nature communications

Article

https://doi.org/10.1038/s41467-025-57145-6

Exploring and expanding the natural chemical space of bacterial diterpenes

Received: 25 October 2024

Accepted: 12 February 2025

Published online: 19 April 2025

Check for updates

Xiuting Wei[1,6], Wenbo Ning[1,6], Caitlin A. McCadden[1], Tyler A. Alsup[1], Zining Li[1,4], Diana P. Łomowska-Keehner[1], Jordan Nafie[2], Tracy Qu[1], Melvin Osei Opoku[1], Glen R. Gillia[1], Baofu Xu[1,3,5], Daniel G. Icenhour[1] & Jeffrey D. Rudolf[1]

Terpenoids are the largest family of natural products but relatively rare in bacteria. Genome mining reveals widespread prevalence of terpene synthases, the enzymes responsible for constructing the hydrocarbon skeletons, in bacteria. Here, we show that 125 (37%) of 334 terpene synthases from 8 phyla, 17 classes, and 83 genera of bacteria are active as diterpene synthases. Isolation and structural elucidation of 28 diterpenes from 31 terpene synthases reveal three previously unreported terpene skeletons, skeletons of natural products from other organisms with unknown biosynthetic pathways, diterpenes that are known in other organisms but to the best of our knowledge not previously seen in bacteria, or new structural and stereochemical isomers of diterpenes. We also identify type I diterpene synthases from myxobacteria and cyanobacteria. This study will help to discover new natural products, advance studies in terpenoid biosynthesis and enzymology, and provide model systems to probe the ecological roles of terpenes.

Terpenoids are the largest and most structurally diverse family of natural products, consisting of over 100,000 known compounds, and are widely used as pharmaceuticals, herbicides, flavors, fragrances, and biofuels[1–5]. Plants, fungi, and marine organisms are well known producers of terpenoids; however, less than 2% of all known terpenoids are of bacterial origin[6]. This disparity is particularly true for diterpenoids, the C20 subclass famous for the anticancer drug Taxol and the plant hormone gibberellin, as less than 1% of diterpenoids are from bacteria (<200 of ~25,000)[1].

Terpene synthases (TSs), the enzymes that begin specialized terpene biosynthesis by converting acyclic and achiral prenyl diphosphates into diverse polycyclic skeletons with numerous stereocenters, are prevalent throughout nature[7,8]. Despite the relatively small numbers of bacterial terpenoids, TSs are widespread in bacterial genomes. In an early genome mining study, over 120 putative bacterial TSs were identified, mostly from 20 actinomycete genomes[9]. A few years later, a hidden Markov model (HMM) identified 262 presumptive TSs from a variety of bacteria[10]. Now, with the ever-growing microbial genomic libraries, a simple search on the UniProt database for the 'terpene cyclase-like 2 family' (IPR034686) reveals over 5000 bacterial proteins. This number does not include other subfamilies of TSs, such as the type II TSs[11] or any of the noncanonical TSs[12]. Overall, this highlights the enormous potential to use TSs to characterize new terpene skeletons, discover new terpenoid natural products, investigate terpene enzymology and evolution, and develop biocatalysts to produce structurally and stereochemically complex hydrocarbons.

TSs act as chaperones that use carbocation chemistry to catalyze complex cyclization reactions[7,8]. They are differentiated into

[1]Department of Chemistry, University of Florida, Gainesville, FL, USA. [2]Biotools, Inc., Jupiter, FL, USA. [3]Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Zhangjiang Hi-Tech Park, Shanghai, China. [4]Present address: Key Laboratory of Drug-Targeting and Drug Delivery System of the Education Ministry and Sichuan Province, Sichuan Engineering Laboratory for Plant-Sourced Drug and Sichuan Research Center for Drug Precision Industrial Technology, West China School of Pharmacy, Sichuan University, Chengdu, China. [5]Present address: Shandong Laboratory of Yantai Drug Discovery, Bohai Rim Advanced Research Institute for Drug Discovery, Yantai, Shandong, China. [6]These authors contributed equally: Xiuting Wei, Wenbo Ning.
e-mail: jrudolf@chem.ufl.edu

https://doi.org/10.1038/s41467-025-57145-6

subclasses based on how they form the initial cation[8,11,12]. Type I TSs use a trinuclear divalent cation cluster to abstract the diphosphate moiety of the substrate and initiate cyclization[8]. Although type I TSs, which are found in plants, fungi, bacteria, and recently corals, have conserved protein folds, they often have little sequence conservation outside of two metal-binding motifs, an Asp-rich motif (e.g., DDxxD) and an NSE/DTE triad (e.g., NxxxSxxxE)[7]. Due to the complexities of terpene cyclization, it is currently impossible to predict the substrate and products of TSs based on protein sequence alone. An additional complicating factor of TS sequence–function relationships is that a single amino acid change, in a well-positioned site, can dramatically alter the product profile of a TS[13]. In this study, we explore and expand the natural chemical space of diterpenes in bacteria by screening 334 uncharacterized TSs for activity.

Results and discussion

Genome mining for bacterial terpene synthases

We set out to map the natural chemical space of bacterial diterpenes by screening a library of putative TSs. There are 152 characterized canonical type I TSs from bacteria: >78% (119) are from the Actinomycetota phylum, 36 are diterpene synthases (di-TSs) that act on GGPP (as opposed to di-TSs that act on precyclized diphosphates, e.g., *ent*-copalyl diphosphate), and only 27 are associated with known natural products. We first investigated the overall distribution of putative TSs in bacteria. Using the Enzyme Function Initiative Taxonomy Sunburst tool[14], there are 133,434 proteins that fall into the 'isoprenoid synthase domain superfamily' (IPR008949) in the UniProt database. Narrowing the search to IPR034686, a type I TS subfamily that is mainly found in bacteria and fungi, we identified 5802 putative TSs. Of the >40 formally recognized bacterial phyla[15], 13 include TSs from the IPR034686 subfamily with Actinomycetota the most significant representative with 79.5% of the proteins; other phyla with >1% include Bacteroidota (8.6%), Pseudomonadota (5.0%), Myxococcota (3.2%), and Cyanobacteriota (2.6%) (Fig. 1A, B). When the total number of genomes of each phylum is considered, Actinomycetota, Myxococcota, and Cyanobacteriota have the highest number of TSs per genome (Supplementary Data 1). Interestingly, there are no type I di-TSs known from myxobacteria or cyanobacteria (Fig. 1B).

Using a combination of genome mining, phylogenetics, sequence similarity networking, and biosynthetic gene cluster (BGC) analysis, we selected 334 TSs from 8 phyla, 17 classes, and 83 genera of bacteria for functional analysis (Fig. 1C, D, Supplementary Fig. 1). The phylogenetic

