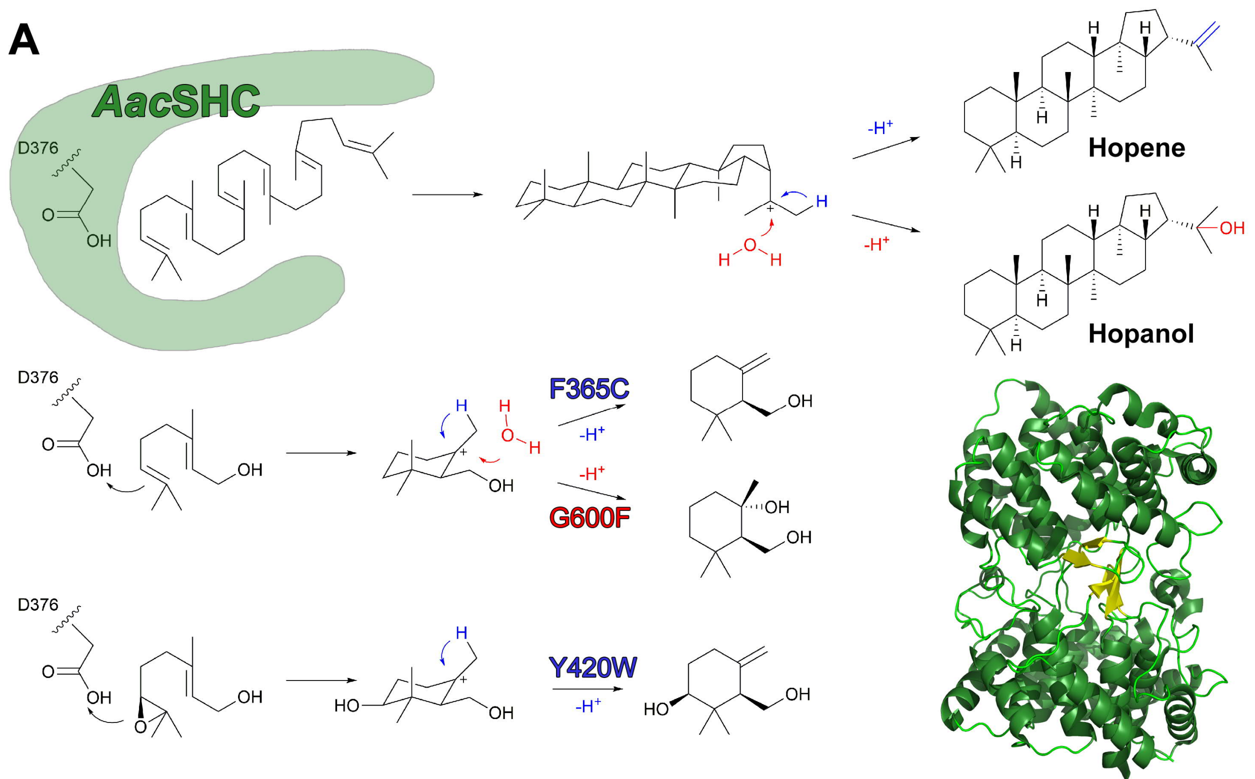
A
Dimer
P78Q/I79F
β4-β5
PgaE
Monomer
Jadomycin A
[O]
JadH
NADPH
1/2 O2
Prejadomycin
PgaE
LanE
NADPH
1/2 O2
PgaE
NADPH
1/2 O2

