Genome Mining and Evolutionary Analysis Reveal Diverse Type III Polyketide Synthase Pathways in Cyanobacteria

Joachim Steen Larsen, Leanne Andrea Pearson, and Brett Anthony Neilan*

School of Environmental and Life Sciences, University of Newcastle, Newcastle, New South Wales, Australia

*Corresponding author: E-mail: brett.neilan@newcastle.edu.au.

Accepted: 14 March 2021

Abstract

Cyanobacteria are prolific producers of natural products, including polyketides and hybrid compounds thereof. Type III polyketide synthases (PKSs) are of particular interest, due to their wide substrate specificity and simple reaction mechanism, compared with both type I and type II PKSs. Surprisingly, only two type III PKS products, hierridins, and (7.7)paracyclophanes, have been isolated from cyanobacteria. Here, we report the mining of 517 cyanobacterial genomes for type III PKS biosynthesis gene clusters. Approximately 17% of the genomes analyzed encoded one or more type III PKSs. Together with already characterized type III PKSs, the phylogeny of this group of enzymes was investigated. Our analysis showed that type III PKSs in cyanobacteria evolved into three major lineages, including enzymes associated with 1) (7.7)paracyclophane-like biosynthesis gene clusters, 2) hierridin-like biosynthesis gene clusters, and 3) cytochrome b5 genes. The evolutionary history of these enzymes is complex, with some sequences partitioning primarily according to speciation and others putatively according to their reaction type. Protein modeling showed that cyanobacterial type III PKSs generally have a smaller active site cavity (mean = 109.035 Å³) compared with enzymes from other organisms. The size of the active site did not correlate well with substrate size, however, the “Gatekeeper” amino acid residues within the active site were strongly correlated to enzyme phylogeny. Our study provides unprecedented insight into the distribution, diversity, and molecular evolution of cyanobacterial type III PKSs, which could facilitate the discovery, characterization, and exploitation of novel enzymes, biochemical pathways, and specialized metabolites from this biosynthetically talented clade of microorganisms.

Key words: type III PKS, evolution, cyanobacteria, (7.7)paracyclophane, hierridin, cytochrome b5.

Significance

Type III polyketide synthases (PKSs) are specialized biosynthetic enzymes found primarily in plants. Surprisingly, only two examples have been characterized from cyanobacteria (blue-green algae), which are prolific producers of polyketides and other natural products. This study demonstrated that type III PKSs are widespread in cyanobacteria and have evolved into three lineages, including one that is closely related to plant enzymes. The phylogeny of cyanobacterial type III PKSs, together with their putative structure and their association with diverse tailoring enzymes suggests that they are involved in the production of a wide range of type III polyketides.

Introduction

Polyketide synthases (PKSs) are specialized enzymes that synthesize polyketides, a large group of natural products, including antibiotics, immunosuppressants, toxins, and a variety of other biologically active compounds (Jenke-Kodama et al. 2005; Shimizu et al. 2017). Cyanobacteria (blue-green algae) are a particularly rich source of polyketides and hybrid compounds thereof, including microcystins, anatoxin, cylindrospermopsin, (7.7)paracyclophanes, and hierridins (Nakamura et al. 2012; Pearson et al. 2016; Costa et al. 2019). The native biological function of these compounds is yet to be resolved, however, they are believed to confer an ecological advantage
PKSs catalyze the decarboxylative Claisen condensation of malonyl-coenzyme A (CoA), methylmalonyl-CoA and/or ethylmalonyl-CoA units onto a substrate molecule forming polyketides in an analogous fashion to fatty acid synthesis (Lai and Cronan 2003) and are classified into three groups according to their structure. Type I PKSs are large, multimodular enzymes widespread in bacteria (including cyanobacteria) that carry out condensation reactions in a stepwise manner. Type II PKSs are aggregates of monofunctional proteins, found mostly in actinobacteria, that carry out condensation reactions in an iterative way. Finally, type III PKSs are small homodimeric enzymes, found mostly in plants, which carry out condensation reactions in an iterative manner (Flores-Sanchez and Verpoorte 2009; Katsuyama and Ohnishi 2012). Type III PKSs are particularly interesting because of their relatively simple reaction mechanism and wide substrate specificity, giving rise to a diverse range of products (Katsuyama and Ohnishi 2012). Most polyketides identified in cyanobacteria, are produced by type I PKSs however, a select few, including the cytotoxic (7.7)paracyclophanes and antiplasmodial hierridins, are produced by type III PKSs (Bui et al. 2007; Costa et al. 2019) (fig. 1).

Cyanobacterial PKSs are encoded within large gene clusters (ranging from 3 to 60 kb in size), usually alongside genes for the modification and transport of the corresponding natural product. Occasionally, these gene clusters encode other biosynthetic enzymes, including fatty acid synthases (FAS) and non-ribosomal peptide synthetases (NRPS) (Tillett et al. 2000; Satou et al. 2013). In these cases, the different biosynthetic enzymes (FASs, PKSs, NRPSs etc.) work together to form a single product. Several such “hybrid” clusters have been identified in cyanobacteria, including the cyanotoxins microcystin and cylindrospermopsin, which are produced by both NRPSs and type I PKSs (Tillett et al. 2000; Mihali et al. 2008), and the alkylresorcinol dimer antibiotics, (7.7)paracyclophanes, which are produced by type I PKSs and type III PKSs (Nakamura et al. 2012). Hybrid type I and III polyketides are rare in nature, and the (7.7)paracyclophanes are the only such compounds to be isolated from cyanobacteria. Recently an analogue of these compounds, dediacloromonostrocyclophane D, was identified in extracts of the Vietnamese marine snail Planaxis sulcatus (Thanh et al. 2020). However, it is unclear whether it is produced by the snail, or a cyanobacterium occupying the same ecological niche.

In total, seven (7.7)paracyclophanes have been identified (fig. 1), however, only three of these have been associated with a gene cluster; the mercyclophanes (May et al. 2017), the carbamatocyclophanes (Preisitsch, Heiden, et al. 2016), and the cylindrocyclophanes (Nakamura et al. 2012). All share the same skeleton, consisting of two p-substituted aromatic rings (resorcinols) bridged between two C12 ester chains. The biosynthesis of cylindrocyclophanes, the best characterized (7.7)paracyclophane, starts with the attachment of a decanoic acid (C10 ester) to a free-standing acyl carrier protein (ACP) unit (CylB). The decanoyl-ACP thioester is putatively chlorinated by a novel halogenase (CylC) (Nakamura et al. 2017) and elongated by two type I PKSs (CylD and CylH). The ketide chain is then transferred to a type III PKS (CylI), which carries out the final elongations and cyclization of the aromatic ring, giving rise to an alkyl resorcinol intermediate. An alkylating enzyme (CylK) then dimerizes two alkyl resorcinol intermediates via a head-to-tail mechanism, substituting the chlorination with the resorcinol group from the other intermediate to form the final (7.7)paracyclophane (Nakamura et al. 2012, 2017). Linear compounds known as cylindrofridins (fig. 1) may be produced in the absence of CylC or CylK (Preisitsch, Niedermeyer, et al. 2016).

Until recently, it was believed that type III PKSs only used CoA-activated substrates, unlike type I and type II PKSs, which use ACP-bound substrates. However, recent studies have demonstrated that some type III PKSs, including CylI, can use both CoA- and ACP-bound substrates (Chemler et al. 2012; Nakamura et al. 2012; Preisitsch, Heiden, et al. 2016; May et al. 2017). As only a few examples have been enzymatically characterized, it is unclear whether this relaxed substrate specificity is a common feature of type III PKSs.

Another family of compounds synthesized by type III PKSs in cyanobacteria is the hierridins, O-methylated monoalkylresorcinols featuring a long-aliphatic chain (fig. 1). First isolated from Phormidium ectocarpi in 1998, hierridins were shown to have antiparasomal and antitumor activity (Papendorf et al. 1998; Leão et al. 2013). Unlike the (7.7)paracyclophanes, the proposed biosynthesis of hierridins occurs through a single-type III PKS (HidC) (Costa et al. 2019). Starting with a C14 acyl-ACP thioester, HidC carries out three elongations with malonyl-CoA before release and cyclization of the resorcinol ring. Thereafter, the structure undergoes reduction and methylation (by HidA and HidB), forming hierridan B. Chlorination of the ring by a yet unknown mechanism gives rise to hierridan C.