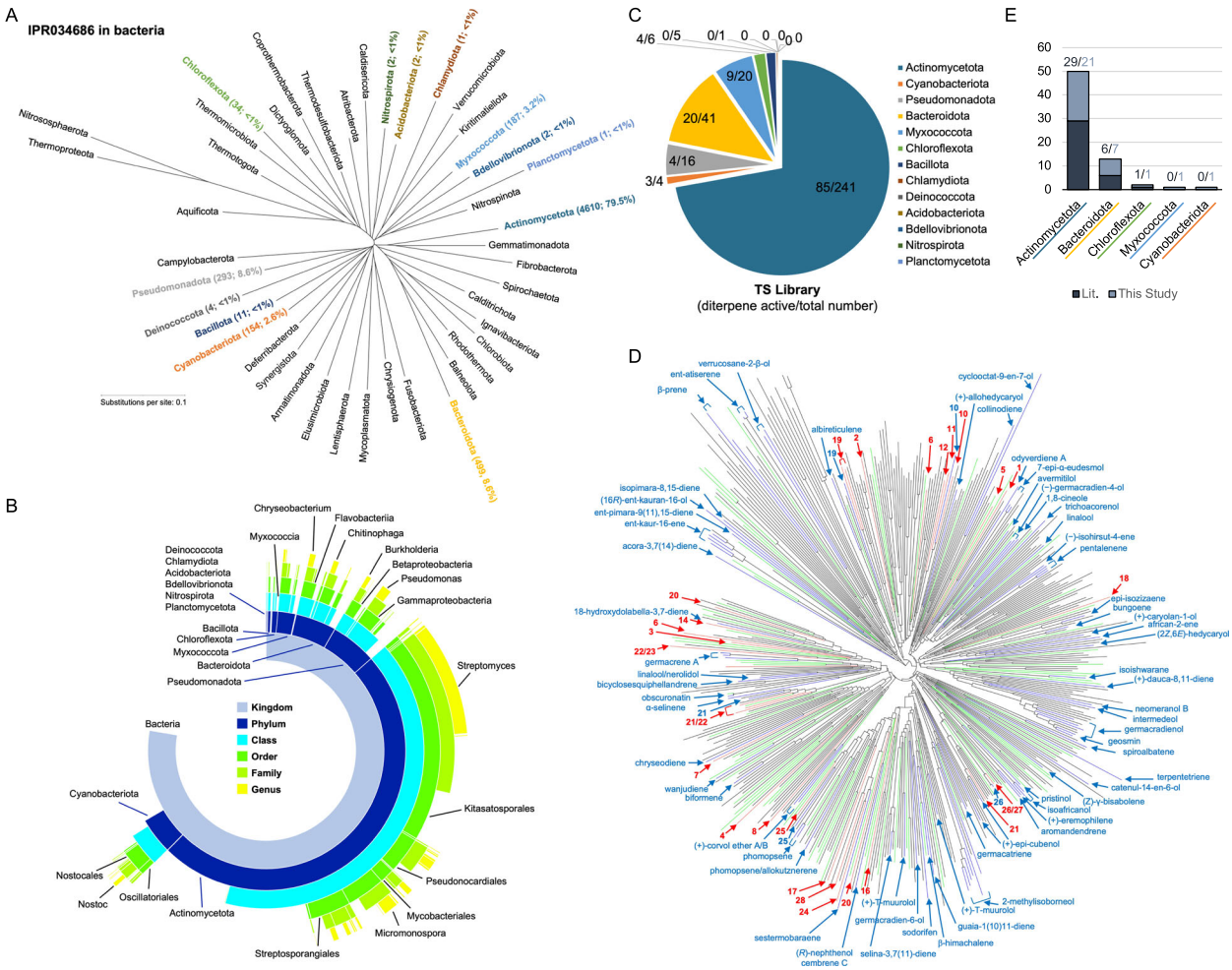

Fig. 1 | Terpene synthases in bacteria. A TSs from the subfamily IPR034686 are found in 13 bacterial phyla. The phylogenetic tree of 43 phyla was created using 16S rRNA of representative bacteria (Supplementary Table 32). Phyla with at least one putative TS from IPR034686 are bolded and colored with the total number and overall percentage listed. **B** A sunburst chart highlighting the phylogenetic breakdown of TSs within the 13 bacterial phyla. Selected taxonomic descriptions are listed. **C** The overall TS library described in this study categorized by phyla (color-coded as in **A**). The numbers in the pie chart represent the number of TSs that showed di-TS activity and the total number within each phylum. **D** Phylogenetic analysis of type I TSs from bacteria showing broad distribution of di-TSs. The blue lines represented previously characterized TSs, red lines are the 31 TSs characterized in this study, green lines are the other 94 TSs that showed di-TS activity but not studied here, and black lines are the remaining 209 TSs in our bacterial TS library. A larger tree is shown in Supplementary Fig. 1. **E** The total count of di-TSs from bacteria reported in the literature and the number of di-TSs characterized in this study.

https://doi.org/10.1038/s41467-025-57145-6

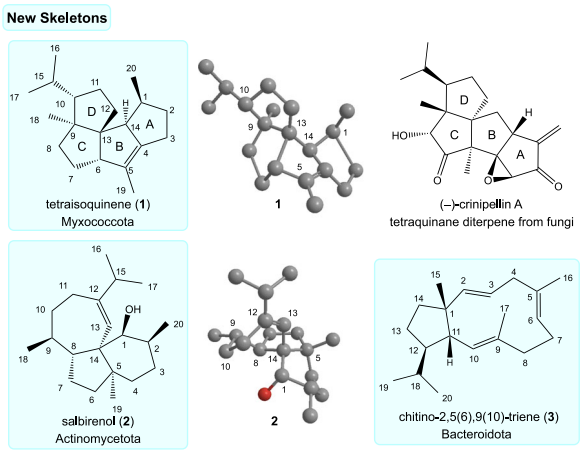

Fig. 2 | Previously unreported skeletons (1–3) isolated from three different phyla of bacteria. 3D representations of **1** and **2** are shown. The phyla encoding the responsible TSs are listed. The fungal crinipellins are structurally the closest known skeleton to **1**.

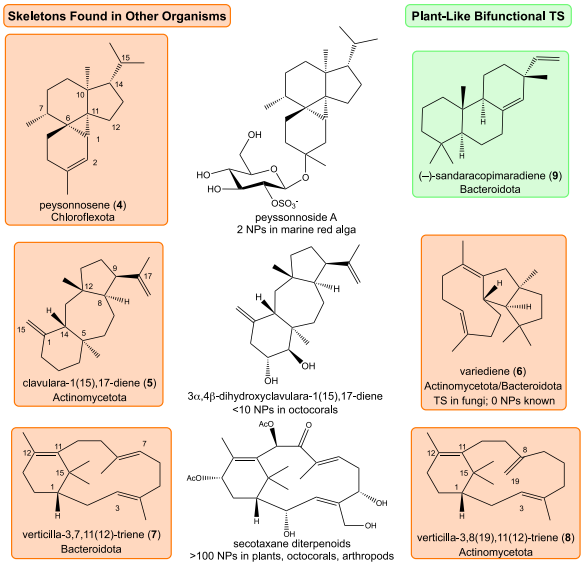

Fig. 3 | Diterpene skeletons (4–9), known in other organisms, now identified in bacteria. Bacterial TSs are the first known to form **4** and **5**; TSs in fungi and plants were known to produce **6** and **9**, respectively. The phyla encoding the responsible TSs are listed.

distribution of selected TSs was intentionally broad and paralleled the overall phylogenetic representation (Fig. 1B, C). To expedite functional characterization, we used an *E. coli* heterologous expression system to detect products without the need for protein purification or substrate synthesis. We synthesized the 334 TSs as codon-optimized genes and cloned them into an engineered *E. coli* strain that overproduces geranylgeranyl diphosphate (GGPP)[16]. Terpene products were initially detected via TLC or HPLC of organic extracts from the TS library in *E. coli*. In a first-pass screen, with no optimization of the genetic system, expression, or culture conditions, we identified 125 positive hits (37%, Fig. 1C). We then prioritized 31 TSs for large-scale fermentations for diterpene isolation and structural elucidation (Fig. 1E, Supplementary Fig. 2). TSs were chosen based on sequence and phylogenetic diversity (Supplementary Data 2), location in a unique BGC (Supplementary Data 3), and/or high initial yields. We isolated 28 diterpenes and determined their structures using NMR, GC-MS, and vibrational circular dichroism (VCD). We organized the bacterial diterpenes into four categories below, although several diterpenes fit into more than one category.

