

OSMAC guided bioprospecting of Atlantic sponges reveals novel bacterial species and evidence for the antimicrobial thiazole alkaloid agrochelin II

Sam E. Williams,^{1,2} Henry L. Stennett,¹ Andrew J. Devine,³ Zhongshu Song,³ Catherine R. Back,¹ Luoyi Wang,^{3,4} Judith Mantell,⁵ Chris Neal,⁵ Mark A. Jepson,⁵ Angela E. Essex-Lopresti,⁶ Jonathan D. Sellars,⁷ James E. M. Stach,⁸ Christine L. Willis,³ Paul Curnow,¹ Paul R. Race⁸

AUTHOR AFFILIATIONS See affiliation list on p. 18.

ABSTRACT The deep sea remains a major reservoir of underexplored microbial diversity and biosynthetic novelty. Here, we describe three bacterial species isolated from Atlantic sponges, including *Stappia quadratibracata* sp. nov., *Bacillus crepusculi* sp. nov., and *Psychrobacter noctis* sp. nov. Genome sequencing, phylogenomics, and phenotypic characterization confirmed their novelty and biosynthetic potential. A targeted 'One Strain Many Active Compounds' (OSMAC) screen revealed previously silent antibacterial activity from *Stappia quadratibracata* sp. nov., only when grown with the carbon source succinate. Metabolomics and molecular networking analysis indicated that this activity was attributable to an unstable thiazole alkaloid with spectral data closely related to, but distinct from, the epimeric siderophores agrochelin and massiliachelin. Spectroscopic studies, in tandem with comparative analysis of the biosynthetic gene cluster for this metabolite, are consistent with agrochelin II, a previously unreported thiazole alkaloid diastereoisomer. Agrochelin II exhibits iron-enhanced antibacterial activity against *Staphylococcus aureus*, underscoring the ecological role of iron acquisition in microbial competition. Our findings highlight the value of OSMAC-guided bioprospecting in uncovering antimicrobial metabolites from sponge-associated bacteria.

IMPORTANCE Bioactive microbial natural products remain the preeminent source of new lead compounds for drug development. Due to the increasingly high levels of strain and compound rediscovery from terrestrial environments, the deep-ocean is increasingly considered an attractive starting point for bioprospecting programs, which seek to isolate and characterize novel chemical scaffolds. Here, we use a combination of genomics, metabolomics, and chemical analysis, to establish the biosynthetic potential of three bacterial species isolated from deep-ocean Atlantic sponges and report the discovery and characterization of a new antimicrobial thiazole alkaloid, agrochelin II. Our findings demonstrate the usefulness of integrated cultivation–screening–metabolomics–genomics pipelines for microbial metabolite discovery and identify the genus *Stappia* as a hitherto neglected source of bioactive natural products.

KEYWORDS antibiotics, deep-sea, natural products, microbiome, metabolomics, OSMAC, genome mining, siderophores

Underexplored environments such as the deep ocean, caves, and hot springs have been suggested as potential sources of microorganisms that produce new antibiotics (1–3). The conditions of these environments are inhospitable to most life forms, with microbial occupants of these niches frequently possessing metabolic innovations that promote resilience (4–6). These innovations may include the ability

Editor Knut Rudi, Norwegian University of Life Sciences, Ås, Norway

Address correspondence to Sam E. Williams, sam.williams@sund.ku.dk, or Paul R. Race, paul.race1@newcastle.ac.uk.

The authors declare no conflict of interest.

See the funding table on p. 18.

Received 16 September 2025

Accepted 20 January 2026

Published 17 February 2026

Copyright © 2026 Williams et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International license.

to biosynthesize natural products with novel structures and functions (7, 8). Marine organisms have historically yielded clinically important natural products, emphasizing their continued potential in drug discovery (9–11). Bull et al. posited that chemical diversity follows from biological diversity, thus by characterizing new species, we increase the probability of discovering new antimicrobial natural products (12). Broadening bioprospecting beyond well characterized gifted taxa, e.g., the actinomycetes, may also help to identify previously overlooked microorganisms and their associated bioactive metabolites (13, 14).

Accessing the full spectrum of chemical diversity from environmental strains is an ongoing challenge in natural product discovery (15, 16). The “One Strain Many Active Compounds” (OSMAC) approach is an experimental strategy that can be employed to identify the chemical diversity produced by a single strain (17). In this method, the composition of media or carbon source is adjusted in an effort to activate biosynthetic gene clusters (BGCs) that are not expressed under standard laboratory conditions. Importantly, this approach has proven to be of particular value in the identification of antimicrobial natural products (17). By applying this method to underexplored bacterial taxa, it may be possible to expedite the discovery of new antimicrobial natural products, by leveraging the additive benefits of targeting historically neglected microbes using OSMAC.

As one of Earth’s largest biomes, the deep-sea and abyssal oceanic zones encompass depths beyond ~200 m and feature low temperatures, high-pressure, oligotrophic and dark conditions (18). These unique conditions have led to specialized metabolic adaptations in deep-sea microbial communities, providing an underexplored and distinct genetic resource for biodiscovery (19, 20). As part of an ongoing antibiotic discovery program, we have isolated hundreds of bacterial strains from deep-sea sponges (21). The sponges that comprise the “Bristol Sponge Microbiome Collection” (BISECT) were recovered from an area of the equatorial Atlantic not previously prospected, at depths of up to 3 km and comprising diverse microbial communities (21–23). Although the average depth of the ocean is 1.2 km (24), most sponge bioprospecting studies have to date sampled tropical waters of less than 200 m deep, a consequence of the ease of sample recovery from these locations (25). In contrast, at least two-thirds of the BISECT sponges were collected at depths exceeding 1 km (21). It has been widely noted that the extreme conditions of pressure, temperature, and salinity found in deep-sea habitats have driven the evolution of diverse and unusual bacteria, which constitute intriguing targets for the discovery of new natural products (5, 26, 27). To this end, we have been engaged in a taxonomic identification of culturable bacteria from the BISECT collection.

Herein, we describe three culturable strains isolated from BISECT sponges and propose the names *Stappia quadratibracata* sp. nov., *Bacillus crepusculi* sp. nov., and *Psychrobacter noctis* sp. nov. We have examined the biosynthetic potential of these bacteria and by systematically adjusting culture conditions activated previously silent anti-methicillin-resistant *Staphylococcus aureus* (MRSA) activity in *S. quadratibracata*. This activity is shown to be attributable to the previously unreported thiazole alkaloid agrochelin II, whose chemical structure has been elucidated using an approach combining spectroscopic studies with bioinformatic analysis. Collectively, our findings demonstrate the usefulness of employing OSMAC to unlock the biosynthetic potential of deep-sea bacteria, including those from less well-studied genera.

RESULTS

Isolation and ecology

Three bacterial strains were isolated from individual deep-sea sponges: 28M-7^T from sponge B0078, 28A-2^T from B01641, and 28M-43^T from B01661. Sponges were collected at depths between 868 and 1,483 m in the twilight (mesopelagic, 150–1,000 m) and midnight (bathypelagic, 1,000–4,000 m) zones of the equatorial Atlantic at the Vayda and Knipovich seamounts (28). Seawater samples were analyzed at seven sites (1,150–2,307

m depth) on the seamounts and ranged between 34.93–35.00 practical salinity units, 165.77–236.25 $\mu\text{mol/kg}$ oxygen, and 2.80–5.23°C. The data describing each of these sponges are presented in Table 1.

Phylogeny

We sequenced the genomes of the three strains using a hybrid approach (Illumina and Oxford Nanopore) and generated high-quality, biologically accurate assemblies (Table S1). To resolve their taxonomic novelty, we conducted comprehensive genomic comparisons including digital DNA-DNA hybridization (dDDH), average nucleotide identity (ANI), and GTDB-based phylogenomic placement (29–33).

Strain 28M-7^T

We extracted the full 1,481 bp 16S rRNA gene from the 28M-7^T genome which had high similarity to several *Stappia* species (Table S2). Whole-genome-based phylogenetic analysis (TYGS) was used to determine 28M-7^T's closest relatives and calculate the dDDH. None of the related strains had a dDDH $\geq 70\%$ or an ANI $\geq 95\%$, the thresholds above which two strains are considered the same species (Table S3). We generated a phylogenetic tree of the *Stappia* and *Hongsooonwoonella* type strains including 28M-7^T, which revealed that this strain clusters with *Stappia indica*, a strain also isolated from the deep-sea (34) (Fig. 1). *Stappia sediminis* and *Stappia albiluteola* clustered separately with *Hongsooonwoonella zoysiae* and have been classified by the GTDB as members of the *Hongsooonwoonella* genus. The GTDB-Tk classified strain 28M-7^T (*Stappia* sp014252955) as a novel species, with its closest relative being *S. indica*. The environmental, diatom-associated strain *Stappia* sp. ARW1T was also identified as a representative of this same species (35, 36). Interestingly, the 28M-7^T assembly contained two contigs (168,820 bp and 69,494 bp) with high BLAST similarity to sequences from *Paracoccus binzhouensis* (37). These contigs were predicted to be plasmids by the sequence classifier tool PlasFlow v 1.1 (38). Both encoded complete MPF-type conjugation systems, along with key plasmid maintenance genes including *repA*, *parA*, and *copG*, confirming their conjugative nature, and suggesting one or more acquisition events. Phenotypic and physiological profiling data for 28M-7^T are shown in Table S4. An electron micrograph of a 28M-7^T cell is shown in Fig. S1.

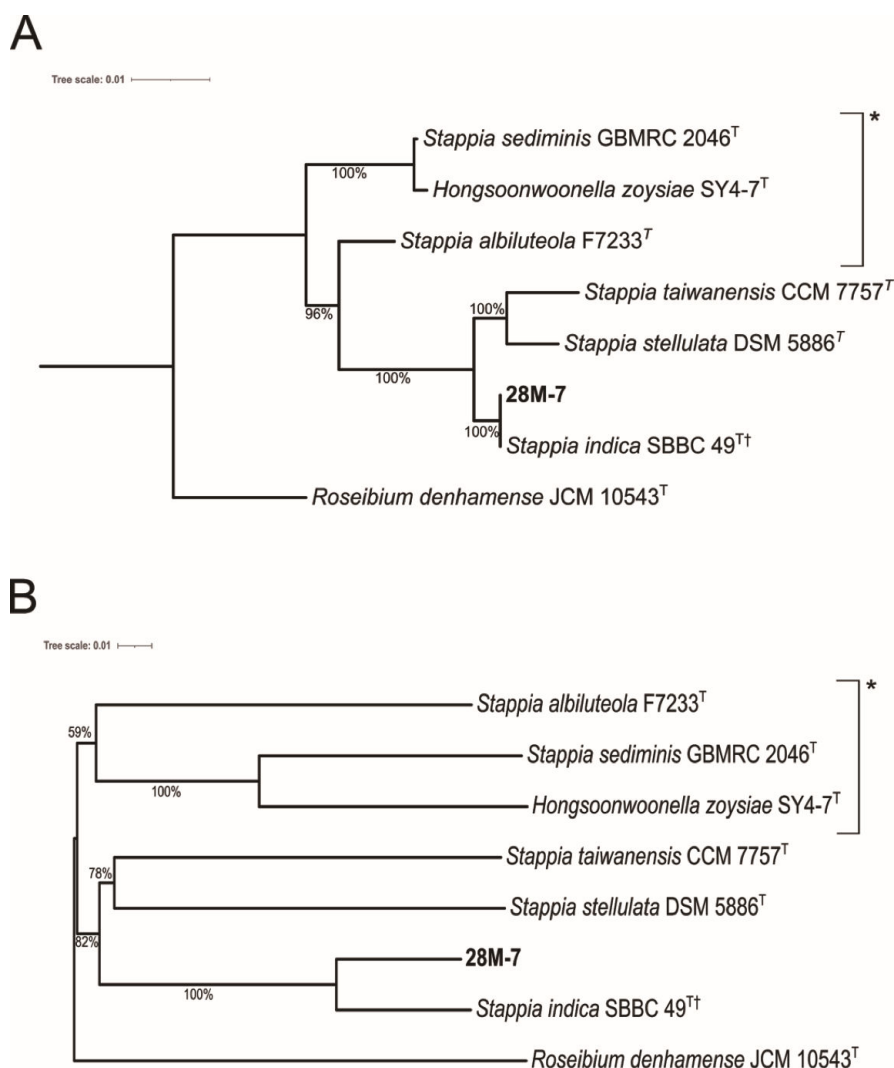
Strain 28A-2^T

A complete 1,544 bp 16S gene was extracted from the 28A-2^T genome. It was found that its closest relatives were all members of the *Bacillus pumilus* group (Table S5). These species share 16S rRNA gene sequence identities of $>99.5\%$ (40). There have been several attempts to delineate the phylogenetics of this group based on DNA gyrase subunit B (*gyrB*) (41), concatenated housekeeping genes (40), and whole genome sequences (42). We produced phylogenetic trees for strain 28A-2^T and the other closely related members of the *B. pumilus* group based on 16S rRNA gene and whole genome similarity (Fig. 2). None of the closely related strains had a dDDH $\geq 70\%$ or an ANI $\geq 95\%$ indicating 28A-2^T represents a novel species within the *B. pumilus* group (Table S6). The GTDB-Tk classified

TABLE 1 Characteristics of the sponges from which the bacteria described in this study were isolated^a

Sponge	Class	Depth (m)	Region	Coordinates
B0078	Demosponge	971	Knipovich	5° 37' 30.0" N 26° 57' 29.0" W
B01641	Hexactinellid	868	Vayda	14° 53' 03.0" N 48° 07' 27.0" W
B01661	Demosponge	1,483	Vayda	14° 51' 47.0" N 48° 14' 29.0" W

^aThe class of each sponge was estimated based on its morphology. The depths, region, and coordinates shown are those at which each sponge was collected from the sea floor.