The functional diversity of type III PKSs and their relatively sparse distribution in plants, fungi, and bacteria compared with type I and II PKSs, make them an interesting group of enzymes to investigate from both an evolutionary and biosynthetic perspective. Here, we mined publicly available cyanobacterial genomes for type III PKS gene clusters and examined their composition, arrangement, and conservation. We also performed a detailed phylogenetic analysis of the encoded type III PKSs and modeled representative enzymes from each clade. The results are discussed in terms of the distribution, diversity, evolution, and putative function of these under-explored biosynthesis pathways.
**HIERRIDIN**

- **A**: $R_1=H$, $R_2=C_{15}$
- **B**: $R_1=H$, $R_2=C_{13}$
- **C**: $R_1=Cl$, $R_2=C_{13}$

**Cylindrofridin**

- **B**: $R=H$
- **C**: $R=Ac$

**Ribocyclophane**

- **A**: $R=O-\beta-D$-glucopyranosyl
- **B**: $R=OH$
- **C**: $R=H$
- **D**: $R=OAc$

**Merocyclophane**

- **A**: $R_1=R_2=CH_3$, $R_3=OH$, $R_4=H$
- **B**: $R_1=R_2=CH_3$, $R_3=R_4=(=O)$
- **C**: $R_1=CH_3$, $R_2=CH_2OH$, $R_3=OH$, $R_4=H$
- **D**: $R_1=R_2=CH_2OH$, $R_3=OH$, $R_4=H$

**Nostocyclophane**

- **A**: $R_1=R_2=OCH_3$, $R_3=R_4=\beta-D$-glucopyranosyl
- **B**: $R_1=R_2=OCH_3$, $R_3=H$, $R_4=\beta-D$-glucopyranosyl
- **C**: $R_1=OCH_3$, $R_2=OH$, $R_3=R_4=H$
- **D**: $R_1=R_2=OCH_3$, $R_3=R_4=H$

**Carbamidocyclophane**

- **A**: $R_1=R_2=CHCl_2$, $R_3=Carbamoyl$
- **B**: $R_1=CHCl_2$, $R_2=CH_2Cl$, $R_3=Carbamoyl$
- **C**: $R_1=CHCl_2$, $R_2=CH_3$, $R_3=Carbamoyl$
- **D**: $R_1=CH_2Cl$, $R_2=CH_3$, $R_3=Carbamoyl$
- **E**: $R_1=R_2=CH_3$, $R_3=Carbamoyl$
- **F**: $R_1=R_2=CHCl_2$, $R_3=H$
- **G**: $R_1=R_2=CHCl_2$, $R_3=Ac$
- **H**: $R_1=R_2=CH_3$, $R_3=H$
- **I**: $R_1=CH_3$, $R_2=CH_2Cl$, $R_3=H$
- **J**: $R_1=R_2=CH_2Cl$, $R_3=Carbamoyl$
- **K**: $R_1=CH_3$, $R_2=CHCl_2$, $R_3=H$
- **L**: $R_1=CH_2Cl$, $R_2=CHCl_2$, $R_3=H$

**Fig. 1.**—Compounds produced by cyanobacterial type III PKSs. Structures reproduced from Martins et al. (2019).
Materials and Methods

Genome Mining and Gene Cluster Comparisons
Cyanobacterial genomes were downloaded from GenBank (National Centre for Biotechnology Information, NCBI) and mined for type III PKS biosynthesis gene clusters using both the Antibiotics and Secondary Metabolite Analysis Shell (antiSMASH version 5.1) set to default (relaxed strictness) with all extra features enabled (Blin et al. 2019) and Basic Alignment Search Tool (BLAST, NCBI) set to default with a cut-off E-value of 10−20. Query sequences used for BLAST were ArsB (ZP_00418325), Cabl (AM848450), CHS (P30075), CyU (AF96143), HidC (QBC65480), MerE (AQ28564), ORAS (EAA31191), PKS18 (P9WP53), THNS (Q9FCA7), and TotC1 (ATL73040). Because the quality of the prediction by antiSMASH is highly dependent on the quality of the input data (i.e., gene clusters may not be detected if they are scattered over multiple small contigs), only RefSeq genomes were used.

Mined and reference gene clusters were organized into a similarity network using the Biosynthetic Genes Similarity Clustering and Prospecting Engine (BiG-SCAPE) (Navarro-Munoz et al. 2020) and compared using Clinker v 0.0.12 (Gilchrist and Chooi 2021).

Type III PKS genes and their inferred primary peptide sequences were extracted from the mined gene clusters for subsequent phylogenetic analysis and protein modeling. Previously characterized type III PKSs (reference sequences) were identified through a literature search and downloaded from GenBank (supplementary table 1, Supplementary Material online). 16 s rRNA gene sequences from the same species were also extracted from the cyanobacterial genomes using RNAmmer version 1.2 (Lagesen et al. 2007). In cases where RNAmmer could not identify 16 s rRNA gene sequences, they were downloaded from the SILVA database (https://www.arb-silva.de/search/, last accessed September 17, 2020) (Yilmaz et al. 2014).

Phylogenetic Tree Construction
A multiple sequence alignment of putative type III PKS peptide sequences and reference sequences was performed using MUSCLE (Edgar 2004) with default settings. The alignment was trimmed to remove large gaps (final length = 775 nucleotides). The tree was constructed using MrBayes with default settings and the best fitting model, JC, which was identified by JModelTest2.1.10. MrBayes was run until the standard deviations were below 0.01.

Protein Modeling
One cyanobacterial enzyme from each of the clades identified in the phylogenetic study was modeled using the online software, I-TASSER (Yang et al. 2015). The model with the highest Template modeling (TM) score was chosen for further analysis. The TM-Score is a means to measure how good the model is based on a given crystal structure; the higher the score, the better the model (Zhang and Skolnick 2004). The models were compared with published crystal structures for type III PKSs found in the Protein Data Bank (PDB) (supplementary table 4, Supplementary Material online). Models were visualized using PyMol2.0. The active site sizes of all the investigated enzymes were predicted using the online software CASTp3.0 with default settings (http://sts.bioe.uic.edu/castp, last accessed September 11, 2020).

Results and Discussion
Distribution and Diversity of Cyanobacterial Type III PKSs
To better understand the distribution and diversity of type III PKSs pathways in cyanobacteria, 517 cyanobacterial RefSeq genomes were downloaded from GenBank and mined using antiSMASH (Blin et al. 2019) and BLAST. Both methods returned the same results. Eighty-seven of the genomes analyzed (~16.8%) harbored one or more type III PKS gene clusters. Genomes of Moorea producens PAL-8-15-08-1, Nostoc sp. 3335mg, Cyanobium gracile PCC 6307, and Microcystis aeruginosa PCC 7005 harbored two type III PKS gene clusters, taking the total number of identified cyanobacterial type III PKS gene clusters to 91 (results summarized in supplementary table 2, Supplementary Material online). By comparison, 34% of fungal genomes (Navarro-Munoz and Collemare 2020) and 100% of plant genomes (Austin and Noel 2003) encode type III PKS pathways. Sequence similarity networks grouped the clusters into 60 families, where 49 were singletons (i.e., single cluster families). Clusters from Prochlorococcus and Microcystis formed the two largest families comprising 10 and 11 members, respectively (supplementary fig. 1, Supplementary Material online). Representative type III PKS gene clusters are depicted in figure 2.