Discovery and structural elucidation of three previously unreported diterpene skeletons

Tetraisoquinene (**1**), a diterpene with a 5/5/5/5-fused tetracyclic skeleton, was produced by a TS (TiqS) from the myxobacterium *Melittangium boletus* (Fig. 2). Myxobacteria are well-established producers of natural products but only a few diterpenoids have been isolated from the phylum[6]. To the best of our knowledge, this is the first di-TS characterized from myxobacteria. The GC-MS of **1** showed a molecular ion peak at *m/z* 272.2499 (Supplementary Fig. 3), supporting a molecular formula of $C_{20}H_{32}$ and a diterpene with five degrees of unsaturation. Using 1D and 2D NMR (Supplementary Note 1, Supplementary Table 1, Supplementary Figs. 4–11), 1 was determined to possess an angular tetraquinane skeleton. As diterpenes are particularly well suited for analysis by vibrational circular dichroism (VCD), a powerful technique for the determination of absolute configuration in solution phase without derivatization[17–19], we used VCD to determine the absolute configuration of **1** to be 1*S*,6*S*,9*S*,10*S*,13*S*,14*S* (Supplementary Fig. 12). The only other tetraquinane diterpenes from nature are the crinipellins, basidiomycete fungal diterpenoids with antibacterial, antifungal, anticancer, and anti-inflammatory properties[20]. However, the connections between the A and B rings of **1** and the crinipellins are not the same (Fig. 2). To the best of our knowledge, the TS responsible for tetraquinane formation is unknown, and no biosynthetic studies have been reported; only chemical synthetic methods to access the crinipellin skeleton are known[21].

Salbirenol (**2**), isolated from *E. coli* expressing SalS from *Streptomyces albireticuli* (Actinomycetota), is a diterpene alcohol with a 7/5/6-tricyclic skeleton (Fig. 2). The GC-MS (*m/z* 290.2607; Supplementary Fig. 13) and NMR spectra (Supplementary Note 1, Supplementary Table 2, Supplementary Figs. 14–20) of **2** supported its planar structure. Using VCD in comparison with the calculated spectra of nearly all 32 possible stereoisomers, the absolute configuration of **2** was determined to be 1*S*,2*S*,5*R*,8*R*,9*S*,14*S* (Supplementary Fig. 21).

Chitino-2,5(6),9(10)-triene (**3**), a 5/11-bicyclic diterpene, was one of the major products of ChtS from *Chitinophaga japonensis* (Bacteroidota; Fig. 2). The GC-MS (*m/z* 272.2498; Supplementary Fig. 22) and NMR spectra of **3** (Supplementary Note 1, Supplementary Table 3, Supplementary Figs. 23–29) revealed its bicyclic skeleton. After determining its relative configuration by 2D NOESY, the absolute configuration of **3** was determined to be 1*R*,11*S*,12*R* by VCD (Supplementary Figs. 30–31). At first glance, **3** appears to be a dolabellatriene; however, the positions of its methyl groups on the 11-membered ring confirms it is a previously unreported skeleton, one that has no corresponding natural products in nature, and is likely formed via a different cyclization mechanism than the dolabellatrienes.

Bacteria produce diterpene skeletons found in other organisms

The next five diterpenes are all known skeletons but have not been previously seen in bacteria, to the best of our knowledge, and either do not have characterized TSs responsible for their formation, associated natural products, or both. Peysonnosene (**4**) was identified as the major product of a TS (PeyS) from the chloroflexota bacterium *Anaerolineaceae* sp. De novo structural elucidation by GC-MS, 1D and 2D NMR spectroscopy (Supplementary Note 1, Supplementary Table 4, Supplementary Figs. 32–39) revealed that **4** shares its 5/6/3/6-fused tetracyclic skeleton with the peyssonnosides, unusual diterpene glycosides isolated from the marine red alga *Peyssonnelia* sp. (Fig. 3). The relative configuration of **4**, which was supported by limited NOESY correlations, was proposed to be the same as that of the peyssonnosides. We confirmed its absolute configuration, 1*S*,6*S*,7 *R*,10*S*,11*S*,14*S*, by VCD (Supplementary Fig. 40). The peyssonnosides have low mM activities against MRSA, the malarial parasite *Plasmodium berghei*, and

the marine fungus *Dendryphiella salina*, with no cytotoxicity against human keratinocytes[22]. While the total synthesis of the peyssonnosides has been completed[23,24], there are no biosynthetic reports of these natural products. Interestingly, *Anaerolineaceae* bacteria are known to occur in marine sediments[25], providing a potential link between their diterpene production and algal natural products.

Clavulara-1(15),17-diene (**5**) is a 6/7/5 tricycle, produced by CvdS from *Nocardia vulneris* (Fig. 3). The structure of **5**, including its absolute configuration, was determined by GC-MS, NMR, and VCD (Supplementary Note 1, Supplementary Table 5, Supplementary Figs. 41–50). We named **5** after a small family of coral diterpenoids, the clavularanes, that share the same planar skeleton (Fig. 3)[26]. However, the presumed biosynthesis of the third ring in the clavularanes, and related dolastanes, stems from oxidized dolabellanes (5/11-bicyclic skeletons)[27,28] and no coral TSs are yet known to produce either the dolabellane or clavularane skeletons[29].

Variediene (**6**), a 5/5/9 tricyclic diterpene, was found from two distinct bacterial TSs (Fig. 3, Supplementary Table 6, Supplementary Figs. 51–57). *Ps*VS from *Prauserella shujinwangii* (Actinomycetota) and *Od*VS from *Olivibacter domesticus* (Bacteroidota) share only 20% sequence identity. Three fungal chimeric TSs, containing both polyprenyl synthase and cyclase domains, from ascomycetes are known to produce variediene[30–32]. To the best of our knowledge, this is the first report of a bacterial variediene synthase and they share essentially no sequence identity (<20%) with any of the fungal versions. The genuine natural products from the fungal variediene BGCs are still unknown.

The verticillenes are 6/12-bicyclic diterpenes that share a biosynthetic relationship with the taxane and phomactin families of natural products[33]. Verticillene natural products are known in plants, corals, and insects, and verticillene isomers have been isolated from taxadiene synthase engineered variants[34,35]. Here, we identified two verticillene synthases, *Cn*VrtS from *Chitinophaga niabensis* and *S*VrtS from *Streptomyces* sp. TLI_235 produce 1*S*-verticilla-3,7,11(12)-triene (**7**) and 1*S*-verticilla-3,8(19),11(12)-triene (**8**), respectively (Fig. 3, Supplementary Tables 7–8, Supplementary Figs. 58–73). The NMR of **7** matched previous literature, confirming its structure. The NMR of **8**, in comparison with that of **7**, clearly showed its exocyclic C-8-C-19 olefin. The absolute configuration of **8** was confirmed by VCD (Supplementary Fig. 73).

While genome mining for unique TSs from bacteria, we identified a putative tridomain, bifunctional TS from *Chitinophaga japonensis*. TSs with both type I and II TS activity, are prominent in plants and fungi[36], and while they have been evolutionarily proposed in bacteria[37], there were no known examples when we began this study. The longer than typical length, 785 amino acids, piqued our interest in this TS, which was previously annotated as a prenyltransferase/squalene oxidase-like protein but showed similarity to plant bifunctional TSs. Expression in *E. coli* yielded (-)-sandaracopimaradiene (**9**, Fig. 3, Supplementary Table 9, Supplementary Figs. 74–80), clearly indicating type II activity forming *n*-copalyl diphosphate from GGPP and subsequent type I cyclization to **9**. There are known sets of discrete type II and type I TSs from Actinomycetota that form (iso)pimaradienes[38], and at least one bacterial diterpenoid, gifhornenolone B[39], formed from this skeleton. At the time of our discovery, no natural bifunctional TS had been identified in bacteria. During preparation of this manuscript, the function of this TS, named *Chj*DCS, was reported[40]. This finding opens the door to understanding TS evolution in terrestrial organisms.

Isomers of known skeletons are prevalent in bacteria

We identified several diterpenes with known planar skeletons that have alkenes at different positions or are diastereomers of previously known diterpenes (Fig. 4). Two *Streptomyces* spp. TSs, *Sa*VenA and *Ss*VenA, produce the 5/5/6/7-tetracyclic scaffold seen in venezuelaene A (**10**)[41]. With sequence identities of ~58% to venezuelaene A synthase (VenA), it is not surprising that *Sa*VenA forms **10** (Supplementary Table 10, Supplementary Figs. 81–84) and *Ss*VenA forms the nearly identical venezuelaene A2 (**11**; Supplementary Table 11, Supplementary Figs. 85–94). Interestingly, another homologous TS from *Streptomyces alkaliterrae*, *Salk*S, with 49% sequence identity to VenA, forms the 5/7/6-tricyclic odyverdiene B2 (**12**; Supplementary Table 12, Supplementary Figs. 95–104). Similar diterpenes were seen in *Streptomyces* sp. ND90[42], but that di-TS is less similar with only 24% identity to *Salk*S. The absolute configuration of **12**, which was determined by VCD (Supplementary Fig. 104), is likely not the same as the previously reported odyverdiene B, whose absolute configuration was not reported[42], given differences in their NMR spectra.