*GTDB assigned taxonomy: *Hongsoonwoonella*

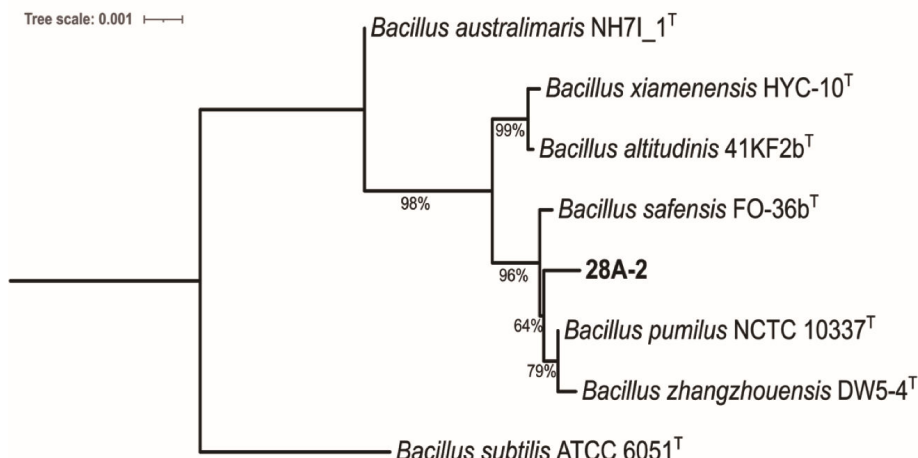
FIG 1 (A) 16S rRNA gene-based phylogenetic tree of 28M-7^T. (B) Whole genome phylogenetic tree for strain 28M-7^T which we propose as *Stappia quadratibracata* sp. nov. Trees were constructed with TYGS (31) and prepared using iTOL (39). GBDP pseudo-bootstrap values from 100 replications are shown, bootstrap values below 50% are not shown. Branch lengths are scaled in terms of GBDP d5 distance formula. The trees were rooted using *Roseibium denhamense* JCM 10543^T, another member of the family Stappiaceae, as the outgroup. [†]*Stappia indica* SBBC 49 included as GTDB species representative as no genome is available for the type strain *S. indica* B106^T. The whole genome tree had an average branch support of 83.8% and a ∂ statistic of 0.266.

strain 28A-2^T as a novel species (*Bacillus* sp014764715), with *Bacillus zhangzhouensis* its closest relative. Phenotypic and physiological profiling data for 28A-2^T are shown in Table S7. An electron micrograph of a 28A-2^T cell is shown in Fig. S2.

Strain 28M-43^T

For strain 28M-43^T a complete 1,534 bp 16S rRNA gene was successfully extracted from the genome. It shared 16S rRNA gene sequence identity of $\geq 98.7\%$ with several *Psychrobacter* species (Table S8). We submitted the 28M-43^T genome assembly to the TYGS server to identify closely related strains and determine taxonomy. No related strains shared dDDH values $\geq 70\%$ or ANI values $\geq 95\%$ (Table S9). We generated a phylogenetic

A



B

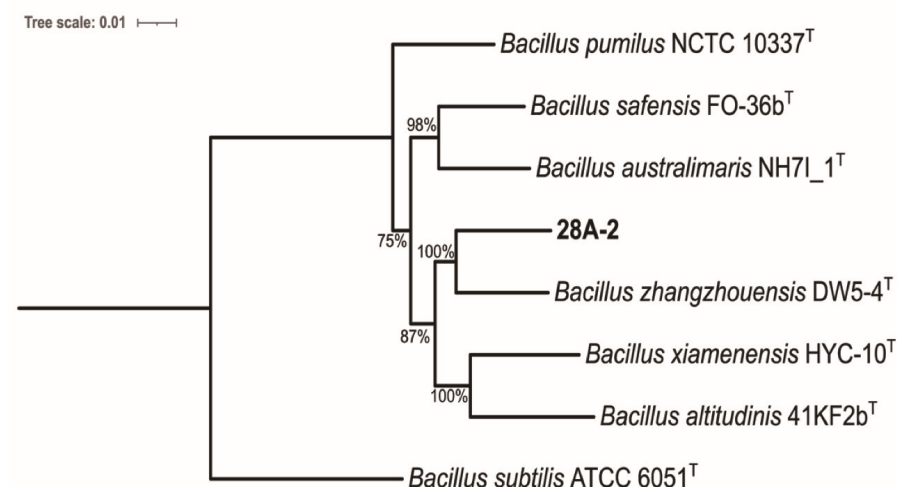


FIG 2 (A) 16S rRNA gene-based phylogenetic tree of strain 28A-2^T. (B) Whole genome phylogenetic tree for strain 28A-2^T, which we propose as *Bacillus crepusculi* sp. nov. Trees were constructed with TYGS (31) and prepared using iTOL (39). GBDP pseudo-bootstrap values from 100 replications are shown, values under 50% are not shown, branch lengths are scaled in terms of GBDP d5 distance formula. The tree was rooted using *Bacillus subtilis* ATCC 6051^T as the outgroup. The whole genome tree had an average branch support of 92% and a δ statistic of 0.264.

tree for 28M-43^T and the 13 closest *Psychrobacter* relatives of 28M-43^T (Fig. 3). The GTDB-Tk also classified strain 28M-43^T as a novel species (*Psychrobacter* sp014770435), with *Psychrobacter nivimaris* being its closest relative. Phenotypic and physiological profiling data for 28M-43^T are shown in Table S10. An electron micrograph of a 28M-43^T cell is shown in Fig. S3.

Antibacterial activity screening of isolated strains

To investigate the ability of 28M-7^T, 28A-2^T, and 28M-43^T to biosynthesize antimicrobial metabolites, antimicrobial activity screening was undertaken using a panel of clinically relevant bacterial pathogens. *Bacillus* 28A-2^T inhibited the growth of Gram-positive strains in these assays, including multidrug-resistant *Staphylococcus aureus* (Fig. 4A). Neither of the other two strains exhibited any antibacterial activity.

To explore the potential of activating silent BGCs, we designed a small OSMAC screen by changing the carbon source in the minimal M9 agar used for their cultivation (Table S11). *Stappia* 28M-7^T displayed antibiotic activity against *S. aureus* Newman only when grown with 10 g/L succinate (Fig. 4B). Interestingly, this activity was found to be further enhanced by the presence of ferric iron (Fig. 4C). *Psychrobacter* 28M-43^T did not display antibacterial activity under any condition tested and was, therefore, deprioritized for further study.

Isolation of active compounds

Strain *Bacillus* 28A-2^T

To investigate the antibiotic activity observed for *Bacillus* 28A-2^T, we prepared organic extracts from bacteria grown on both agar plates and in liquid media. Agar plate extracts exhibited broad spectrum antimicrobial activity (Fig. S4) while broth extracts did not. We attempted activity-guided fractionation of the agar extracts using high-performance liquid chromatography (HPLC) but were unable to isolate the active constituent(s) in sufficient quantity to enable further characterization. We, therefore, adopted a bioinformatic approach using antiSMASH 7.0 to identify BGCs within the 28A-2^T genome. Eleven BGCs were identified, of which two possess significant sequence identity to known clusters associated with antimicrobial natural products, namely, the peptide antibiotic bacilysin (43) and a lichenysin-type biosurfactant (44). The anti-Gram-positive activity of strain 28A-2^T is likely attributable to bacilysin biosynthesis, as this compound has a more limited spectrum of activity than lichenysin, an observation in keeping with our antibacterial screening data (45, 46). That this activity was only observed in solid-state cultures is consistent with previous work showing that *Bacillus* species frequently redirect metabolic flux toward bioactive secondary metabolites, including lipopeptide biosurfactants, under surface-associated growth conditions (47).

Strain *Stappia* 28M-7^T

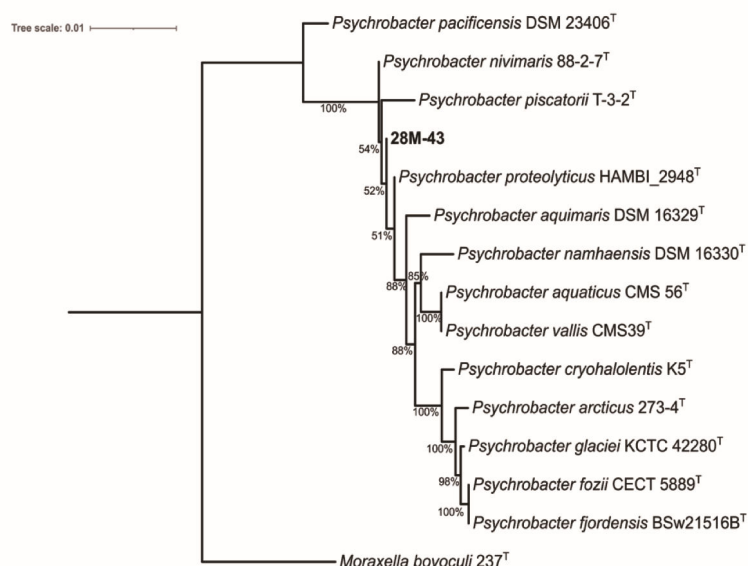
To identify the natural product responsible for the antibiotic activity of *Stappia* 28M-7^T, we first prepared a small-scale crude extract from cultures grown in M9 media supplemented with 10 g/L succinate. This extract was subjected to LC-MS/MS and the resulting data analyzed with feature-based molecular networking using the Global Natural Products Social Molecular Networking (GNPS) platform (48). The dominant metabolite m/z 467.205 [M+H]⁺ was present as two isomers, grouped into a network comprising 14 nodes (Fig. 5A). This m/z value corresponds to the reported molecular weight of the thiazole alkaloid siderophores massiliachelin (1) and agrochelin (2), previously isolated from *Massilia* and *Agrobacterium* spp., respectively (C₂₃H₃₄N₂O₄S₂; MW 466). Other metabolites within the network varied by ± 14 Da, consistent with CH₂ carbon chain length variations, with an abundant metabolite at 481.22 [M+H]⁺ (+14.0) and a less abundant metabolite of 453.2 [M+H]⁺ (−14.0), both directly linked to the central node. Furthermore, two low abundance features (m/z 325.0705) were identified in the extract with GNPS library matches to pyochelin (CCMSLIB00005724305; cosine score 0.87) although these did not cluster with the main network.

As no MS/MS reference spectra is available for massiliachelin or agrochelin, we used MetFrag (49) to compare the spectra with an *in silico* fragmentation prediction of agrochelin (NPA017357). This resulted in a final score of 1.0, indicating high agreement between predicted and observed fragments (Fig. 5B). Further analysis using Modifinder (50) identified the m/z 481.22 fragments matching those of m/z 467.205, with a +14 Da modification localized to the salicylic acid moiety, consistent with a previously unreported extended alkyl chain variant. Although Modifinder suggests a terminal +14 Da extension, odd-chain fatty acids have not been reported from the *Stappia* genus. An alternative is methyl esterification at the salicylate hydroxyl group. We did not detect a dimethyl ester, but neither can be definitively excluded.