Hybrid Type I/Type III PKS Gene Clusters
Twenty-one (7.7)paracyclophanes-like biosynthesis gene clusters were identified in the cyanobacterial genomes analyzed. (7.7)paracyclophanes are the best-studied type III PKS metabolites from cyanobacteria. As mentioned above, they are
synthesized through an unusual hybrid type I/III PKS system, which has not been identified in other phyla. To elucidate the genetic basis for (7.7)paracyclophanes diversity, we compared the mined gene clusters, including those responsible for the biosynthesis of carbamidocyclophane \( (\text{Nostoc} \; sp. \; CAVN2) \), cylindrocyclophane \( (\text{Cylindrospermum lichen forme} \; \text{UTEX B} \; 2014) \), and merocyclophane \( (\text{Nostoc} \; sp. \; \text{UIC10110}) \), and the orphan type I/III PKS gene cluster from \( \text{M. aeruginosa} \; \text{PCC} \; 7806 \) (MKS, “\text{Microcystis Ketide Synthase}”). Although no metabolite has been matched to the MKS cluster, its transcription appears to be inversely co-regulated with the production of the hepatotoxin microcystin, making this gene cluster of special interest (Makower et al. 2015).

The (7.7)paracyclophane-like biosynthesis gene clusters display similar overall composition and organization (fig. 2). The carbamidocyclophane \( (\text{cab}) \) and cylindrocyclophane \( (\text{cly}) \) biosynthesis gene clusters share particularly high sequence homology (~80% identity). The merocyclophane \( (\text{mer}) \) biosynthesis cluster differs with regard to the type I PKS \( (\text{merC}) \) module organization, having unique dehydrogenase (DH) and enoylreductase (ER) domains. The Type I PKS, \( \text{merD} \), also has a thioesterase (TE) in place of the terminal ACP found in \( \text{cabD} \) and \( \text{clyD} \). However, the main biosynthesis genes, including the type III PKS \( (\text{merE}) \), are otherwise well conserved. In contrast, the genes encoding tailoring enzymes vary from cluster to cluster, indicating horizontal gene transfer or gene loss has shaped these clusters. For instance, the carbamidotransferase gene \( (\text{cabL}) \), which carries out the attachment of the carbamido group to the final carbamidocyclophane is unique to the carbamidocyclophane biosynthesis cluster, whereas the two oxidoreductase genes \( (\text{clyL} \) and \( \text{clyM} \)
might be important for the production of cylindrocyclophe
phane D and E (fig. 1). The cyl cluster also contains a glyco-
syltransferase gene homologue (cylN), putatively involved in
the production of ribocyclophephane; glycosylated analogous of
cylindrocyclophaneinos (May et al. 2018). Strikingly, the
AMP-binding enzyme (cabA/cylA/merl) is differentially lo-
cated in the merocyclophane biosynthesis cluster compared
with the cylindrocyclophane and carbamidocyclophane bio-
synthesis clusters, indicative of a genomic rearrangement in
the former. Quite strikingly, the mer cluster also has fewer
tailoring enzymes than the other (7.7)paracyclophane bio-
synthesis clusters. This suggests that a deletion at the 3' end might have occurred, resulting in the differential loca-
tion of the AMP-binding enzyme.

The MKS cluster is more divergent from the characterized
(7.7)paracyclophane biosynthesis gene clusters than they are
from each other. This could reflect taxonomic divergence be-
tween Microcystis and the Nostoc/Cylindrospermum lineage,
and the possible production of different type III PKS products in
the Microcystis genus. However, parts of the second type I
PKS gene (mksE), the type III PKS gene (mksG), and the up-
stream methyltransferase (mksH) are well conserved in the
Microcystis and Nostoc/Cylindrospermum lineages. Interestingly,
mksC encodes an AMP-binding enzyme (homologue of cabA, cylA, and merA) together with an ACP domain
(homologue of the free-standing ACP encoded by cabB, cylB, and merA). This suggests that there has been a recombination event between the AMP-binding enzyme and the free-
standing ACP in mksC. Equally interesting is the absence of
cylC (halogenase) and cylK (alkylating enzyme) homologues in
the MKS gene cluster, which are highly conserved in the other
non-Microcystis clusters. CylC and CylK are important for the
dimerization of the final ketide chain in the biosynthesis of
(7.7)paracyclophane to form the characteristic cyclic structure.
The absence of these two genes in the MKS cluster suggests that it does not produce a dimeric, cyclic structure-like (7.7)par-
acyclophane, but perhaps a linear analogue. Interestingly, the
Microcystis clusters all contain three transporters, two encoding
resistance nodulation cell division (RND) family transporters and
one encoding an major facilitator superfamily (MFS) trans-
porters. This suggests that the metabolites produced by these
clusters have an extracellular function. As no type III PKS prod-
ucts have been isolated and identified from Microcystis, further
research is needed to verify the compound produced by the
MKS cluster. All the MKS clusters from Microcystis have the
same composition and structural organization.

Detailed comparison of (7.7)paracyclophane-like biosynthe-
sis gene cluster families using Clinker v 0.0.12 can be found in
supplementary figure 2, Supplementary Material online.

Hierridin-Like Biosynthesis Gene Clusters

Forty-seven hierridin-like biosynthesis gene clusters were iden-
tified in the cyanobacterial genomes analyzed (fig. 2). As
mentioned above, hierridins are synthesized by a single type
III PKS. The cyclized polyketide is then modified by a methyl-
transferase and an oxidoreductase to produce hierridin B. In
Cyanobium sp. LEGE 06113, the hierridin skeleton is chlori-
nated by an unknown mechanism to produce hierridin C
(fig. 1) (Costa et al. 2019). So far, this gene cluster is the only
of its class to have been associated with a compound, however, as we have shown, multiple hierridin-like biosynthe-
sis gene clusters are present in cyanobacteria, particularly in the
Prochlorococcus and Synechococcus genera, suggesting that production of hierridins and related compounds is com-
mon in marine species (supplementary table 2, Supplementary Material online).

Compared with the (7.7)paracyclophane-like biosynthesis
gene clusters, the hierridin-like clusters are mostly well con-
served in terms of composition and organization (fig. 2). The
major difference being two different patterns of gene ar-
rangement: hidA>hidB, >hidC and hidC>hidA>hidB, with the
former arrangement being the more conserved. Besides
encoding HidA (methyltransferase), HidB (oxidoreductase),
and HidC (type III PKS), the hierridin-like biosynthesis gene
clusters encode a variety of potential stand-alone tailoring
enzymes. For instance, the cluster from Leptolyngbya sp.
PCC 7375 encodes a second methyltransferase and a reduc-
tase, suggesting that this species produces a poly-
methylated compound, which has been reduced further than hierridin B.

Two hierridin-like biosynthesis gene clusters were identified
in the genome of Nostoc sp. 335 mg. One of the clusters encodes two methyltransferases and an oxygenase, but lacks
the oxidoreductase, suggesting a deletion might have oc-
curred. The other cluster encodes the three main biosynthesis
genes but no additional tailoring enzymes. The presence of
two gene clusters in Nostoc sp. 335 mg suggests that dupli-
cation and deletion events occurred, resulting in the duplica-
tion of the methyltransferase and loss of the oxidoreductase.
Uniquely, the hierridin-like biosynthesis gene cluster from
Raphidiopsis mediterranea Skuja FSS1-150/1 encodes a pre-
nyltransferase, suggesting that this species produces a novel
tenylated compound.

Detailed comparison of hierridin-like biosynthesis gene
cluster families using Clinker v 0.0.12 can be found in
supplementary figure 2, Supplementary Material online.

Gene Clusters Encoding Type III PKSs and Cytochrome b5

Sixteen gene clusters encoding type III PKS and cytochrome
b5 genes were identified in the genomes of Nostocales and
Chroococcales species (fig. 2). These gene clusters were poorly
conserved in terms of sequence homology and structural or-
organization, compared with the (7.7)paracyclophane and
hierridin-like biosynthesis gene clusters.