Dolabellatrienes, the 5/11-bicyclic precursors to the dolabellane family of natural products known in algae, fungi, marine invertebrates[27], are rare in bacteria. Only a few are known and are often implicated as intermediates in more complex cyclization cascades[43]. *Sa*VenA, the venezuelaene A (**10**) synthase described above also yields 1*S*,3*E*,7*E*,11*R*,12*S*-dolabella-3,7,18-triene (**13**; Supplementary Table 13, Supplementary Figs. 105–107), previously reported as a shunt product from several VenA mutated variants[44]. *Cb*DS, from *Chitinophaga barathri*, forms **14** (Supplementary Table 14, Supplementary Figs. 108–115), a diastereomer of **13**; a relative configuration of 1*R**,11*S**,12*S** for **14** is supported by NOESY. Finally, we identified di-TS from cyanobacteria: *Sh*DS from *Scytonema hofmannii*, which was shown to form 1*R*,7*E*,10*E*,12*R*-dolabella-4(16),7,10-triene (**15**; Supplementary Table 15, Supplementary Figs. 116–124), a possible precursor to the coral diterpenoids stolonitriene and stolonidiol[45].

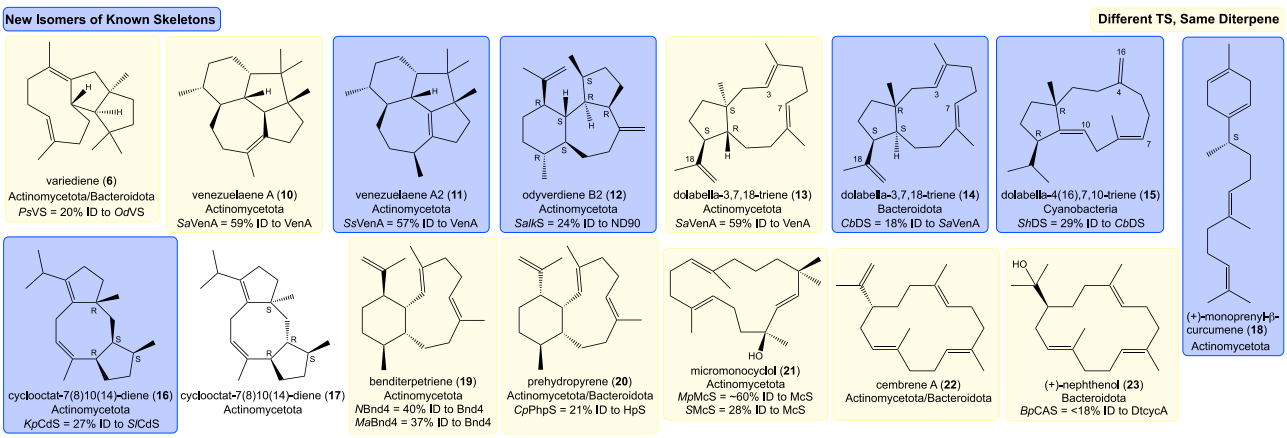

Fig. 4 | Structural and stereochemical isomers of bacterial diterpenes (10–27). The phyla encoding the responsible TSs and protein sequence identities to known TSs or TSs in this study are listed. Previously unreported and reported diterpenes are shown in blue and yellow boxes, respectively.

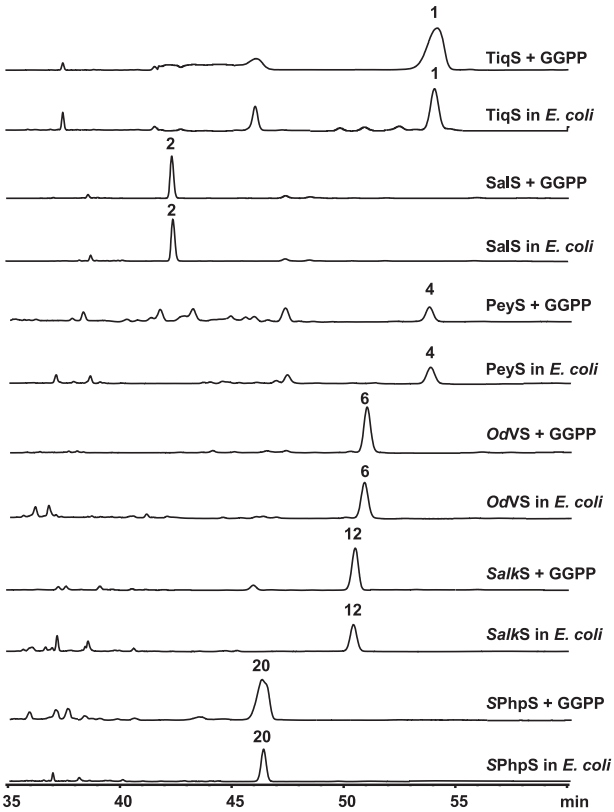

Fig. 5 | HPLC analysis of in vitro reactions of six TSs with GGPP in comparison to extracts obtained from expressing these genes in our GGPP-production system in *E. coli*. Absorbance was measured at 210 nm.

*Kp*CdS from *Kibdelosporangium phytohabitans* (Actinomycetota), with 19%–28% identities to the dolabellatriene synthases above, formed a 5/8/5-tricyclic skeleton that likely originates from a dolabellyl intermediate. The structure of this diterpene (**16**; Supplementary Table 16, Supplementary Figs. 125–134) matched that of the previously reported cyclooctat-7(8),10(14)-diene (**17**) from *Streptomyces lactacystinaeus* OM-6159[42], but its NMR was not identical. Since the absolute configuration of **17** was not reported, we obtained *slt18_1078*, which encodes this di-TS (here named *Sl*CdS), and isolated **17** (Supplementary Table 17, Supplementary Figs. 135–143). Comparison of the NOESY data and VCD analysis concluded their absolute configurations: 2*S*,3*S*,6 *R*,7*Z*,10*Z*,11*R*-**16** from *Kp*CdS and 2 *R*,3*S*,6 *R*,7*Z*,10*Z*,11*S*-**17** from Slt18_1078. These two enzymes share 27% identity yet construct diastereomeric 5/8/5-tricyclic diterpenes.

(+)-Monoprenyl-b-curcumene (**18**) was produced by *St*MpcS from *Saccharopolyspora terrae* (Supplementary Table 18, Supplementary Figs. 144–150). Monocyclic **18** is the C20 version of β-curcumene produced by plant sesqui-TSs and mutants of *epi*-isozizaene synthase[46,47] and tetraprenyl-β-curcumene from YtpB in *Bacillus*[48]. *St*MpcS shares 30% sequence identity with *epi*-isozizaene synthase, but no significant sequence similarity with the plant sesqui-TSs or the non-canonical "large TS" YtpB. Tetraprenyl-β-curcumenes are often precursors for a 2nd cyclization reaction, but the BGC encoding *St*MpcS does not have a type II TS nearby suggesting different downstream tailoring modifications (Supplementary Data 3).

Distinct TSs construct the same diterpenes

As described above for the variediene (**6**) synthases from different bacterial phyla and fungi, it is not unprecedented for TSs with disparate sequences, often from different phyla or genera, to form identical products. Here, we list several examples of known diterpenes isolated in this study (Fig. 4).