Isolation and activity screening of the *Stappia* 28M-7^T metabolite

To enable compound isolation and to validate the source of the antibacterial activity of *Stappia* 28M-7^T, we scaled up cultures (5 × 1 L) under identical growth conditions to those employed above. An organic extract was prepared from the culture supernatant and analyzed by LC-MS. Surprisingly, this revealed only a single dominant peak at *m/z* 467.205, indicative of altered metabolic flux within the bacterium during large scale

A



B

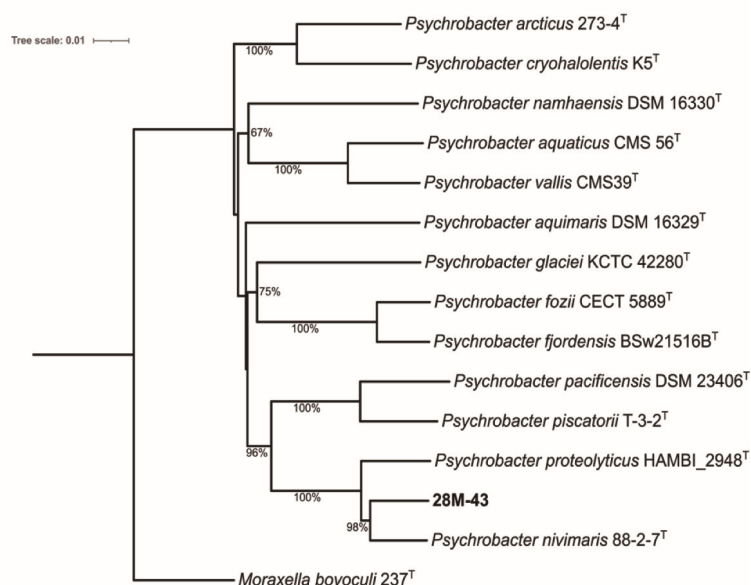


FIG 3 (A) 16S rRNA gene-based phylogenetic tree of strain 28M-43^T. (B) Whole genome similarity tree for strain 28M-43^T, which we propose as *Psychrobacter noctis* sp. nov. Trees were constructed with TYGS (31) and prepared using iTOL (39). GBDP pseudo-bootstrap values from 100 replications are shown, with branch lengths scaled in terms of GBDP d5 distance formula. The trees were rooted using *Moraxella bovoculi* 237^T as the outgroup. The phylogenomic tree had an average branch support of 80.5% and a Δ statistic of 0.231.

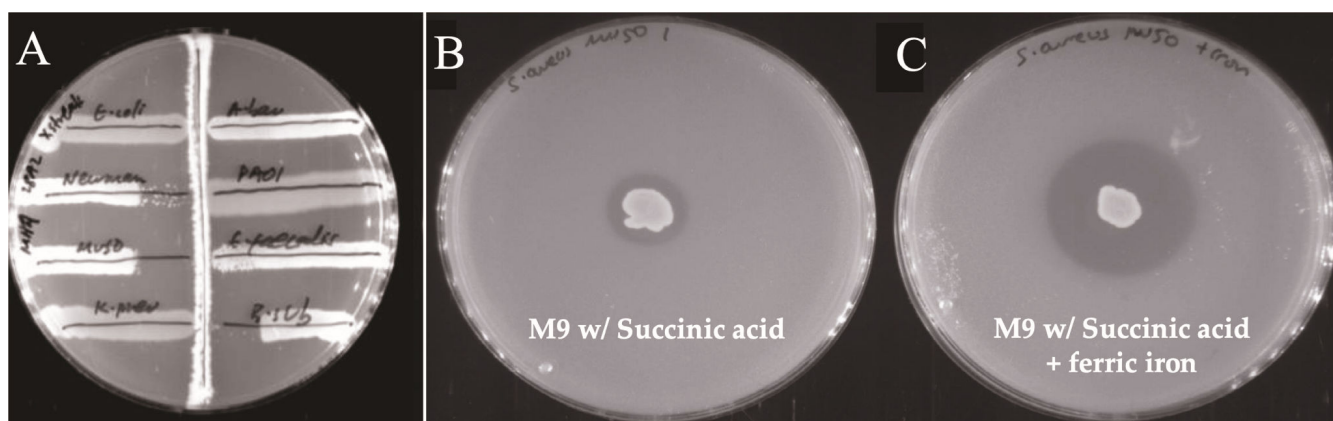


FIG 4 Antibacterial screening results. (A) Cross-streak assay on Mueller-Hinton agar showing the antibacterial activity of strain 28A-2^T. Clockwise from the top left, *Escherichia coli* BW25113, *Acinetobacter baumannii* ATCC 19606, *Pseudomonas aeruginosa* PAO1, *Enterococcus faecalis* UB591, *Bacillus subtilis*, *Klebsiella pneumoniae* NCTC 5055, *Staphylococcus aureus* Mu50, and *Staphylococcus aureus* Newman are shown as streaks perpendicular to the central streak of strain 28A-2^T. (B) Soft agar overlay for strain 28M-7^T grown on M9 agar with 10 g/L succinate for 7 days, overlaid with *S. aureus* Newman. (C) Soft agar overlay for strain 28M-7^T grown on M9 agar with 10 g/L succinate and ferric iron for 7 days, overlaid with *S. aureus* Newman.

fermentation. This change in metabolite profile could be a consequence of altered aeration, nutrient availability, or other factors. Purification by reverse-phase HPLC yielded sufficient quantity of the identified metabolite to enable NMR analysis and antibacterial testing.

We tested the purified compound for antibacterial activity and, in cases where activity was observed, determined the minimum inhibitory concentration (MIC; Table 2). The compound was tested both in isolation and following the addition of an equimolar quantity of ferric iron. The compound was found to exhibit antibacterial activity against *S. aureus* Newman and methicillin–vancomycin-resistant *S. aureus* Mu50 (51). Notably, this activity was enhanced fourfold when the compound was preincubated with iron, suggesting improved potency or increased cellular uptake in the iron bound form. Inhibition of growth was not observed against the Gram-negative species *E. coli*, *K. pneumoniae*, *A. baumannii*, and *P. aeruginosa*.

Structure elucidation of the *Stappia* 28M-7^T metabolite

¹H NMR analysis of the purified 28M-7^T metabolite indicated it to be closely related to the thiazole alkaloid siderophores massiliachelin (1) and agrochelin (2), in accord with our MS data (Fig. 6; Table S12). Indeed, full structural analysis by ¹H, ¹³C, and 2D NMR revealed a compound with connectivity consistent with the 2D structures of these known metabolites (Tables S12 and S13). Importantly, however, our data were not entirely in keeping with those reported for either compound (52, 53) although they more closely resembled agrochelin (2). Additionally, the compound was found to be unstable, rapidly degrading into a mixture with poorly resolved ¹H NMR signals. In an attempt to prevent degradation, we supplemented the product sample with ammonium iron (II) sulfate, in line with previous studies of the structurally related siderophores pyochelin I (4) and II (5), from *Pseudomonas aeruginosa* (Fig. 6) (54). Here, the equilibrium between pyochelin I (4) and pyochelin II (5) is driven toward pyochelin I (4) by the addition of divalent metal ions (55, 56), but no improvement in stability was observed here. The observed discrepancies between our data and those reported for massiliachelin/agrochelin could be due to subtle concentration or pH effects or may indicate that this compound is a novel diastereoisomer. We believed this compound may represent agrochelin II (3), with stereochemistry similar to that of enatio-pyochelin. This uncertainty prompted genome mining to identify the corresponding biosynthetic gene cluster.

Genome mining of *Stappia* 28M-7^T

In an effort to resolve the structure of the *Stappia* 28M-7^T metabolite, we employed an approach analogous to that used by Dietrich et al. (53), marrying bioinformatic predictions of metabolite chemistry with spectroscopic data. This approach required us to formally identify the BGC responsible for the biosynthesis of the natural product under investigation. The antiSMASH 7.0 platform identified nine BGCs in the 28M-7^T genome (Table 3), of which one (BGC 5) could be unambiguously assigned to the

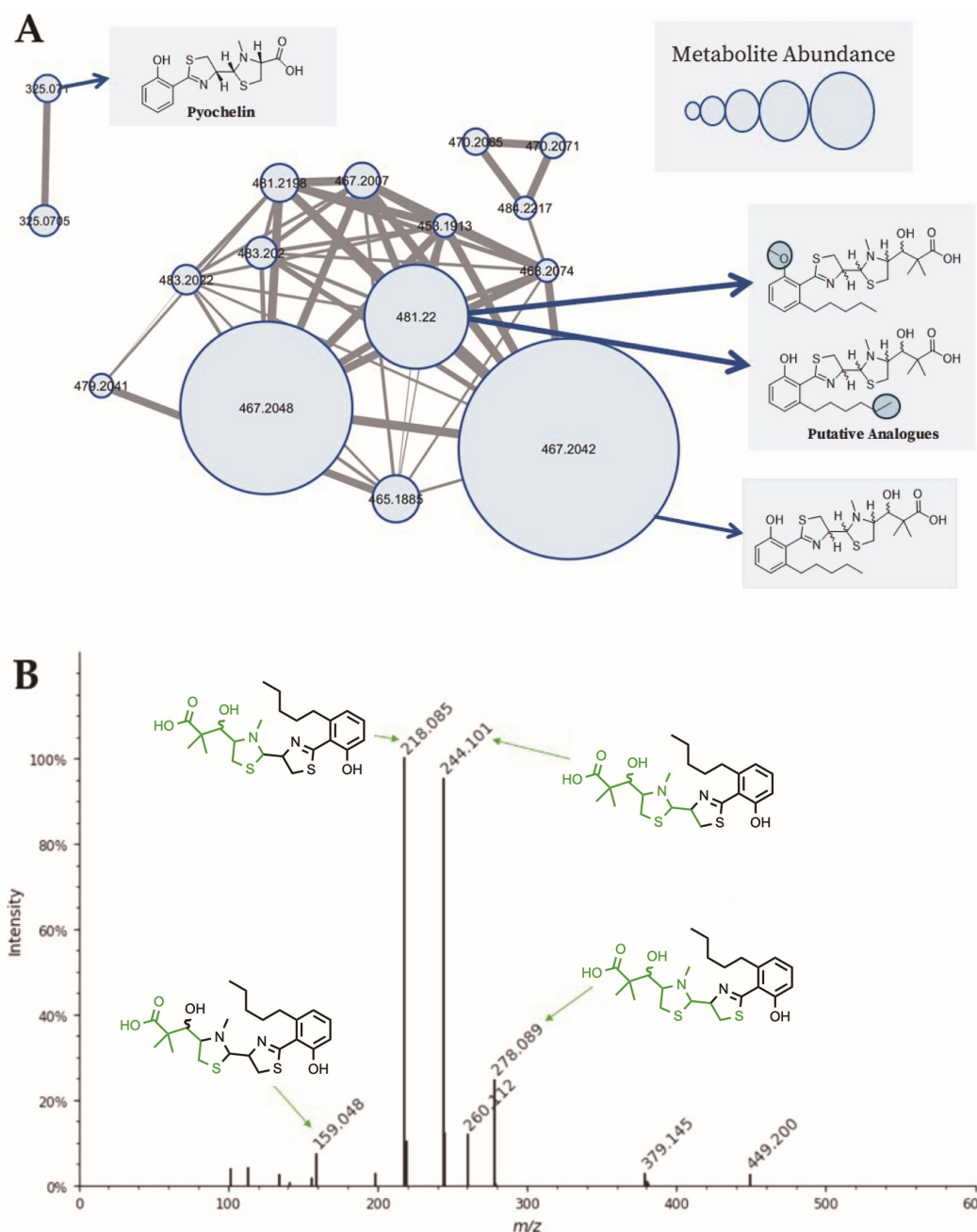


FIG 5 Feature-based molecular networking and *in silico* fragmentation analysis of the dominant metabolite from *Stappia* 28M-7^T. (A) GNPS molecular network of selected MS/MS features from *Stappia* 28M-7^T. The dominant metabolite isomer nodes m/z 467.20 $[M+H]^+$ in a 14-node subnetwork believed to be agrochelin/massiliachelin. Related features at m/z 481.22 and 453.20 likely represent CH_2 chain length variants. Nodes are sized based on metabolite abundance based on the peak area. (B) MetFrag *in silico* fragmentation of agrochelin (NPA017357) against the experimental spectrum for m/z 467.205, with major fragments and assigned substructures shown in green.

TABLE 2 Antibacterial activities of the purified *Stappia* 28M-7^T metabolite before and after incubation with equimolar ammonium iron(II) sulfate (Fe)^a

Strain	Minimum inhibitory concentration (μg/mL)			
	Thiazole alkaloid		Ampicillin	Eq. Fe
	Alone	With eq. Fe		
<i>S. aureus</i> Newman	256	64	8	1,024
<i>S. aureus</i> Mu50	256	64	512	1,024

^aAmpicillin was used as a positive control and ammonium iron(II) sulfate as a negative control. MIC assays were performed in triplicate. 861 μg/mL of ammonium iron(II) sulfate is equimolar to 1,024 μg/mL of the isolated metabolite, the value given for “Eq. Fe”.

biosynthesis of a massiliachelin/agrochelin type thiazole alkaloid. The assignment was based on the distinctive domain architecture of this BGC as compared to that reported previously (53). The antiSMASH 7.0 pipeline also detected a *trans*-AT PKS/NRPS hybrid BGC (BGC 2) homologous to the *Stappia indica* sesbanimide C biosynthetic gene cluster, the only reported natural product from this genus (57). Notably, low-abundance features corresponding to sesbanimide A (*m/z* 328.1391 [M+H]⁺) and sesbanimide C (*m/z* 314.1598 [M+H]⁺) were both observed in our metabolomics data.

