Nematophilic bacteria as a source of novel macrocyclised antimicrobial non-ribosomal peptides

By

Ryan Musumba Awori

Dissertation

zur Erlangung des Doktorgrades

der Naturwissenschaften

vorgelegt beim Fachbereich für Biowissenschaften (15)

der Johann Wolfgang Goethe-Universität

in Frankfurt am Main

von

Ryan Musumba Awori

aus Nakuru, Kenia

Frankfurt am Main 2020

D30

vom Fachbereich Biowissenschaften (15) der Johann Wolfgang Goethe-Universität
als Dissertation angenommen

Dekan:	Prof. Dr. Sven Klimpel
Gutachter:	Prof. Dr. Helge Bode
	Prof. Dr. Jörg Soppa
Datum der Disputation:	09.03.2021

“Tune your ear to wisdom,
and concentrate on understanding.
Cry out for insight,
and ask for understanding.
Search them out as you would for silver;
seek them like hidden treasures.”
Traditional Proverb

ABSTRACT

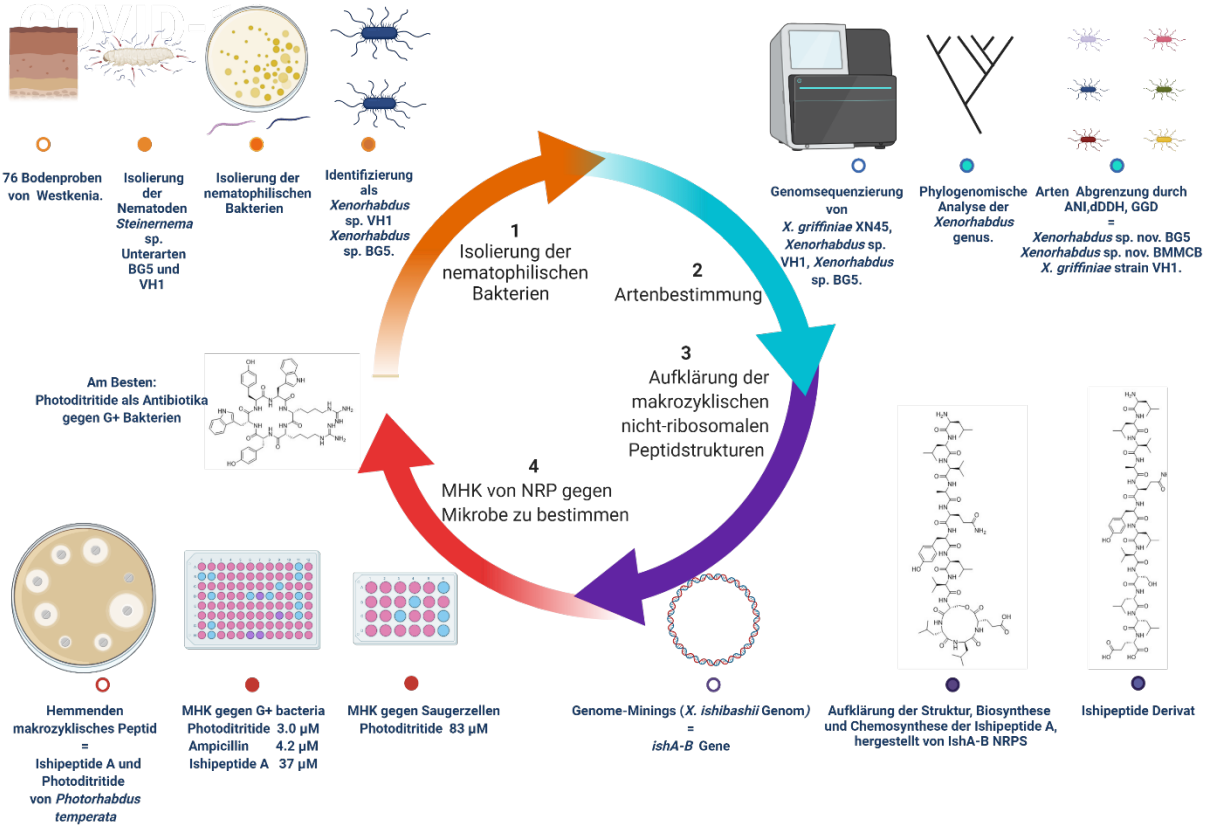
A solution to ineffective clinical antimicrobials is the discovery of new ones from under-explored sources such as macrocyclic non-ribosomal peptides (NRP) from nematophilic bacteria. In this dissertation an antimicrobial discovery process –from soil sample to inhibitory peptide– is demonstrated through investigations on six nematophilic bacteria: *Xenorhabdus griffiniae* XN45, *X. griffiniae* VH1, *Xenorhabdus* sp. nov. BG5, *Xenorhabdus* sp. nov. BMMCB, *X. ishibashii* and *Photorhabdus temperata.* To demonstrate the first step of bacterium isolation and species delineation, endosymbionts were isolated from *Steinernema* sp. strains BG5 and VH1 that were isolated directly from soil samples in Western Kenya. After genome sequencing and assembly of novel *Xenorhabdus* isolates VH1 and BG5, species delineation was done via three overall genome relatedness indices. VH1 was identified as *X. griffiniae* VH1, BG5 as *Xenorhabdus* sp. nov. BG5 and *X. griffiniae* BMMCB was emended to *Xenorhabdus* sp. nov. BMMCB. The nematode host of *X. griffiniae* XN45, *Steinernema* sp. scarpo was highlighted as a putative novel species. To demonstrate the second step of genome mining and macrocyclic non-ribosomal peptide structure elucidation, chemosynthesis and biosynthesis, the non-ribosomal peptide whose production is encoded by the *ishA-B* genes in *X. ishibashii* was investigated. Through a combination of refactoring the *ishA-B* operon by a promoter exchange mechanism, isotope labelling experiments, high resolution tandem mass spectrometry analysis, bioinformatic protein domain analysis and chemoinformatic comparisons of actual to hypothetical mass spectrometry spectra, the structures of Ishipeptides were elucidated and confirmed by chemical synthesis. Ishipeptide A was a branch cyclic depsidodecapeptide macrocyclised via an ester bond between serine and the terminal glutamate. It chemosynthesis route was via a late stage macrolactamation and linearised Ishipeptide B was synthesised via solid phase iterative synthesis. Ishipeptides were not *N*-terminally acylated despite being biosynthesised from the IshA protein that had a C-starter domain. It was highlighted that more than restoration of the histidine active site of this domain is required to restore *N*-terminal acylation activity.

To demonstrate the final step of determination of antimicrobial activity, minimum inhibitory concentrations of Ishipeptides and Photoditritide from *Photorhabdus*

temperata against fungi and bacteria were determined. None were antifungal while only the macrocyclic compounds were inhibitory, with Ishipeptide A inhibitory to Gram-positive bacteria at 37 μM. The cationic Photoditritide, a cyclic hexapeptide macrocyclised via a lactam bond between homoarginine and tryptophan, was 12 times more inhibitory (3.0 μM), even more effective than a current clinical compound, Ampicillin (4.2 μM). For both, macrocyclisation was hypothesised to contribute to antimicrobial activity. Ultimately, this dissertation demonstrated not only nematophilic bacteria as a source of novel macrocyclic antimicrobial non-ribosomal peptides but also a process of antimicrobial discovery–from soil sample to inhibitory peptide– from these useful bacteria genera. This is significant for the fight against antimicrobial resistance.

ZUSAMMENFASSUNG

Grafische Zusammenfassung



Resistenz gegen Antimikrobiotikum (AMR) ist eine Herausforderung für die ganze Welt. Zudem verschlimmert sie andere gesundheitliche Herausforderungen. Zum Beispiel kann sich die Situation sehr kranker COVID-19 Patienten durch Infektionen mit Antibiotikum-resistenten Mikroben weiter verschlechtern. Für Länder, die viele Tuberkulose Patienten und wenig molekulare diagnostische Maschinen haben, ist es jetzt schwieriger Antibiotikum-resistente Tuberkulose Infektionen zu entdecken, weil alle molekulare Maschinen jetzt für COVID-19 Diagnosen benutzt werden. AMR wird noch schlimmer (circa 1917 Tote pro Tag), wenn wir keine Forschung betreiben, um Lösungen zu finden. Eine Lösung für das Problem der Antibiotikaresistenzen ist die Entwicklung neuer Antibiotika aus bisher wenig untersuchten Quellen, wie makrozyklische nicht-ribosomale Peptide (NRP) von nematophilischen Bakterien. In dieser Arbeit wird anhand der Untersuchung von sechs nematophilischen Bakterien – nämlich *Xenorhabdus griffiniae* XN45, *X. griffiniae* VH1, *Xenorhabdus* sp. nov. BG5,

Xenorhabdus sp. nov. BMMCB, *X. ishibashii* and *Photorhabdus temperata* – ein antimikrobielles Suchverfahren von der Bodenprobe bis zum hemmenden Peptid vorgestellt.

Xenorhabdus haben eine Symbiose mit *Steinernema* Nematoden, *Photorhabdus* mit *Heterohabditis* und *Serratia* mit beiden Insectivora Gruppen von *Oschieus* Nematoden und *Ceanorhabditis briggisae.* Alle diese Nematoden töten Insekten(EPN) und die nematophilische Bakterien helfen dabei, das zu machen. Nicht alle *Serratia* sind nematophilische Bakterien und es gibt wenig bekannte *Serratia-Oschieus/Ceanorhabditis* Symbiosen. Aber alle 44 bekannten *Xenorhabdus* und *Photorhabdus* würden von Nematoden isoliert haben. Daher liegt der Fokus dieser Dissertation nur auf *Xenorhabdus* und *Photorhabdus* als nematophilische Bakterien. Beide sind Gram-negativ Bakterien von *Morganellaceae*, *Enterobacterales* und *Gammaproteobacteria*. Zwei wichtige Besonderheiten sind: *Xenorhabdus* sind die einzigen *Enterobacterales*, die Katalase-negativ sind und *Photorhabdus* sind die einzigen irdischen Bakterien, die Biolumineszenz haben. Heute gibt es 26 *Xenorhabdus* und 18 *Photorhabdus*, die von sehr verschiedenen Orten isoliert werden können. Wegen des Klimas, konnte man von tropischen Afrikanischen Orten viele nematophilische Bakterien isolieren. Aber nur *X. griffiniae* XN45, X. *griffiniae* BMMCB und *X. khoisanae* wurden von hier isoliert. Daher ist es wichtig mehr Forschung in diesen Orten zu machen.

Die phylogenetische Diversität von *Xenorhabdus* ist höher als für *Photorhabdus.* Daher ist es schwieriger *Photorhabdus* Arten zu trennen. Um das zu richtig machen, sollte man phylogenomische Techniken nutzen. Sie sind Average Nucelotide Identity (ANI),Genom-Genom Distanz(GGD), digitale DNA-DNA-Hybridisierung (dDDH) und Pan-genom Analysen. ANI trennt zwei Arten, wenn sie niedriger als 95.1% Identitaten haben. dDDH trennt zwei Arten wenn sie niedriger als 70% Identitaten haben und GGD trennt zwei Arten wenn sie mehr als 0.0361 Distanz haben. Zwei Unterarten können keine gleiche Arten sein, wenn es mehr als 1% Differenz zwischen ihrem GC-content gibt. Die Pan-genom Technik ist sehr genau aber braucht mindestens fünf Genome pro Arten um zu funktionieren. Es kann für die Identifikation von neuen *Photorhabdus*

Arten richtig gut funktionieren. Aber für neue *Xenorhabdus* Arten ist ANI, dDDH, GGD ausreichend, weil sie eine nicht zu niedrige phylogenetische Diversität haben.

In der Natur steigt bei *Photorhabdus* und *Xenorhabdus* Bakterien die Fekundität von Ihren Nematoden an, durch die Herstellung von Naturstoffen mit verschiedenen ökologischen Funktionen. Sie sind: Naturstoffe die Insekten töten (Rhabduscin, Xpt toxin, Mcf toxin, Xax toxin), Naturstoffe die Protozoon töten (Xenortide, Rhabdopeptide, Xenoamicin, Amabctin), Naturstoffe die Plins töten (Cabalinasin, Biocornutin, Xenocoumacin 2, EP-20, GP-19), Naturstoffe die andere Bakterien töten (Odilorhabdins, Nematophin, Xenohabdin, Xenocoumacin1, PAX peptides, Xenematides, Xenobactin), Naturstoffe die Insekten vertreiben (SDF, ADF) und Naturstoffe die Symbiose helfen (NilC). Eine paar von diesen Naturstoffen sind NRP. Was genauso ist ein NRP? Ein NRP besteht aus zwei oder mehr verbundenen Aminosäuren, die von einem Nicht-ribosomal Peptid Synthetase (NRPS) Protein hergestellt sind. Wegen dieser Herstellung haben NRP diese Besonderheiten: Sie erhalten D-Aminosäuren; Sie können über 500 verschiedene Aminosäuren/Aminosäurederivative erhalten; Während der Synthese im NRPS können sie Modifikationen wie *N*- Terminal Acylierung, Methylierung, Formylierung, Makrozyklisierung bekommen. Jeder NRPS hat mindestens ein Modul, das mindestens A, C und T Domain hat. Um ein NRP herzustellen, wird zuerst die ganze NRPS von *apo* bis *holo* konvertiert. Dann wird die A domain der richtigen Aminosäure ausgewählt. Dann wird die A Domain dieser Aminosäure aktiviert. Danach wird die T Domain mit der aktivierten Aminosäure(AAS) durch einen phosphopantethienyl co-Faktor(PPT) verbunden, um die AAS zu transportieren. Die C domain wird AAS verbinden, die von der zwei T Domains die C Domain inzwischen ist und der letzte Schritt erzeugt ein Peptid. Zum Schluss trennt die TE Domain das Peptid von der NRPS. Die TE domain kann auch eine Makrozyklisierung machen. Das Ergebnis sind makrozyklisierte antimikrobielle NRP von nematophilischen Bakterien: Taxalllaids, Chaiyaphumins, Szentiamide, Xenematide, Xenobactin, PAX lipopeptides, Xenoamicin, Ambactin, Phototemtide.

Mit Hilfe der Bioinformatik Genom Analyse kann man schnell herausfinden, welche NRP ein Bakterium herstellen kann, wie groß es ist und ob es unbekannt ist. Dann

kann man durch Isotopenmarkierung, hochauflösender Tandem-Massenspektrometrie, bioinformatischer Proteindomänenanalyse und chemoinformatischer Vergleiche gemessener mit hypothetischen massenspektrometrischen Spektren die Strucktur erklären. Danach kann man die minimalen Hemmkonzentration (MHK) vom Peptid finden.