Gene clusters encoding type III PKSs and cytochrome
genes (cytochrome P450 or cytochrome b5) have previously
been identified in mycobacteria. In *M. tuberculosis* H37Rv, the encoded type III PKS (PKS18) is an alkylpyrone synthase, which uses long aliphatic acids as substrates (C₆–C₂₀), extending them two or three times to form triketide and tetraketide pyrones, respectively (Saxena et al. 2003). Although the activity of the associated cytochrome b5 is unclear, it is likely involved in the oxidation of the type III PKS product, giving rise to myolic acid, a membrane lipid found in mycobacteria (Kolattukudy et al. 1997; Saxena et al. 2003). In the cyanobacterial genomes investigated here, cytochrome b5 is encoded just upstream of the type III PKS genes (or downstream in the case of *Chamaesiphon minutus* PCC 6605), suggesting a role in biosynthesis. As no cyanobacterial cytochrome b5-associated gene clusters have been experimentally characterized, their products remain unknown. However, it seems unlikely that they are involved in the synthesis of myolic acid, as in *M. tuberculosis*, as this compound has not been isolated from any cyanobacterial species (Rajeshwari and Rajashkekar 2011).

The cyanobacterial type III PKS gene clusters associated with cytochrome b5 also encode a combination of the three transporter families; RND, MFS, and ATP-binding cassette (ABC) transporters, whereas only RND and MFS transporters are present in the *Microcystis* (7.7)paracyclophane-like biosynthesis gene clusters. Interestingly, the cluster from *Rivularia* sp. PCC 7116 encodes all families, whereas the cluster from *Microcystis aeruginosa* SP777 lacks the MFS transporter, indicating a deletion in this strain. The presence of the transporters suggests that the produced compounds have an extracellular role. In contrast, the *M. tuberculosis* H37Rv cluster does not encode a transporter.

Interestingly, the gene cluster from *Rivularia* sp. PCC 7116 encodes a stand-alone hydroxylase. To our knowledge, only one other stand-alone hydroxylase gene (*srsC*) has been associated with a type III PKS cluster (the *srsABC* cluster from *S. griseus* NBRC 13350, involved in the production of alkylre sorcinols) (Funabashi et al. 2008). However, hydroxylation of some point during biosynthesis or chain termination, as in *Cylindrospermum licheniforme* UTEX B 2014 (Nakamura et al. 2012), MerE (merocyclophane synthase) from *Nostoc* sp. UIC 10110 (May et al. 2017), Cabl (carbamo cyclophane synthase) from *Nostoc* sp. CAVN2 (Preisitsch, Heiden, et al. 2016), and HidC (hierrdin synthase) from *Cyanobium* sp. LEGE 06113 (Costa et al. 2019). Other type III PKS reference sequences from plants, bacteria and fungi were also included in the analysis. The FAS, FabH, from *E. coli* was used as the outgroup (supplementary table 1, Supplementary Material online). Parts of the multiple sequence alignment upon which the tree was based are shown in supplementary material (supplementary fig. 5, Supplementary Material online). A second phylogenetic analysis based on the 16S rRNA genes of the same species was also conducted to infer speciation (supplementary fig. 6, Supplementary Material online).

The type III PKS sequences partitioned into three major clades: two bacterial clades (Clades 1 and 2) and a clade comprising mostly eukaryotic and cyanobacterial sequences (Clade 3) (fig. 3). Our overall tree is largely in agreement with a former study by Shimizu et al. (2017), which demonstrated that the evolution of type III PKS sequences is complex, with some sequences partitioning primarily according to speciation (analogous to 16S rRNA gene phylogeny, supplementary fig. 6, Supplementary Material online) and others putatively according to their reaction type (supplementary table 3, Supplementary Material online).

Clade 1 comprised actinobacterial type III PKSs (e.g., DpgA, Ken2, VemA) encoded within hybrid gene clusters (fig. 3), involved in the production of (3,5-dihydroxyphenyl)acetyl-CoA, a precursor of glycopeptide antibiotics. Enzymes within this clade initiate biosynthesis of the ketide chain, via condensation of four malonyl-CoAs, before transfer to two type I PKSs for further elongation. Cyclization occurs via a CB–C3 Claisen reaction (Wu et al. 2012; Zeng et al. 2012; Chen et al. 2017) (supplementary table 3, Supplementary Material online). A single cyanobacterial type III PKS from *Moorea producens* PAL-8-15-08-01 was present within this clade, suggesting that this species is capable of producing glycopeptide antibiotics or similar compounds. However, the gene cluster from this species is not similar to any of the other bacterial gene clusters that partitioned in this clade (fig. 2).

Clade 2 comprised a variety of bacterial and cyanobacterial sequences, which partitioned further into two subclades: Clades 2.1 and 2.2. Clade 2.1 comprised actinobacterial enzymes (e.g., RppA, STS, THNS) involved in tetrahydroxy naphthalene production, a precursor of flaviolin. Enzymes in...
this clade almost exclusively use malonyl-CoA starter substrates and C3–C8 or C2–C7 aldol condensation cyclization mechanisms, the only exception being Gcs, which uses either methylmalonyl-CoA or ethylmalonyl-CoA before carrying out a lactonization to produce germicidin (Chemler et al. 2012) (supplementary table 3, Supplementary Material online).

Clade 2.2 was divided into two subclades: Clade 2.21 and 2.22. Clade 2.21 comprised characterized enzymes from actinobacteria (e.g., SrsA, PKS10, PKS11) and cyanobacteria (e.g., Cabl, Cyl, MerE) involved in the synthesis of alkylresorcinols (supplementary table 3, Supplementary Material online), as well as 25 uncharacterized enzymes from cyanobacteria. Within Clade 2.22, actinobacterial and cyanobacterial sequences formed distinct monophyletic groups. The actinobacterial enzymes are known to use long aliphatic esters as substrates and carry out either a lactonization or C2–C7 aldol condensation with either malonyl- or methylmalonyl-CoA (Saxena et al. 2003; Sirakova et al. 2003). The only exception is RppA, which is...
similar to the enzymes found in Clade 2.1, in terms of its substrate and extender specificities (Ghimire et al. 2008). The previously characterized cyanobacterial enzymes in Clade 2.22 use malonyl-CoA to extend dodecanoyl-CoA (C_{12} thioester obtained from a type I PKS) and cyclize it via a C2–C7 aldol condensation (Kang et al. 2012; Nakamura et al. 2012; Luo et al. 2014). Significantly, nearly all the cyanobacterial sequences within this clade 2.2 originate from hybrid clusters containing genes encoding both type I and type III PKSs (fig. 2), where the type III PKS carries out the last reaction in the corresponding pathway. The exception is the Gloeocapsa sp. PCC 73106 enzyme, which was mined from a cluster containing a single type III PKS. None of the characterized actinobacterial enzymes from clade 2.1 or clade 2.2 originate from hybrid clusters. The enzymes mined from Microcystis (including MksG from the MKS cluster) formed a monophyletic group, and their associated biosynthesis gene clusters displayed a different architecture to those encoding the other cyanobacterial type III PKSs in Clade 2.22, that is, they lacked genes encoding halogenase and alkylation enzymes (cylC and cylK homologues) (Nakamura et al. 2017) (fig. 2). This double gene deletion in Microcystis could foreseeably enable the production of a novel monomeric polyketide similar to the cylindrofridins (fig. 1) (Preisitsch, Niederreyer, et al. 2016).

Clade 3 was the largest and most diverse clade in the tree, comprising mainly actinorhizal and cyanobacterial type III PKSs. The clade partitioned into two subclades: Clade 3.1 and 3.2. Clade 3.1 comprised uncharacterized type III PKSs encoded alongside cytochrome b5 (fig. 2). Except for the alkylpyrone biosynthesising PKS18 from Mycobacterium tuberculosis, all sequences within this subclade originated from cyanobacteria. PKS18 uses long aliphatic esters as substrates and extends those with malonyl-CoA, followed by a lactonization to form the final pyrones (Saxena et al. 2003; Sankaranarayanan et al. 2004) (supplementary table 3, Supplementary Material online). The uncharacterized cyanobacterial enzymes in this clade are likely to incorporate similar substrates and reaction mechanisms. Interestingly, no alkylpyrones have been isolated from cyanobacteria; all type III PKS products isolated from cyanobacteria to date belong to the alkylresorcinol group (Martins et al. 2019). However, the close relationship of Clade 3.1 cyanobacterial enzymes with PKS18 from M. tuberculosis suggests that pyrone production by cyanobacteria could be possible. The function of cytochrome b5 in the biosynthesis of Clade 3.1 polyketides is yet to be verified, however, as mentioned previously, it may play a role in the synthesis of mylic acid, using the products of the type II PKS (Kolattukudy et al. 1997; Saxena et al. 2003).