We used CORASON to identify massiliachelin/agrochelin BGCs in the 20 publicly available *Stappia* and *Hongsooonwoonella* genomes from NCBI and identified the cluster in five other *Stappia* strains (Fig. 7). By aligning the homologous genes in these BGCs, we were able to identify the nine genes conserved between all six BGCs, which we propose as the minimal massiliachelin/agrochelin BGC. This locus has been deposited in MiBIG (BGC0002764) (58). A full manual annotation of the massiliachelin/agrochelin BGC and surrounding genes in the strain 28M-7^T genome is shown in Table S14. To further explore the distribution of this cluster and briefly review the biosynthetic potential of the genus, we performed a gene cluster family (GCF) analysis using BiG-SCAPE2 (59). Seventeen assemblies were included, including four metagenome assembled genomes, restricted to GTDB-validated *Stappia* strains representing 10 species (Table S15). This analysis revealed the six strains containing the agrochelin BGC were split across four GTDB species clusters: “*Stappia quadratibracata*” sp. nov., *Stappia indica_A*, *Stappia sp900185725*, and *Stappia sp002722295* (Fig. S5). Interestingly, the massiliachelin BGC did not cluster with this GCF, likely due to a missing *agrB* homolog in the MiBIG accession (BGC0002687). In total, there were 181 BGCs across the 17 strains, consisting of 52 GCFs and 22 singletons. Notably, only two GCFs clustered with known BGCs in MiBIG, the sesbanimide BGC from *Stappia indica* PHM037 (57) and the ectoine BGC from the *Methylococcus marina* (60).

BGC analysis and evidence for the biosynthesis of agrochelin II

With the *Stappia* 28M-7^T massiliachelin/agrochelin BGC sequence in hand, we next proceeded to assign the stereochemistry of the associated thiazole alkaloid using a bioinformatics-led approach. The presence of a methyltransferase-like epimerization (MT_e) domain in the PKS-NRPS hybrid proteins responsible for the biosynthesis of thiazoline-containing siderophores is required for a D-configured ring (61). Conserved domain searching (62) revealed that the MT_e domain of AgrA from *Stappia* 28M-7^T is likely non-functional, due to the presence of an eight amino acid insertion within the active site region of this domain, consistent with a 14S configuration derived from incorporation of L-cysteine into the thiazoline ring. The presence of only a single adenylation domain for L-cysteine in AgrA further supports a 15S configuration equivalent to that of enantio-pyochelin (6) and thiazostatin (8) (Fig. 6) (63). The ketoreductase domain of AgrB has an aspartic acid at position 94 (Fig. S6), which is known, or predicted, to catalyze reduction resulting in a D-hydroxy (64); a proline at position 141 further supports this B-type stereochemistry (65). Intriguingly, the ketoreductase domains of HMWP1 (yersinibactin), MicC (micacocidin), and AgrA (*Stappia* 28M-7^T BGC) have a tyrosine to histidine substitution at residue 147. Tyrosine is part

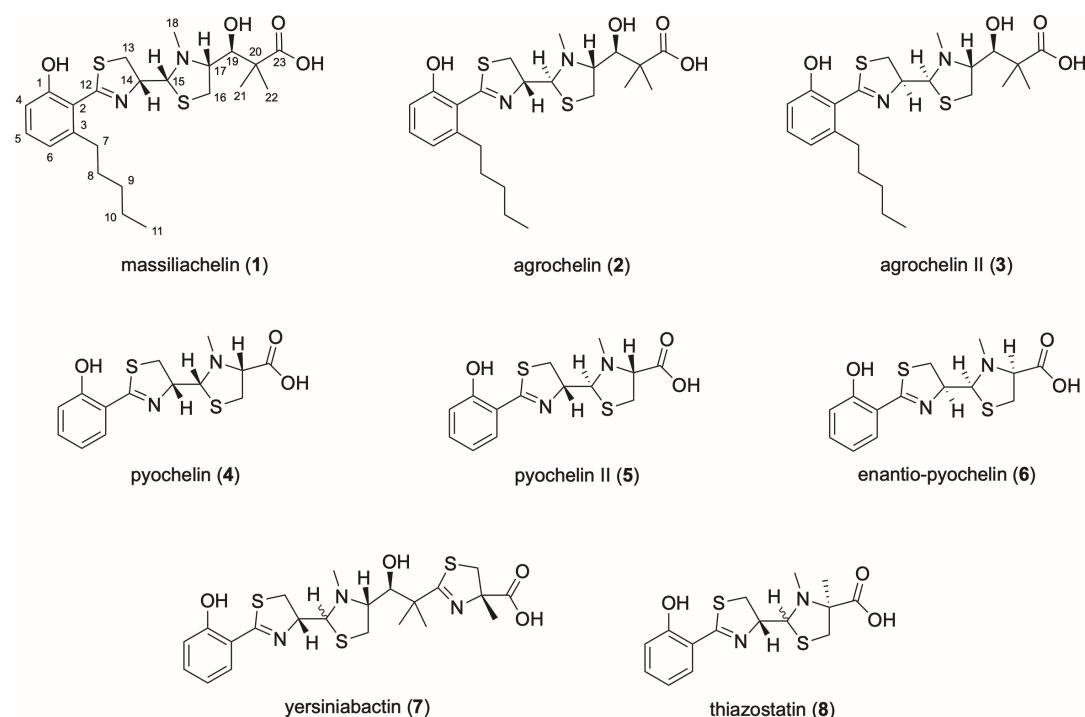


FIG 6 Comparative chemical structures of thiazole alkaloid natural products.

of the catalytic triad of ketoreductase domains required for activity, and these ketoreductase domains are predicted to be inactive (64–66). This observation conflicts with our spectroscopic data and suggests that ketoreduction at C19 is achieved via an alternative mechanism. Collectively, our spectroscopic and bioinformatics analyses are consistent with the assignment of the identified bioactive *Stappia* 28M-7^T metabolite as the previously unreported natural product agrochelin II (3) (Fig. 6).

DISCUSSION

Here, we report the isolation and characterization of three bacterial species—“*Bacillus crepusculi*” sp. nov., “*Stappia quadratibracata*” sp. nov., and “*Psychrobacter noctis*” sp. nov.—from deep-sea sponges from the equatorial Atlantic Ocean. These findings form part of a wider collaborative study seeking to investigate the composition of the deep-sea sponge microbiome and its usefulness as a source of bioactive natural products (21).

TABLE 3 BGCs identified in the genome of strain *Stappia* 28M-7^T by antiSMASH 7.0^a

BGC	Type	Size (kb)	Predicted product
1	NRPS	41.7	Unknown
2	Thiopeptide, <i>trans</i> -AT PKS, T3PKS, NRPS	119	Sesbanimide R
3	RiPP-like	11.4	Unknown
4	T1PKS	46.6	Capsular lipooligosaccharide
5	NRPS, T1PKS	77.7	Agrochelin and massiliachelin
6	PKS-like, ectoine, NRPS-independent-siderophore	14.3	Ectoine and desferrioxamine relative
7	Terpene	20.9	Carotenoid
8	NAGGN	14.8	NAGGN
9	Phosphonate	41.6	Phosphinothricin tripeptide

^aAssignment of predicted product is based on our own manual annotations as informed by the published literature. NRPS, nonribosomal peptide synthetase; *trans*-AT PKS, *trans*-acyltransferase polyketide synthase; T1/3PKS, type I/III polyketide synthase; RiPP-like, unspecified ribosomally synthesised and post-translationally modified peptide product; NAGGN, *N*-acetylglutaminylglutamine.

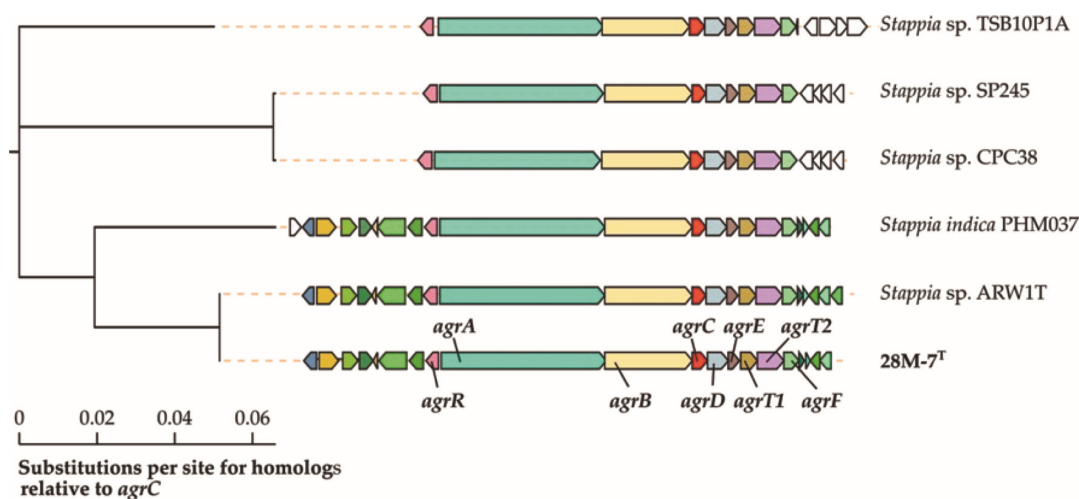


FIG 7 Cladogram for the minimal massiliachelin/agrochelin BGC from strain *Stappia* 28M-7^T compared to those of its closest relatives, generated using CORASON, with *agrC* as the query gene. Homologous genes are displayed using the same color. The annotated genes for the 28M-7^T massiliachelin/agrochelin BGC are shown in Table S13.

To explore the secondary metabolic potential of the three strains described herein antibacterial screening, in combination with OSMAC cultivation was undertaken, an experimental approach which proved effective in establishing the biosynthetic potential of the isolated bacteria. Both *Bacillus* 28A-2^T and *Stappia* 28M-7^T were found to biosynthesize metabolites with antibacterial activities. While the active metabolites from 28A-2^T could not be recovered in sufficient quantities to enable their formal identification, genome analysis identified BGCs encoding known antimicrobial compounds, findings consistent with our *in vitro* screening data (43, 44). In contrast, growth of *Stappia* 28M-7^T on succinate as a carbon source yielded a dominant metabolite, with spectroscopic and genomic data consistent with the previously unreported thiazole alkaloid agrochelin II, a close structural relative of the siderophores massiliachelin (1) and agrochelin (2). Intriguingly, complimentary untargeted metabolomics studies of *Stappia* 28M-7^T identified a number of additional distinct yet closely related analogs of agrochelin II, present within the fermentation media. The molecular masses of these compounds were consistent with both extended chain and truncated thiazole alkaloid analogs, along with the structurally related compound pyochelin. These findings contrast with those of Steinmetz et al., in their studies of *Massilia* sp. NR 4-1, where although multiple shunt products were identified, each retained the salicylate-carbon chain core, but were otherwise incomplete (67). The production of a broader portfolio of metabolites in *Stappia* 28M-7^T indicates a degree of promiscuity in the starter unit selection machinery of the associated biosynthetic pathway, consistent with previous mutasynthesis studies of both the yersiniabactin and pyochelin synthetases (68, 69). The biological relevance of relaxed starter unit selectivity in these pathways remains to be unambiguously established; however, such findings hint at the potential of these systems for use in the production of functionally optimized “non-natural” analogs.

Given the inherent chemical complexities common to natural product scaffolds, unambiguous structure elucidation remains a significant technical challenge. To address this issue, in our studies of *Stappia* 28M-7^T, bioinformatic analysis has been successfully employed to resolve the structure of the dominant bioactive metabolite produced by this bacterium, agrochelin II. We propose that this approach be considered of broader utility in informing natural product structure elucidation, in particular in instances where compound instability and/or structural complexity precludes structure determination by spectroscopic means alone. The *Stappia* MTe domain harbors an eight-residue insertion near the predicted enzyme active site, which is not observed in the pyochelin, yersiniabactin, micacocidin, or massiliachelin BGCs. Disruption of MTe activity precluding L- to

D-cysteine epimerization accounts for the 14S configuration observed in agrochelin II. While further investigation is needed to formally validate this hypothesis, it provides a plausible biosynthetic rationale for the stereochemical divergence of agrochelin II.

The genus *Stappia* remains a largely unexplored source of bioactive natural products, with studies to date having focused predominantly on the ecology, sulfur metabolism, and physiology of this bacterium (70, 71). As a consequence, the findings outlined herein constitute one of the few reported attempts to establish the usefulness of this bacterium as a source of novel secondary metabolites. Despite their widespread occurrence, comparatively little is known about the phylogenetic and physiological relationships amongst *Stappia* isolates, or how environmental factors influence their distribution and biological activities. The paucity of studies exploring the biosynthetic potential of *Stappia* is particularly surprising given that this genera resides within the *Rhodobacteraceae*/*Roseobacter* clade, which contains a number of well-studied, biosynthetically gifted bacteria (72, 73). The absence of previous studies demonstrating “hit” identification in *Stappia* has undoubtedly disincentivized researchers to investigate natural product biosynthesis in this genera, further compounded by complexities associated with growth and cultivation, which are generally considered to be more challenging than for more historically favored producers (36). However, given the findings of this study, we suggest that *Stappia* ought to be reconsidered as an attractive target for future investigations seeking to identify clinically relevant bioactive metabolites.