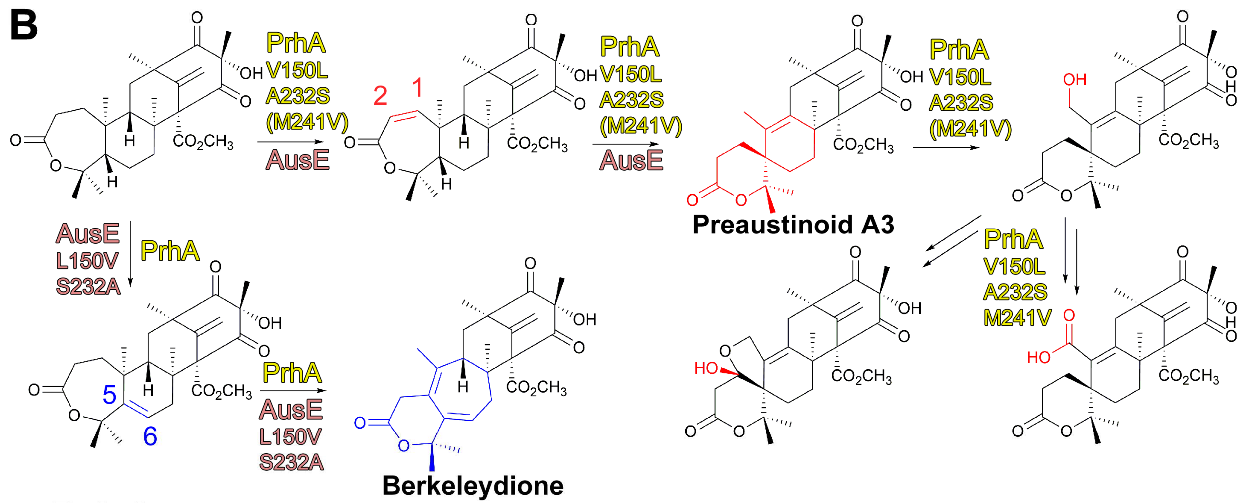


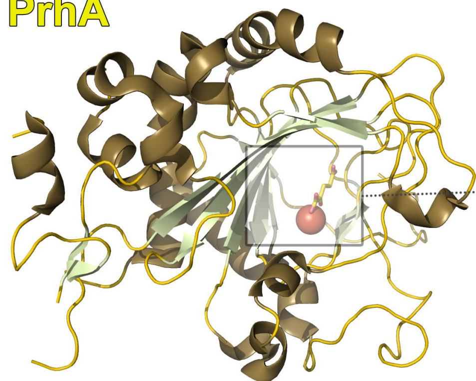
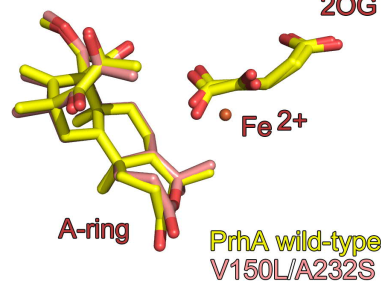
B
LanV
S192/I192
LanV
LanE
NADPH
UrdMred
NADPH
LanGT2
UrdGT2
Landomycin A
Urdamycin A