In dieser Dissertation habe ich an sechs nematophilischen Bakterien geforscht. Das Ziel war die Prozessschritte, wie man ein antimikrobielles NRP von nematophilischen Bakterien finden kann, zu demonstrierten. Diese Schritte waren 1) der Isolierung der Bakterien und Artenbestimmung 2) Genome-Minings und der Aufklärung der makrozyklischen nicht-ribosomalen Peptidstrukturen 3) MHK von NRP gegen Mikrobe zu bestimmen.

Als erster Schritt der Isolierung der Bakterien und Artenbestimmung wurden Endoymbionten aus den *Steinernema* sp. Unterarten BG5 und VH1 isoliert, die wiederum aus 76 Bodenproben entnommen wurden, die im Oktober 2018 in Westkenia gesammelt wurden. Nach der Genomsequenzierung der *Xenorhabdus* sp. Unterarten BG5 & VH1 und *X. griffiniae* XN45 wurden die Arten durch die drei Indizes ANI,GGD und dDDH abgegrenzt. *Xenorhabdus* sp. unterart VH1 wurde als *X. griffinae* VH1; *Xenorhabdus* sp. Unterart BG5 als *Xenorhabdus* sp. nov. BG5 identifiziert und *X. griffinae* BMMCB wurde zu *Xenorhabdus* sp. nov. BMMCB korrigiert. Der Nematodenwirt von *X. griffiniae* XN45, *Steinernema* sp. scarpo ist möglicherweise eine neue Nematodenspezies. Um den zweiten Schritt des Genome-Minings und der Aufklärung der makrozyklischen nicht-ribosomalen Peptidstrukturen, der Chemo- und Biosynthese vorzuführen, wurde das nicht-ribosomale Peptid untersucht, dessen Produktion durch die *ishA-B* Gene in *X. ishibashii* kodiert wird. Durch Veränderung des *IshA-B* Operon durch einen Promoter-Austausch-Mechanismus in Kombination mit Isotopenmarkierung, hochauflösender Tandem-Massenspektrometrie, bioinformatischer Proteindomänenanalyse und chemoinformatischer Vergleiche gemessener mit hypothetischen massenspektrometrischen Spektren, wurden die Strukturen der Ishipeptide erhellt und durch chemische Synthese bestätigt. Ishipeptide A war ein verzweigt-zyklisches Dodecadepsipeptid mit einer Makrozyklisierung durch eine Esterbindung zwischen Serin (D-Ser^{9}) und dem terminalen Glutamat (L-Glu$^{12.}$).

Die chemosynthetische Route verlief über eine Makrolaktamation in einem späten Stadium und linearisiertes Ishipeptide B wurde durch iterative Merrifield-Synthese synthetisiert. Ishipeptide waren nicht am N-Terminus acyliert, obwohl sie aus einem IshA-Protein mit einem C_{start} Domain biosynthetisiert worden waren.Es stellte sich heraus, dass mehr als die Wiederherstellung des Hystidin-aktiven Zentrums erforderlich ist, um die Acylierungsaktivität des N-Terminus wiederherzustellen. Ishipeptide A ist nicht stabil und konnte bis Ishipeptide B(linear derivativ) verwenden. Ishipeptide B kann eine Methylierung an der C-Terminal erhalten. Alle drei existieren in MS spectra.

Die letzte Schritt war die Ermittlung der minimalen Hemmkonzentration (MHK) der Ishipeptide A, B und Photoditritide zu bestimmen, die aus *Photorhabdus temperata* isoliert worden waren. Keine wirkten antimykotisch. Alle drei hat keine Inhibition gegen *Candida lusitaniae* gehabt. Nur die makrozyklischen Stoffe wirkten hemmend.Ishipeptide B wirkte nicht hemmend gegen *Micrococcus luteus* gehabt während Ishipeptide A eine MHK von 37 µM hat. Das kationische Photoditritide, ein Hexapeptid mit einer Makrozyklisierung durch Laktambindung zwischen Homoarginin (D-Har^1) und Tryptophan (L-Trp^6), war 12-mal mehr hemmend und sogar effektiver als der aktuelle klinische Wirkstoff Ampicillin. Photoditritide hat eine MHK von 3.0 µM gegen *M. Luteus* und Ampicillin hat eine MHK von 4.2 µM. Es wird angenommen, dass die Makrozyklisierung zur antimikrobiellen Wirkung beiträgt. Diese Arbeit zeigte nicht nur, dass nematophilische Bakterien als Quelle neuartiger makrozyklischer antimikriobiell wirksamer nicht-ribosimaler Peptide dienen können, sondern demonstrierte auch einen methodischen Ansatz für die Suche nach Antibiotika – von der Bodenprobe bis zum hemmenden Wirkstoff – aus diesen sehr nützlichen Bakteriengattungen. Die ist für den Kampf gegen Antibiotikresistenz von Bedeutung.

ACKNOWLEDGMENTS

I wish to acknowledge Prof. Dr. Helge Bode as my supervisor, first reviewer and proof-reader of this dissertation. I am grateful for his faithful supervision since our first meeting in November 2016. I am grateful to Prof. Dr. Jörg Soppa for agreeing to be the second reviewer of this dissertation.

This research was supported by the German Academic Exchange Service (DAAD) under program number 57299294 and reference number 91653288. This research was also supported by the Kenya National Research Fund under a 2017 multidisciplinary grant program titled “Drug Development of antibiotics: *Xenorhabdus* bacteria from Kenya”. I acknowledge the supervision of the Principal investigator Dr. Nelson Amugune and co-investigators Dr. Dan Masiga and Dr. Charles Waturu through my participation in this program. The following activities were directly supported through this program: fieldwork in Western Kenya, nematode isolation, bacteria isolation from nematodes and genome-sequencing.

It was an honour to be a student at the Institute of Molecular Biosciences. I am grateful to Sophie Moktadir and Bruni Schönberger for their swift support for all matters administrational, and Peter Grün, Andreas Lück & Stefanie Lambert for all matters technical. It has been edifying to my doctoral studies to have worked with the following scientists, post-doctoral fellows and fellow students: Dr. Carsten Kegler, Dr. Jürgen Breitenbach, Dr. Edna Bode, Dr. Nicolas Tobias, Dr. Yi-Ming Shi, Dr. Yan-Ni Shi, Dr. Zhengyi Qian, Dr. Kenan A. J. Bozhüyük, Dr. Gina Grambitter, Dr. Antje Heinrich, Sebastian Wenski, Desalegne Abebew Syit, Andreas Tietze, Lukas Kreling, Jonas Watzel, Moritz Drechsler, Janik Kranz, Margaretha Westphalen, Helena Vural, Siyar Kavakli, Nick Neubacher, Nadya Abbood, Yuanyuan Zuo and Alexander Rill. A special thanks to Dr. Lei Zhao with whom I worked most closely with, specifically on the Ishipeptides and Photoditritide sections of this dissertation. I acknowledge that Dr. Nicholas Tobias assembled the genomes of strains VH1 and BG5 and Peninnah Wambua extracted the DNA from *X. griffiniae* XN45 while James Kabii facilitated its transportation. I acknowledge that Lorine Nyongesa, Mr. Simuyu and Dr. Nelson Amugune of the School of Biological Sciences, University of Nairobi assisted in soil sample collection in Western Kenya. Laboratory support for isolation of nematodes

from soils and isolation of bacteria from nematodes was provided by Mr. Nderitu of Horticultural Research Institute, Thika, Kenya and Mr. Wachira of the Genetics Lab, School of Biological Sciences, University of Nairobi, Kenya, respectively.
I wish to thank Nicola Witbooi, Nikolaus von Bargen and Holger Fink for grammar checking the German texts. I am utterly grateful to my Frankfurt family (Small group, ICF Oberursel, Frankfurt Connect, River of Life) for their support that was in ways that words cannot accurately express. I am blessed to have had my wife with me, every step of the way. Thank you for your steadfast presence. Lastly, I am grateful to God for this chance to have pursued this PhD degree.

Table of Contents

INTRODUCTION

Antimicrobial resistance in our world today

Antimicrobial resistance (AMR) is the inability of current clinical compounds that previously inhibited the growth of bacteria, fungi, protozoa, and viruses, to exert the same effect due to genetic mutations within these microorganisms. As of June 2014, the containment of infections by antimicrobial-resistant pathogens was classified as a top global health priority by the World Health Organisation[1]. Today, the top global health priority is the emergence and subsequent containment of the novel SARS-CoV-2 virus and its resultant coronavirus disease 2019(COVID-19), which was classified as a global pandemic on March 11th, 2020. Yet the control of COVID-19 and antimicrobial-resistant pathogen infections is intertwined[2]. For example, as much as global health measures –such as increased sterilisation of patient rooms and hospital equipment in health care settings[3], increased hand washing and sanitisation of surfaces in public settings, reduced social gatherings[4] and global travel[3]– to reduce SARS-CoV-2 infections have inadvertently[5] reduced microbial infections[3,6], the COVID-19 pandemic may increase spread of antimicrobial-resistant pathogens in the following ways.

First, despite bacterial and fungal co-infections in hospitalized COVID-19 patients being only seven percent, 14% was observed in those in intensive care units(ICU) including those by carbapenem-resistant & extended-spectrum beta-lactamase *Kleibsella pneumoniae* and methicillin-resistant *Staphylococcus aureus,*[7] all of which are classified as urgent or serious global health threats[1,5]. The high levels of indiscriminate, unrecommended use of antimicrobials as prophylaxis on COVID-19 patients[7] further engenders the emergence of such antimicrobial resistant pathogens. Relatedly, increased indiscriminate use of antimicrobials on COVID-19 patients has resulted in an increase in *Clostridioides difficile* co-infections[8].

Second, the increased demand for diagnosis of SARS-CoV-2 by nucleic acid amplification tests may lead to a decrease in capacity to diagnosis multi-drug resistant *Mycobacterium tuberculosis* in high-burden low-resource countries, as previously used molecular platforms such as Cepheid GeneXpert[9] are being repurposed for SARS-CoV-2 screening[10]. This reduced capacity for antimicrobial stewardship may increase

spread of drug-resistant *Mycobacterium.* Taken together, we see AMR as a pervasive problem that makes other health problems worse.

The actual cost of human lives lost to AMR is similar to 1917 per day and this is predicted to increase unless deliberate interventions are continually undertaking[11]. One such intervention is development of new antimicrobials such as novel non-ribosomal peptides from nematophilic bacteria. Thus, the aim of this dissertation was to demonstrate an antimicrobial discovery process –from soil sample to inhibitory compound– through investigations on six nematophilic bacteria: *Xenorhabdus griffiniae* XN45, *X. griffiniae* VH1, *Xenorhabdus* sp. nov. BG5, *Xenorhabdus* sp. nov. BMMCB, *X. ishibashii* DSMZ 22670 and *Photorhabdus temperata* Meg1.

What are nematophilic bacteria?

Nematophilic "nematode loving" bacteria are members of Kingdom Monera that are natural obligate symbionts of members of Phylum Nematoda. Three such genera are *Xenorhabdus*, *Photorhabdus* and *Serratia,* which are symbionts of *Steinernematidiae*, *Heterorhabditidae* and *Rhabditidae* members respectively. Each of these three families of the order *Rhabditida* contain entomopathogenic nematodes(EPNs) –all members of *Steinernema* genus, all members of *Heterorhabditis* genus, all members of Insectivora-group of *Oscheius* genus and *Ceanorhabditis briggsae.* However, it is their *Xenorhabdus*, *Photorhabdus* or *Serratia* symbionts that contribute in large part to this trait through both septicaemia and toxaemia[12–15]. Hence, *Xenorhabdus, Photorhabdus* and a few *Serratia* species are also termed as entomopathogenic bacteria.

Whereas *Serratia* symbionts form associations with *Oscheius* and *Caenorhabditis* nematodes, *Xenorhabdus* and *Photorhabdus* symbionts are more genus specific, associating only with *Steinernema* and *Heterorhabditis* species respectively with even more specificity seen in the former as a *Steinernema* species associates only with one *Xenorhabdus* species. Though *Xenorhabdus* and *Serratia* are the most-studied, other bacteria such as entomopathogenic *Pseudomonas* sp.[16] and *Alcaligenes* sp.[17] respectively associate with *Steinernema* and *Oscheius* nematodes.

This classification of *Serratia* as nematophilic bacteria has some caveats. First, not all *Serratia* species are nematode symbionts[18] and only a few *Oscheius-Serratia/Ceanorhabditis-Serratia* associations are known. Second, for some of these

associations *Serratia* were only facultative nematode symbionts[13,19]. Nonetheless, this is an emergent field of study worthy of mention and one can predict many more associations will be discovered, if only based on the ubiquitous and astronomical abundance of nematodes within the global top soils[20].

All forty-four characterised *Xenorhabdus* and *Photorhabdus* species were natural nematode intestinal symbionts as highlighted in Table 1 (although the name *P. asymbiotica* means 'lack of a symbiont' as it was first isolated from a human patient in Texas, the nematode host of *P. asymbiotica* subsp. australis –reclassified as *P. australis*– was later discovered as *H. gerrardi*[21]). Thus, this dissertation focused on these two genera as nematophilic bacteria.

Described species of *Xenorhabdus* and *Photorhabdus*

Xenorhabdus and *Photorhabdus* are both Gram-negative, rod shaped, peritrichously flagellated, facultative anaerobes of family *Morganellaceae*, order *Enterobacterales* and class *Gammaproteobacteria*[22] uniquely characterised by having both primary and secondary variants and endosymbiosis with EPNs. Other distinguishing traits include *Photorhabdus* being the only terrestrial bioluminescent bacterium genus and *Xenorhabdus* being the only member of *Enterobacterales* that does not produce catalase.