Benditerpetriene (**19**), a 2*E*,6*E*-*cis*-eunicellane, which is known in benditerpenoic acid and aridacin biosynthesis[49,50] was produced by two additional TSs (*N*Bnd4, *Ma*Bnd4) from *Nocardia* sp. SYP-A9097 and *Mycobacteroides abscessus* (Supplementary Table 19, Supplementary Figs. 151–153). *N*Bnd4 and *Ma*Bnd4 share 40% and 37% sequence identity with Bnd4, respectively, and while *N*Bnd4 is found in a BGC similar to that of benditerpenoic acid and the aridacins[49], it is intriguing to consider why the pathogenic *M. abscessus* encodes a eunicellane synthase.

Prehydropyrene (**20**), a diastereomer of benditerpetriene (**19**), was found to be the major product of *S*PhpS and *Cp*PhpS (Supplementary Table 20, Supplementary Figs. 154–160). Although **20** is known as a neutral intermediate in the biosynthesis of hydropyrene[51], to the best of our knowledge, this is the first report of a prehydropyrene synthase. *S*PhpS, from *Streptomyces* sp. TLI_053, shares no significant sequence identity with hydropyrene synthase (HpS) from *Streptomyces clavuligerus*; *Cp*PhpS, from *Chryseobacterium populi* (Bacteroidota), shares 21% sequence identity with HpS.

Micromonocyclol (**21**), a diterpene alcohol with a rare 15-membered ring originally identified from McS from *Micromonospora marina*[52], was rediscovered from three TSs (Supplementary Table 21, Supplementary Figs. 161–166). *Mp*McS and *Mb*McS are from two *Micromonospora* spp. with 62% and 61% sequence identities to McS; *S*McS is from *Streptomyces* sp. CB01580 and only shares 28% identity.

Several 14-membered monocyclic diterpenes were identified from bacterial TSs. *Mp*McS and *Mb*McS, as well as *Ml*CAS from *Micromonospora* sp. Liam0 and *Bp*CAS from *Belliella pelovolcani* (Bacteroidota) produced cembrene A (**22**; Supplementary Table 22, Supplementary Figs. 167–173); *Bp*CAS also formed (+)-nephthenol (**23**; Supplementary Table 23, Supplementary Figs. 174–176) but does not show any significant identity (<18%) to DtcycA, a producer of **23**[53]. *Ny*ICS, from *Nocardia yunnanensis*, forms isocembrene C (**24**; Supplementary Table 24, Supplementary Figs. 177–182).

Spiroviolene (**25**), a spirocyclic triquinane diterpene from *Streptomyces violens* and known in fungi[54,55], was produced by *S130*SvS, a TS with 50% ID to SvS from *S. violens* (Supplementary Table 25, Supplementary Figs. 183–189). Spata-13,17-diene (**26**) and cneorubin Y (**27**) were found from a TS (SpS) with high homology to the known spatadiene synthase (Supplementary Tables 26–27, Supplementary Figs. 190–194)[56]. In some cases, we identified the GGPP elimination product b-springene (**28**), such as from the *Streptomyces* sp. Tü2975 TS (Supplementary Table 28, Supplementary Figs. 195–200). We hypothesized that this is not a di-TS and should be screened for preference to other prenyl diphosphates. Gratifyingly, this TS was recently characterized as the sesterterpene synthase, StSS, responsible for producing sesterviolene[57].

Diterpene products verified by in vitro characterization

To verify that the TSs are catalyzing the same reactions in vitro as they do in our heterologous expression system, we subcloned TS genes encoding TiqS, SalS, PeyS, *Od*VS, *Salk*S, and *S*PhpS with N-terminal His_6 tags and purified them from *E. coli* (Supplementary Fig. 201). For each of the six TSs, the major products matched with those produced in *E. coli* (Fig. 5), supporting that the isolated products are genuine bacterial diterpenes. It remains to be seen if expression in their native hosts yields identical compounds; although recent characterization of AlbS, a *trans*-eunicellane synthase from *Streptomyces albireticuli*, in vitro and via heterologous expression in both *E. coli* and *Streptomyces albireticuli* confirmed the same enzymatic product[58,59].

In conclusion, we used genome mining to identify and screen a large library of bacterial TSs for diterpene activity. This screen nearly doubles the number of characterized bacterial diTSs. It is highly probable that many of the remaining TSs are active but are selective for different prenyl lengths (i.e., monoterpene, sesquiterpene, or sesterterpene synthases); these studies are being conducted now.

Libraries like this one will be essential for advancing our understanding of how these enzymes spatially and stereochemically control cyclization and future approaches to predicting TS function.

It is evident that the bacterial terpenome[6], the collection of genomically-encoded terpenoid natural products, is much larger than previously thought. Diterpenes that are either previously unobserved in nature or those seen, to the best of our knowledge, for the first time in bacteria are likely precursors to more complex diterpenoid natural products. We characterized type I diterpene synthases from myxobacteria and cyanobacteria and also identified functional TSs in organisms, including human pathogens, that were not known to produce terpenoids. It is also tempting to speculate the possibility that bacterial symbionts may be producing some of the known terpenoids in eukaryotic systems[60]; however, the discovery of coral, and very recently arthropod, TSs support that at least some, if not all, are eukaryotic natural products[61–63]. Very little is known about the ecological roles of bacterial terpenoids, although recent studies suggest volatile terpenes enable interspecies communication[64,65]. Finding similar compounds in multiple phyla or organisms may provide insights into some of these difficult-to-answer questions.

Methods

General experimental procedures

All ^{1}H, ^{13}C, and 2D NMR (^{1}H-^{13}C HSQC, ^{1}H-^{1}H COSY, ^{1}H-^{13}C HMBC, and ^{1}H-^{1}H NOESY) experiments were run in $CDCl_3$ or C_6D_6 on a Bruker AVANCE III Ultrashield 600, Bruker AVANCE III HD 600, or Bruker AVANCE III 800. All NMR chemical shifts were referenced to residual solvent peaks or to $Si(CH_3)_4$ as an internal standard. Terpene production was monitored by thin-layer chromatography (TLC) and high-performance liquid chromatography (HPLC). TLC was performed with 0.25 mm silica gel plates (60 F_{254}) using short-wave UV light to visualize, and I_2 or $KMnO_4$ and heat as developing agents. HPLC was performed on an Agilent 1260 Infinity LC equipped with a Restek Roc-C18 column (150 mm × 4.6 mm, 5 μm). Preparative HPLC was carried out on an Agilent 1260 Infinity LC equipped with an Agilent Eclipse XDB-C18 column (250 mm × 21.2 mm, 7 μm). GC-MS analysis was run using a Thermo Scientific Orbitrap Exploris spectrometer with a Rxi-5MS column (Restek Corp, 30 m × 0.25 mm i.d. and 0.251 μm film). Optical rotations were measured using a JASCO P-2000 polarimeter. IR data were measured by PerkinElmer Spectrum Two FT-IR Spectrometer. VCD data were collected on a BioTools, Inc. (Jupiter, FL) ChiralIR 2X Dual PEM FT-VCD spectrometer.

Bacterial strains, plasmids, and chemicals

Strains, plasmids, and PCR primers used in this study are listed in Supplementary Tables 29–31. All genes encoding bacterial TSs screened in this study were synthesized by the U.S. Department of Energy Joint Genome Institute as part of a User Proposal for Functional Genomics (part of proposal: 10.46936/10.25585/60008123). Genes were codon-optimized for *E. coli* and cloned into pET28a via Gibson assembly. PCR was conducted by PCR primers were obtained from Sigma-Aldrich. Q5 high-fidelity DNA polymerase and restriction endonucleases were purchased from NEB and used by following the protocols provided by the manufacturers. DNA gel extraction and plasmid preparation kits were purchased from Omega Bio-Tek. DNA sequencing was conducted by Genewiz. Common chemicals and media components were purchased from standard commercial sources.