Clade 3.2 was divided into Clades 3.21 and 3.22. Clade 3.21 comprised euukaryotic type III PKSs, including fungal and plant sequences, which partitioned further into distinct monophyletic groups. This topology suggests that type III PKSs in Eukaryota evolved from a common ancestor before the differentiation of the superkingdom, which is supported by previous studies (Mallika et al. 2011; Shimizu et al. 2017; Navarro-Muñoz and Collemare 2020). The fungal enzymes in Clade 3.21 are known to synthesize a variety of related compounds, including tri- and tetraketide pyrones (e.g., ORAs, CsyA), tetra- and pentaketide resorcinols (e.g., ORAS), and (alkenyl)-α pyrone (e.g., CsyB), using either malonyl- or ethylmalonyl-CoAs as extender units, and cyclizing the ketide with either lactonization or C2–C7 aldol condensation (Shishima et al. 2005; Furuta et al. 2007; Seshime, Juwadi, Kitamoto, Ebizuka, and Fujii 2010; Seshime, Juwadi, Kitamoto, Ebizuka, Nonaka, et al. 2010; Hashimoto, Ishida, et al. 2013; Hashimoto, Seshime, et al. 2013). Most of the plant enzymes on the other hand (e.g., OLS, BAS, STS, CHS), synthesize flavonoid precursors from p-coumaroyl-CoA and use a variety of cyclization mechanisms (Flores-Sanchez and Verpoorte 2009), with a few showing a different substrate specificity and product formation (Eckermann et al. 1998; Mizuuchi et al. 2009) (supplementary table 3, Supplementary Material online).

Subclade 3.22 was comprised primarily of cyanobacterial type III PKSs (49 in total), including the hierridin biosynthesis enzyme, HiDc (alkylresorcinol synthase). It was further divided into Clades 3.221 and 3.222. Clade 3.221 surprisingly included two Azotobacter type III PKSs (ArsB and ArsC) as well as unknown enzymes from cyanobacteria. It is speculated that ArsB and ArsC use ACP-bound long aliphatic thioesters rather than CoA-activated substrates in vivo, although they can use both substrates in vitro (Furuta et al. 2006). Intramolecular cyclization occurs via C3–C8 aldol condensation and lactonization, respectively (supplementary table 3, Supplementary Material online). As the branches in Clade 3.221 are long, suggesting significant evolutionary distance, it is difficult to predict whether the cyanobacterial enzymes share a similar reaction mechanism to ArsB or ArsC.

Clade 3.222 was comprised exclusively of cyanobacterial sequences from hierridin-like biosynthesis gene clusters, which partitioned further into clades, 3.2221 and 3.2222. The gene clusters associated with Clade 3.2221 enzymes have a different genetic architecture to those associated with Clade 3.2222 enzymes, which closely resemble the hidABC gene cluster from Cyanobium sp. LGE06113 (Costa et al. 2019) (fig. 2). Several of the subclade 3.2221 gene clusters encode two methyltransferases (HiDb homologs), suggesting that the metabolites produced undergo an extra methylation step compared with the products of Clade 3.2222 pathways. Interestingly, the gene cluster associated with the type III PKS from R. mediterranea Skuja FSS1-150/1 (Clade 3.2221) also encodes a prenyltransferase (fig. 2), therefore this gene cluster could produce a completely novel compound. Prenyltransferases were not identified in any of the other gene clusters linked to this clade, suggesting that it was acquired through horizontal gene transfer. The only characterized enzyme in Clade 3.2222, HiDc, putatively uses a C_{14}
acyl-ACP thioester as a starter substrate, extending it with malonyl-CoA before cyclization through a C2-C7 aldol condensation (supplementary table 3, Supplementary Material online) (Costa et al. 2019). HidC-like PKS genes seem to be more abundant in marine cyanobacteria like Prochlorococcus and Synechocystis, whereas those encoding (7.7)paracyclophane synthases (Clade 2.222) are more prevalent in terrestrial species (supplementary table 2, Supplementary Material online). This suggests that the two different pathways could have evolved to facilitate ecological niche colonization by the producing organisms. However, our results also show that some terrestrial species, including Nostoc contain derivatives of the hid cluster. This could be due to horizontal gene transfer from one ecotype to another, as has been shown between type I PKS gene clusters in Streptomyces (Vicente et al. 2018).

In summary, our phylogenetic analysis suggests that modern cyanobacterial type III PKSs comprise three major lineages, including 1) enzymes associated with (7.7)-paracyclophane-like biosynthesis gene clusters (Clade 2.222), 2) enzymes associated with hierridin-like biosynthesis gene clusters (Clade 3.222), and 3) enzymes associated with cytochrome b5-encoding biosynthesis gene clusters (Clade 3.1) (fig. 3). All characterized cyanobacterial type III PKSs synthesize alkylresorcinols from long aliphatic acids (C_{12}–C_{16}) via a C2-C7 aldol cyclization mechanism, and it is likely that most of the uncharacterized enzymes also use this reaction mechanism, except for those in Clade 3.1, which probably use lactonization like their nearest homologues. The gross topology of the tree suggests that several different forces have shaped the evolution of type III PKSs in cyanobacteria, including vertical and horizontal gene transfer as well as duplication and gene loss events (this is also evident in the architecture of the corresponding gene clusters, fig. 2). For example, the proximity of cyanobacterial HidC (hierridin biosynthesis)-like enzymes (Clade 3.22) to the eukaryotic enzymes (Clade 3.21), suggests that an ancestral type III PKS was transferred from cyanobacteria to the progenitor of eukaryotes and its structure was conserved. Symbiogenesis, the evolution of eukaryotic organelles (e.g., chloroplasts) from prokaryotic organisms (e.g., cyanobacteria), lends further support to this theory (Cavalier-Smith 2002). Interestingly, nearly all plant type III PKSs investigated thus far are encoded within the nuclear genome not the chloroplast. However, it is possible that like many other genes originating from cyanobacteria, type III PKSs in plants were transferred to the nucleus during endosymbiosis (Timmis et al. 2004).

Gene duplication likely gave rise to multiple type III PKSs within single cyanobacterial species, as observed in Microcystis and Nostoc genomes. For example, in Nostoc sp. 335 mg, two hierridin-like biosynthesis gene clusters have evolved; most likely due to a duplication, followed by a methyltransferase duplication and introduction of an oxygenase (fig. 2). However, in most cases where two type III PKS clusters are present within a single organism, they are phylogenetically distinct, suggestive of horizontal gene transfer between phyla (fig. 3) (Zhaxybayeva et al. 2006). Horizontal transfer of genes between cyanobacteria and distantly related bacterial phyla could also explain the close phylogenetic relationship between cyanobacterial type III PKSs from Clade 3.21 and PKS18 from Mycobacterium tuberculosis, as well as cyanobacterial type III PKSs from Clade 3.22 and ArsC from Azotobacter vinelandii. Similarly, this might also explain the presence of lone cyanobacterial sequences in certain clades (e.g., Moorea producens PAL-8-15-08-1 in Clade 1 and Mastigocoleus testarum BC008 in Clade 2.1).

Presently, only two families of type III PKS metabolites have been characterized from cyanobacteria: the hierridins and the (7.7)paracyclophanes. However, our phylogenetic tree shows that at least three families are present, suggesting the potential repertoire of type III PKS products is more diverse than originally surmised. Characterized cyanobacterial type III PKSs appear to use very specific substrates in vivo (Bui et al. 2007; Nakamura et al. 2012; Preisitsch, Heiden, et al. 2016; Costa et al. 2019). This is comparable to bacterial and plant type III PKSs that show a rather rigid substrate specificity in vivo (Ghimire et al. 2008; Ghosh et al. 2008; Nakano et al. 2009; Nualkaew et al. 2012; Meslet-Cladiere et al. 2013). However, in vitro studies on purified bacterial type III PKSs suggest that they actually have a flexible substrate specificity (Zheng et al. 2001; Funabashi et al. 2008; Chemler et al. 2012; Hashimoto et al. 2014). Future in vitro studies might similarly reveal a broader substrate repertoire for cyanobacterial enzymes.