Table 1. Nematophilic bacteria associated with entomopathogenic nematodes

Species	Nematode Host of Isolation	Geographic Region of Type Strain Isolation
X. beddingii[23]	*S. longicaudum*	Tasmania, Australia[24]
X. bovienii[23]	*S. affinie*	Tasmania, Australia[24]
	S. intermedium	
	S. kraussei	
	S. feltiae	
X. budapestensis[25]	*S. bicornutum*	Szabadka, Serbia[25]
X. caballinasii[26]	*S. riobrave*	Texas, USA[26]
X. doucetiae[26]	*S. diaprepesi*	Martinique, Carribean[27]
X. ehlersii[25]	*S. serratum*	Sourthern China[25]

X. griffiniae[26]	*S. hermaphroditum*	Indonesia[26]
X. hominickii[26]	*S. karii*	Kirinyaga, Kenya[28]
	S. monticolum	
X. indica[25]	*S. thermophilum*	New Delhi, India[29]
X. innexi[25]	*S. scapterisci*	Uruguay[25]
X. japonica[30]	*S. kushidai*	Hamakita, Japan[30]
X. ishibashii[31]	*S. aciari*	Haimen, China[32]
X. khoisanae[33]	*S. khoisanae*	Western Cape Province, South Africa[33]
X. koppenhoeferi[26]	*S. scarabaei*	New Jersey, USA[26]
X. kozodoii[26]	*S. arenarium*	Russia[34]
	S. apuliae	
X. magdalanensis[35]	*S. australe*	Isla Magdalena, Chile[35]
X. mauleonii[26]	*Steinernema* sp.	St. Vincent, Carribean[27]
X. miraniensis[26]	*Steinernema* sp.	Mirani, Australia[26]
X. nematophila[23]	*S. carpocapsae*	Virginia, USA[36]
X. poinarii[23]	*S. glaseri*	North Carolina, USA[37]
	S. cubanum	
X. romanii[26]	*S. puertoricense*	Puetro Rico, USA[26]
X. stockiae[26]	*S. siamkayai*	Thailand[26]
X. szentirmaii[25]	*S. rarum*	Cordoba, Argentina[25]
X. vietnamensis[38]	*S. sangi*	Xuanmy, Vietnam[39]
X. eapokensis[40]	*S. eapokense*	EaPok, Vietnam[40]
X. thuoxuangnensis[40]	*S. sangi*	Thuong Xuan, Vietnam[40]
S. nematodiphila[41]	*O. chongmingensis*	Chongming Islands, China[42]
S. marcescens[19]	*O. carolinensis*	Raleigh, USA[43]
	O. safricana	North West Province, South Africa[44]

Serratia sp. strain TEL[45]	*O. basothovii*	Suikerbosrand Nature Reserve, South Africa[45]
Serratia sp. strain N19[46]	*O. microvilli*	Chongming Island, China[46]
Serratia sp. strain SCBI[13]	*C. briggsae*	Mpumalanga Province, South Africa[47]
P. bodei[48]	*H. beicherriana*	Liaoning Province, China[48]
P. caribbeanensis[48]	*H. bacteriophora*	Basse Terre, Guadeloupe Islands [27]
P. hainanesis[48]	*Heterohabditis* sp.	Hainan Island, China[49]
P. kleinii[48]	*H. georgiana*	Ohio, USA[50]
P. namnaonensis[48]	*H. baujardi*	Nam Nao, Thailand[51]
P. noeniputensis[48]	*H. noenieputensis*	Mbombela, South Africa[52]
P. tasmaniensis[48]	*H. zealandica*	Nicholls Rivulet, Australia[49]
P. thracensis[48]	*H. bacteriaphora*	Kirklareli, Turkey[53]
P. stackebrandtii[48]	*H. bacteriophora* strain GPS11	Atwood, USA[54]
P. kayaii[48]	*H. bacteriophora*	Aksaray, Turkey[53]
P. asymbiotica[55]	*H. gerrardi*	San Antonio, USA[56]
P. cinerea[48]	*H. downesi*	Asotthalom, Hungary[57]
P. akhurstii[48]	*H. indica*	Grande Terre, Guadeloupe Islands[27]
P. luminescens[48]	*H. bacteriophora*	Brecon, Australia[58]
P. heterohabditis[59]	*H. zealandica*	Brits, South Africa[60]
P. khanii[48]	*H. bacteriophora*	Clayton, USA[61]
P. laumondii[48]	*H. bacteriophora* strain HP88	Trindad, Trindad and Tobago[27]
P. temperata[48]	*H. megidis*	Nachodka, Russia[49]

Despite this taxonomic relatedness, the similar ecological niche of the two genera is due more to convergent evolution than phylogenetic distance[62]. Phylogenetically,

Xenorhabdus is more divergent with more characterised species while *Photorhabdus* is less divergent despite having fewer species[63]. Twenty-six *Xenorhabdus* species that were isolated from twenty-six nematode symbionts have been described to date (Table 1). However, 84 *Steinernema* species have been described[64] highlighting that at most –because of species like *X. bovienii, X. hominickii, X. kozodoii* and *X. poinarii* that are symbionts of more than one nematode– 58 novel *Xenorhabdus* species could be added to the genus from respective under-investigated yet described steinernematids. This prediction can be mathematically supported by calculating the mean α exponent of Heap's Law of the *Xenorhabdus* pangenome with user-friendly tools like the Panweb web-server that uses a PGAP platform[65]. A value less than one indicates that the pangenome is open, meaning that more novel species need be added to the genus so as to complete it[66].

As of 2018, the use of nucleotide sequence-based techniques for species delineation became the gold standard in bacterial taxonomy[67] and is useful for identification of new field isolates or emendation of already described taxon. For example, *Xenorhabdus* sp. strain BMMCB was described as an *X. griffiniae* species[68], but we[69] demonstrated that its nucleotide identities values for the recombinase A(recA), phosphoserine transferase(serC) and small sub unit rRNA(SSU) genes, with those of the type species were below the accepted threshold for same species assignment (97% for protein coding genes[38], 98.7% for SSU gene[70]). Moreover it[71] was demonstrated that the nematode host of *Xenorhabdus* sp. strain BMMCB, *S. khoisanae* BMMCB, was not *S. khoisanae* but a novel species based on its low nucleotide identities values(<95%) for the internal transcribed spacer region (ITS) with those of most closely-related species. This indicates that *Xenorhabdus* sp. strain BMMCB is not an *X. griffiniae* strain but may be a novel species.

The two most powerful nucleotide sequence-based techniques are average nucleotide identities (ANI) and digital DNA-DNA hybridisation (dDDH) that both delineate species by calculate how related two genomes are. The threshold for two strains to be classified as one species is >95%[72] and >70%[67] for ANI and dDDH respectively. Both of these were used to recently reclassify the *Photorhabdus* genus into 18 species[48]. However, strains S8-52, S9-53 and S10-54, identified as *P. kleinii,* had ANI values of 96.7% with

the *P. bodeii* type strain and *P. temperata* Meg1 had ANI values of 96.3% with the *P. thracensis* type strain demonstrating the difficulties[27,38] of species delineation in *Photorhabdus* due to its low phylogenetic diversity. As a single nucleotide-based threshold cannot work across all bacterial species[73,74], the use of pangenome analysis for species definitions and delineations, as was done with the *Prochloroccus* genus[75], can be a useful technique for *Photorhabdus* systematics when sufficient genome sequences–at least five per species–[66] are available.

Nematode-bacterium lifecycle and bacterial small molecules

The nematode-bacterium lifecycle begins with soil-dwelling infective third larval stage juvenile nematodes (IJ3) preying on an insect. Anatomically, they are J3 nematodes with a retained second larval stage cuticle that seals both mouth and anus rendering the nematodes into a non-feeding, developmentally-arrested, perennation-like stage[76], in which they can remain viable for more than 6 months. Steinernematids IJ3 gain entry into a prey only through natural openings while heterohabditids can additionally gain entry by piercing into the haemocoel using a bursa[77] that only they possess among the rhabditids[78]. Once within, the IJ3 undergo “recovery”[12] whereby they shed their second larval stage cuticle and release (via regurgitation for *Heterohabditis*[79] but defecation for *Steinernema*[80]) into the haemocoel, their bacterial gut symbionts.

For *Steinernema*, *Xenorhabdus* would have been previously localised in a receptacle[81] at the anterior gut while for *Heterohabditis, Photorhabdus* would have previously lined the entire gut[82]. Detection of L-proline concentrations >4.8mM in the haemolymph triggers upregulation of bacterial secretion of specialised metabolites of various ecological functions[83]. Despite the following grouping of biomolecules from both *Xenorhabdus* and *Photorhabdus* according to similarity of ecological function, their biosynthesis is species specific[84].

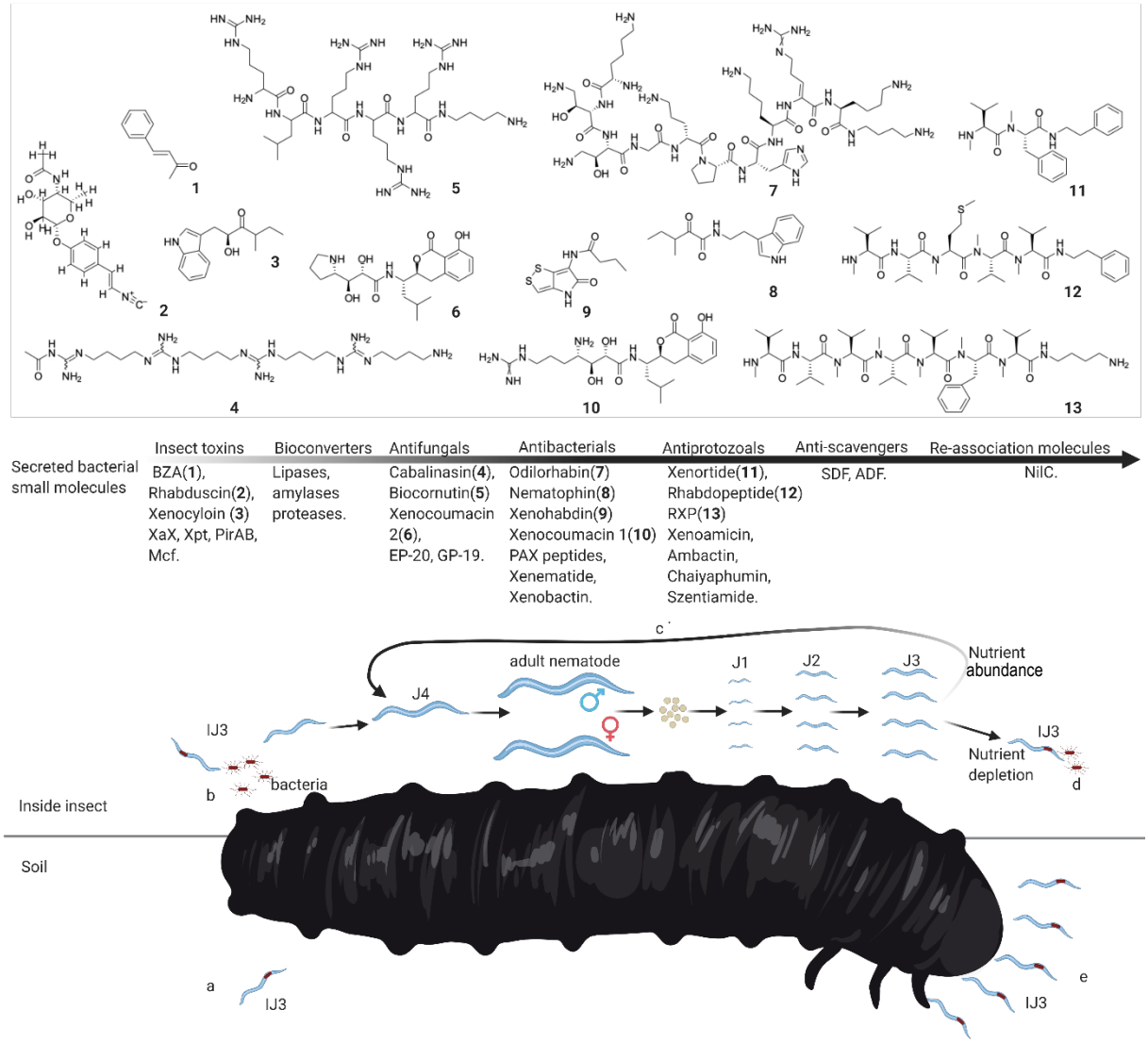


Figure 1. *Xenorhabdus-Steinernema* lifecycle and selected biomolecules that contribute towards nematode fecundity. a) Free living infective third-stage juvenile (IJ3) nematodes seek out insects and gain entry through natural openings such as spiracles and b) once within the haemocoel, nematodes shed cuticle thus exiting non-feeding state and release *Xenorhabdus* gut symbionts into the insect body by defection. Upon release into the insect, bacteria secrete a range of small molecules that increase the fecundity of the nematode. Topmost arrow indicates putative sequence of secreted molecules. c) Nematodes go through 1-3 complete lifecycles thus increasing their numbers and d) upon depletion of nutrients each J3 re-

associates with 1-3 *Xenorhabdus* bacteria, re-enter a non-feeding state and leave insect cadaver as IJ3. e) From only one nematode of each sex infecting an insect prey, millions can leave the cadaver. J1, J2, J3, J4= first, second, third, fourth larval stage juvenile nematodes, respectively. BZA= benzylideneacetone. Created with BioRender.com

The first grouping is insecticidal toxins and these can be divided into insect immune-suppressors via inhibition of phenoloxidase pathway: 1,2-benzenedicarboxylic acid(PA)[85], benzylideneacetone (BZA)[86] rhabduscin[83,87], 1,3-dihydroxy-2-(isopropyl)-5-(2-phenylethenyl)benzene[88]; haemocyte pore-forming complexes: *Xenorhabdus* particulate toxins(Xpt)[89,90], Toxin complex toxins(Tc)[91], *Xenorhabdus* α-xenorhabdolysin toxins (Xax)[92,93]; apoptosis inducers: Make caterpillar floppy toxins(Mcf)[94], PaTox[95] toxins and those with yet unknown modes of action: PirAB[96].
Another ecological function of secreted metabolites is bioconversion by enzymes such as lipases, proteases, amylases, and proteases –their respective genes are enriched in *Xenorhabdus* and *Photorhabdus* genomes[62]– creating a rich nutrient pool for the bacteria. To defend this from colonisation by microbial competitors, a broad spectrum of antimicrobials is produced. These include antifungals: Biocornutin[97], Cabalinasin[98], EP-19, GP-20[99], Xenocoumacin[100,101]; antibacterials: Darobactin[102], Xenematide[103,104], Xenocoumacin[105], Photoditritide[106], Nemaucin[107], Xenobactin[108], Odilorhabdins[109], Xenorhabdin[110,111], PAX peptides[112,113]: antiprotozoals: Phototemtide[114], Xenoamicin[115], Szentiamide[116], Ambactin[117], Chaiyaphumin[118], RXPs[119], Xenortide[120] and cytotoxic agents: Fabclavines[121], Rhabdopeptides[80,81], Phenylethylamine (PEA) derivatives[124].
Recovered IJ3s leverage this nutrient-filled enclosed environment to moult to fourth larval stage juvenile nematodes(J4) nematodes then adult nematodes, which lay eggs after mating in the case of all steinernematids expect *S. haemaphroditum* that lay eggs without mating due to its hermaphroditic nature and this is similar to the androdioecious heterohabditids whose adult females are also self-fertilised. Uniquely, *Heterohabditis* adult females lay eggs into their uterus which hatch and develop to first larval stage juvenile nematodes (J1) nematodes through *endotokia matricida*[12]. Both steinernematids and heterohabditids moult from J1 through to J4 then adults that lay eggs thus beginning another lifecycle and this continues until nutrients are depleted

within the cadaver. Notably, both *Steinernema* and *Heterohabditis* infected cadavers are themselves protected from consumption by non-microbial competitors such as ants by the bacterial production of scavenger deterring factors[125,126], a group of which is the non-proteinaceous, heat-stable, extracellular ant deterring factors(ADF) that is produced by *X. nematophila* and *P. luminescens*[127].