Finally, our findings provide further support for the usefulness of deep-sea bioprospecting as an approach for the discovery of bioactive microbial natural products, including those with antimicrobial activities. This adds to a growing body of published literature, including from our own laboratories, which indicates that the distinctive environmental conditions experienced by deep-sea microbes translate into the acquisition of metabolic novelty (21–23). The value of integrating OSMAC screening into discovery platforms is also demonstrated, with succinate supplementation shown to be an effective strategy for the activation of silent BGCs, consistent with previous reports of succinate-induced siderophore production (74, 75). Collectively, these findings provide compelling evidence for the benefit of incorporating succinate into OSMAC screens as a general mechanism for BGC activation. Interestingly, iron-associated agrochelin II displayed enhanced antibacterial activity, potentially through increased cellular uptake or enhanced potency (76), which to our knowledge has not previously been reported for structurally related bioactive natural products. On the basis of this finding, we suggest value in revisiting previously reported thiazole alkaloids to establish if iron supplementation constitutes a general strategy of enhancing the antimicrobial activities of this family of compounds.

In summary, this study expands the known chemical space of *Stappia* natural products and marine siderophores, and demonstrates the value of integrated cultivation–screening–metabolomics–genomics pipelines for the discovery of bioactive secondary metabolites.

Descriptions of new species

***Description of Stappia quadratibracata* sp. nov.**

Stappia quadratibracata: qua.dra.ti.bra.ca’ ta. L. fem. nom. n. *quadrata*, square or squared, L. fem. adj. *bracata*, wearing trousers or breeched, N.L. fem. adj. *quadratibracata*, square-trousered, for the character SpongeBob SquarePants, referencing the sponge from which the type strain was isolated.

Cells are Gram-stain negative, aerobic, flagellated rods approximately $0.9\text{--}1.3 \times 0.3\text{--}0.5\text{ }\mu\text{m}$ in size. On marine agar 2216, colonies are pinkish-beige, shiny, viscous, opaque, round, and convex with entire margins after 72 h growth at 28 °C. Growth occurs on unbuffered nutrient agar at 15–37°C, at pH 5–12, and with 0%–8% (wt/vol) NaCl. Using the API 20NE system, cells are positive for oxidase, nitrate respiration, aesculin hydrolysis, gelatinase, and β -galactosidase activities, and negative for denitrification, indole production, fermentation, arginine dihydrolase, and urease activities. With the API 20NE

system, positive for nitrate respiration, denitrification, and oxidase activities; negative for indole production, fermentation, arginine dihydrolase, urease, esculin hydrolysis, gelatinase, and β -galactosidase activities. D-Glucose, D-mannitol, adipic acid, malic acid, citric acid, and phenylacetic acid are assimilated as sole carbon sources; L-arabinose, D-mannose, *N*-acetylglucosamine, D-maltose, gluconic acid, and capric acid are not.

The type strain is 28M-7^T (= NCIMB 15403^T = DSM 112057^T), which was isolated from a deep-sea sponge from the Knipovich seamount in the Atlantic Ocean. The genome size of the type strain is 5.13 Mb, and its genomic DNA G + C content is 66.7 mol %.

Description of *Bacillus crepusculi* sp. nov.

Bacillus crepusculi: cre' pus.cul.i. L. gen. n. *crepusculi* of the twilight, referring to the "twilight zone" of the ocean from which the type strain was isolated.

Cells are Gram-stain positive, aerobic, spore-forming rods approximately 2.7–3.5 $\times$ 0.6–1.0 μ m in size and motile by means of peritrichous flagella. Exhibits sliding surface motility on media with 1% agar or less. Inhibits the growth of *S. aureus*, *E. faecalis*, and *B. subtilis* after growth on Mueller-Hinton or actinomycetes isolation agar for 3 days. On Luria-Bertani agar (LBA), colonies are opaque, off-white, shiny, circular, and flat with undulate margins after 48 h growth at 28 °C. Growth occurs on unbuffered LBA at 15–45°C, at pH 5–12, and with 0%–10% (wt/vol) NaCl. Using the API 20NE system, cells are positive for oxidase, nitrate respiration, aesculin hydrolysis, gelatinase, and β -galactosidase activities, and negative for denitrification, indole production, fermentation, arginine dihydrolase, and urease activities. With the API 20NE system, D-Glucose, L-arabinose, D-mannose, D-mannitol, *N*-acetylglucosamine, D-maltose, gluconate, malate, and citrate are assimilated as sole carbon sources, but capric acid, adipic acid, and phenylacetate are not.

The type strain is 28A-2^T (= NCIMB 15401^T = DSM 111873^T), which was isolated from a deep-sea sponge from the Vayda seamount in the Atlantic Ocean. The genome size of the type strain is 3.70 Mb, and its genomic DNA G + C content is 41.7 mol %.

Description of *Psychrobacter noctis* sp. nov.

Psychrobacter noctis: noc' tis. L. gen. n. *noctis* of the night, referring to the "midnight zone" of the ocean from which the type strain was isolated.

Cells are Gram-stain negative, aerobic, non-spore-forming cocci approximately 1.2–1.6 μ m in diameter arranged in chains. On nutrient agar, colonies are opaque, cream, shiny, circular, and convex with entire margins after 72 h growth at 28°C. Growth occurs on unbuffered LBA at 4–34°C, at pH 5–12, and with 0%–12% (wt/vol) NaCl. Using the API 20NE system, cells are positive for oxidase activity and negative for nitrate respiration, denitrification, indole production, fermentation, arginine dihydrolase, urease, aesculin hydrolysis, gelatinase, and β -galactosidase activities. With the API 20NE system, D-Glucose, L-arabinose, D-mannose, D-mannitol, *N*-acetylglucosamine, D-maltose, gluconate, capric acid, adipic acid, malate, citrate, and phenylacetate are not used as sole carbon sources.

The type strain is 28M-43^T (= NCIMB 15404^T = DSM 112058^T), which was isolated from a deep-sea sponge from the Vayda seamount in the Atlantic Ocean. The genome size of the type strain is 3.12 Mb and its genomic DNA G + C content is 43.0 mol %.

MATERIALS AND METHODS

Isolation and antimicrobial screening of bacteria

The sponges, Demosponge B0078 (971 m, Knipovich Seamount), Hexactinellid B01641 (868 m, Vayda Seamount), and Demosponge B01642 (1,483 m, Vayda Seamount) were processed, and bacteria were isolated from them as previously described (21). In short, we rinsed the sponges thrice with artificial seawater to remove loosely associated

bacteria and then ground each with artificial seawater in a sterile pestle and mortar. We made serial dilutions of the resulting solutions, spread the dilutions on a range of agar-based media, and incubated them at 4°C and 28°C for 2–10 weeks. We selected individual colonies and re-streaked them to yield axenic cultures, which we used to inoculate Microorganism Preservation System Protect Cryotubes (Technical Service Consultants Ltd, Heylwood, UK) for storage at –70°C. Strains 28M-7^T and 28M-43^T were isolated on marine agar 2216 (BD Difco) at 28°C, and 28A-2^T was isolated on actinomycetes isolation agar (Sigma-Aldrich) at 28°C.

The bacteria isolated from the sponges were screened for antibacterial activity with a soft agar overlay assay, as previously described (21). Initially, we grew the sponge bacteria on agar under the same conditions that they were isolated under and overlaid the cultures with suspensions of bacteria pathogens suspended in soft agar. We then selected a subset of our sponge bacteria, including strains 28M-7^T, 28A-2^T, and 28M-43^T for an OSMAC screen, as described in Williams et al. (21). Briefly, we grew 90 strains on M9 agar (Serva Electrophoresis GmbH, Heidelberg, Germany) with a range of different carbon sources (colloidal chitin, dextrose, glycerol, mannitol, soluble starch, and succinate) for a week and then overlaid them with pathogens in soft agar. The sponge isolates which fell within a zone of inhibition were grown as monocultures on the relevant agar and screened again in triplicate—those which continued to produce zones of inhibition were considered active. A summary of the OSMAC screening results for 28M-7^T is provided in Table S11, and the compositions of the media used in this study are shown in Table S16.

Genome sequencing and assembly

For genome sequencing, each strain was cultured in 10 mL of broth culture in a shaking incubator for three days at 28°C and 180 rpm. The liquid broth used for this step was equivalent to the agar composition used for the original isolation. Cells were harvested by centrifugation and genomic DNA extracted using a GenElute Bacterial Genomic DNA Kit (Sigma-Aldrich, St. Louis, MO, USA). The concentration and purity of the genomic DNA was assessed spectroscopically by measuring the absorbance at 230, 260, and 280 nm, and gel electrophoresis on a 0.8% agarose gel was used to assess the degree of DNA shearing.

Sequencing was performed by the Genomics Facility at the University of Bristol. One microgram of DNA was used to generate a paired end TruSeq library. This was indexed with a Nextera XT Index Kit prior to shotgun sequencing on an Illumina MiSeq platform (Illumina, San Diego, USA). Illumina reads were trimmed with Trim Galore v 0.4.4 (77) using a PHRED score of 20. In parallel, 1 µg of the same genomic DNA sample was also sequenced with a R9 MinION flow-cell indexed with the EXP-NBD103 kit (Oxford Nanopore Technologies, Oxford, UK) following the manufacturer's instructions. Nanopore reads were basecalled with Albacore v 1.2.5 with adaptor trimming and demultiplexing via Porechop (Porechop, RRID: SCR_016967). Hybrid genome assemblies were generated with the Unicycler pipeline v 0.4.6 using default parameters in “normal” mode, first performing a de Bruijn graph assembly using the Illumina paired-end reads with SPAdes (78) and then using a semi-global aligner to align long Nanopore reads into a contiguous single assembly graph (79). Assemblies were quality checked with CheckM v 1.1.6 (80) and BUSCO v 5.3.2 (81). PlasFlow v 1.1.0 (38) was used to check for likely plasmids in assemblies via the web platform (<https://usegalaxy.eu/>) (82). Mobilome-related genes, including conjugation systems and plasmid maintenance genes, were annotated using Prokka v1.13 (83).

Taxonomy

Genome assemblies were annotated with Prokka v 1.13 (83). Nucleotide sequences for complete 16S rRNA genes were used for nucleotide BLAST searches of the NCBI 16S rRNA database. The Genome Taxonomy Database Toolkit (GTDB-Tk v 2.4.1) was used to assign taxonomy by classifying the assemblies in the GTDB reference tree v.

R220 (32, 33). For taxonomic placement with recognized type strains, we used the Type Strain Genome Server v 401 (TYGS) (31) pipeline. TYGS implements the Genome BLAST Distance Phylogeny (GBDP) method to infer whole genome phylogenetic trees (29). Intergenomic distances were calculated with the GBDP d5 distance formula, and trees inferred with FastME 2.1.4 with subtree pruning and regrafting postprocessing (84, 85). Phylogenetic tree branch support was assessed using 100 GBDP pseudo-bootstrap replicates, and robustness of treelikeness was assessed using δ statistic, where low δ values indicate strong treelikeness and high branch support (86). Only validly published type strains were included in phylogenetic analyses, consistent with standard practice for novel species descriptions. Closest reference strains were selected based on TYGS and GTDB-Tk results. Trees were rooted using an appropriate outgroup and visualized using the Interactive Tree of Life v 5.7 (iTOL) (39). ANI values were calculated using fastANI v 1.31 (30).

The API 20NE kit (BioMérieux, Marcy-l'Étoile, France) was used to assess enzymatic activities and carbon source utilization for isolated strains, following the manufacturer's instructions and repeating the assays in triplicate. The pH of the standard agars was adjusted to 5, 6, 7, 8, 9, 10, 11, and 12. The NaCl content of the standard agars was adjusted to 0%, 2%, 4%, 6%, 8%, 10%, and 12% (wt/vol). Standard agar plates were incubated at 4°C, 15°C, 20°C, 28°C, 37°C, 42°C, and 45°C and incubated for 3 days (or a month for 4°C). All culturing was performed in triplicate.

For transmission electron microscopy (TEM), 5 μ L of the liquid culture after 72 h growth was drop cast on to a carbon-coated formvar TEM grid pre-dipped in 2.5% glutaraldehyde. After 5 min, excess material was wicked away and grids washed twice in distilled water before imaging with an FEI Tecnai12 BioTwin TEM with an FE CETA camera.