### Protein Modeling Highlights the Structural Diversity of Cyanobacterial Type III PKSs

The structure of type III PKS active sites is believed to provide clues to their substrate specificity and catalytic mechanism (Ferrer et al. 1999; Jez et al. 2001). Therefore, we modeled and compared the structures of cyanobacterial type III PKSs (n = 7) to previously elucidated crystal structures of bacterial and plant enzymes (n = 3) (supplementary table 4, Supplementary Material online) from our phylogenetic analysis (fig. 3). The size and topology of the active site cavities and their constitutive residues were closely examined in each case. The catalytic triad of active site residues, Cys–Asn–His (green residues in fig. 4), as well as the overall topology, were conserved across all the cyanobacterial type III PKSs modeled when compared with known crystal structures. Interestingly, the hydrophobic nonpolar “Gatekeeper” residues identified in plant CHS, Phe215, and 265 (numbering from CHS from Medicago sativa (blue residues in fig. 4), were only partially conserved in the bacterial sequences. Residues corresponding to Phe265 in CHS were particularly heterogeneous across the enzymes examined, and phylogenetic correlations were observed (supplementary fig. 5, Supplementary Material online). Although Clade 3.2222 enzymes, including
**Fig. 4**—Protein models of representative type III PKSs from cyanobacteria, other bacteria, plants, and fungi. Structures were obtained through I-TASSER and visualized using PyMol2. A cross-section of the type III PKSs is displayed, with the active site cavity (dark gray) highlighted. Green residues are the conserved active site residues, Cys–His–Asn. Blue residues correspond to the so-called “Gatekeeper” residues identified in CHS from *Medicago sativa*. The size of the active site cavity is given below the protein model. The crystal structures used to infer the models are listed in supplementary table 4, Supplementary Material online.
HidC, which is predicted to use acyl-ACP thioesters as substrates to form alkylresorcinols (Costa et al. 2019), had a conserved methionine (hydrophobic nonpolar) substitution at this position, all Clade 3.1 enzymes had a cysteine (sulfur-containing polar) residue. The substrate and product for this group are still unknown. Finally, Clade 2.222 enzymes, which use C\textsubscript{12} thioesters obtained from a type I PKS to produce alkylresorcinols (Costa et al. 2019), had a glutamine or an asparagine (polar uncharged) residue at this position. A possible explanation for these confounding results is that CylI, and closely related characterized Clade 2 cyanobacterial enzymes, have evolved a different mechanism to tether and process their substrates. For instance, the characterized cyanobacterial type III PKSs from clade 2.222 receive their substrates from the ACP-module rather than taking up the substrate directly from a CoA or from a single ACP unit (Nakamura et al. 2012).

Next, the size of the active sites was investigated. The size for each enzyme was predicted using the online software CASTp. The size of the active sites analyzed ranged between 6.731 Å\textsuperscript{3} (Leptolyngbya sp. PCC 7375) and 276.945 Å\textsuperscript{3} (PKS11 from M. tuberculosis H37Rv), with a mean of 153.217 Å\textsuperscript{3} (fig. 4). The cyanobacterial active sites were generally smaller (mean = 109.035 Å\textsuperscript{3}) than other bacterial (mean = 261.083 Å\textsuperscript{3}) and CHS from plants (220.957 Å\textsuperscript{3}) cavities. Previous studies have demonstrated a link between the size of the active site cavity and the size of the substrates handled by the type III PKSs (Jez et al. 2000). As most of the cyanobacterial enzymes are predicted to use long aliphatic acid substrates, it was surprising that many had relatively small active site cavities, compared with their bacterial counterparts, regardless of their phylogeny (figs. 3 and 4). Notably, the active sites of Clade 2 enzymes from C. licheniforme UTEX B 2014 (CylI, 80.045 Å\textsuperscript{3}) and M. aeruginosa PCC 7806 (MksG, 81.946 Å\textsuperscript{3}) are predicted to be approximately half the size of the average type III PKS enzyme cavity. We expected CylI (and its homologues) to have a large active site like that of PKS11 from M. tuberculosis H37Rv (Clade 2; 276.945 Å\textsuperscript{3}), which uses similar substrates. Similarly, ArsC from A. vinelandii has an active site of 271.044 Å\textsuperscript{3} whereas HidC is predicted to be 136.628 Å\textsuperscript{3}; only half the size (fig. 4). The only difference between these cyanobacterial and non-cyanobacterial type III PKSs is the cyclization event; where ArsC and PKS11 make pyrones, HidC and CylI cyclize their products into resorcinols. This might indicate that the larger active site is needed for the formation of the pyrone structure, whereas the smaller active site might be better suited to the production of resorcinols. This has also been shown in trORAS, where a mutation in Phe252 to glycine disrupted resorcinol production and increased pyrone production (Rubin-Pitel et al. 2008). As glycine is smaller than phenylalanine, this mutation is bound to increase the size of the active site. Exactly how the cyclization occurs is still unclear, likewise, it is unknown if this happens inside the enzyme or as a spontaneous event after the release of the ketide chain from the enzyme (Austin, Bowman, et al. 2004; Li et al. 2011; Wu et al. 2012; Satou et al. 2013). These results suggest that cyclization occurs while the ketide chain is within the active site. However, further research is needed to verify this.

Most of the predicted active site cavities of the cyanobacterial enzymes from Clade 3 were larger than those from Clade 2, except for the Leptolyngbya (6.731 Å\textsuperscript{3}) and Rivularia sp. PCC 7116 (65.187 Å\textsuperscript{3}) enzymes. However, they were still generally smaller than their noncyanobacterial counterparts. The extremely small size of the Leptolyngbya sp. PCC 7375 enzyme cavity suggests that it might not be functional or produces a very small polyketide, as the size of the active site of PKSs is generally proportional to the number of cyclization events (Austin, Izumikawa, et al. 2004; Rubin-Pitel et al. 2008; Wu et al. 2012). The small active site of the Rivularia sp. PCC 7116 enzyme suggests that it produces resorcinols rather than pyrones, despite its phylogenetic proximity to the pyrone synthase (PKS18) from M. tuberculosis H37Rv. On the other hand, the type III PKS from M. aeruginosa SPC777 has a relatively large active site cavity, which might be able to sustain pyrone formation (Sankaranarayanan et al. 2004). Likewise, the putative type III PKS from Prochlorococcus MIT 1306 is predicted to have a large active site (233.542 Å\textsuperscript{3}) and, much like plant and bacterial type III PKSs, could potentially sustain pyrone formation. Hence, pyrone formation by cyanobacterial type III PKSs should not be excluded even though to date, no pyrones have been isolated from cyanobacteria (Martins et al. 2019). In summary, protein modeling highlights the significant structural diversity present among cyanobacterial type III PKSs however, further research is needed to confirm if this structural diversity translates to functional diversity.

Conclusions
This study demonstrated that gene clusters encoding type III PKS are present in 16.8% of sequenced cyanobacterial genomes. Comparison of type III PKS biosynthesis gene cluster architecture and phylogenetic analysis showed that type III PKSs in cyanobacteria evolved into three major lineages, including enzymes associated with 1) (7.7)paracyclophane-like biosynthesis gene clusters, 2) hieridin-like biosynthesis gene clusters, and 3) cytochrome b5 genes. The evolutionary history of these enzymes is complex, with some sequences partitioning primarily according to speciation and others putatively according to their reaction type. Protein modeling indicated that cyanobacterial type III PKSs have a similar overall topology to plant and bacterial enzymes, although their active site cavities are usually smaller. Our results also showed that the size of the active site cavity does not necessarily correspond to the size of the substrate used by type III PKSs or the number of extensions catalyzed. Although the catalytic triad of amino acids (Cys–Asn–His) is conserved in cyanobacteria,
one of the so-called “Gatekeeper” amino acid residues (Phe265 in CHS) is not, suggesting that it might not be important for substrate specificity in cyanobacterial type III PKSs. Overall, our protein modeling highlights the structural diversity of type III PKSs from different taxa.