Upon nutrient depletion, J3 nematodes commence transformation to IJ3s by reassociating with bacterial symbionts which can be as few as one per nematode in the case of *Xenorhabdus* associations[81]. Moreover, a highly species-specific re-association occurs in *Xenorhabdus*-*Steinernema* complexes and this is attributed to the NilC protein[128] in *X. nematophila.* By retaining the second larval stage cuticle rendering it into a non-feeding state, J3s complete transformation to IJ3 that then emigrate the cadaver in search of new insect prey. Evidently the fecundity of the nematode is increased by the production of bacterial secondary metabolites.

Biotechnologically, two of these, the Odilorhabdins[109] from *X. nematophila* and Tapinarof[129] from *P. luminscens* and *P. asymbiotica* are furthest down the therapy development pipeline with the former having completed testing in mice models as an antibiotic against *Escherichia coli* and *Kleibsella pnumoniae*[130] and Tapinarof Phase 2b clinical trials as an anti-inflammatory against dermatitis[131]. These exemplify nematophilic bacteria as a source of novel therapies. Of note is that Odilorhabdins are a type of biomolecule to which seven[132] clinical antimicrobials belong i.e non-ribosomal peptides.

What are non-ribosomal peptides?

Non-ribosomal peptides (NRP) are chains of amino acids(aa) formed from catalytic reactions of large energy-dependent multifunctional proteins called non-ribosomal peptide synthetases (NRPS) that are organised into modules, which are further divided into domains. Fundamentally similar to peptides made by ribosomes in that they are both two or more amino acids linked by amide bonds, NRPs differ in the following ways due to their biosynthesis from NRPS: incorporation of D-amino acid, nonproteinogenic and modified amino acid residues; assembly from a 500-plus[133] pool of residues instead of only 20; incorporation of non-amino acid residues and modifications at

N[115,134] and *C*[119] -terminals during initial peptide synthesis. These results in NRPs being structurally diverse in comparison to their ribosomally produced counterparts[135].

Whereas in ribosomal peptide synthesis, a tRNA synthetase selects its corresponding amino acid and activates it, thus converting it to its amino acyl adenylate, then loads it to tRNA to form an amino acyl tRNA, which then forms a complex together with elongation factor and guanosine triphosphate (GTP) that then moves into large ribosomal subunit (LSU) acceptor site, where peptide bond formation between amino acids of adjacent complexes occurs thus forming a peptide, NRP synthesis occurs solely in the NRPS –these are themselves predominantly found in fungi[136] and bacteria[135]. However, within modules of these enzymes are domains with analogous functions to tRNA synthetase, amino acyl tRNA and LSU acceptor site, and these are expounded upon in the next section.

Non-ribosomal peptide synthetase machinery

The subset of an NRPS responsible for the incorporation of a single amino acid residue into a growing peptide chain is called a module. Two out of the three NRPS types are classified based on how many times they use a module in the biosynthesis of a single peptide molecule. Type one NRPS only use a module once to synthesise a complete peptide molecule while type two use a module repeatedly to synthesise one complete molecule such as the InxA, InxB, and InxC synthetases that produce the potent antiprotozoal rhabdopeptides from *X. innexi*[119]. Thus for type one NRPS, the number of modules in an NRPS is equal to the number of amino acids in the final biosynthesized peptide (except for those that exhibit module skipping) and the order of the modules reveals the amino acid sequence of the final linearised peptide. Each module is further divided into the previously noted enzymatic subunits called domains with the basic three being: condensation (C) domain, adenylation (A) domain and peptidyl carrier protein/thiolation (T) domain, arranged in that order. When the arrangement of domains within modules deviates from this order, the NRPS falls under the final classification –type three NRPS[135].

For all three types of NRPS, a way constituent domains are identified is by first identifying NRPS-encoding genes in a genome sequence –called biosynthetic gene clusters(BGC) because they are found in one genomic locus and encode all proteins

responsible for the biosynthesis of a specific biomolecule and its derivatives– then translating these into the suitable frame and searching for known, domain-specific motifs that are then used to demarcate respective domains. For example, C1-C7 motifs[134] and the core-T motif[135] are used to identify C and T domains respectively. Alternatively, phylogenetic relationships to known domain sequences can be used to identify unknown cognate domains, as is done for C domains with the NaPDoS[137] bioinformatic tool. Alternatively, sequence alignments of known domains can be used to create Hidden Markov models[138] that detect distantly-related unknown homologous domains in a probabilistic manner[139]. Identification of domains in turn leads to demarcation of a module, which cumulatively results in *in silico* prediction of NRPs biosynthesised from a bacterium/fungus by genome analysis using one-stop-shop tools such as PRISM,[140] NRPS predictor[141] and antiSMASH[142].

Analogous to tRNA synthetase, A domain (ca. 550 aa[143]) functions to select a specific amino acid, which it converts to its amino acyl–*O*–adenosine monophosphate (amino acyl adenylate) using energy from the breakdown of an adenosine triphosphate(ATP). What specifies the amino acid an A domain activates is the amino acid residues found at positions 235, 236, 239, 278, 299, 301, 322, 330, 331, 517 of the A domain's primary sequence[144]. These ten selectivity-conferring residues (A1-A10) are referred to as the non-ribosomal[145]/Stachelhaus code[144] and are flanked by highly conserved residues that are the actual catalytic sites. The use of Stachelhaus codes in *in silico* prediction tools, has greatly hastened structure elucidation of non-ribosomal peptides[146].

Before the A domain selects it cognate amino acid, the NRPS would have been primed –converted from its *apo* to *holo* form– by the attachment of a 4' phosphopantetheinyl cofactor(PP arm) to the *β*-hydroxy of a highly conserved serine residue of the T domain in an ATP-consuming, Mg^{2+} dependent, 4-phosphopantetheinyl transferase(PPTase) catalysed reaction. Accordingly, deletion of the PPTase gene in *X. cabanillasii* resulted in total abolition of NRP production[147].

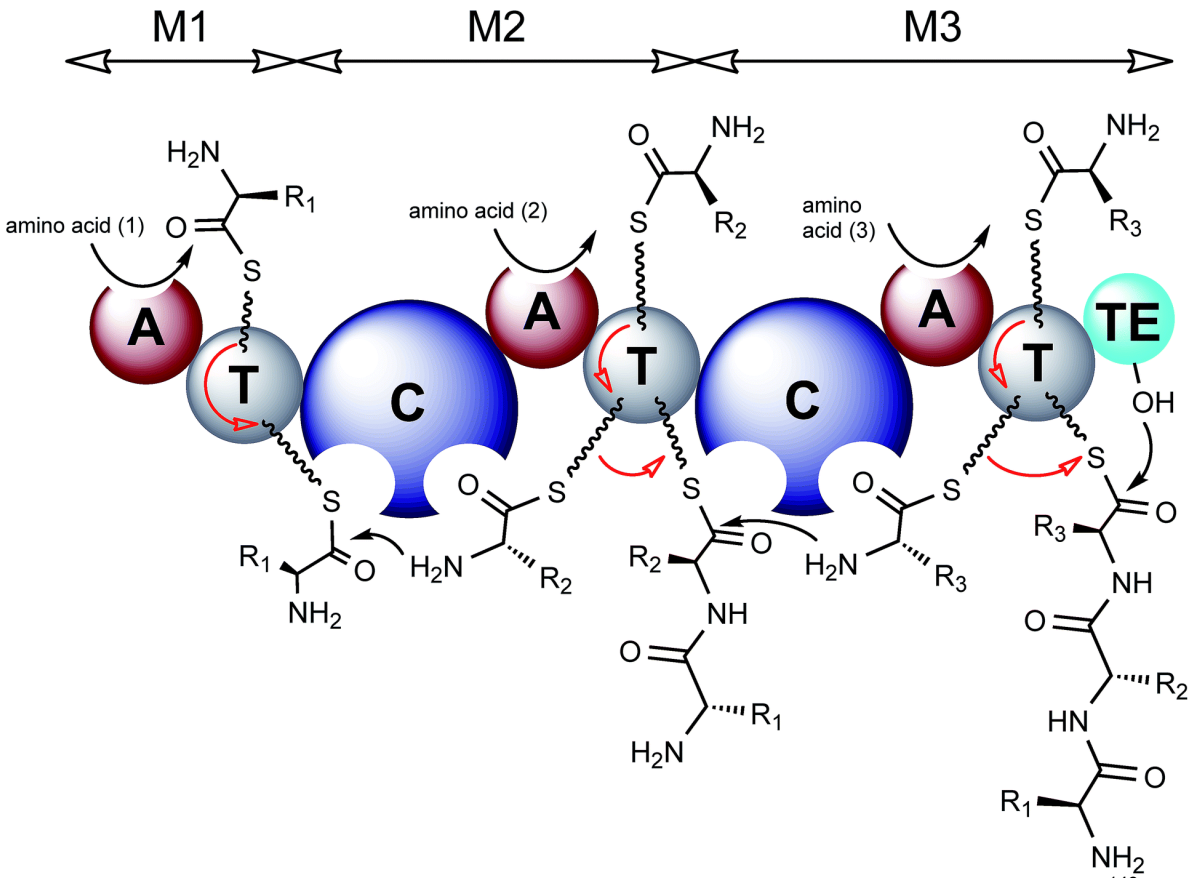


Figure 2: Model of NRPS domains essential for biosynthesis as reproduced from [148]. M= module; A, T, C, TE = adenylation, thiolation, condensation and thioestrase domains. Each C domain has an acceptor (bottom right semi-circle) and donor (bottom-left semi-circle) site. Phosphopantetheinyl cofactor (PP arm) arm is attached to T domain at one end and thioestered to amino acid adenylate/ peptidyl at the other end. PP arms are attached to T domains prior to any peptide bond formation. In addition to these essential domains, tailoring domains can also be present. Examples are C_{start} domain that is only found before the A domain of M1, epimerisation domains that are found adjacent to a C Domain and methylation domains that are embedded in an A domain. Authorised reproduction is under a CC license (https://creativecommons.org/licenses/by/3.0/legalcode).

Analogous to tRNA, T domains (ca. 80 aa residues[143]) have their PP arm covalently bind amino acyl adenylates(aaa) via thioesterification[146] and subsequently transport them to a catalytic site, the C domain. Functionally analogous to the LSU acceptor site, the C domain (ca. 450 aa residues[143]) bears acceptor and donor catalytic sites[149] that condense amino acids into peptides. Specifically, C domain catalyses amide bond

formation between two activated amino acids attached to the two T domains the C domain is found in between by catalysing a reaction where the *α*-amino group is the nucleophile/Lewis base as it forms a bond by donating both its electrons to its reaction partner – the amino acyl adenylate of the upstream peptidyl intermediate[135]. This is an example of a nucleophilic attack[150].

Apart from these three basic domains, other domains may also be found in an NRPS. A thioesterase (TE) domain demarcates the end of the NRPS and if present is vital for peptide synthesis as demonstrated by deletion of TE domain of plipastatin synthetase, which resulted in truncated enzymes unable to synthesise the lipopeptide[151]. TE domains catalyse peptide chain release and peptide macrocyclisation from the final module of the NRPS. They do this in two ways, the first of which is hydrolysis of the thioester attaching the peptidyl to the PP arm, and this results in release of a linear peptide. The other is by a nucleophilic attack on the terminal acyl group by either the hydroxyl or amino group of the side chain of a constituent amino acid resulting a branch cyclic peptide, or by the *α*-amino group of the *N*-terminal resulting in a head-to-tail cyclised peptide. This is what creates macrocyclised NRPs. When a hydroxyl group is the nucleophile an ester bond is formed creating a depsipeptide/macrolactone[151] while when an amino group is the nucleophile a lactam bond is formed creating a macrolactam. Linear peptide release from the NRPS can alternatively be catalysed by reductase (R) domains through reduction of electrons from the thioester attaching the peptidyl to the PP arm[143]. When two electrons are reduced an aldehyde is formed while reduction of four electrons forms an alcohol[146].

When present, an epimerization (E) domain is adjacent to a C domain and coverts the terminal L- amino acyl of the peptidyl intermediate to its D- enantiomer. Methylation domains can be found embedded in an A domain and catalyse the methylation of the *N*-terminals, sulphur groups of cysteine as well as *O*-methylation[146]. Formylation domains (F) catalyse *N*-terminal formylation[116].

If present, C starter(C_{start}) domains[134] are the first domain of an NRPS –this results in the initiation (first) module having C, A, T instead of only A, T domains– catalyse *N*-terminal acylation of a peptide resulting in lipopeptides, and potent clinical antibiotics such as daptomycin[152] polymyxin B[153] & colistin[154] are biosynthesised through this

mechanism. Re-engineering an NRPS to biosynthesise a lipopeptide derivative of its native NRP would thus focus on manipulating this domain either through site directed mutagenesis to generate desired amino acid residues in an active site[155] or entire domain swapping[156] of the native C_{start} domain with one that loads a different fatty acid, and both techniques can be accomplished through polymerase chain reaction(PCR) mediated overlap extension[157]. As lipopeptides are currently the most vital class of clinical antibiotics (Daptomycin[158], Colistin[159], Oritavancin[160], Dalvabancin[161], Telavancin[162]) due to their activity against Gram-positive infections[158,163] caused by antimicrobial resistant pathogens such as methicillin-resistant *S. aureus*[164] and vancomycin-resistant *Enterococcus*[163,165], C_{start} domain studies are vital to drug development of antimicrobial non-ribosomal peptides.

Macrocyclised antimicrobial non-ribosomal peptides from *Xenorhabdus* and *Photorhabdus*

Photorhabdus and *Xenorhabdus* bacteria are a rich source of specialised metabolites[166] including non-ribosomal peptides[167] and an ecological reason for this is the afore described endosymbiosis[168]. The following section gives a review of macrocyclised antimicrobial non-ribosomal peptides produced from these bacteria with a focus on how their physiochemical properties –it is these properties more than the amino acid composition that contribute to antimicrobial activity[169,170] although the latter affects the former– may contribute to their inhibitory effect.