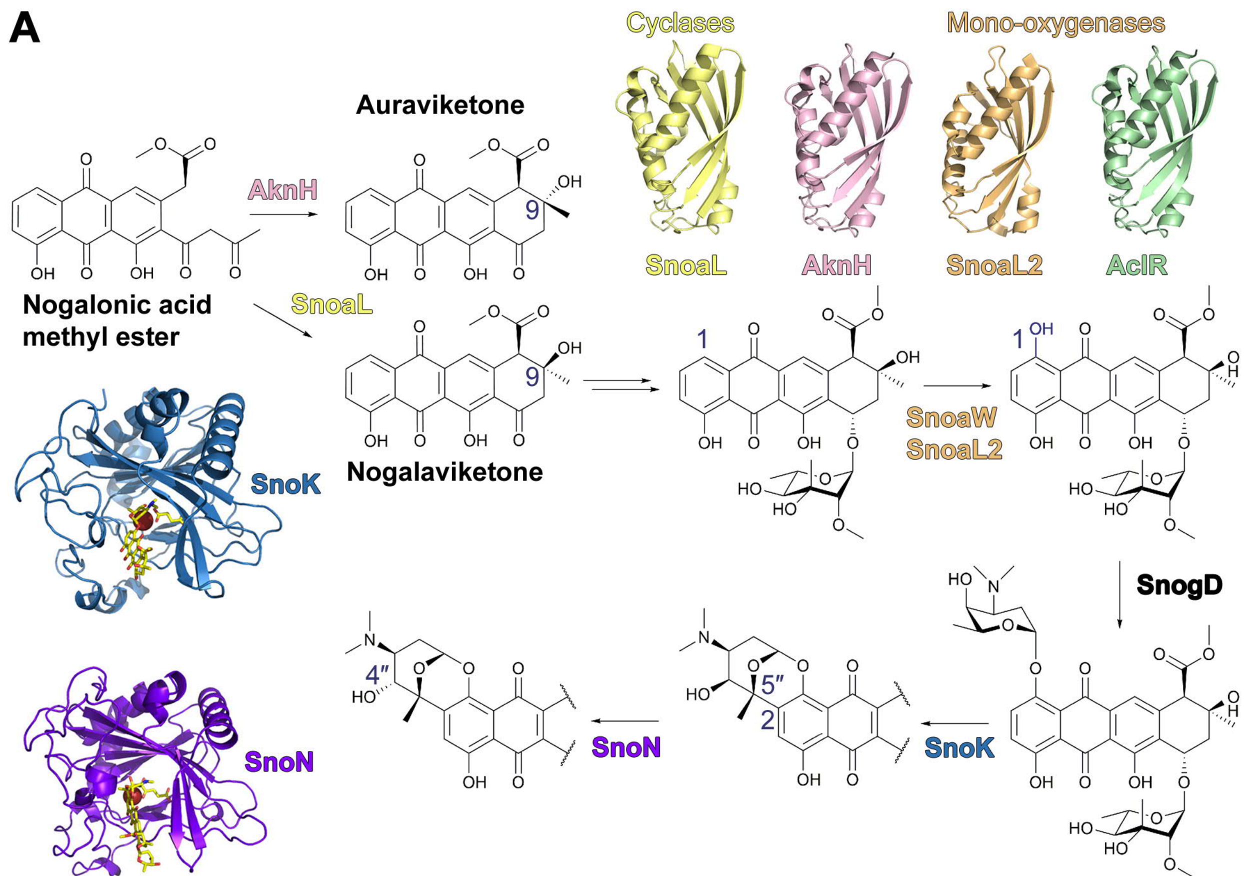


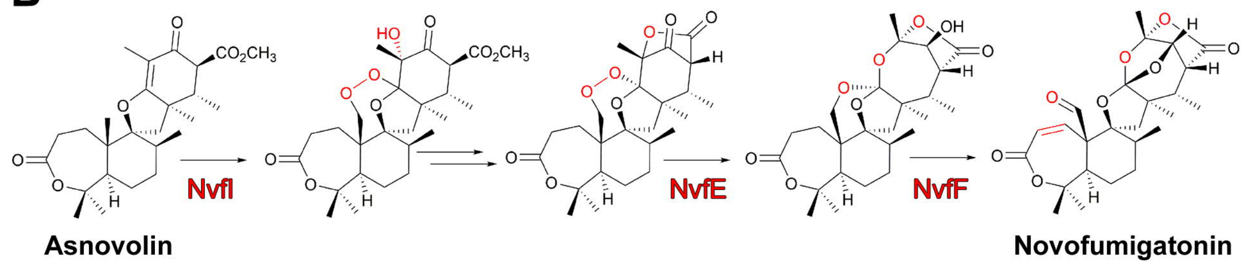
C
UrdGT2
I58
I62
LanGT2
A62
I58
LanGT2S8Ac
I58
I62
α3