As only four cyanobacterial type III PKS gene clusters have been experimentally characterized, the repertoire of compounds is largely unexplored, and future studies are needed to identify the prevalence of different type III PKS metabolites from cyanobacteria, and their biological function. As the previously identified products have diverse bioactivities (antiplasmodial, cytotoxic, UV blocking), cyanobacteria remain a potentially rich, yet untapped source of biomedically relevant type III polyketides, including potentially prenylated, hydroxylated, and sulfonated type III PKSs and their novel biosynthetic enzymes.

**Supplementary Material**

**Supplementary data** are available at Genome Biology and Evolution online.

**Acknowledgment**

This research was funded by the Australian Research Council (Grant No. LP140100642).

**Data Availability**

Nucleotide and protein sequences are available in the GenBank Nucleotide Database at https://www.ncbi.nlm.nih.gov/genbank/ and can be accessed with accession numbers provided in **supplementary table 1, Supplementary Material online**. 3D (PyMOL) protein models will be shared on reasonable request to the corresponding author.

**Literature Cited**

Austin MB, Bowman ME, Ferrer J-L, Schroeder J, Noel JP. 2004. An aldol switch discovered in stilbene synthases mediates cyclization specificity of type III polyketide synthases. Chem Biol. 11(9):1179–1194.

Austin MB, Izumikawa M, et al. 2004. Crystal structure of a bacterial type III polyketide synthase and enzymatic control of reactive polyketide intermediates. J Biol Chem. 279(43):45162–45174.

Austin MB, Noel JP. 2003. The chalcone synthase superfamily of type IIIs polyketide synthases. Nat Prod Rep. 20(1):79–110.

Blin K, et al. 2019. antiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. Nucleic Acids Res. 47(W1):W81–W87.

Brown DW, et al. 2015. Identification of a 12-gene fusaric acid biosynthetic gene cluster in Fusarium species through comparative and functional genomics. Mol Plant Microbe Interact. 28(3):319–332.

Bui HTN, Jansen R, Pham HTL, Mundt S. 2007. Carbamidocyclophanes A–E, chlorinated paracyclophanes with cytotoxic and antibiotic activity from the Vietnamese cyanobacterium Nostoc sp. J Nat Prod. 70(4):499–503.

Cavaler-Smith T. 2002. Chloroplast evolution: secondary symbiogenesis and multiple losses. Curr Biol. 12(2):R62–R64.

Chemler JA, et al. 2012. Biochemical and structural characterization of germicidin synthase: analysis of a type III polyketide synthase that employs acyl-ACP as a starter unit donor. J Am Chem Soc. 134(17):7359–7366.

Chen R, et al. 2017. Genome mining and activation of a silent PKS/NRPS gene cluster direct the production of totopotersamides. Org Lett. 19(20):5697–5700.

Costa M, et al. 2019. Structure of hieridin C, synthesis of hieridins B and C, and evidence for prevalent alkylresorcin biosynthesis in picocya-

bacteria. J Nat Prod. 82(2):393–402.

Eckermann S, et al. 1998. New pathway to polyketides in plants. Nature 396(6709):387–390.

Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32(5):1792–1797.

Ferrer JL, Jez JM, Bowman ME, Dixon RA, Noel JP. 1999. Structure of chalcone synthase and the molecular basis of plant polyketide biosynthesis. Nat Struct Biol. 6(8):775–784.

Flores-Sanchez U, Verpoorte R. 2009. Plant polyketide synthases: a fasci-

nating group of enzymes. Plant Physiol Biochem. 47(3):167–174.

Funa N, Ozawa H, Hirata A, Horinouchi S. 2006. Phenolic lipid synthesis by type III polyketide synthases is essential for cyst formation in Azotobacter vinelandii. Proc Natl Acad Sci U S A. 103(16):6356–6361.

Funa N, Awakawa T, Horinouchi S. 2007. Pentaketide resorcylic acid syn-
thesis by type III polyketide synthase from Neorurpora crassa. J Biol Chem. 282(19):14476–14481.

Funabashi M, Funa N, Horinouchi S. 2008. Phenolic lipids synthesized by type III polyketide synthase confer penicillin resistance on Streptomonas griseus. J Biol Chem. 283(20):13983–13991.

Gan N, et al. 2012. The role of microcystins in maintaining colonies of bloom-forming Microcystis spp. Environ Microbiol. 14(3):730–742.

Ghimire GP, Oh TJ, Liou K, Sohng JK. 2008. Identification of a cryptic type III polyketide synthase (1,3,6,8-tetrahydroxynaphthalene syn-
thesis) from Streptomonas peucetius ATCC 27952. Mol Cells. 26(4):362–367.

Ghosh R, et al. 2008. Dissecting the functional role of polyketide synthases in Dickeyostelium discoideum: biosynthesis of the differentiation regu-

lating factor 4-methyl-5-pentylbenzene-1,3-diol. J Biol Chem. 283(17):11348–11354.

Gilchrist CLM, Chooi YH. 2021. Clinker & clustermap.js: automatic genera-

tion of gene cluster comparison figures. Bioinformatics. Advance Access published January 18, 2021, doi: 10.1093/bioinformatics/ tbat007.

Gu L, et al. 2009. Polyketide decarboxylative chain termination preceded by o-sulfonation in curacin a biosynthesis. J Am Chem Soc. 131(44):16033–16035.

Hai Y, Chen M, Huang A, Tang Y. 2020. Biosynthesis of mycotoxin fusaric acid and application of a PLP-dependent enzyme for chemoenzymatic synthesis of substituted l-pipecolic acids. J Am Chem Soc. 142(46):19668–19677.

Hashimoto M, Ishida S, Seshime Y, Kitamoto K, Fuji i I. 2013. Aspergillus oryzae type III polyketide synthase CsyB uses a fatty acyl starter for the biosynthesis of csypyrone B compounds. Bioorg Med Chem Lett. 23(20):5637–5640.

Hashimoto M, Nonaka T, Fuji i I. 2014. Fungal type III polyketide synthases. Nat Prod Rep. 31(10):1306–1317.

Hashimoto M, Seshime Y, et al. 2013. Identification of csypyrone B2 and B3 as the minor products of Aspergillus oryzae type III polyketide synthase CsyB. Bioorg Med Chem Lett. 23(3):650–653.

Huelsenbeck JP, Ronquist F. 2001. MRBAYES: Bayesian inference of phy-
logenetic trees. Bioinformatics 17(8):754–755.

Humble AV, Gadd GM, Codd GA. 1997. Binding of copper and zinc to three cyanobacterial microcystins quantified by differential pulse po-
larography. Water Res. 31(7):1679–1686.
Makower AK, et al. 2015. Transcriptomics-aided dissection of the intra-cellular and extracellular roles of microcystin in Microcystis aeruginosa. FEBS J. 276(8):2391–2401.

Jenke-Kodama H, Sandmann A, Müller R, Dittmann E. 2005. Evolutionary implications of bacterial polyketide synthases. Mol Biol Evol. 22(10):2027–2039.

Jez JM, Bowman ME, Noel JP. 2001. Structure-guided programming of polyketide chain-length determination in chalcone synthase. Biochemistry 40(49):14829–14838.

Jez JM, et al. 2000. Structural control of polyketide formation in plant-specific polyketide synthases. Chem Biol. 7(12):919–930.

Kang H-S, et al. 2012. Merocyclophanes A and B, antiproliferative cyclophanes from the cultured terrestrial Cyanobacterium Nostoc sp. Phytochemistry 79:109–115.

Katsuyama Y, Ohrishi Y. 2012. Type III polyketide synthases in microorganisms. Methods Enzymol. 515:359–377.

Kolattukudy PE, Fernandes ND, Azad AK, Fitzmaurice AM, Sirakova TD. 2011. Evolutionary implications and classification and phylogenomics. ChemBioChem. 18(1):50–65.