Bioinformatics

The phylogenetic trees were constructed via the phylogenetic tree builder function in MEGA11 using Muscle for multiple sequence alignment, neighbor joining algorithm for tree creation, and bootstrap analysis using 1000 replicates[66]. Bootstrap values, ranging from 0 (near the root) to 1 with half of the nodes >0.3, were removed from the figures for clarity. The tree of bacterial phyla (Fig. 1A) was built from a multiple sequence alignment of the 16S rRNA sequences of 42 bacterial phyla (Supplementary Table 32). The tree of bacterial TSs was constructed using the amino acid sequences of 377 uncharacterized bacterial TSs and 141 characterized bacterial di-TSs from the literature. The phylogenetic trees were then visualized and annotated using iTOL[67]. Each TS sequence was additionally manually analyzed to identify the presence of highly conserved type I TS motifs (Supplementary Data 2). Protein sequence alignment was performed using ClustalW (http://www.clustal.org/) and the results were displayed using outputs created with ESPript 3.0 (http://espript.ibcp.fr/ESPript/ESPript/).

Terpene synthase screen

Plasmids harboring synthesized TS genes in pET28a were isolated using NEB Turbo *E. coli* grown in lysogeny broth (LB) with kanamycin (50 mg L^{-1}) following manufacturer's protocols. Each plasmid was individually transformed into *E. coli* BL21 Star (DE3) competent cells already containing pJR1064b, a pCDF-Duet vector harboring genes that produce GGPP in vivo via an alternative isoprenoid precursor pathway[58]. Transformants were cultivated in LB containing kanamycin (50 mg L^{-1}) and streptomycin (50 mg L^{-1}). The cells were cultured overnight were subsequently inoculated into 50 mL of Terrific Broth (TB) media at 37 °C with shaking at 250 rpm until reaching an optical density at 600 nm (OD_{600}) of 1.0. Isopropyl β-D-1-thiogalactopyranoside (IPTG, 0.5 mM final concentration) and isoprenol (4 mM, final concentration) were added to the culture and the culture was shaken at 250 rpm for 48 h at 28 °C. For analysis, 1 mL of culture was extracted using 0.5 mL of acetonitrile, saturated with solid NaCl, after vigorous vortexing and centrifugation at 21,300 × *g* for 5 min. The organic layer, containing the terpene products of each TS, was initially analyzed by TLC or HPLC. For HPLC analysis, products were detected at 210 nm using a linear gradient with flow rate of 1 mL min^{-1}: 5% acetonitrile/water (0–5 min); 5% to 95% acetonitrile/water (5–35 min at 5% min^{-1}); 95% acetonitrile/water (hold 35 min) at 35 °C.

Isolation and purification of terpenes

Large-scale fermentations of each target TS were conducted as described above, but in 6–12 L. To isolate the diterpene products, cultured cells were harvested by centrifugation at 5000 × *g* for 10 min at 25 °C and subsequently transferred to a glass beaker. The cell pellets were extracted with acetone (2× volume of the pellet) and the organic phase was extracted with the same volume of hexanes three times. The hexanes extractions were combined and concentrated in vacuo at room temperature. The resulting extract was redissolved in hexanes and purified by silica chromatography, employing an isocratic hexanes mobile phase. Fractions containing the target products were combined and subjected to further purification using preparative LC. For preparative LC, products were detected at 210 nm using a linear gradient with flow rate of 20 mL min^{-1}: 5% acetonitrile/water (0–5 min); 5% to 95% acetonitrile/water (5–15 min at 9% min^{-1}); 95% acetonitrile/water (hold 65 min). Isolated yields and spectroscopic data of **1**–**28** are reported in Supplementary Note 2 and 3, respectively.

GC-MS analysis

For GC-MS analysis, purified diterpenes were dissolved in hexanes at a concentration of 0.1 mg mL^{-1}. The source, transfer line, and injection port were set to 250 °C, respectively, and the carrier gas flow rate was set at 1.2 mL min^{-1}. Products were measured with an electron ionization of 70 eV and mass scan range was from *m/z* 30–500 @ 30000 resolution with a temperature gradient as follows: 50 °C (0–3 min), ramp to 280 °C @ 10 °C min^{-1} (hold 5 min). Kovats retention index (RI) values were calculated for all products in this study using the GC-MS conditions above in comparison with C7–C30 saturated alkanes (Sigma).

VCD measurements and calculations

To a small vial containing 2.5–44 mg of diterpene was added 100–125 µL of $CDCl_3$. $CDCl_3$ (Cambridge Isotope Labs Silver Foil) used to dissolve diterpenes was run through a small plug of activated basic alumina immediately before use. The resulting solution was transferred to a liquid IR cell (BaF_2, 100 µm cell path) and placed in the measurement chamber. The spectrometer was set to 4 cm^{-1} resolution with PEM (both 1 and 2) maximum frequency set to 1400 cm^{-1}. The sample was then measured for 4–18 h in 1-h blocks. The IR data from the first block were solvent and water vapor subtracted, then offset to zero at 2000 cm^{-1}. The VCD data blocks were averaged, and from this was subtracted a previously measured solvent baseline (18-h avg.). Finally, the VCD spectrum was offset to zero at 2000 cm^{-1}. The VCD noise data were block averaged and used without further processing. For detailed methods on VCD calculations and calculation data, see Supplementary Method 1 and Supplementary Data 4, respectively.

In vitro terpene synthase assays

For each in vitro enzyme reaction, 40 µM TS was incubated at 37 °C in 50 mM Tris-HCl, pH 8.0, containing 10 mM GGPP and 10 mM $MgCl_2$ at a total volume of 100 µL. The reactions were incubated for 1 h at 37 °C, quenched with equal volume of acetonitrile, and saturated with solid NaCl. After separating the two phases by vortexing, the organic phase was analyzed by HPLC and/or GC-MS. For HPLC and GC-MS analysis, the methods described above were used. For detailed methods on cloning design for protein production and protein purification, see Supplementary Method 2 and 3, respectively.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The protein sequences used in this study are available in the UniProt or NCBI databases under accession code A0A250IBE0 (TiqS), A0A1Z2LDZ7 (SalS), A0A562T706 (ChtS), A0A2E1AC30 (PeyS), A0A0C1AYW6 (CvdS), A0A1H7UDD2 (*Od*VS), A0A2T0LNX7 (*Ps*VS), A0A1N6E2E2 (*Cn*VrtS), A0A2A3HMH3 (*S*VrtS), A0A562SP41 (*Chj*DCS), A0A1I1IX82 (*Sa*VenA), A0A7M3M6K6 (*Ss*VenA), A0A5P0YYU2 (*Salk*S), A0A3N4M544 (*Cb*DS), A0A139WXP3 (*Sh*DS), A0A0N9I8A9 (*Kp*CdS), A0A097ZQD0 (*Sl*CdS), A0A4R4VDA4 (*St*MpcS), A0A6I2FRA3 (*N*Bnd4), WP_186376445 [https://www.ncbi.nlm.nih.gov/protein/WP_186376445.1/] (*Ma*Bnd4), A0A1H1T2Q7 (*S*PhpS), J2T8D5 (*Cp*PhpS), A0A495JMV3 (*Mp*McS), A0A0M8XX89 (*Mb*McS), A0A1Q5GJH1 (*S*McS), A0A3N1A8E7 (*Ml*CAS), A0A1N7KE47 (*Bp*CAS), A0A386ZKV4 (*Ny*ICS), A0A552RHR4 (*S130*SvS), A0A3M8F3X6 (SpS), A0A6G9FDK2 (*St*SS). Plasmids generated in this study are available upon request. The raw NMR data of **1**–**8**, **10**–**20**, **22**, **23**, and **25** were deposited in the Natural Product Magnetic Resonance Database under accession codes NP0341974, NP0342000, NP0342001, NP0342002, NP0342003, NP0342004, NP0342005, NP0342006, NP0342007, NP0342008, NP0342009, NP0342010, NP0342011, NP0342012, NP0342013, NP0342014, NP0342015, NP0342016, NP0342017, NP0350556, NP0080347, and NP0016132. The GC-MS fragmentation data of **1**–**28** were deposited in the Mass Spectra for Chemical Ecology database [https://leopard.tu-braunschweig.de/servlets/MCRFileNodeServlet/dbbs_derivate_00054495/MACE_r05.mace].