Taxalllaids

These are cyclic lipodepsipeptides biosynthesised from a type one NRPS TaxA-B of *X. indica*[171]. The first domain of TaxA is a C_{start} domain and thus acylates the *N*-terminal with a fatty acid while the last domain on TaxB is a thioesterase and thus catalyses macrocyclisation through a nucleophilic attack of *β*-hydroxy group of L-Thr^1 on the *C*-terminal group creating a lipodepsipeptide. Diversity of taxalllaid derivatives is due to variation in fatty acid length (*iso* fatty acids of 2 or 3 carbons) and aa residue$^{position\ 3}$ (Phe/Leu)[171]. The primary peptide sequence of Taxalllaids is L-Thr^1 L-Ala^2 L-Phe/L-Leu^3 D-Leu^4 D-Leu^5 L-Leu^6 L-Ala^7 and thus have a net charge of (0), total hydrophobic ratio of at least 85% and probably do not form *α*-helices due to the short 7 aa residue chain length. Derivatives with L-Phe and 3 carbon fatty acids are more hydrophobic than

those with L-Leu and 2 carbon fatty acids because the hydrophobicity of L-Phe (1.19) is higher than that of L-Leu (1.06)[172] while longer fatty acids are more hydrophobic. Hydrophobicity of Taxalllaids may have contributed to antimicrobial activity against *Plasmodium falciparum* as the activity(in IC_{50}) decreased with decreasing hydrophobicity as follows: B(L-Phe & 3 carbon fatty acid) 2.87 μM; E (L-Phe & 2 carbon fatty acid) 2.9 μM and D(L-leu & 2 carbon fatty acid) 5.17 μM. For the weak antimicrobials, the linear derivative G was inactive against *Trypanasoma brucei* and *P. falciparum* while its macrocyclic derivative A inhibited both at IC_{50} 7.19 μM and 16.6 μM respectively. Against *Leishmania donovani* and *T. cruzi,* A was 2-fold & 4-fold more potent than its linear derivative. This demonstrated that macrolactonation contributed to antimicrobial activity in these highly hydrophobic depsipeptides.

Chaiyaphumins

These[118] are cyclic lipodepsipeptides macrolactoned by an ester bond between *β*-hydroxy of L-Thr and the terminal carboxyl and *N*-terminally acylated by 4 different fatty acids. The primary peptide sequence is L-Thr^1-D-Phe^2-D-Ala^3-L -Pro^4-L-Tyr^5 and thus the peptide has a net charge of (0) and total hydrophobic ratio of at least 60%. Diversity of Chaiyaphumins is caused by variation in fatty acids at *N*-terminal with A having phenylacetic acid (PAA), B n-butyrate, C n-propionate and D acetate. Accordingly, A was the most hydrophobic derivative and had the higher antimicrobial activity against *P. falciparum* and *T. brucei* (IC_{50}: 0.61 μM and 5.11 μM respectively) when compared to the only other tested derivative B (IC_{50}: *P. falciparum* 15.4 μM; *T. brucei* 77.8μM). Both were not toxic to mammalian cells at high concentrations (IC_{50}: A 92 μM; B 151 μM) demonstrating these hydrophobic lipodepsipeptides as strong selective antimicrobials against *P. falciparum.*

Szentiamide

These[116] are branch cyclic depsipeptides biosynthesised from *X. szentirmaii* and that are macrocyclised by an ester bond between *β*-hydroxy of L-Thr and the terminal carboxyl and *N*-terminally formylated. The primary peptide sequence is D-Leu^1-L-Thr^2-D-Phe^3-D-Val^4-L-Tyr^5-L-Trp^6 conferring the peptide with a net charge of (0) and total hydrophobic ratio of at least 60%. Szentiamide was inhibitory to *P. falciparum, L.*

donovani, T. brucei (IC_{50}:1.2 μM, 11.6 μM, 12.8 μM respectively) non-inhibitory to bacteria and lowly toxic to mammalian cells(IC_{50}:66.7 μM)[173] indicating these hydrophobic branch cyclic depsipeptides as antimicrobials selective against protozoa.

Xenematides

In these[103,174,175] cyclic lipodepsipeptides –biosynthesised from *X. nematophila*– that consists of Xenematide A-G, diversity was due to the low substrate specificity of A domain of modules two and three of the Xnc NRPS that specified Phe/Trp/Leu/*α*-aminoheptanoic(Aha) acid and Phe/Trp respectively[103,175]. Conversely, all seven had the same Phenylacetic acid (PAA) moiety at their *N*-terminal indicating a high substrate specificity of the C_{start} domain. This in contrast to the lowly substrate-specific C_{start} domain of taxalllaid[171], chaiyaphumin[118] and plipastatin[151] biosynthesis that diversify these lipopeptides by acylating different lipid moieties.

The primary peptide sequence is L-Thr^1-L-Trp/L-Phe/L-Leu/L-Aha^2-Trp/Phe^3-*β*-Ala^4 and thus has a net charge of (0) and has a total hydrophobic ratio of at least 75%. In terms of bioactivity, Xenematides possessed antibacterial activity against the Gram-positive *Bacillus subtilis* in the following decreasing order: A, G, F, B (MIC: 15, 26, 54, 109 μM respectively) and against Gram-negative *Pseudomonas* sp. only A and G (MIC 15, 26 μM respectively) were active. The highly hydrophobic residue L-Aha may have increased hydrophobicity of G that in turn led to its high activity due to better insertion into cell membranes. Relatedly, the high activity of Xenematide A could be due its 50% composition of tryptophan residues. These preferentially interact with cell membranes surfaces[176]enhancing the interaction of the peptide with the cell membrane and thus may contribute to antibacterial activity in these cyclic lipodepsipeptides.

Xenobactin

This[108] is a cyclic hexadepsipeptide biosynthesised by *Xenorhabdus* sp. PB30.3. The presence of D amino acid residues, *N*-terminal acyl moiety and macrolactonation indicate that it is an NRP biosynthesised by a C_{start} domain-containing, six-module NRPS. Macrocyclisation was via an ester bond between L-Thr^1 and L-Leu^6 and *N*-terminal acylation was with an acetate moiety. The amino acid sequence of was L-Thr-L-Trp-L-Thr-D-Val-D-Ile-L-Leu thus had a net charge of (0) and total hydrophobic ratio of

at least 66%. Xenobactin inhibited Gram-positive *M. luteus* (MIC: 84.6 µM) and was a broad-spectrum antiprotozoal, active against *T. brucei*, *T. cruzi* and *P. falciparum* (IC_{50} 41.8; 76.5, 88.6 µM respectively). Conversely, it was non-inhibitory to fungi, Gram-negative bacteria and nontoxic to mammalian cells indicating a selective antimicrobial activity of these hydrophobic depsipeptides.

Phototemtide

Phototemtide A-D[114] are cyclic lipopeptides biosynthesised from the *P. temperata* PttB-C synthetase that functionally requires an MbtH protein[177]. PttB synthetase had a C_{start} domain that was responsible for diversity among Phototemtides as A, B, C, D differed in the fatty acid moiety at the *N*-terminal that is, 3-hydroxydecanoic acid, 3-hydroxyhexanoic acid, 3-hydroxyheptanoic acid and 3-hydroxynonanoic acid respectively. Unique was the macrocyclisation of Phototemtides that was *C*-*O* esterification between a hydroxy of the *N*-terminal fatty acid with a terminal carboxy of the peptide. The amino acid sequence was Gly^1-L-Val^2-D-Val^3-L-Thr^4-L-Ile^5 giving the lipopeptide a total hydrophobic ratio of at least 60% and a net charge of (0). Phototemtide A was selectively inhibitory to protozoa–no inhibition of fungi, bacteria nor mammalian cells– of the following species: *P. falciparum*, *T. brucei*, *T. cruzi* (IC_{50} was 9.8 µM, 62 µM, 83 µM respectively). Thus this cyclic lipopeptide displayed selectivity antiprotozoal activity, similar to the cyclic lipodepsipeptides Szentiamide, Chaiyaphumin A and B.

Xenoamicin

Among the depsipeptides isolated form *Xenorhabdus* species, Xenoamicin[115] is not only the largest but also the most widespread with detected production from four[178] (*X. innexi, X. nematophila*, *X. mauleoni*, *X. doucetiae*) out of the 26 characterised species of the genus. The lead compound of its class Xenoamicin A[115] was biosynthesised from the XabA-D synthetases that contain both a C_{start} domain and double TE domain resulting not only in *N*-terminal acylation but also macrocyclisation that was by an ester bond between D-Thr^6 and L-Val^{13}. Diversity among Xenoamicin A-H was due to variation in the fatty acid moiety at the *N*- terminal indicating a low substrate specificity for this C_{start} domain. Methylation of D-Ala^2 and the promiscuity of A domain of module

10 further derivated the Xenoamicins. The sequence of Xenoamicin A is L-Pro[1]-D-Ala[2]-L-Val[3]-D-Leu[4]-L-Ile[5]-D-Thr[6]-L-Val[7]-D-Val[8]-L-Val[9]-D-Ala[10]-*β*-Ala[11]-L-Pro[12]-L-Val[13] and thus has net charge of (0) and total hydrophobic ratio of at least 75%. Despite its amino acid residue length being more than seven and being composed of alternating D and L amino acids, Xenoamicin has a non-terminal proline residue (L-Pro[12]) and this makes it unlikely to form a helix secondary structure upon contact with a cell membrane[179]. Nonetheless, Xenoamicin A was potent as an antiprotozoal with activity against *P. falciparum*, *T brucei, T. cruzi, L. donovani* (IC_{50}: 1.8 μM, 4.9 μM, 23.5 μM, 38.6 μM respectively). It was non-inhibitory to bacteria, yeast fungi and lowly toxic to mammalian cells (IC_{50}: 52 μM) indicating a selective antiprotozoal activity of these large hydrophobic depsipeptides.

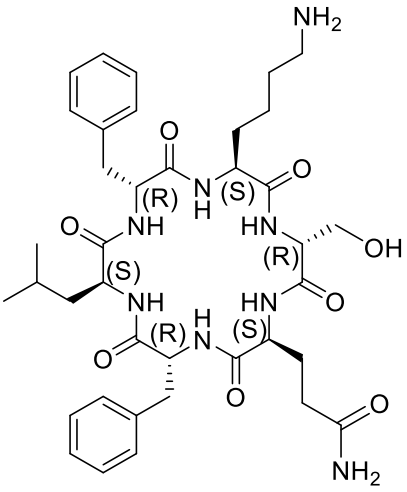
PAX lipopeptides

The synthesis of PAX lipopeptides from *X. nematophila* is encoded by the *PaxA-C* operon[113]. These lipopeptides are cyclised by an amide bond between side chain amino group of D-Lys[3] and terminal carboxyl of L-Lys[7] and *N*-terminally acylated by 3-hydroxy fatty acids derivatives. Diversity is increased by variation of amino acid of position two, thus the primary peptide sequence is Gly[1]-L-Lys/L-Arg[2]-D-Lys[3]-L-Lys[4]-D-lys[5]-D-Lys[6]-L-Lys[7] making it completely hydrophilic with a net charge of (+6). A degree of hydrophobicity is required to penetrate the phospholipid bilayer for most antibacterial antimicrobial peptides[170] and the fatty acid moiety in the PAX peptide could contribute towards this. In fact, the fatty acid moiety is structurally responsible for the difference in MIC against *M. luteus*[112] between PAX 3(MIC:3.125μM) and PAX 5(MIC:1.56 μM) as both had the same peptide sequence, but 3-hydroxypetandecenol, 3-hydroxytetradec-7-enol fatty acids respectively. This difference in MIC was similarly observed against *Fusarium oxysporum* (PAX3, PAX5:1.56μM, 0.78μM). PAX peptides had no toxicity towards mammalian cells[112], demonstrating selective activity of these highly cationic macrolactam lipopeptides.

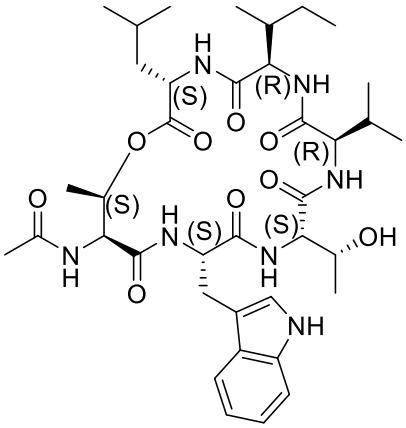
Ambactin

Ambactin[117] from *X. miraniensis* is bioactive macrolactam i.e it was cyclised by a nucleophilic attack of the *α*-amino group of D-Ser[1] on the terminal carboxyl. It had a

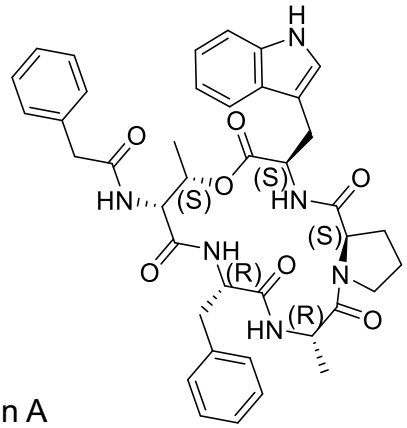
low total hydrophobic ratio of 50% due to its primary structure of D-Ser[1]-L-Gln-[2]D-Phe[3]-L-Leu[4]-D-Phe[5]-L-lys[6] and a net charge of (1+) due to the lysine residue. One anomaly was its lack of a lipid moiety at the *N*-terminal despite being biosynthesised from the C_{start} domain-containing AmbS synthetase. Through multiple sequence alignment analysis of C_{start} domain polypeptides, the authors[117] attributed the non *N*-terminal acylation to a mutation in an active site, which in functional C_{start} domains contained His-His while in Ambactin was His-Pro. A successful NRPS re-engineering of the C_{start} domain active site back to His-His did not result in lipidated ambactin derivatives. In spite of this anomaly ambactin had antimicrobial activity inhibiting *T. brucei* at IC_{50}: 7.98 μM. This cationic antimicrobial peptide most likely utilised electrostatic interactions with the anionic cell membrane of Trypanosomes to effect its inhibitory effect. Whether the macrolactams from nematophilic bacteria tend to utilise a cationic nature to effect their antimicrobial activity is a question that can only be answered by discovery of more antimicrobial macrolactams.