The three strains described in this study are publicly available from the NCIMB (UK) and DSMZ (Germany) culture collections under the following accessions: *Stappia quadratibracata* 28M-7^T (= NCIMB 15403^T = DSM 112057^T), *Bacillus crepusculi* 28A-2^T (= NCIMB 15401^T = DSM 111873^T), and *Psychrobacter noctis* 28M-43^T (= NCIMB 15404^T = DSM 112058^T)

Metabolomics

For metabolomics analysis, a primary culture of *Stappia* 28M-7^T was grown in M9 media plus 10 g/L succinate (Sigma-Aldrich) (pH 7.0) for 4 days at 28°C at 180 rpm. Five milliliters of this primary was inoculated into 100 mL of the same media in a 250 mL Erlenmeyer flask and incubated for a further 8 days under the same conditions. Cells were harvested by centrifugation (4,000 $\times g$, 20 min, 4°C), and the remaining supernatant was mixed with an equal volume of ethyl acetate (Sigma-Aldrich). The aqueous phase was discarded and anhydrous MgSO₄ (Sigma-Aldrich) was added to the organic phase to bind any remaining water. The organic phase was filtered and concentrated *in vacuo* until dry, before redissolving the remaining extract in 1 mL LC-MS grade methanol (Sigma-Aldrich). LC-MS/MS data were acquired in electrospray ionization (ESI) positive mode on a Waters XEVO G2-XS QToF (Waters, Milford, MA, USA) with a C18 reverse-phase column (Phenomenex, 150 $\times$ 15 mm, 5 μ m; Phenomenex, Torrance, CA, USA). A linear gradient of 5% MeCN to 95% MeCN with 0.1% formic acid over 20 min at a flow rate of 0.25 mL/min (Sigma-Aldrich). Waters .raw data files were converted to mzML format using MSConvert (v 3.0), and data were processed in MZmine (mzmine.io v 4.7.8) (87) using the mzmine wizard with default parameters for the instrument and applying a strict MS¹ intensity cut-off of 1×10^5 . Feature-based molecular networking was performed on GNPS2 (Workflow: 2025.07.09) (88). The molecular network was visualized using Cytoscape (v 3.10) (89). MetFrag (v 2.6.10) was used to generate *in silico* fragmentation via <https://msbi.ipb-halle.de/MetFrag/> (49), using the agrochelin structure from NPAtlas (NPA017357) (90). Finally to identify modifications within the network, Modifinder (50) was run on the GNPS2 platform using the agrochelin SMILES to identify chemical modifications of related nodes.

Natural product isolation

For culture scale-up, a primary 300 mL culture of 28M-7^T was grown in M9 media plus 10g/L succinate adjusted to pH 7.0 for 4 days at 28°C on a shaking table at 180 rpm. Fifty milliliters of this primary culture was inoculated into each of 5 × 1 L culture broth in 2.5 L glass flasks and grown on for a further 8 days. Cells were harvested by centrifugation, and the supernatant was mixed with an equal volume of ethyl acetate (Sigma-Aldrich). The aqueous phase was discarded, and organic phase was dried with anhydrous MgSO₄, filtered, and concentrated *in vacuo* redissolved in 2 mL methanol (Sigma-Aldrich). LC-MS data were obtained on a Waters LCMS system comprising Waters 2767 autosampler, Waters 2545 HPLC pump, Waters 2998 Diode Array detector, Waters 2424 ELS detector, and Waters Quatro Micro mass spectrometer. HPLC grade H₂O and MeCN were added with 0.05% formic acid as the solvent system. Analytical LC-MS data were obtained using a Phenomenex Luna column (C18, 250 × 4.60 mm, 5 μm) at a flow rate of 1 mL/min, with a gradient of 5% MeCN to 95% MeCN over 20 min. Preparative HPLC purification was carried out using a Phenomenex Kinetex column (C18, 250 × 21.20 mm, 5 μm) at a flow rate of 16 mL/min, with a gradient of 30% to 40% of MeCN over 20 min. HR-ESI-MS data were obtained on a Bruker Daltonics micrOTOF II instrument. NMR data were collected on a Bruker Cryo500 NMR spectrometer.

Natural product antibacterial assays

To determine the minimal inhibitory concentrations (MIC) of agrochelin II, we followed the broth dilution protocol outlined by the European Committee for Antimicrobial Susceptibility Testing (91). We performed the assay in triplicate, using ampicillin as a positive control, and ammonium iron(II) sulfate and methanol as negative controls. We tested purified agrochelin II alone and preincubated with equimolar ammonium iron(II) sulfate. We determined MIC values for a panel of clinically relevant bacteria: *Staphylococcus aureus* Newman and Mu50, *Pseudomonas aeruginosa* PA01, *Klebsiella pneumoniae* NCTC 5055, *Acinetobacter baumannii* ATCC 19606, and *Escherichia coli* BW55113.

Genome mining

Gene cluster analysis of the isolated strains used antiSMASH v.7.0 (92) which utilizes MiBIG 3.0 (93). CORASON (94) was used to define the core agrochelin BGC which was further annotated and defined based on gene conservation within *Stappia*. For the comparative gene cluster family analysis of the *Stappia* genus, only genomes which were assigned to the GTDB *Stappia* genus were included (Table S15). Genomes were annotated with antiSMASH (95) (v.8.0.4) and clustered into gene cluster families with BIG-SCAPE v.2.0.0 run in mixed mode (59) and with MiBIG 4.0 reference BGCs (58). The GCF cluster cut-off value was set to 0.35, and the cluster network was exported and visualized in Cytoscape (v 3.10) (89). Amino acid sequences of the MTe domain were extracted and aligned by MAFFT –auto (v 7.490) (96).

ACKNOWLEDGMENTS

We thank the Bristol Genomics Facility for performing Illumina sequencing. We thank Laura Robinson, Chief Scientist for expedition JC094, and Kate Hendry for the initial collection of the sponge samples and for facilitating subsequent access.

This work was funded by BBSRC grants BB/T001968/1, BB/L01386X/1, and BB/M025624/1 and through the award of PhD studentships to H.L.S. (EPSRC Centre for Doctoral Training in Synthetic Biology; EP/L016494/1 and Dstl) and S.E.W. (MRC GW4 BIOMED DTP; MR/N013794/1). Postdoctoral funding was also awarded to S.E.W. (Novo Nordisk Foundation Postdoctoral Fellowship; NNF22OC0079021).

Conceptualization, H.L.S., S.E.W., C.L.W., P.C., and P.R.R.; methodology, H.L.S., S.E.W., A.J.D., Z.S., C.R.B., J.M., and M.A.J.; formal analysis, S.E.W., H.L.S., A.J.D., Z.S., L.W., J.E.M.S., C.L.W., and P.R.R.; investigation, S.E.W., H.L.S., Z.S., L.W., C.R.B., J.M., C.N., and M.A.J.;

resources, M.A.J., C.L.W., P.C., and P.R.R.; data curation, S.E.W. and H.L.S.; writing, S.E.W., H.L.S., J.D.S., J.E.M.S., C.L.W., and P.R.R.; visualization, S.E.W., H.L.S., and J.E.M.S.; supervision, A.E.E-L., C.L.W., P.C., and P.R.R.; project administration, P.C. and P.R.R.; funding acquisition, A.E.E-L., C.L.W., P.C., and P.R.R. All authors have read and agreed to the published version of the manuscript.

AUTHOR AFFILIATIONS

¹School of Biochemistry, University of Bristol, Bristol, United Kingdom

²Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Lyngby, Denmark

³School of Chemistry, University of Bristol, Bristol, United Kingdom

⁴Institute of Microbiology, Chinese Academy of Sciences, Beijing, China

⁵Wolfson Bioimaging Facility, University of Bristol, Bristol, United Kingdom

⁶Defence Science and Technology Laboratory, Salisbury, United Kingdom

⁷School of Pharmacy, Newcastle University, Newcastle upon Tyne, United Kingdom

⁸School of Natural and Environmental Sciences, Newcastle University, Newcastle upon Tyne, United Kingdom

AUTHOR ORCIDs

Sam E. Williams  <http://orcid.org/0000-0002-6949-426X>

Andrew J. Devine  <http://orcid.org/0000-0002-4200-5542>

Catherine R. Back  <http://orcid.org/0000-0002-9835-9171>

Luoyi Wang  <http://orcid.org/0000-0003-2032-7206>

Judith Mantell  <http://orcid.org/0000-0002-9061-8933>

Jonathan D. Sellars  <http://orcid.org/0000-0003-1430-012X>

Christine L. Willis  <http://orcid.org/0000-0002-3919-3642>

Paul Curnow  <http://orcid.org/0000-0001-9889-0248>

Paul R. Race  <http://orcid.org/0000-0003-0184-5630>

FUNDING

Funder	Grant(s)	Author(s)
Medical Research Council	MR/N0137941/1	Sam E. Williams
Engineering and Physical Sciences Research Council	EP/L016494/1	Henry L. Stennett
Biotechnology and Biological Sciences Research Council	BB/L01386X/1	Paul R. Race
Biotechnology and Biological Sciences Research Council	BB/M025624/1	Paul R. Race
Biotechnology and Biological Sciences Research Council	BB/T001968/1	Paul R. Race
Novo Nordisk Fonden	NNF22OC0079021	Sam E. Williams

AUTHOR CONTRIBUTIONS

Sam E. Williams, Data curation, Formal analysis, Investigation, Methodology, Validation, Visualization, Writing – original draft, Writing – review and editing | Henry L. Stennett, Formal analysis, Investigation, Methodology, Validation, Visualization, Writing – review and editing | Andrew J. Devine, Formal analysis, Investigation, Methodology, Writing – review and editing | Zhongshu Song, Formal analysis, Investigation, Methodology | Catherine R. Back, Formal analysis, Investigation, Writing – review and editing | Luoyi Wang, Formal analysis, Investigation, Writing – review and editing | Judith Mantell, Formal analysis, Investigation, Writing – review and editing | Chris Neal, Formal analysis,

Investigation, Writing – review and editing | Mark A. Jepson, Formal analysis, Supervision, Writing – review and editing | Angela E. Essex-Lopresti, Conceptualization, Funding acquisition, Methodology, Project administration, Supervision, Writing – review and editing | Jonathan D. Sellars, Formal analysis, Validation, Visualization, Writing – review and editing | James E. M. Stach, Formal analysis, Investigation, Methodology, Writing – original draft, Writing – review and editing | Christine L. Willis, Conceptualization, Methodology, Project administration, Resources, Supervision, Writing – original draft, Writing – review and editing | Paul Curnow, Conceptualization, Funding acquisition, Methodology, Project administration, Resources, Supervision, Visualization, Writing – original draft, Writing – review and editing | Paul R. Race, Conceptualization, Formal analysis, Funding acquisition, Methodology, Project administration, Resources, Supervision, Validation, Visualization, Writing – original draft, Writing – review and editing

DATA AVAILABILITY

Genome and 16S rRNA gene GenBank accession numbers for the reported strains are as follows: *Stappia quadratibracata* 28M-7^T genome: [JACMIA000000000.1](https://doi.org/10.1093/jac/dkz000); assembly: [GCA_014252955.1](https://doi.org/10.1093/gca/014252955.1); 16S rRNA gene: [MT999484.1](https://doi.org/10.1093/mt/999484.1); *Bacillus crepusculi* 28A-2^T genome: [JACXXE000000000.1](https://doi.org/10.1093/jac/dkz000000.1); assembly: [GCA_014764715.1](https://doi.org/10.1093/gca/014764715.1); 16S rRNA gene: [MW003702.1](https://doi.org/10.1093/mw/003702.1); *Psychrobacter noctis* 28M-43^T genome: [CP061739.1](https://doi.org/10.1093/cp/061739.1), assembly: [GCA_014770435.1](https://doi.org/10.1093/gca/014770435.1); 16S rRNA gene: [MW003704.1](https://doi.org/10.1093/mw/003704.1). The mass spectrometry data supporting this study have been deposited in the MassIVE repository under accession number [MSV000098784](https://doi.org/10.1093/msv/000098784). The MS2 spectra of agrochelin was deposited in the GNPS library under the accession [CCMSLIB00016341469](https://doi.org/10.1093/ccmslib/00016341469).

ADDITIONAL FILES

The following material is available [online](#).

Supplemental Material

Supplemental material (AEM01877-25-s0001.pdf). Tables S1 to S16, Fig. S1 to S6, and compound spectra.