References

1. *Dictionary of Natural Products*. Accessed 15 August 2024. http://dnp.chemnetbase.com.
2. Yan, Y. et al. Resistance-gene-directed discovery of a natural-product herbicide with a new mode of action. *Nature* **559**, 415–418 (2018).
3. Avalos, M. et al. Biosynthesis, evolution and ecology of microbial terpenoids. *Nat. Prod. Rep.* **39**, 249–272 (2022).
4. Huang, M. et al. Terpenoids: natural products for cancer therapy. *Exp. Opin. Investig. Drugs* **21**, 1801–1818 (2012).
5. Mewalal, R. et al. Plant-derived terpenes: a feedstock for specialty biofuels. *Trends Biotechnol.* **35**, 227–240 (2017).
6. Rudolf, J. D., Alsup, T. A., Xu, B. & Li, Z. Bacterial terpenome. *Nat. Prod. Rep.* **38**, 905–980 (2021).
7. Dickschat, J. S. Bacterial terpene cyclases. *Nat. Prod. Rep.* **33**, 87–110 (2016).
8. Christianson, D. W. Structural and chemical biology of terpenoid cyclases. *Chem. Rev.* **117**, 11570–11658 (2017).
9. Cane, D. E. & Ikeda, H. Exploration and mining of the bacterial terpenome. *Acc. Chem. Res.* **45**, 463–472 (2012).
10. Yamada, Y. et al. Terpene synthases are widely distributed in bacteria. *Proc. Natl. Acad. Sci. USA* **112**, 857–862 (2015).
11. Pan, X., Rudolf, J. D. & Dong, L.-B. Class II terpene cyclases: structures, mechanisms, and engineering. *Nat. Prod. Rep.* **41**, 402–433 (2024).
12. Rudolf, J. D. & Chang, C.-Y. Terpene synthases in disguise: enzymology, structure, and opportunities of non-canonical terpene synthases. *Nat. Prod. Rep.* **37**, 425–463 (2020).
13. Wilderman, P. R. & Peters, R. J. A single residue switch converts abietadiene synthase into a pimaradiene specific cyclase. *J. Am. Chem. Soc.* **129**, 15736–15737 (2007).
14. Oberg, N., Zallot, R. & Gerlt, J. A. EFI-EST, EFI-GNT, and EFI-CGFP: enzyme function initiative (EFI) web resources for genomic enzymology tools. *J. Mol. Biol.* **435**, 168018 (2023).
15. Oren, A. & Garrity, G. M. Valid publication of the names of forty-two phyla of prokaryotes. *Int. J. Syst. Evol. Microbiol.* **71**, 005056 (2021).
16. Xu, B., Ning, W., Wei, X. & Rudolf, J. D. Mutation of the eunicellane synthase bnd4 alters its product profile and expands its prenylation ability. *Org. Biomol. Chem.* **20**, 8833–8837 (2022).
17. Nafie, L. A., Dukor, R. K. & Freedman, T. B. Vibrational circular dichroism. In *Handbook of Vibrational Spectroscopy* (eds Chalmers, J. M. & Griffiths, P. R.) 731–744 (Wiley, 2002).
18. He, Y., Bo, W., Dukor, R. K. & Nafie, L. A. Determination of absolute configuration of chiral molecules using vibrational optical activity: a review. *Appl. Spectrosc.* **65**, 699–723 (2011).
19. Klein Hendges, A. P. P. et al. Phytotoxic neocassane diterpenes from eragrostis plana. *J. Nat. Prod.* **83**, 3511–3518 (2020).
20. Kupka, J., Anke, T., Oberwinkler, F., Schramm, G. & Steglich, W. Antibiotics from Basidiomycetes. VII. Crinipellin, a New Antibiotic from the Basidiomycetous Fungus Crinipellis Stipitaria (Fr.) Pat. *J. Antibiot.* **32**, 130–135 (1979).
21. Kotha, S. & Fatma, A. Synthetic approaches to natural and unnatural tetraquinanes. *Asian J. Org. Chem.* **11**, e202100595 (2022).
22. Chhetri, B. K. et al. Peyssonnosides A–B, unusual diterpene glycosides with a sterically encumbered cyclopropane motif: structure elucidation using an integrated spectroscopic and computational workflow. *J. Org. Chem.* **84**, 8531–8541 (2019).
23. Chesnokov, G. A. & Gademann, K. Concise total synthesis of peyssonnoside A. *J. Am. Chem. Soc.* **143**, 14083–14088 (2021).
24. Xu, B., Liu, C. & Dai, M. Catalysis-Enabled 13-Step Total Synthesis of (−)-Peyssonnoside A. *J. Am. Chem. Soc.* **144**, 19700–19703 (2022).
25. Bedoya-Urrego, K., Alzate, J. F. Phylogenomic discernments into anaerolineaceae thermal adaptations and the proposal of a candidate genus mesolinea. *Front. Microbiol.* **15** https://doi.org/10.3389/fmicb.2024.1349453 (2024).
26. Braekman, J. C. et al. Chemical studies of marine invertebrates—XXVIII: diterpenes from *Clavularia inflata* (coelenterata, octocorallia, stolonifera). *Tetrahedron* **34**, 1551–1556 (1978).
27. Amaya García, F. et al. Semisynthesis of dolabellane diterpenes: oxygenated analogues with increased activity against Zika and Chikungunya viruses. *J. Nat. Prod.* **84**, 1373–1384 (2021).