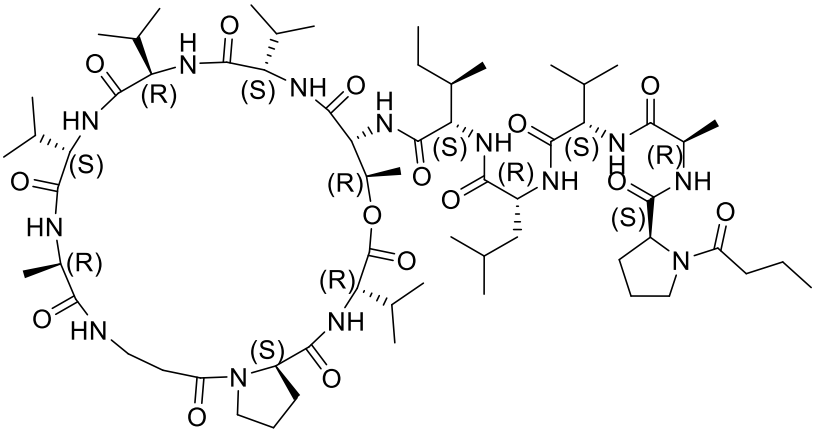
Ambactin



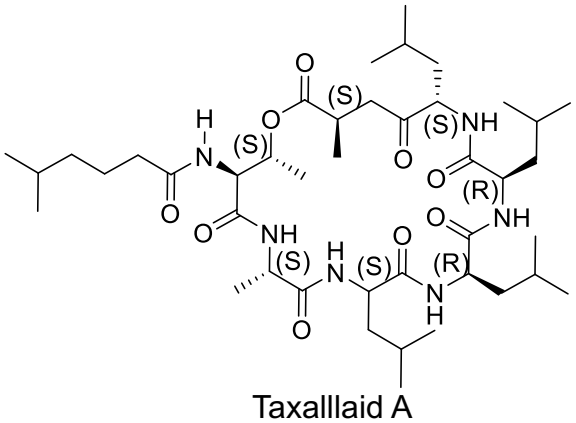
Xenobactin



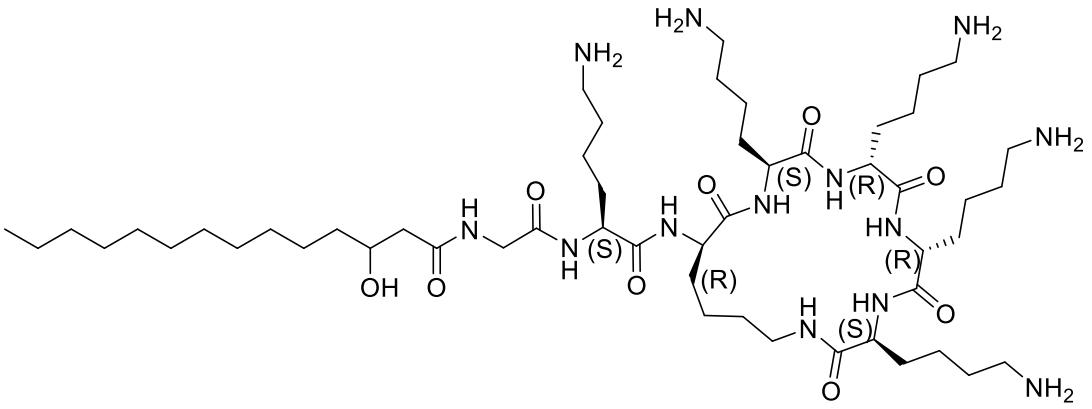
Chaiyaphumin A



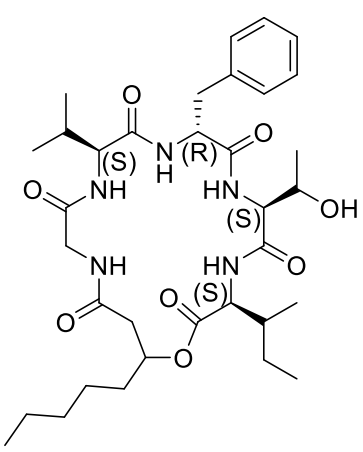
Xenoamicin A



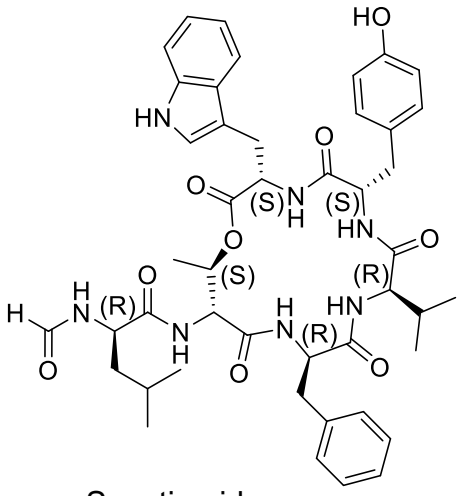
Taxalllaid A



PAX Peptide 1



Phototemtide



Szentiamide

Figure 3: Macrocyclised antimicrobial non-ribosomal peptides classes from nematophilic bacteria.

Objectives of this study

General Objective

To demonstrate an antimicrobial discovery process, from soil sample collection to detection of inhibitory activity, through investigations on nematophilic bacteria.

Specific Objectives

1. Isolation of novel *Xenorhabdus* strains from soils of Western Kenya.
2. Conclusive species delineation of isolated *Xenorhabdus* strains.
3. Comparative genome analysis of isolated *Xenorhabdus* strains with *X. griffiniae* XN45.
4. Structure elucidation, biosynthesis, and chemosynthesis of putative macrocyclic non-ribosomal peptides whose synthesis is encoded by *ishA & ishB* genes of *X. ishibashii* DSMZ 22670.
5. Determination of antimicrobial activity of macrocyclised non-ribosomal peptides Ishipeptide A from *X. ishibashii* and Photoditritide from *P. temperata* Meg1.

MATERIALS AND METHODS

Collection of field soil samples

Fieldwork was carried out from 16.10.2018-04.11.2018 in Western & Rift Valley provinces of Kenya as part of the 2017 Kenya National Research Fund Multi-disciplinary grant research project titled “Drug development of antibiotics from novel sources: *Xenorhabdus* bacteria from Kenya”. Ten localities were selected for collection of soil samples: Nandi Hills, Tinderet, Fort Tenan, Kakamega, Gisambai, Vihiga, Kisumu, Bungoma, Kaimosi and Mt. Elgon. Within each, collection points were selected from cultivated lands, fallow lands, forests, crop edges, shoreline, swamps and riverine areas. This resulted in a total of 76 soil collection points. To collect a soil sample, vegetation was first cleared from the topsoil. Then using a digging fork, soil was dug out to a depth of not more than 60 cm. Using a collection spade, soil was scooped into a measuring cup to an amount of ca. 500g. Twigs, branches, and stones were removed before the soil sample was placed in labelled cotton bags. Dug out soil was returned to the hole and soil samples were then transported at room temperature to Entomopathogenic Nematology Laboratory of Horticultural Research Institute,Kenya and Genetics laboratory of School of Biological Sciences, University of Nairobi, Kenya. Geographic coordinates, altitude, and a description of the soil collection point are provided in Table S1.

Isolation of nematodes from soils samples

To isolate EPNs from soils, a soil sample soil was first spread out on a tray and crumbled to facilitate ease of movement for putative EPNs. Soil was then redistributed into transparent polyethylene teraphthalate (PET) plastic containers of 20 cm diameter and 5 cm depth. To bait EPNs from soil, two-three *Galleria mellonella* larvae were buried in the soil in a hole of ca. 1 cm diameter and 5 cm depth. In total, about 15 *G. mellonella* larvae buried in 5 holes were used as bait per container. After a maximum of seven days, containers were checked. A number of samples had larvae that were alive while a few had dead larvae suggesting death from EPNs. These were collected and further assessed for the following characteristics that typify an EPN infection: limp cadaver, tan or red in colour and no smell or sign of putrefaction. Samples BG5 and

VH1 had dead cadavers that were either light-red or tan in colour. BG5 was clay soil collected from fallow land that was highly vegetative and ten meters from a stream with many ants observed at the collection point. VH1 was collected from land cultivated with cabbages. Other soil samples with dead cadavers were KS8, TN10, BY8, KR2, BY10, VH4, ST4, KS9 and ND14.

To isolate putative EPNs from these cadavers, a modified White trap[180] was made as follows. Clean PET containers of the aforementioned size were filled with distilled water to depth of ca 0.4 cm. A clean petri-dish was placed upside down into the container such that the petri-dish surface was raised from the bottom of the PET container. Clean cotton cloths of the same size as the petri-dish were placed on this raised surface. Selected cadavers were placed onto the cotton cloths. To allow putative EPNs to emigrate from the cadavers to the water, a part of the cloth was let to dip into the distilled water. PET containers were covered and kept for seven days. To assess for the presence of putative EPNs, distilled water was observed daily under a dissecting microscope. Presence of white motile, ca. 1mm long nematodes indicated putative EPNs.

For positive samples, contaminants such as cadaver tissue debris were separated from nematodes by a series of sedimentation and decanting using distilled water. Nematodes were stored in contamination-free distilled water –to a depth of not more than 0.4 cm– in clear plastic containers in the dark. Stored EPN nematode cultures were named after their soil collection point.

Isolation of bacteria from nematodes

To isolate bacteria, nematode isolates from collection point BG5 and VH1 were selected. This was done as follows. Instead of discarding cadavers from which BG5 and VH1 nematodes had been baited with, they were surface sterilised and dissected under aseptic conditions. A light-yellow, viscous, heterogenous, fluid was aseptically obtained and streaked onto nutrient agar supplemented with 0.0025% (w/v) bromothymol blue and 0.004% (w/v) 2,3,5 triphenyl tetrazolium chloride (NBTA). This was incubated at 30°C for 96 h. Only colonies that had the following observed morphologies were selected for further pure culture techniques: blue/yellow pigment, irregular margins, umbonate shape and visible swarming patterns. On these pure

cultures, a catalase test was performed as follows. A loopful of bacteria colony was collected from an agar plate culture and placed into a sterile Petri dish. Using a pipette, two drops of 20%(v/v) H_2O_2 was added. Absence of bubble production indicated a catalase-negative isolate and these were presumptively identified as *Xenorhabdus* species. They were named *Xenorhabdus* sp. strain BG5 and strain VH1.

Genome sequencing and assembly

Previously, we isolated *X. griffiniae* XN45 from *Steinernema* sp. Scarpo that was originally isolated from Muran'ga District in Kenya[69]. Thus, in addition to *Xenorhabdus* sp. strains VH1 and BG5, this strain was selected for genome sequencing and assembly in order to aid in species delineation of the novel isolates and comparative genome analysis.

DNA from strain XN45 was extracted with FastDNA Spin Kit for Soil® (Mp Bio, California USA) to yield concentration of 20 ng/µl and ultraviolet(uv) absorbance ratio at $260_{nm}/280_{nm}$ (A260/280) >1.8. From this, only 0.1 ng of DNA was used to prepare a library using a Nextera XT kit (Illumina, California USA). Sequencing was done by CeGaT GmBH, Tuebingen, Germany on a NovaSeq® 6000 platform with the following parameters for short insert paired reads of 100bp and targeted coverage of 100x. Output data was raw sequence reads in fastq.gz format (2.902 GB), which had an Illumina standard Phred score(offset +33) and adapter sequences already removed. In terms of quality, 91.32% of reads had a Q30 value. Genome assembly was done with Spades v3.8.1/3.10.1[181] with thresholds for minimum contig length and coverage set at 1000bp and 5x respectively. For VH1 and BG5, DNA was isolated with Gentra® Puregene® DNA extraction kit (Qiagen, Venlo Netherlands) to yield samples of 1µg/µl concentration and A260/280 ratios of >1.8. Genome sequencing and assembly was done as previously described[166], which was a workflow similar to the aforementioned. For assembled genomes of strains VH1, BG5, *X. griffiniae* XN45 from this study and *X. griffinae* strain BMMCB from[68], characteristics such as completeness, contamination, N50, L50, length and guanine-cytosine (GC) content as well as genome annotation with the RAST-k algorithm[182] were determined using the comprehensive genome analysis tool of the PATRIC platform[183].

Phylogenomic reconstruction

Twenty-six fasta files of selected genomes (Table S2) were used as input data for a whole genome-based taxonomic analysis on the Type strain genome server platform (TYGS) [184]. Output data –downloaded from TYGS server on 2020-08-26– was obtained as briefly described below. MASH algorithm[185] was used to quickly calculate intergenomic relatedness and determine the strains with the smallest distances. All pairwise comparisons and inference of intergenomic distances among the set of genomes were conducted using GBDP 'trimming' algorithm and distance formula d_5 [186]. 100 distance replicates were calculated each. Genome-genome distance calculator (GGDC) 2.1 formula was further used to calculate dDDH values and confidence intervals[186]. Intergenomic distances were then used to infer a balanced minimum evolution tree with branch support via FASTME 2.1.4 including SPR post-processing [187]. Branch support was inferred from 100 pseudo-bootstrap replicates each. Rooting of trees was done at midpoint while visualisation and graphics editing was done with iTOL and Inkscape respectively[188,189]. A similar workflow was used with thirteen genomes of *Photorhabdus*.

Minimum thresholds for two strains to be classified as one species and sub-species were 70% dDDH and 79% dDDH respectively[184,190]. To calculate ANI values between species most closely related to strains VH1&BG5, the orthoANI algorithm[191] was used within the OAT software package, which was also used to obtain genome-genome distances (GGD) 2.1 values.

Pangenome analysis

Twenty-six genomes of *Xenorhabdus* species were used to construct a pangenome of the genus using the anvio pangenome workflow[192]. After fasta files were obtained from the Genbank database, draft genomes contigs were concatenated in Geneious® R8.1.9 Software suite (Biomatters Limited, New Zealand) to remove deflines and thus enable suitability for anvio workflow processing. Gene prediction was done with Prodigal[193] and genomes were then converted to .db files. On these, hidden Markov model profiling was done. A .txt file that had the names of the each of the genomes and file paths to its .db file per line was created and this was used as data input for the

final anvio pangenome analysis command. To create clusters of orthologous groups (protein clusters) protein-protein BLAST[194] coupled with the MCL cluster algorithm[195] –although slower and computationally more demanding– was used instead of Diamond[196] in order to increase the accuracy of computed protein clusters. MCL inflation was maintained at 2.0. A similar workflow was used for creating the *X. griffiniae* clade and *Photorhabdus* pangenomes. The strain names, accession numbers and number of predicted genes of each genome are listed in Table S2. To estimate whether pangenomes were open or closed the mean α value was determined for each using the P-GAP platform running on the Panweb server[65]. Briefly, RAST-k annotated genomes were used as data input on the Panweb server[65] and the following parameters were selected for clustering genes into one orthologous group: minimum 80% nucleotide similarity, minimum 80% coverage with gene family algorithm. A similar workflow was used for *Photorhabdus* pangenome analysis.