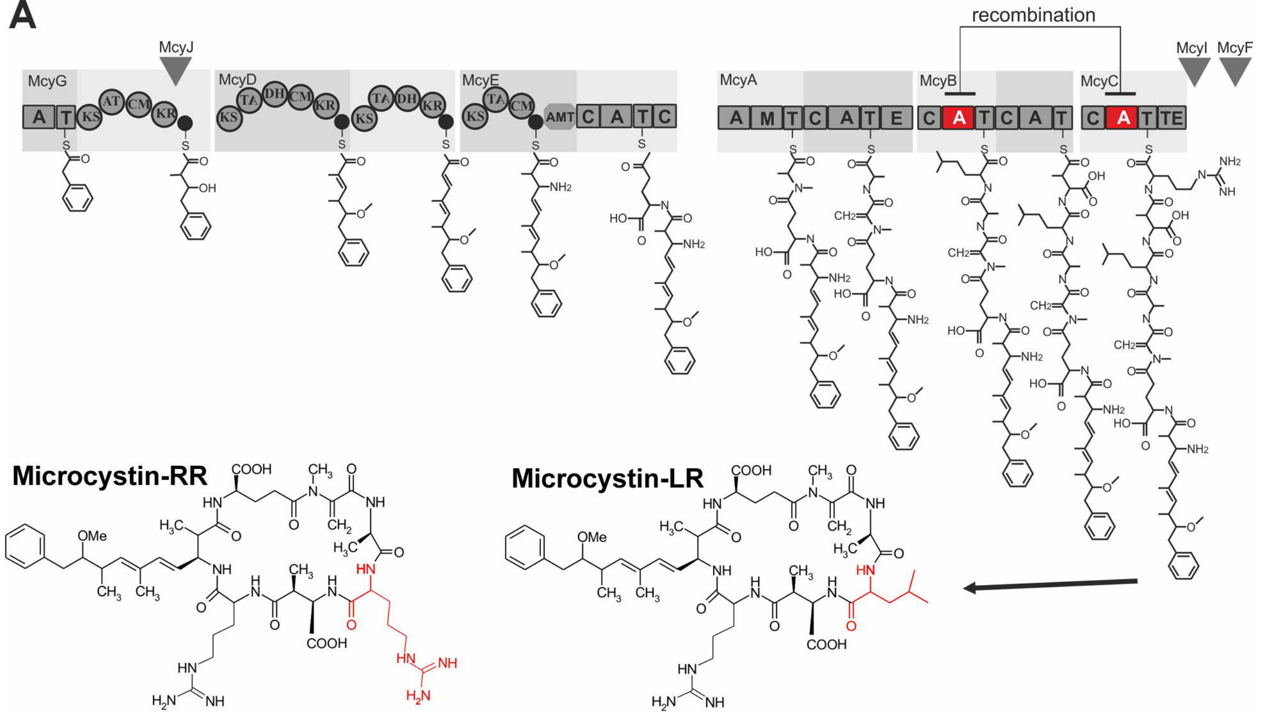

A
AacSHC
D376
-H+
Hopene
Hopanol
F365C
G600F
Y420W

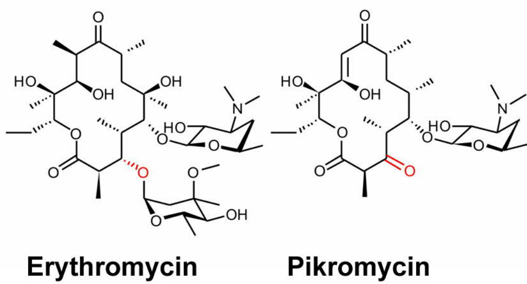

B
PrhA
V150L
A232S
(M241V)
AusE
Preaustinoid A3
L150V
S232A
M241V
Berkeleydione

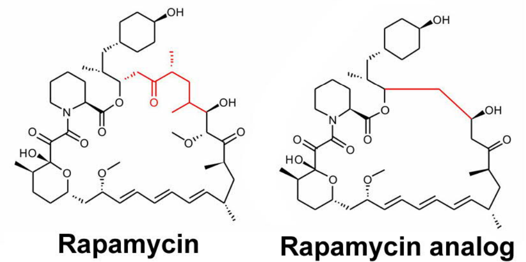

PrhA

Preaustinoid A1
2OG
Fe 2+
A-ring
PrhA wild-type
V150L/A232S

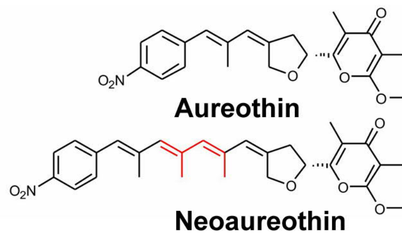

A
Cyclases
Mono-oxygenases
Auraviketone
AknH
SnoaL
AknH
SnoaL2
AclR
Nogalonic acid methyl ester
SnoaL
Nogalaviketone
SnoaW
SnoaL2
SnoK
SnogD
SnoN
SnoK
SnoN