REFERENCES

- Cheeptham N, Saiz-Jimenez N. 2015. New sources of antibiotics: caves, p 213–228. In Sanchez S, Demain AL (ed), *Antibiotics: current innovations and future trends*. Caister Academic Press.
- Tortorella E, Tedesco P, Palma Eposito F, January GG, Fani R, Jaspars M, de Pascale D. 2018. Antibiotics from deep-sea microorganisms: current discoveries and perspectives. *Mar Drugs* 16:355. <https://doi.org/10.3390/md16100355>
- Mahajan GB, Balachandran L. 2017. Sources of antibiotics: hot springs. *Biochem Pharmacol* 134:35–41. <https://doi.org/10.1016/j.bcp.2016.11.021>
- Coker JA. 2019. Recent advances in understanding extremophiles. *F1000Res* 8:8. <https://doi.org/10.12688/f1000research.20765.1>
- Rampelotto PH. 2013. Extremophiles and extreme environments. *Life* (Basel) 3:482–485. <https://doi.org/10.3390/life3030482>
- Stennett HL, Tiwari K, Williams S, Curnow P, Race P. 2020. The extremophilic pharmacy: drug discovery at the limits of life, p 43–72. In Lee NM (ed), *Biotechnological applications of extremophilic microorganisms*. De Gruyter.
- Srinivasan R, Kannappan A, Shi C, Lin X. 2021. Marine bacterial secondary metabolites: a treasure house for structurally unique and effective antimicrobial compounds. *Mar Drugs* 19:530. <https://doi.org/10.3390/md19100530>
- Carroll AR, Copp BR, Grkovic T, Keyzers RA, Prinsep MR. 2025. Marine natural products. *Nat Prod Rep* 42:257–297. <https://doi.org/10.1039/d4np00067f>
- Chen J, Jia Y, Sun Y, Liu K, Zhou C, Liu C, Li D, Liu G, Zhang C, Yang T, et al. 2024. Global marine microbial diversity and its potential in bioprospecting. *Nature* 633:371–379. <https://doi.org/10.1038/s41586-024-07891-2>
- Voser TM, Campbell MD, Carroll AR. 2022. How different are marine microbial natural products compared to their terrestrial counterparts? *Nat Prod Rep* 39:7–19. <https://doi.org/10.1039/d1np00051a>
- Karthikeyan A, Joseph A, Nair BG. 2022. Promising bioactive compounds from the marine environment and their potential effects on various diseases. *J Genet Eng Biotechnol* 20:14. <https://doi.org/10.1186/s43141-021-00290-4>
- Bull AT, Ward AC, Goodfellow M. 2000. Search and discovery strategies for biotechnology: the paradigm shift. *Microbiol Mol Biol Rev* 64:573–606. <https://doi.org/10.1128/MMBR.64.3.573-606.2000>
- Pidot SJ, Coyne S, Kloss F, Hertweck C. 2014. Antibiotics from neglected bacterial sources. *Int J Med Microbiol* 304:14–22. <https://doi.org/10.1016/j.ijmm.2013.08.011>
- Goh KM, Shahar S, Chan K-G, Chong CS, Amran SI, Sani MH, Zakaria II, Kahar UM. 2019. Current status and potential applications of underexplored prokaryotes. *Microorganisms* 7:468. <https://doi.org/10.3390/microrganisms7100468>
- Bentley SD, Chater KF, Cerdeño-Tárraga A-M, Challis GL, Thomson NR, James KD, Harris DE, Quail MA, Kieser H, Harper D, et al. 2002. Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). *Nature* 417:141–147. <https://doi.org/10.1038/417141a>
- Romano S, Jackson SA, Patry S, Dobson ADW. 2018. Extending the “One strain many compounds” (OSMAC) principle to marine microorganisms. *Mar Drugs* 16:244. <https://doi.org/10.3390/md16070244>
- Bode HB, Bethe B, Höfs R, Zecek A. 2002. Big effects from small changes: possible ways to explore nature’s chemical diversity. *Chembiochem* 3:619–627. [https://doi.org/10.1002/1439-7633\(20020703\)3:7<619::AID-CBIC619>3.0.CO;2-9](https://doi.org/10.1002/1439-7633(20020703)3:7<619::AID-CBIC619>3.0.CO;2-9)

18. Fulke AB, Sharma N, Nadekar J. 2025. Darkness to discovery: a comprehensive mini-review on culturable and non-culturable microbial diversity from deep sea. *Microb Ecol* 88:77. <https://doi.org/10.1007/s00248-025-02527-y>
19. Khan S, Wang T, Arifeen MZU, Huang S. 2025. Exploring the bioactive potential of deep-sea microorganisms: a review of recent discoveries. *Bioorg Chem* 161:108521. <https://doi.org/10.1016/j.bioorg.2025.108521>
20. Zhang C, Peng Y, Liu X, Wang J, Dong X. 2024. Deep-sea microbial genetic resources: new frontiers for bioprospecting. *Trends Microbiol* 32:321–324. <https://doi.org/10.1016/j.tim.2024.01.002>
21. Williams SE, Stennett HL, Back CR, Tiwari K, Ojeda Gomez J, Challand MR, Hendry KR, Spencer J, Essex-Lopresti AE, Willis CL, Curnow P, Race PR. 2020. The Bristol sponge microbiome collection: a unique repository of deep-sea microorganisms and associated natural products. *Antibiotics (Basel)* 9:509. <https://doi.org/10.3390/antibiotics9080509>
22. Williams SE, Varliero G, Lurgi M, Stach JEM, Race PR, Curnow P. 2024. Diversity and structure of the deep-sea sponge microbiome in the equatorial Atlantic Ocean. *Microbiology (Reading, Engl)* 170:001478. <https://doi.org/10.1099/mic.0.001478>
23. Williams SE, Back CR, Best E, Mantell J, Stach JEM, Williams TA, Race PR, Curnow P. 2023. Discovery and biosynthetic assessment of “*Streptomyces ortus*” sp. nov. isolated from a deep-sea sponge. *Microb Genom* 9. <https://doi.org/10.1099/mgen.0.000996>
24. Costello MJ, Cheung A, De Hauwere N. 2010. Surface area and the seabed area, volume, depth, slope, and topographic variation for the world’s seas, oceans, and countries. *Environ Sci Technol* 44:8821–8828. <https://doi.org/10.1021/es1012752>
25. Steinert G, Busch K, Bayer K, Kodami S, Arbizu PM, Kelly M, Mills S, Erpenbeck D, Dohrmann M, Wörheide G, Hentschel U, Schupp PJ. 2020. Compositional and quantitative insights into bacterial and archaeal communities of South Pacific deep-sea sponges (Demospongiae and Hexactinellida). *Front Microbiol* 11:716. <https://doi.org/10.3389/fmicb.2020.00716>
26. Sogin ML, Morrison HG, Huber JA, Welch DM, Huse SM, Neal PR, Arrieta JM, Herndl GJ. 2006. Microbial diversity in the deep sea and the underexplored “rare biosphere”. *Proc Natl Acad Sci USA* 103:12115–12120. <https://doi.org/10.1073/pnas.0605127103>
27. Grassle JF. 1989. Species diversity in deep-sea communities. *Trends Ecol Evol* 4:12–15. [https://doi.org/10.1016/0169-5347\(89\)90007-4](https://doi.org/10.1016/0169-5347(89)90007-4)
28. Aristegui J, Agustí S, Middelburg JJ, Duarte CM. 2005. Respiration in the mesopelagic and bathypelagic zones of the oceans, p 181–205. In Giorgio P, Williams P (ed), *Respiration in aquatic ecosystems*. Oxford University Press, USA.
29. Meier-Kolthoff JP, Auch AF, Klenk H-P, Göker M. 2013. Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics* 14:60. <https://doi.org/10.1186/1471-2105-14-60>
30. Jain C, Rodriguez-R LM, Phillippy AM, Konstantinidis KT, Aluru S. 2018. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat Commun* 9:5114. <https://doi.org/10.1038/s41467-018-07641-9>
31. Meier-Kolthoff JP, Göker M. 2019. TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. *Nat Commun* 10:2182. <https://doi.org/10.1038/s41467-019-10210-3>
32. Chaumeil PA, Mussig AJ, Hugenholtz P, Parks DH. 2019. GTDB-Tk: a toolkit to classify genomes with the genome taxonomy database. *Bioinformatics* 36:1925–1927. <https://doi.org/10.1093/bioinformatics/bt z848>
33. Parks DH, Chuvochina M, Rinke C, Mussig AJ, Chaumeil P-A, Hugenholtz P. 2022. GTDB: an ongoing census of bacterial and archaeal diversity through a phylogenetically consistent, rank normalized and complete genome-based taxonomy. *Nucleic Acids Res* 50:D785–D794. <https://doi.org/10.1093/nar/gkab776>
34. Lai Q, Qiao N, Wu C, Sun F, Yuan J, Shao Z. 2010. *Stappia indica* sp. nov., isolated from deep seawater of the Indian Ocean. *Int J Syst Evol Microbiol* 60:733–736. <https://doi.org/10.1099/ijso.0.013417-0>
35. Mayali X, Samo TJ, Kimbrel JA, Morris MM, Rolison K, Swink C, Ramon C, Kim YM, Munoz-Munoz N, Nicora C, Purvine S, Lipton M, Stuart RK, Weber PK. 2023. Single-cell isotope tracing reveals functional guilds of bacteria associated with the diatom *Phaeodactylum tricornutum*. *Nat Commun* 14:5642. <https://doi.org/10.1038/s41467-023-41179-9>
36. Padaki VG, Mayali X, Weber PK, Giovannoni SJ, Abraham K, Jacobs K, Collart L, Halsey KH. 2025. Bacterial volatile organic compound specialists in the phycosphere. *ISME J* 19:229. <https://doi.org/10.1093/ismej/wraf229>
37. Wang F, Gong XF, Meng D, Liu YL, Gu PF, Fan XY, Huang ZS, Du ZJ, Li Q. 2021. *Paracoccus binzhouensis* sp. nov., isolated from activated sludge. *Arch Microbiol* 203:3007–3013. <https://doi.org/10.1007/s00203-021-02286-7>
38. Krawczyk PS, Lipinski L, Dziembowski A. 2018. PlasFlow: predicting plasmid sequences in metagenomic data using genome signatures. *Nucleic Acids Res* 46:e35–e35. <https://doi.org/10.1093/nar/gkx1321>
39. Letunic I, Bork P. 2021. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res* 49:W293–W296. <https://doi.org/10.1093/nar/gkab301>
40. Liu Y, Lai Q, Dong C, Sun F, Wang L, Li G, Shao Z. 2013. Phylogenetic diversity of the *Bacillus pumilus* group and the marine ecotype revealed by multilocus sequence analysis. *PLoS One* 8:e80097. <https://doi.org/10.1371/journal.pone.0080097>
41. Wang L-T, Lee F-L, Tai C-J, Kasai H. 2007. Comparison of *gyrB* gene sequences, 16S rRNA gene sequences and DNA-DNA hybridization in the *Bacillus subtilis* group. *Int J Syst Evol Microbiol* 57:1846–1850. <https://doi.org/10.1099/ijso.0.64685-0>
42. Espariz M, Zuljan FA, Esteban L, Magni C. 2016. Taxonomic identity resolution of highly phylogenetically related strains and selection of phylogenetic markers by using genome-scale methods: the *Bacillus pumilus* group case. *PLoS One* 11:e0163098. <https://doi.org/10.1371/journal.pone.0163098>
43. Walker JE, Abraham EP. 1970. The structure of bacilysin and other products of *Bacillus subtilis*. *Biochem J* 118:563–570. <https://doi.org/10.1042/bj1180563>
44. Yakimov MM, Kröger A, Slepak TN, Giuliano L, Timmis KN, Golyshin PN. 1998. A putative lichenysin A synthetase operon in *Bacillus licheniformis*: initial characterization. *Biochim Biophys Acta* 1399:141–153. [https://doi.org/10.1016/S0167-4781\(98\)00096-7](https://doi.org/10.1016/S0167-4781(98)00096-7)
45. Kenig M, Abraham EP. 1976. Antimicrobial activities and antagonists of bacilysin and anticapsin. *J Gen Microbiol* 94:37–45. <https://doi.org/10.1099/00221287-94-1-37>
46. Yakimov MM, Timmis KN, Fredrickson HL. 1995. Characterization of a new lipopeptide surfactant produced by thermotolerant and halotolerant subsurface *Bacillus licheniformis* BAS50. *Appl Environ Microbiol* 61:1706–1713. <https://doi.org/10.1128/aem.61.5.1706-1713.1995>
47. Valdés-Velasco LM, Favela-Torres E, Théatre A, Arguelles-Arias A, Saucedo-Castañeda JG, Jacques P. 2022. Relationship between lipopeptide biosurfactant and primary metabolite production by *Bacillus* strains in solid-state and submerged fermentation. *Bioresour Technol* 345:126556. <https://doi.org/10.1016/j.biortech.2021.126556>
48. Nothias L-F, Petras D, Schmid R, Dührkop K, Rainer J, Sarvepalli A, Protzyuk I, Ernst M, Tsugawa H, Fleischauer M, et al. 2020. Feature-based molecular networking in the GNPS analysis environment. *Nat Methods* 17:905–908. <https://doi.org/10.1038/s41592-020-0933-6>
49. Ruttkies C, Schymanski EL, Wolf S, Hollender J, Neumann S. 2016. MetFrag relaunched: incorporating strategies beyond *in silico* fragmentation. *J Cheminform* 8:3. <https://doi.org/10.1186/s13321-016-0115-9>
50. Shahneh MRZ, Strobel M, Vitale GA, Geibel C, Abiead YE, Garg N, Wagner B, Forchhammer K, Aron A, Phelan VV, Petras D, Wang M. 2024. Modifinder: tandem mass spectral alignment enables structural modification site localization. *J Am Soc Mass Spectrom* 35:2564–2578. <https://doi.org/10.1021/jasms.4c00061>
51. Avison MB, Bennett PM, Howe RA, Walsh TR. 2002. Preliminary analysis of the genetic basis for vancomycin resistance in *Staphylococcus aureus* strain Mu50. *J Antimicrob Chemother* 49:255–260. <https://doi.org/10.1093/jac/49.2.255>
52. Acebal C, Cañedo LM, Puentes JLF, Baz JP, Romero F, Calle FDL, Grávalos MDG, Rodríguez P. 1999. Agrochelin, a new cytotoxic antibiotic from a marine *Agrobacterium*. Taxonomy, fermentation, isolation, physico-chemical properties and biological activity. *J Antibiot* 52:983–987. <https://doi.org/10.7164/antibiotics.52.983>
53. Dietrich J, Kage H, Nett M. 2019. Genomics-inspired discovery of massiliachelin, an agrochelin epimer from *Massilia* sp. NR 4-1. *Beilstein J Org Chem* 15:1298–1303. <https://doi.org/10.3762/bjoc.15.128>
54. Schlegel K, Taraz K, Budzikiewicz H. 2004. The stereoisomers of pyochelin, a siderophore of *Pseudomonas aeruginosa*. *Biomaterials* 17:409–414. <https://doi.org/10.1023/B:BIOM.0000029437.42633.73>