Elucidation of genomic islands in Strain BG5

To highlight putative genomic islands flanked by transposases, an annotated record of the BG5 genome was concatenated and used as the reference genome in BRIG[197] and compared to genomes of VH1 and XN45 by the BLAST algorithm[194] utilising an NCBI-blast 2.4.0+ bin library. Selected rings to be visualised were for BG5 genome GC content (ring 1) and skew (ring 2), VH1 draft genome (ring 4), XN45 draft genome (ring 5) and loci of CDS annotated as transposases on BG5 genome (ring 6). Output visualisations were obtained as .svg files and enhanced in Inkscape[198].

To obtain genomic coordinates of putative genomic islands and extract them, BG5 contigs bearing them were aligned against complimentary contigs in XN45 using the progressive Mauve 2.3.1 algorithm[199] plug in Geneious® R8.1.9 Software under the following settings: do a full alignment; compute locally collinear blocks (LCB), automatically calculate the seed weight and minimum LCB. Prophinder[200] was used to identify genomic coordinates of putative prophage loci. To compare if GC content of these regions differed significantly from that of the genome, the GC content of each position in a genomic island or prophage region was calculated in n Geneious® R8.1.9 Software with setting sliding window= 57 and exported as a .csv file. These were then used to compute a one-sample t-test[201] against a theoretical mean GC content of

43.377993%. For genomic loci that were associated with transposases and had significantly different GC content, the predicted products of coding DNA sequences (CDS) contained therein were compiled.

General chemistry experimental procedures

High performance liquid chromatography- electrospray ionization-high resolution mass spectrometry (HPLC-ESI-HRMS) was acquired using a Bruker Impact II Ultra-High Resolution Qq-Time-Of-Flight(UHR-QqTOF) mass spectrometer coupled to Thermo Scientific™ UltiMate™ 3000 HPLC system equipped with an Acquity UPLC BEH C18 column (130 Å, 2.1 mm × 50 mm, 1.7 µm). Semi-preparative HPLC was performed on an Agilent 1260 system with a Cholester column (10 ID × 250 mm).

Bacterial strains, *in silico* and molecular methods

Bacterial strains obtained for investigations on putative macrocyclic non-ribosomal peptides whose synthesis is encoded by *ishA & ishB* genes were *X. ishibashii* DSMZ 22670[31] and the aminolevulinic acid (ALA) auxotroph *E. coli* strain ST18 DSMZ 22074[203]. To elucidate biosynthetic gene clusters in *X. ishibashii* DSMZ 22670, the draft genome (Master record Genbank accession number: NJAK00000000.1) was submitted as data input in an antiSMASH version 5[142] analysis under default options (cluster detection: relaxed, known cluster blast: on, active site finder: on, sub cluster blast: on). Output files revealed contig NJAK01000001.1 to contain the hitherto unknown *ishAB* NRPS cluster whose A and C domains within all modules were further analysed to determine putative amino acid building blocks and their stereochemistry, respectively. Using the aforementioned genome sequence, cloning protocols including design of polymerase chain reaction (PCR) primers were developed using Geneious® R8.1.9 Software suite (Biomatters Limited, New Zealand) and oligonucleotides were purchased from Eurofins Genomics, Germany. Total DNA extraction from bacterial strains was done using Gentra® Puregene® DNA extraction kit (QIAGEN GmbH, Germany). To refactor the native promoter of *ishAB* via promoter exchange mechanism[204], enzymatic amplification of a DNA fragment of the first 629 base-pairs (bp) of *ishA* was first done via PCR with primers RMA87 and RMA88 that both contained vector-derived 23 bp overhangs at the 5’ end to facilitate downstream

isothermal ligation independent cloning into plasmid vectors via the hot-fusion method[205].The fragment was amplified in a 25 µL reaction tube containing final concentrations of 0.32 µM each primer, 1.5 mM $MgCl_2$, 200 µM each dNTP, 0.2 U Phusion High Fidelity DNA polymerase (ThermoFischer Scientific, Germany) and 50-150 ng genomic DNA. Cycling conditions were as follows: 98 °C for 30 s then 10 cycles of 98 °C for 6 s, 59 °C for 30 s, 72 °C for 35 s, then 25 cycles of 98 °C for 6 s, 72 °C for 30 s, 72 °C for 40 s. PCR products were visualized on gels constituted from 1% (w/v) agarose in TAE buffer, which had been previously stained with ethidium bromide at final concentration of 0.5 µg/mL and subject to an electric field of 7 V/cm with TAE as electrolyte. Expected PCR products were then column-purified using Stratec Msb Spin Pcrapace kit (Invitek Molecular GmbH, Germany). These were cloned via the hot fusion method[205] into linearized pCEP plasmids[204] by incubation at 50 °C for 1 h in a 5 µL reaction volume. The reaction (5 µL) was directly mixed with electro-competent *E. coli* ST18 cells (50 µL) that were then transformed, via electroporation (1.25 KV, 200 Ω, 25 µF and 1 mm cuvette), by ligated plasmids. Electroporated cells were immediately mixed with LB medium (950 µL) supplemented with ALA (50 µg/mL) and incubated at 37 °C with agitation at 200 rpm for 1 h to facilitate replication of plasmids in positive transformants. This culture was inoculated onto LB agar plates supplemented with kanamycin (50 µg/mL) and ALA (50 µg/mL) and incubated overnight at 37 °C to select for positive transformants. Cultured cell PCR was performed on positive transformants with backbone-vector derived primers v_PCEP_FW and PDS132_rv to confirm clones as carrying plasmid pCEP_JC0094. Positive *E. coli* clones were used as donor cells in diparental mating with *X. ishibashii* wild type strains. Briefly, both donor and recipient cells were grown overnight in appropriate media. Then a sub-culture (5 mL) of each was made by inoculating 1% (v/v) of overnights and each was cultivated until exponential phase. For the donor strain, a cell pellet from 1 mL culture volume was washed thrice by mixing with LB medium (1 mL) then centrifugation at 13,000 *g* for 1 min, to eradicate traces of ALA. For the recipient strain, cells from 5 mL culture volume were pelleted by centrifugation, then mixed with the donor cell pellet in a total LB medium volume of 50 µL. This was dropped on LB agar plates which were incubated at 30 °C for 24 h, transforming recipient cells

to *X. ishibashii* strain pCEP_JC0094. LB agar supplemented with Kanamycin (50 µg/mL) was used to select for positive transformants and cultured cell PCR was done using primers ALPCEP1f and RMAr032 to confirm successful plasmid integration at the targeted genomic locus. Genetically confirmed *X. ishibashii* strain pCEP_JC0094 were used for fermentation of **1** and **2**. To verify selected regions of the *ishA* and *ishB* gene sequences of contig NJAK01000001.1, three regions (Table S4) were PCR-amplified and sequenced (Eurofins Genomics, Germany). These were aligned to NJAK01000001.1 in Geneious® R8.1.9 Software suite (Biomatters Limited, New Zealand) with Geneious alignment with the following settings: global alignment with free end gaps; cost matrix 65% similarity (5.0/-4.0); gap open penalty 12; gap extension penalty 3. This revealed 3 regions with GGGGGGG that had a deletion of one G. Corrected *ishAB* gene nucleotide sequences were deposited in DDBJ under accession numbers LC536431.1 and LC536432.1.

NRPS re-engineering

To re-engineer the NRPS, the protein encoded by *ishA* gene was altered via site-directed mutagenesis by changing its glutamine of position 143 to a histidine. To do this, the exact same procedure for creating a mutant as above was used, except that primer RMA 88b was instead of RMA 88. This primer introduced a point mutation at position 429 from (G) to (T). This mutant was named *X. ishibashii* _pCEP_JCHIS143.

Bacterial fermentation and HPLC-ESI-HRMS analysis

Bacterial fermentation was used to produce compounds **1** and **2** for routine HPLC-ESI-HRMS analysis and compound isolation. *Xenorhabdus-Photorhabdus* production media (XPPM)[106]supplemented with 4% Amberlite® XAD-16 (Sigma Aldrich, Germany) resin beads (w/v) and 0.2% L-arabinose (w/v) in Erlenmeyer flasks five times the media volume, were inoculated with 24 h pre-cultures *of X. ishibashii* strain pCEP_JC0094. These were incubated for 72 h at 30 °C with continuous shaking at 130 rpm. Bacterial cells and supernatants were decanted out and elution of compounds bound to residual XAD-16 resin beads was done by mixing them with MeOH (1:1) and stirring magnetically for ca. 1 h. For routine analysis, this XAD-16 methanol extract was further diluted with MeOH (1:4) and 5 µL injected into an HPLC system under an aqueous

MeCN (supplemented with 0.1% formic acid (v/v)) gradient of 5-95% for 16 mins at a flow rate of 0.4 mL/min. The mass spectrometer was set to choose the 5 most intense precursor ions in each cycle (3s) to fragment over the mass-to-charge ratio (*m*/*z)* range of 100-1400/100-1200. Electron spray source conditions were as follows: end plate offset 500 V, nitrogen gas temperature 200 °C, nitrogen gas flow rate 8 L/min, nebulizer 3.1 bar and capillary voltage 4500 V. Samples were run in positive ion mode and at collision energy of 10 eV, and spectral data was obtained as .d files that were analysed in MzMine[206] (i.e converted mzXML files) and Compass Data analysis 4.3 software (Bruker Daltonik GmbH, Germany), which was also used to calculate the sum formulae for **1** and **2** and *de novo* peptide sequencing of **2**.

Labelling experiments

Isotope labelling experiments combined with detailed HPLC-ESI-HRMS analysis were used to aid structure elucidation as previously described[123]. Labelling experiments were used to verify calculated formulae for **1** & **2** and determine amino acid monomers of **1**. Briefly, to experimentally verify formulae, **1** and **2** were fermented as described above from the following media that had different isotopes of carbon and nitrogen as their main sources: standard medium (^{12}C, ^{14}N) (Sf-900 medium, Gibco Thermofisher, Germany), fully labelled ^{13}C medium (^{13}C, ^{14}N) and fully labelled ^{15}N medium (^{12}C, ^{15}N) (Isogro Sigma Aldrich, Germany). Comparison of the masses via high resolution tandem mass spectrometry ($HRMS^2$) analysis, of compounds **1** and **2** when fermented from each medium revealed mass shifts identical to the exact carbon and nitrogen atoms of the compounds. The amino acid building blocks of compound **1** were similarly determined via labelling experiments[123]. Briefly, by comparing mass of **1** when fermented from sf-900 medium to when fermented from the same medium supplemented with 1 mM of a deuterated amino acid ([D_{10}] L-leucine, [D_4] L-tyrosine, [D_8] L-valine) (Cambridge Isotopes Laboratories, USA) mass shifts that corroborated with the number deuterated hydrogen atoms in a constituent amino acid monomer (+9 for ([D_{10}] L-leucine, +3 for [D_4] L-tyrosine, +7 for [D_8] L-valine) of **1** under analysis**,** were visualised via $HRMS^2$ analysis. Mass shifts of (-5) that showed L-glutamic acid and L-glutamine as constituent monomers of **1**, as they corroborated with the number of carbon atoms in a single molecule of L-glutamic acid or L-glutamine, were

experimentally determined via inverse feeding experiments that entailed fermentation from fully labelled ^{13}C medium supplemented with 1 mM of either ^{12}C L-amino acid.

Compound isolation

For isolation of **1**, 6 L of media was used, and harvested Amberlite™ XAD-16 resin beads(Sigma Aldrich, Germany) (XAD-16) were eluted three times with MeOH (2 L) that was further concentrated by rotary evaporation under reduced pressure to 0.5 L. This was then sieved through 0.4 µm then 0.2 µm pore-size filters to remove carry over particles, and final concentration under reduced pressure yielded 5.65 g of XAD-16 resin methanol extract. To fractionate out compounds of low solubility in MeOH as **1** had demonstrated good solubility in MeOH, the extract was dissolved in 100 ml of MeOH and incubated at 4°C. Only the liquid fraction was collected and concentrated to dryness *in vacuo* to yield 4.62 g. This was dissolved in MeOH (14 mL) and further fractionated by Sephadex LH-20 (Sigma Aldrich, Germany) gel filtration chromatography using MeOH as the eluent to yield 100 fractions, of which 14-24 contained **1** as detected by routine HPLC-MS analysis and were thus pooled. From this pool, a fraction that contained **1** was eluted by semi-preparative HPLC with an isocratic 40% MeCN/H_2O solution containing 0.1% formic acid (v/v) at a flow rate of 3 mL/min for 20 min to yield **1** (5.8 mg). This isolation method gave the production titer of **1** from *X. ishibashii* strain pCEP_0094 as 1mg/L.

Chemical synthesis of 1

Step a - Solid-phase peptide synthesis. The linear sequence was synthesised on the 2-chlorotrityl chloride (2-CTC) resin preloaded with D-leucine on a 25 µmol scale with a Syro Wave peptide synthesizer by using standard 9-fluorenylmethoxycarbonyl (Fmoc) chemistry. The resin was placed in a plastic reactor vessel with a Teflon frit and an amount of 6 equivalents (eq.) of amino acid derivatives (Fmoc-D-Ala-OH, Boc-Leu-OH, Fmoc-Leu-OH, Fmoc-Gln(Trt)-OH, Fmoc-Val-OH, Fmoc-D-Leu-OH, Fmoc-D-Ser-OH, Fmoc-D-Val-OH, Fmoc-D-Tyr(O*t*Bu)-OH, 0.2 M) were activated in situ at room temperature with 6 eq. of *O*-(6-Chlorobenzotriazol-1-yl)-*N*,*N*,*N*′,*N*′-tetramethyluronium hexafluorophosphate (HCTU, 0.6 M) in dimethylformamide (DMF) in the presence of 12 eq. *N*,*N*-diisopropylethylamine (DIPEA, 2.4 M) in *N*-Methyl-2-pyrrolidone (NMP) for

50 min. Fmoc-protecting groups were removed with a solution of 40% piperidine in DMF for 5 min and the deprotection step was repeated for another 10 min with 20% piperidine in DMF. After each coupling and deprotection step, the resin was washed with NMP. After the addition of the final residue, the resin was washed with NMP, DMF, and dichloromethane (DCM) and dried.

Step b - Esterification and cleavage from the resin. For the esterification of serine with glutamic acid, resin was added with 20 eq. Fmoc-Glu(OtBu)-OH, 20 eq. benzoyl chloride (BzCl), 40 eq. triethylamine (Et_3N), and 0.4 eq. 4-(Dimethylamino)pyridine (DMAP) in DCM (3 mL) and stirred at room temperature overnight. After reaction, the Fmoc-protecting group was removed by using 40% piperidine in DMF for 5 min and then 20% piperidine in DMF for 10 min at room temperature. The protected branched peptide was cleaved with 20% hexafluoroisopropanol (HFIP) in DCM at room temperature for 1 h.