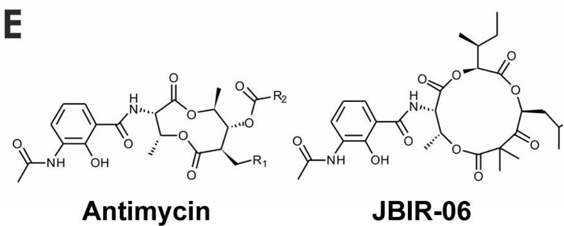

B
NvfI
NvfE
NvfF
Asnovolin
Novofumigatonin

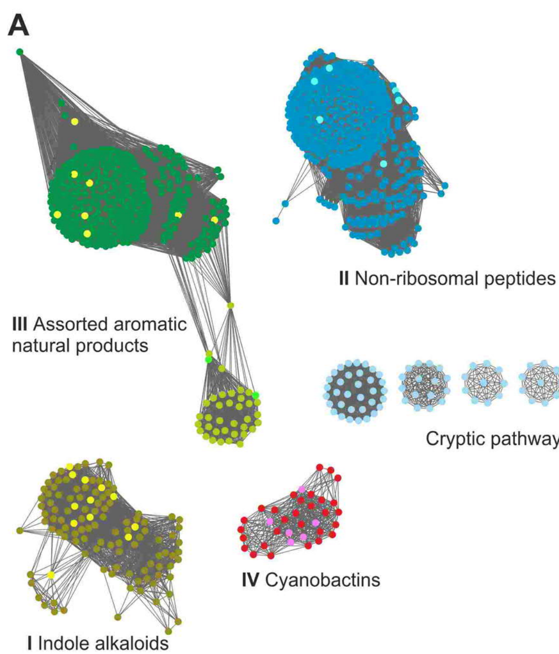

A
recombination
McyJ
McyI
McyF
McyG
McyD
McyE
McyA
McyB
McyC
A T KS AT CM KR
KS TA DH CM KR KS TA DH KR
KS TA CM AMT C A T C
A M T C A T E
C A T C A T
C A T TE
Microcystin-RR
Microcystin-LR