Leao PN, et al. 2013. Antitumor activity of hieridin B, a cyanobacterial secondary metabolite found in both filamentous and unicellular marine strains. PLoS One 8(7):e69562.

Li H, Murphy T, Guo J, Parr T, nalewajko C. 2009. Iron-stimulated growth and microcrystin production of Microcystis novacekii UAM 250. Limnologica. 39(3):255–259.

Li J, Luo Y, Lee JK, Zhao H. 2011. Cloning and characterization of a type III polyketide synthase from Aspergillus niger. Bioorg Med Chem Lett. 21(20):6085–6089.

Luo S, et al. 2014. Carbamidocyclophanes F and G with anti-Mycobacterium tuberculosis activity from the cultured freshwater cyanobacterium Nostoc sp. Tetrohedron Lett. 55(3):686–689.

Makover AK, et al. 2015. Transcriptomics-aided dissection of the intracellular and extracellular roles of microcystin in Microcystis aeruginosa PCC 7806. Appl Environ Microbiol. 81(2):544–554.

Malik S, Sivakumar KC, Soniya EV. 2011. Beta-ketoacyl-acyl carrier protein synthase III (FabH) is essential for bacterial fatty acid synthesis. J Biol Chem. 280(45):37455–37460.

Mayo KM, et al. 2017. Ribocyclophanes A and B, antiproliferative cyclophanes from the cultured terrestrial Cyanobacterium Nostoc sp. CAVN2. Mar Drugs. 14(1):21.

Nakano C, Ozawa H, Akanuma G, Funu N, Horiouchi S. 2009. Biosynthesis of aliphatic polyketides by type III polyketide synthase and methyltransferase in Bacillus subtilis. J Bacteriol. 191(15):4916–4923.

Navarro-Muñoz JC, Collemare J. 2020. Evolutionary histories of type III polyketide synthases in fungi. Front Microbiol. 10: 3018.

Navarro-Munoz JC, et al. 2020. A computational framework to explore large-scale biosynthetic diversity. Nat Chem Biol. 16(1):60–68.

Nualkaew N, et al. 2012. Benzophenone synthase from Garcinia mangostana L. pericarps. Phytochemistry. 77:60–69.

Papendorf O, König GM, Wright AD. 1998. Hieridin B and 2,4-diamino-6-heptadecyl-phenol, secondary metabolites from the cyanobacterium Phormidium ectocarp. with antiplasmodial activity. Phytochemistry 49(8):2383–2386.

Pearson LA, et al. 2016. The genetics, biosynthesis and regulation of toxic specialized metabolites of cyanobacteria. Harmful Algae 54:98–111.

Peric-Concha N, et al. 2005. Ablation of the octC gene encoding a post-polyketide hydroxylase from the oxytetracycline biosynthetic pathway in Streptomyces rimosus results in novel polyketides with altered chain length. J Biol Chem. 280(45):37455–37460.

Preitsch M, Heiden SE, et al. 2016. Effects of halide ions on the carbamidocyclophane biosynthesis in Nostoc sp. CAVN2. Mar Drugs. 14(1):21.

Preitsch M, Niedermeyer TH, et al. 2016. Cylindroclinids A–C, linear cylindrocyclophane-related alkylresorcinols from the cyanobacterium Cylindrospermum stagnale. J Nat Prod. 79(1):106–115.

Rajeshwari KR, Rajashekhar M. 2011. Biochemical composition of seven species of cyanobacteria isolated from different aquatic habitats of Western Ghats, Southern India. Braz Arch Biol Technol. 54(5):849–857.

Ronquiot F, et al. 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst Biol. 61(3):539–542.

Rubin-Pitel SB, et al. 2008. Distinct structural elements dictate the specificity of the type III pentaketide synthase from Neurospora crassa. Chem Biol. 15(10):1079–1090.

Sankaranarayanan R, Saxena P, et al. 2004. A novel tunnel in mycobacterial type III polyketide synthase reveals the structural basis for generating diverse metabolites. Nat Struct Mol Biol. 11(9):894–900.

Satou R, et al. 2013. Structural basis for cyclization specificity of two Azotobacter type III polyketide synthases: a single amino acid substitution reverses their cyclization specificity. J Biol Chem. 288(7):34146–34157.

Saxena P, Yadav G, Mohanty D, Gokhale RS. 2003. A new family of type III polyketide synthases in Mycobacterium tuberculosis. J Biol Chem. 278(45):44780–44790.

Sheshe D, Juvvadi PR, Fujii I, Kitamoto K. 2005. Discovery of a novel superfamily of type III polyketide synthases in Aspergillus oryzae. Biochem Biophys Res Commun. 331(1):253–260.

Sheshe D, Juvvadi PR, Kitamoto K, Ebizuka Y, Fujii I. 2010. Identification of crypyrone B1 as the novel product of Aspergillus oryzae type III polyketide synthase CsyB. Bioorg Med Chem. 18(12):4542–4546.

Sheshe D, Juvvadi PR, Kitamoto K, Ebizuka Y, Nonaka T, et al. 2010. Aspergillus oryzae type III polyketide synthase CsyA is involved in the biosynthesis of 3,5-dihydroxybenzoic acid. Bioorg Med Chem Lett. 20(16):4785–4788.

Shimizu Y, Ogata H, Goto S. 2017. Type III polyketide synthases: functional classification and phylogenomics. ChemBioChem. 18(1):50–65.

Srirakova TD, Dubey VS, Cynamon MH, Kolattukudy PE. 2003. Attenuation of Mycobacterium tuberculosis by disruption of a mas-like gene or a chalcone synthase-like gene, which causes deficiency in dimyocerosyl phthiocerol synthesis. J Bacteriol. 185(10):2999–3008.
Teufel R, et al. 2013. Flavin-mediated dual oxidation controls an enzymatic Favorskii-type rearrangement. Nature 503(7477):552–556.

Thanh NV, et al. 2020. A new [7.7]paracyclophane from Vietnamese marine snail *Planaxis sukatuts* (Born, 1780). Nat Prod Res. 34(2):261–268.

Thompson JD, Higgins DG, Gibson TJ. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22(22):4673–4680.

Tillett D, et al. 2000. Structural organization of microcystin biosynthesis in *Microcystis aeruginosa* PCC7806: an integrated peptide-polyketide synthetase system. Chem Biol. 7(10):753–764.

Timmis KN, Ayliffe MA, Huang CY, Martin W. 2004. Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. Nat Rev Genet. 5(2):123–135.

Vicente CM, et al. 2018. Comparative genomics among closely related *Streptomyces* strains revealed specialized metabolite biosynthetic gene cluster diversity. Antibiotics (Basel, Switzerland) 7:86.

Wu HC, et al. 2012. Chain elongation and cyclization in type III PKS DpgA. ChemBioChem. 13(6):862–871.

Yang J, et al. 2015. The I-TASSER Suite: protein structure and function prediction. Nat Methods. 12(1):7–8.

Yilmaz P, et al. 2014. The SILVA and “All-species Living Tree Project (LTP)” taxonomic frameworks. Nucleic Acids Res. 42(Database issue):D643–D648.

Zeng J, Decker R, Zhan J. 2012. Biochemical characterization of a type III polyketide biosynthetic gene cluster from *Streptomyces toxytricini*. Appl Biochem Biotechnol. 166(4):1020–1033.

Zhanybayeva O, Gogarten JP, Charlebois RL, Doolittle WF, Papke RT. 2006. Phylogenetic analyses of cyanobacterial genomes: quantification of horizontal gene transfer events. Genome Res. 16(9):1099–1108.

Zheng D, Schröder G, Schröder J, Hrazdina G. 2001. Molecular and biochemical characterization of three aromatic polyketide synthase genes from *Rubus idaeus*. Plant Mol Biol. 46(1):1–15.

Zilliges Y, et al. 2011. The cyanobacterial hepatotoxin microcystin binds to proteins and increases the fitness of *Microcystis* under oxidative stress conditions. PLoS One 6(3):e17615.

**Associate editor:** Tal Dagan