55. Rinehart KL, Staley AL, Wilson SR, Ankenbauer RG, Cox CD. 1995. Stereochemical assignment of the pyochelins. *J Org Chem* 60:2786–2791. <https://doi.org/10.1021/jo00114a029>
56. Cuppels DA, Stipanovich RD, Stoessl A, Stothers JB. 1987. The constitution and properties of a pyochelin–zinc complex. *Can J Chem* 65:2126–2130. <https://doi.org/10.1139/v87-354>
57. Kačar D, Cañedo LM, Rodríguez P, González EG, Galán B, Schleissner C, Leopold-Messer S, Piel J, Cuevas C, de la Calle F, García JL. 2021. Identification of trans-AT polyketide clusters in two marine bacteria reveals cryptic similarities between distinct symbiosis factors. *Environ Microbiol* 23:2509–2521. <https://doi.org/10.1111/1462-2920.15470>
58. Zdouc MM, Blin K, Louwen NLL, Navarro J, Loureiro C, Bader CD, Bailey CB, Barra L, Booth TJ, Bozhüyükl KAJ, et al. 2025. MIBiG 4.0: advancing biosynthetic gene cluster curation through global collaboration. *Nucleic Acids Res* 53:D678–D690. <https://doi.org/10.1093/nar/gkae1115>
59. Draisma A, Loureiro C, Louwen NLL, Kautsar SA, Navarro-Muñoz JC, Doering DT, Mouncey NJ, Medema MH. 2025. BiG-SCAPE 2.0 and BiG-SLiCE 2.0: scalable, accurate and interactive sequence clustering of metabolic gene clusters. *bioRxiv*. <https://doi.org/10.1101/2025.08.20.671210>
60. Reshetnikov AS, Khmelenina VN, Mustakhimov II, Kalyuzhnaya M, Lidstrom M, Trotsenko YA. 2011. Diversity and phylogeny of the ectoine biosynthesis genes in aerobic, moderately halophilic methylophilic bacteria. *Extremophiles* 15:653–663. <https://doi.org/10.1007/s00792-011-0396-x>
61. Patel HM, Tao J, Walsh CT. 2003. Epimerization of an L-cysteinyll to a D-cysteinyll residue during thiazoline ring formation in siderophore chain elongation by pyochelin synthetase from *Pseudomonas aeruginosa*. *Biochemistry* 42:10514–10527. <https://doi.org/10.1021/bi034840c>
62. Marchler-Bauer A, Derbyshire MK, Gonzales NR, Lu S, Chitsaz F, Geer LY, Geer RC, He J, Gwadz M, Hurwitz DI, Lanczycki CJ, Lu F, Marchler GH, Song JS, Thanki N, Wang Z, Yamashita RA, Zhang D, Zheng C, Bryant SH. 2015. CDD: NCBI's conserved domain database. *Nucleic Acids Res* 43:D222–D226. <https://doi.org/10.1093/nar/gku1221>
63. Inahashi Y, Zhou S, Bibb MJ, Song L, Al-Bassam MM, Bibb MJ, Challis GL. 2017. Watasemycin biosynthesis in *Streptomyces venezuelae*: thiazoline C-methylation by a type B radical-SAM methylase homologue. *Chem Sci* 8:2823–2831. <https://doi.org/10.1039/C6SC03533G>
64. Reid R, Piagentini M, Rodríguez E, Ashley G, Viswanathan N, Carney J, Santi DV, Hutchinson CR, McDaniel R. 2003. A model of structure and catalysis for ketoreductase domains in modular polyketide synthases. *Biochemistry* 42:72–79. <https://doi.org/10.1021/bi0268706>
65. Caffrey P. 2003. Conserved amino acid residues correlating with ketoreductase stereospecificity in modular polyketide synthases. *ChemBiochem* 4:654–657. <https://doi.org/10.1002/cbic.200300581>
66. Xiang C, Yao S, Wang R, Zhang L. 2024. Bioinformatic prediction of the stereoselectivity of modular polyketide synthase: an update of the sequence motifs in ketoreductase domain. *Beilstein J Org Chem* 20:1476–1485. <https://doi.org/10.3762/bjoc.20.131>
67. Steinmetz T, Lombe BK, Nett M. 2023. Intermediates and shunt products of massiliachelin biosynthesis in *Massilia* sp. NR 4-1. *Beilstein J Org Chem* 19:909–917. <https://doi.org/10.3762/bjoc.19.69>
68. Miller DA, Luo L, Hillson N, Keating TA, Walsh CT. 2002. Yersiniabactin synthetase: a four-protein assembly line producing the nonribosomal peptide/polyketide hybrid siderophore of *Yersinia pestis*. *Chem Biol* 9:333–344. [https://doi.org/10.1016/S1074-5521\(02\)00115-1](https://doi.org/10.1016/S1074-5521(02)00115-1)
69. Ankenbauer RG, Staley AL, Rinehart KL, Cox CD. 1991. Mutasynthesis of siderophore analogues by *Pseudomonas aeruginosa*. *Proc Natl Acad Sci USA* 88:1878–1882. <https://doi.org/10.1073/pnas.88.5.1878>
70. Gupta RS, Mok A. 2007. Phylogenomics and signature proteins for the alpha proteobacteria and its main groups. *BMC Microbiol* 7:106. <https://doi.org/10.1186/1471-2180-7-106>
71. Weber CF, King GM. 2007. Physiological, ecological, and phylogenetic characterization of *Stappia*, a marine CO-oxidizing bacterial genus. *Appl Environ Microbiol* 73:1266–1276. <https://doi.org/10.1128/AEM.01724-06>
72. Simon M, Scheuner C, Meier-Kolthoff JP, Brinkhoff T, Wagner-Döbler I, Ulbrich M, Klenk HP, Schomburg D, Petersen J, Göker M. 2017. Phylogenomics of Rhodobacteraceae reveals evolutionary adaptation to marine and non-marine habitats. *ISME J* 11:1483–1499. <https://doi.org/10.1038/ismej.2016.198>
73. Bentzon-Tilia M, Gram L. 2017. Biotechnological applications of the *Roseobacter* clade, p 137–166. Vol16.
74. Sasirekha B, Srividya S. 2016. Siderophore production by *Pseudomonas aeruginosa* FP6, a biocontrol strain for *Rhizoctonia solani* and *Colletotrichum gloeosporioides* causing diseases in chilli. *Agric Nat Resour* 50:250–256. <https://doi.org/10.1016/j.anres.2016.02.003>
75. Sarvepalli M, Velidandi A, Korrapati N. 2023. Optimization of siderophore production in three marine bacterial isolates along with their heavy-metal chelation and seed germination potential determination. *Microorganisms* 11:2873. <https://doi.org/10.3390/microorganisms11122873>
76. Gräff AT, Barry SM. 2024. Siderophores as tools and treatments. *NPJ Antimicrob Resist* 2:47. <https://doi.org/10.1038/s44259-024-00053-4>
77. Krueger F. 2015. Trim galore. A wrapper tool around Cutadapt and FastQC to consistently apply quality and adapter trimming to FastQ files
78. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Pribelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 19:455–477. <https://doi.org/10.1089/cmb.2012.0021>
79. Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput Biol* 13:e1005595. <https://doi.org/10.1371/journal.pcbi.1005595>
80. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. 2015. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* 25:1043–1055. <https://doi.org/10.1101/gr.186072.114>
81. Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 31:3210–3212. <https://doi.org/10.1093/bioinformatics/btv351>
82. Afgan E, Baker D, Batut B, van den Beek M, Bouvier D, Cech M, Chilton J, Clements D, Coraor N, Grüning BA, Guerler A, Hillman-Jackson J, Hiltmann S, Jalili V, Rasche H, Soranzo N, Goecks J, Taylor J, Nekrutenko A, Blankenberg D. 2018. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Res* 46:W537–W544. <https://doi.org/10.1093/nar/gky379>
83. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30:2068–2069. <https://doi.org/10.1093/bioinformatics/btu153>
84. Lefort V, Desper R, Gascuel O. 2015. FastME 2.0: a comprehensive, accurate, and fast distance-based phylogeny inference program. *Mol Biol Evol* 32:2798–2800. <https://doi.org/10.1093/molbev/msv150>
85. Meier – Kolthoff JP, Auch AF, Klenk H, Göker M. 2014. Highly parallelized inference of large genome - based phylogenies. *Concurr Comput* 26:1715–1729. <https://doi.org/10.1002/cpe.3112>
86. Holland BR, Huber KT, Dress A, Moulton V. 2002. δ plots: a tool for analyzing phylogenetic distance data. *Mol Biol Evol* 19:2051–2059. <https://doi.org/10.1093/oxfordjournals.molbev.a004030>
87. Schmid R, Heuckeroth S, Korf A, Smirnov A, Myers O, Dyrland TS, Bushuiev R, Murray KJ, Hoffmann N, Lu M, et al. 2023. Integrative analysis of multimodal mass spectrometry data in MZmine 3. *Nat Biotechnol* 41:447–449. <https://doi.org/10.1038/s41587-023-01690-2>
88. Aron AT, Gentry EC, McPhail KL, Nothias L-F, Nothias-Esposito M, Bouslimani A, Petras D, Gauglitz JM, Sikora N, Vargas F, et al. 2020. Reproducible molecular networking of untargeted mass spectrometry data using GNPS. *Nat Protoc* 15:1954–1991. <https://doi.org/10.1038/s41596-020-0317-5>
89. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 13:2498–2504. <https://doi.org/10.1101/gr.1239303>
90. Poynton EF, van Santen JA, Pin M, Contreras MM, McMann E, Parra J, Showalter B, Zaroubi L, Duncan KR, Linington RG. 2025. The Natural Products Atlas 3.0: extending the database of microbially derived natural products. *Nucleic Acids Res* 53:D691–D699. <https://doi.org/10.1093/nar/gkae1093>
91. European Committee for Antimicrobial Susceptibility Testing (EUCAST) of the European Society of Clinical Microbiology and Infectious Diseases (ESCMID). 2000. EUCAST definitive document E.DEF 3.1, June 2000: determination of minimum inhibitory concentrations (MICs) of antibacterial agents by agar dilution. *Clin Microbiol Infect* 6:509–515. <https://doi.org/10.1046/j.1469-0691.2000.00142.x>
92. Blin K, Shaw S, Augustijn HE, Reitz ZL, Biermann F, Alanjary M, Fetter A, Terlouw BR, Metcalf WW, Helfrich EJM, van Wezel GP, Medema MH, Weber T. 2023. antiSMASH 7.0: new and improved predictions for

- detection, regulation, chemical structures and visualisation. *Nucleic Acids Res* 51:W46–W50. <https://doi.org/10.1093/nar/gkad344>
93. Terlouw BR, Blin K, Navarro-Muñoz JC, Avalon NE, Chevrette MG, Egbert S, Lee S, Meijer D, Recchia MJ, Reitz ZL, et al. 2023. MIBiG 3.0: a community-driven effort to annotate experimentally validated biosynthetic gene clusters. *Nucleic Acids Res* 51:D603–D610. <https://doi.org/10.1093/nar/gkac1049>
94. Navarro-Muñoz JC, Selem-Mojica N, Mullaney MW, Kautsar SA, Tryon JH, Parkinson EI, De Los Santos ELC, Yeong M, Cruz-Morales P, Abubucker S, Roeters A, Lokhorst W, Fernandez-Guerra A, Cappellini LTD, Goering AW, Thomson RJ, Metcalf WW, Kelleher NL, Barona-Gomez F, Medema MH. 2020. A computational framework to explore large-scale biosynthetic diversity. *Nat Chem Biol* 16:60–68. <https://doi.org/10.1038/s41589-019-0400-9>
95. Blin K, Shaw S, Vader L, Szenei J, Reitz ZL, Augustijn HE, Cedié-Becerra JDD, de Crécy-Lagard V, Koetsier RA, Williams SE, et al. 2025. antiSMASH 8.0: extended gene cluster detection capabilities and analyses of chemistry, enzymology, and regulation. *Nucleic Acids Res* 53:W32–W38. <https://doi.org/10.1093/nar/gkaf334>
96. Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* 30:772–780. <https://doi.org/10.1093/molbev/mst010>