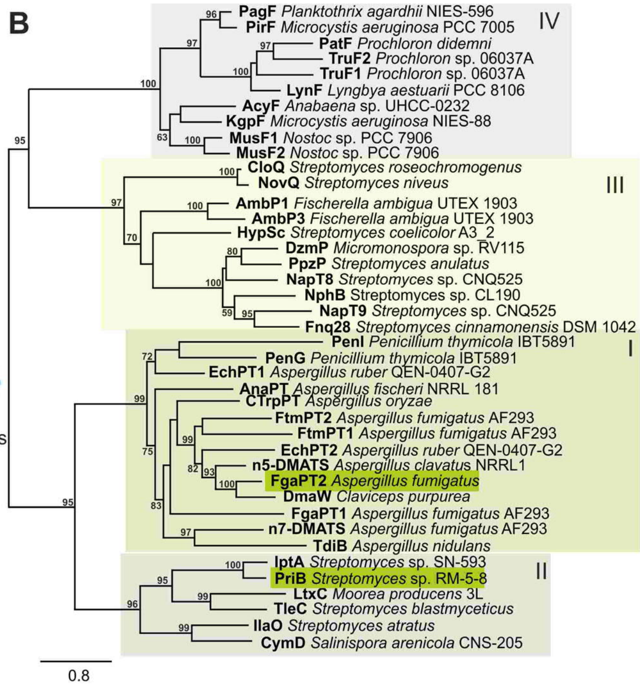

B
Erythromycin
Pikromycin

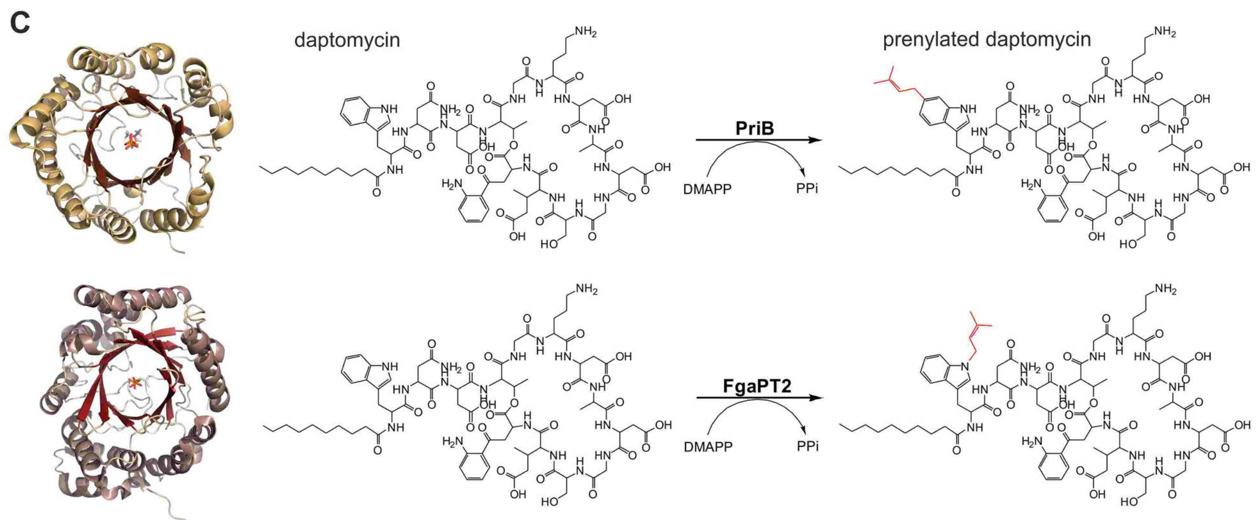

C
Rapamycin
Rapamycin analog


D
Aureothin
Neoaureothin


E
Antimycin
JBIR-06


A
II Non-ribosomal peptides
III Assorted aromatic natural products
Cryptic pathways
I Indole alkaloids
IV Cyanobactins


B
IV
PagF Planktothrix agardhii NIES-596
PirF Microcystis aeruginosa PCC 7005
PatF Prochloron didemni
TruF2 Prochloron sp. 06037A
TruF1 Prochloron sp. 06037A
LynF Lyngbya aestuarii PCC 8106
AcyF Anabaena sp. UHCC-0232
KgpF Microcystis aeruginosa NIES-88
MusF1 Nostoc sp. PCC 7906
MusF2 Nostoc sp. PCC 7906
III
CloQ Streptomyces roseochromogenus
NovQ Streptomyces niveus
AmbP1 Fischerella ambigua UTEX 1903
AmbP3 Fischerella ambigua UTEX 1903
HypSc Streptomyces coelicolor A3_2
DzmP Micromonospora sp. RV115
PpzP Streptomyces anulatus
NapT8 Streptomyces sp. CNQ525
NphB Streptomyces sp. CL190
NapT9 Streptomyces sp. CNQ525
Fnq28 Streptomyces cinnamonensis DSM 1042
I
PenI Penicillium thymicola IBT5891
PenG Penicillium thymicola IBT5891
EchPT1 Aspergillus ruber QEN-0407-G2
AnaPT Aspergillus fischeri NRRL 181
CTrpPT Aspergillus oryzae
FtmPT2 Aspergillus fumigatus AF293
FtmPT1 Aspergillus fumigatus AF293
EchPT2 Aspergillus ruber QEN-0407-G2
n5-DMATS Aspergillus clavatus NRRL1
FgaPT2 Aspergillus fumigatus
DmaW Claviceps purpurea
FgaPT1 Aspergillus fumigatus AF293
n7-DMATS Aspergillus fumigatus AF293
TdiB Aspergillus nidulans
II
IptA Streptomyces sp. SN-593
PriB Streptomyces sp. RM-5-8
LtxC Moorea producens 3L
TleC Streptomyces blastmyceticus
IlaO Streptomyces atratus
CymD Salinispora arenicola CNS-205
0.8


C
daptomycin
prenylated daptomycin
PriB
DMAPP
PPi
FgaPT2
DMAPP
PPi