28. Bian, G. et al. A clade II-D fungal chimeric diterpene synthase from Colletotrichum gloeosporioides produces dolasta-1(15),8-diene. *Angew. Chem. Int. Ed.* **57**, 15887–15890 (2018).
29. Bowden, B. F., Braekman, J. & Mitchell, S. J. Studies of Australian soft corals. XX. A new sesquiterpene furan from soft corals of the family Xeniidae and an examination of clavularia inflata from north Queensland waters. *Aust. J. Chem.* **33**, 927–932 (1980).
30. Qin, B. et al. An unusual chimeric diterpene synthase from *Emericella variecolor* and its functional conversion into a sesterterpene synthase by domain swapping. *Angew. Chem. Int. Ed.* **55**, 1658–1661 (2016).
31. Rinkel, J. et al. A family of related fungal and bacterial Di- and sesterterpenes: studies on fusaterpenol and variediene. *ChemBioChem* **21**, 486–491 (2020).
32. Jiang, L. et al. Norditerpenoids biosynthesized by variediene synthase-associated P450 machinery along with modifications by the host cell *Aspergillus oryzae*. *Synth. Syst. Biotechnol.* **7**, 1142–1147 (2022).
33. Palframan, M. J. & Pattenden, G. The verticillenes. Pivotal intermediates in the biosynthesis of the taxanes and the phomactins. *Nat. Prod. Rep.* **36**, 108–121 (2019).
34. He, S., Abdallah, I. I., van Merkerk, R. & Quax, W. J. Insights into taxadiene synthase catalysis and promiscuity facilitated by mutability landscape and molecular dynamics. *Planta* **259**, 87 (2024).
35. Palframan, M. J., Bandi, K. K., Hamilton, J. G. C. & Pattenden, G. Sobralene, a new sex-aggregation pheromone and likely shunt metabolite of the taxadiene synthase cascade, produced by a member of the sand fly *Lutzomyia Longipalpis* Species Complex. *Tetrahedron Lett.* **59**, 1921–1923 (2018).
36. Gao, Y., Honzatko, R. B. & Peters, R. J. Terpenoid synthase structures: a so far incomplete view of complex catalysis. *Nat. Prod. Rep.* **29**, 1153–1175 (2012).
37. Cao, R. et al. Diterpene cyclases and the nature of the isoprene fold. *Proteins Struct. Funct. Bioinfor.* **78**, 2417–2432 (2010).
38. Ikeda, C., Hayashi, Y., Itoh, N., Seto, H. & Dairi, T. Functional analysis of eubacterial *ent*-copalyl diphosphate synthase and pimara-9(11),15-diene synthase with unique primary sequences. *J. Biochem.* **141**, 37–45 (2007).
39. Shirai, M. et al. Terpenoids produced by actinomycetes: isolation, structural elucidation and biosynthesis of new diterpenes, gifhornenolones A and B from Verrucosispora gifhornensis YM28-088. *J. Antibiot* **63**, 245–250 (2010).
40. Chen, X. et al. Discovery of bifunctional diterpene cyclases/synthases in bacteria supports a bacterial origin for the plant terpene synthase gene family. *Hortic. Res.* **11**, uhae221 (2024).
41. Li, Z. et al. Fragrant Venezuelans A and B with A 5–5–6–7 tetracyclic skeleton: discovery, biosynthesis, and mechanisms of central catalysts. *ACS Catal.* **10**, 5846–5851 (2020).
42. Yamada, Y. et al. Novel terpenes generated by heterologous expression of bacterial terpene synthase genes in an engineered streptomyces host. *J. Antibiot.* **68**, 385–394 (2015).
43. Dickschat, J. S. Bacterial diterpene biosynthesis. *Angew. Chem. Int. Ed.* **58**, 15964–15976 (2019).
44. Li, Z. et al. Molecular insights into the catalytic promiscuity of a bacterial diterpene synthase. *Nat. Commun.* **14**, 4001 (2023).
45. Mason, J. W., Schmid, C. L., Bohn, L. M. & Roush, W. R. Stolonidiol: synthesis, target identification, and mechanism for choline acetyltransferase activation. *J. Am. Chem. Soc.* **139**, 5865–5869 (2017).
46. Köllner, T. G., Schnee, C., Gershenzon, J. & Degenhardt, J. The variability of sesquiterpenes emitted from two zea mays cultivars is controlled by allelic variation of two terpene synthase genes encoding stereoselective multiple product enzymes. *Plant Cell* **16**, 1115–1131 (2004).
47. Blank, P. N. et al. Substitution of aromatic residues with polar residues in the active site pocket of epi-isozizaene synthase leads to the generation of new cyclic sesquiterpenes. *Biochemistry* **56**, 5798–5811 (2017).
48. Sato, T. et al. Sesquarterpenes (C35 terpenes) biosynthesized via the cyclization of a linear c35 isoprenoid by a tetraprenyl-β-curcumene synthase and a tetraprenyl-β-curcumene cyclase: identification of a new terpene cyclase. *J. Am. Chem. Soc.* **133**, 9734–9737 (2011).
49. Zhu, C., Xu, B., Adpressa, D. A., Rudolf, J. D. & Loesgen, S. Discovery and biosynthesis of a structurally dynamic antibacterial diterpenoid. *Angew. Chem. Int. Ed.* **60**, 14163–14170 (2021).
50. Wang, Z. et al. Cytochrome P450 mediated cyclization in eunicellane derived diterpenoid biosynthesis.*Angew. Chem.* **135**, e202312490 (2023).
51. Rinkel, J. et al. Mechanisms of the Diterpene Cyclases β-Pinacene Synthase from Dictyostelium discoideum and Hydropyrene Synthase from Streptomyces clavuligerus. *Chem. Eur. J.* **23**, 10501–10505 (2017).
52. Rinkel, J. & Dickschat, J. S. Characterization of micromonocyclol synthase from the marine Actinomycete Micromonospora marina. *Org. Lett.* **21**, 9442–9445 (2019).
53. Meguro, A., Tomita, T., Nishiyama, M. & Kuzuyama, T. Identification and characterization of bacterial diterpene cyclases that synthesize the cembrane skeleton. *ChemBioChem* **14**, 316–321 (2013).
54. Gu, B., Goldfuss, B. & Dickschat, J. S. Mechanistic characterisation and engineering of Sesterviolene synthase from *Streptomyces violens*. *Angew. Chem. Int. Ed.* **62**, e202215688 (2023).
55. Niu, S. et al. Spirograterpene A, a tetracyclic spiro-diterpene with a fused 5/5/5/5 ring system from the deep-sea-derived fungus Penicillium granulatum MCCC 3A00475. *J. Nat. Prod.* **80**, 2174–2177 (2017).
56. Rinkel, J., Lauterbach, L. & Dickschat, J. S. Spata-13,17-diene synthase-an enzyme with sesqui-, Di-, and Sesterterpene Synthase Activity from *Streptomyces xinghaiensis*. *Angew. Chem. Int. Ed.* **56**, 16385–16389 (2017).
57. Chhalodia, A. K. et al. Functional characterisation of twelve terpene synthases from actinobacteria. *Beilstein J. Org. Chem.* **19**, 1386–1398 (2023).
58. Li, Z. et al. First *trans*-eunicellane terpene synthase in bacteria. *Chem* **9**, 698–708 (2023).
59. Li, Z. et al. Cryptic isomerization in diterpene biosynthesis and the restoration of an evolutionarily defunct P450. *J. Am. Chem. Soc.* **145**, 22361–22365 (2023).
60. Piel, J. Metabolites from symbiotic bacteria. *Nat. Prod. Rep.* **26**, 338–362 (2009).
61. Burkhardt, I., de Rond, T., Chen, P. Y.-T. & Moore, B. S. Ancient plant-like terpene biosynthesis in corals. *Nat. Chem. Biol.* **18**, 664–669 (2022).
62. Scesa, P. D., Lin, Z. & Schmidt, E. W. Ancient defensive terpene biosynthetic gene clusters in the soft corals. *Nat. Chem. Biol.* **18**, 659–663 (2022).
63. Chen, X. et al. Canonical terpene synthases in arthropods: intra-phylum gene transfer. *Proc. Natl Acad. Sci. USA* **19**, e2413007121 (2023).
64. Zaroubi, L. et al. The ubiquitous soil terpene geosmin acts as a warning chemical. *Appl. Environ. Microbiol.* **88**, e00093–22 (2022).
65. Becher, P. G. et al. Developmentally regulated volatiles geosmin and 2-methylisoborneol attract a soil arthropod to *Streptomyces* bacteria promoting spore dispersal. *Nat. Microbiol.* **5**, 821–829 (2020).
66. Tamura, K., Stecher, G. & Kumar, S. MEGA11: molecular evolutionary genetics analysis version 11. *Mol. Biol. Evol.* **38**, 3022–3027 (2021).
67. Letunic, I. & Bork, P. Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool. *Nucl. Acids Res.* **52**, W78–W82 (2024).

Acknowledgements
This work is supported in part by the National Institutes of Health Grant R35 GM142574 and the University of Florida (J.D.R.). The synthesis of the TS genes (part of proposal: 10.46936/10.25585/60008123) was conducted by the U.S. Department of Energy Joint Genome Institute, a DOE Office of Science User Facility, operated under contract no. DE-AC02-05CH11231 (J.D.R.). T.A.A. was supported in part by a Chemistry-Biology Interface Research Training Program Grant T32 GM136583. D.G.I. was supported in part by a National Science Foundation Graduate Research Fellowship under Grant No. DGE-2236414. We thank the University of Florida Mass Spectrometry Research and Education Center, which is supported by the NIH (S10 OD021758-01A1) and Jodie Johnson and Katie Heiden for GC-MS support. We thank Sandra Loesgen for helpful discussions. We thank Ethan Xu for discussions about VCD and connecting the authors with BioTools, Inc.

Author contributions
Conceptualization, J.D.R.; methodology, X.W., W.N., C.A.M., T.A.A., D.P.Ł.-K., J.N., B.X., and J.D.R.; investigation, X.W., W.N., C.A.M., T.A.A., D.P.Ł.-K., J.N., T.Q., M.O.O., G.R.G., B.X., D.G.I., and J.D.R.; data analysis, X.W., W.N., C.A.M., T.A.A., Z.L., D.P.Ł.-K., J.N., and J.D.R.; writing—original draft, X.W., W.N., and J.D.R.; writing—review & editing, X.W., W.N., J.N., and J.D.R.; visualization, X.W., W.N., J.N., and J.D.R.; supervision, J.D.R.; funding acquisition, J.D.R.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-025-57145-6.

Correspondence and requests for materials should be addressed to Jeffrey D. Rudolf.

Peer review information *Nature Communications* thanks Haruo Ikeda and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.