Step c - Cyclisation. The peptide was cyclised in solution assisted by microwave irradiation (30 min, 25 W, 75 °C) by using 2 eq. O-(7-azabenzotriazol-1-yl)-*N*,*N*,*N*,*N*-tetramethyluronium hexafluorophosphate (HATU), 2 eq. 1-hydroxy-7-azabenzotriazole (HOAt), and 4 eq. DIPEA in DMF (c = 1 mM).

Step d - Deprotection. The cyclized product was fully deprotected by incubation with 95% trifluoroacetic acid (TFA) and 2.5% triisopropylsilane (TIPS) in water at room temperature for 2 h. After evaporation under reduced pressure, the residue was dissolved in ethyl acetate and the solution was washed with saturated $NaHCO_3$. The organic layer was concentrated *in vacuo*. The aforementioned semi-preparative HPLC method was used to purify the peptide to give **1**. The structure of **1** was confirmed by HRMS.

Chemical synthesis of 2

Step a - Loading of first amino acid on the 2-CTC resin. A solution of Fmoc-Glu(O*t*Bu)-OH (255 mg, 0.6 mmol, 3 eq.) and DIPEA (306 µL, 1.8 mmol, 9 eq.) in 4 mL dry DCM was placed in a plastic reactor vessel filled with 2-CTC resin (125 mg, 0.2 mmol, 1.0 eq.). The resulting mixture was incubated at room temperature overnight.

The remaining free binding sites were capped upon incubating twice with DCM/MeOH/DIPEA (80:15:5) for 10 min at room temperature. The resin was washed several times with DMF, MeOH, and DCM, and treated with 20% piperidine in DMF to remove the Fmoc-protecting group. The combined filtrates were used to determine the actual loading of the resin at $\lambda_{301\ nm}$. Afterwards, the resin was washed with DCM and dried.

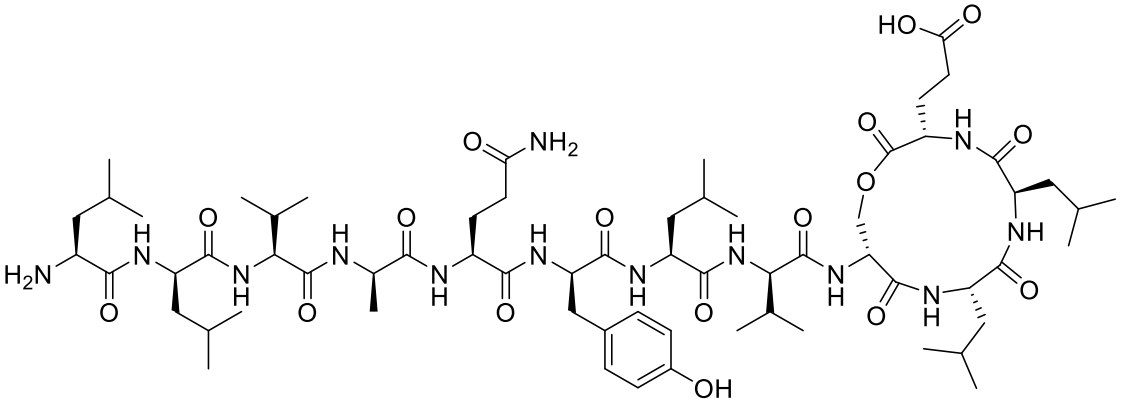
Step b - solid-phase peptide synthesis. The linear sequence was synthesised on the preloaded Glu (O*t*Bu)-2-CTC resin on a 25 µmol scale with a Syro Wave peptide synthesizer by using standard Fmoc chemistry. The resin was placed in a plastic reactor vessel with a Teflon frit and an amount of 6 eq. of amino acid derivatives (Fmoc-D-Ala-OH, Fmoc-Leu-OH, Fmoc-Gln(Trt)-OH, Fmoc-Val-OH, Fmoc-D-Leu-OH, Fmoc-D-Ser-OH, Fmoc-D-Val-OH, Fmoc-D-Tyr(O*t*Bu)-OH, 0.2 M) were activated in situ at room temperature with 6 eq. of HCTU (0.6 M) in DMF in the presence of 12 eq. DIPEA (2.4 M) in NMP for 50 min. Fmoc-protecting groups were removed with a solution of 40% piperidine in DMF for 5 min and the deprotection step was repeated for another 10 min with 20% piperidine in DMF. After each coupling and deprotection step, the resin was washed with NMP. After the addition of the final residue, the resin was washed with NMP, DMF, and DCM and dried.

Step c - cleavage of peptide from the resin. A total of 1 mL 95% TFA and 2.5% TIPS in water were added to the peptidyl resin (25 µmol) and the mixture was agitated for at least 2 h at room temperature. The resin was removed by filtration and washed twice with TFA. The solution was concentrated *in vacuo*. The peptide was purified by Agilent HPLC system to give **2**. The structure of **2** was confirmed by HRMS.

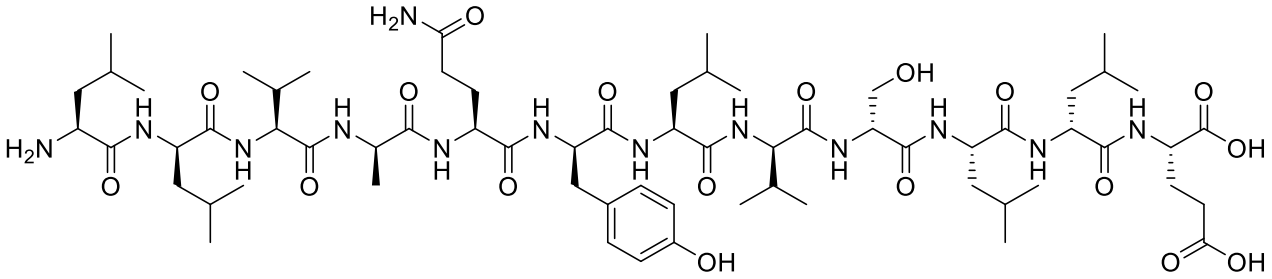
Antimicrobial susceptibility testing

Antibacterial and antifungal activities were investigated for **1, 2** isolated herein and **3,** which we[106] previously isolated. Antibacterial activity was assessed against the reference bacterium[207] *Micrococcus luteus* for **1**,**2** and **3**. As **3** was highly inhibitory against *M. luteus*, it was further tested against a Gram-negative strain, *Esherichia coli*. All tests were via broth microdilution method based on the CLSI standard[208] of the following specifications: Rosewell Park Memorial Institute (RPMI) medium

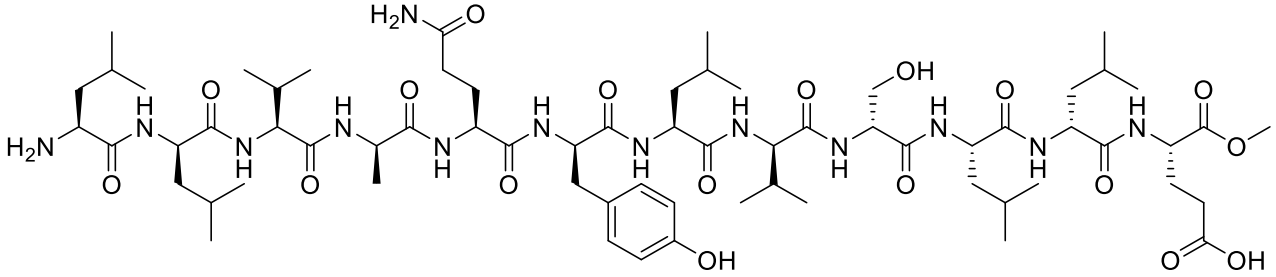
supplemented with Luria-Bertani(LB) medium (10%) (v/v)[209] in round-bottom 96-well plates (100 µL per well), a dilution range of 100 µg/mL to 0.0488 µg/mL of test compounds, inoculum concentration per well of $1.2*10^5$ CFU/mL with incubation at 30 °C for 21 h. Positive control used was Ampicillin (K029.3, Carl Roth GmbH). Antifungal testing for **1**,**2** and **3** was done against a representative yeast fungus (*Candida lusitaniae*) with the broth microdilution method based on the EUCAST antifungal broth microdilution standard[210] with the following specifications: medium as Yeast-Potato-Dextrose (YPD) (100 µL per well), 96-well round-bottom plates, a dilution range of 100 µg/mL to 0.0488 µg/mL of test compounds, inoculum concentration per well of $0.45*10^5$ CFU/mL with incubation at 30 °C for 48 h. Positive control used was Amphotericin B (A1907, Biochemica). For both assays, media with microorganism and excipient only (1% (v/v) Dimethyl Sulfoxide-DMSO) and media with the least compound concentration (0.0488 µg/mL) but no microorganism was included as controls for microbial growth and media sterility, respectively. Experiments were done in duplicate with three reproductions.



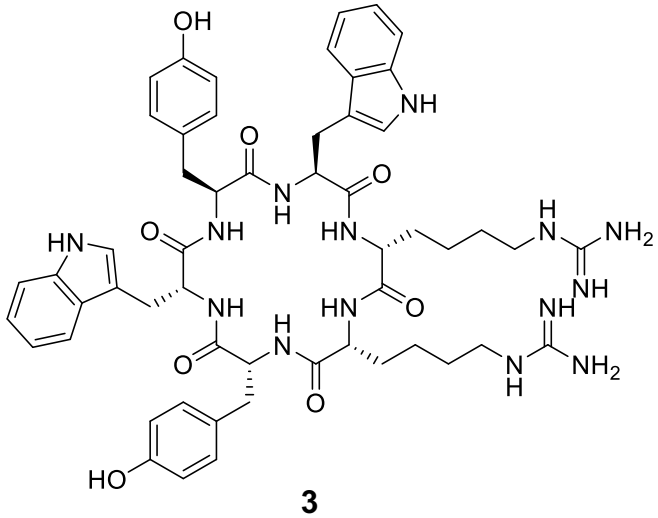
1



2



2-Me



3

RESULTS

Seventy-six soil samples were collected from points, all within a tropical climate and located between latitudes 0.882906 and -0.21363, longitudes 34.3763 and 35.31924, and altitudes 1117m and 2744m. Twenty-four samples (Table S1) were investigated for EPNs and they were from riverine (6/24), fallow land (4/24), cultivated land (5/24), forests (2/24), tropical rain forests (1/24) and lake shorelines (4/24). EPNs were isolated from samples VH1 and BG5 (Figure 4). VH1 was located at 0.06293, 34.72903 and altitude 1624m and was red volcanic soil on cabbage-cultivated land (Figure S1). BG5 was located at 0.48044, 34.40836 and altitude 1239m and was clay soil on a crop edge in riverine land (Figure S2). Bacteria were isolated from these two EPNs. Then, using morphological characteristics of being pigmented umbonate colonies with irregular margins and visible swarming patterns coupled with them being catalase-negative, they were designated genera as follows: *Xenorhabdus* sp. strain BG5 and *Xenorhabdus* sp. strain VH1. Assembled genomes of *Xenorhabdus* sp. strain BG5, *Xenorhabdus* sp. strain VH1 and *X. griffiniae* XN45 all had coverage >50x and N50 values of 102,633 bp, 29,298 bp, 45,296 bp respectively (Table 2). Genomes of *Xenorhabdus* sp. BG5, *Xenorhabdus* VH1, *X. griffiniae* XN45 and *X. griffiniae* BMMCB were all 100% complete with estimated contamination below 0.7% when calculated using the *X. bovienii* SS-2004 complete genome as a reference strain. Annotated genomes for the four had 3827, 4232, 4160, 4318 coding-DNA sequences (CDS) respectively and >30% in each encoded hypothetical proteins (Table 3).

Phlyogenomically, the most closely related species to *Xenorhabdus* sp. strain VH1 and *Xenorhabdus* sp. strain BG5 was *X. griffiniae* XN45 at 0.0 and 0.059 genome-genome distances, respectively. *X. griffiniae* BMMCB did not cluster with *X. griffiniae* XN45 and was most closely related to *Xenorhabdus* sp. strain BG5. The other most closely related species to these three were type strains of *X. elhersii, X. ishibashii, X eapokensis* and *X. thuongxuanensis* (Figure 5). *Xenorhabdus* sp. strain VH1 and *X. griffiniae* XN45 had, between them, orthoANI, dDDH, and GGD of 99%, 99.0% and 0 respectively. *Xenorhabdus* sp. strain BG5 and *X. griffiniae* XN45 had, between them, orthoANI, dDDH and GGD values of 94.24%, 67.3% and 0.059, respectively. They

were thus designated species as follows: *X. griffiniae* VH1 and *Xenorhabdus* sp. nov. BG5. *X. griffiniae* BMMCB and *Xenorhabdus* sp. nov. BG5 had between them, orthoANI, dDDH and GGD values of 92.25%, 57.1% and 0.08, respectively (Table 5). It was emended to *Xenorhabdus* sp. nov. BMMCB. *Steinernema* sp. scarpo, the nematode host of *X. griffiniae* XN45 was demonstrated as a probable new species due to 93.2% nucleotide similarity of the internally transcribed spacer (ITS) region sequences to its mostly closely related species (Table 4). The pangenome of 25 species of *Xenorhabdus* had a total of 12,306 protein encoding genes, 1217 of which were present in all species (Figure 6). It was an open pangenome with its mean α exponent of Heap's Law= 0.2735. For *Photorhabdus* its pangenome of 13 species had 9691 protein coding genes 1933 of which were core genes. It was also an open pangenome with its mean α exponent of Heap's Law= 0.3967. Phylogenomic reconstruction of the *Photorhabdus* genus revealed low species diversity.

The pangenome containing *X. griffiniae* XN45, *X. griffiniae* VH1 and *Xenorhabdus* sp. nov. BG5 had 4057 protein coding genes, 767 and 502 of which were *X. griffiniae* and *Xenorhabdus* sp. nov. B5 specific, respectively (Figure 9). When compared to *X. griffiniae* genomes, the *Xenorhabdus* sp. nov. BG5 had 16 IS-transposase associated genomic islands (Figure 10) that had GC content that was significantly different from the rest of the genome (Table 6). Seven prophage loci were also identified that had significantly different GC content (Table S3). Within IS-transposase associated genomic islands 62% and 23% identified CDS encoded unknown proteins and transposases, respectively (Figure 11). Single counts of genes encoding the following products made up the rest: anguibactin system regulator, thioredoxin, formalaldehyde forming aminobutanoate oxidase, 5(hydroxymethyl) furfural oxidase, phosphoethanolamine transferase, aspartate carbomolytransferase, ornithine monooxygenase, LgrE dehydrogenase, tryptophan decarboxylase, mersacidin decarboxylase, Methionine tRNA ligase and Fis DNA binding protein.

The *ishAB* BGC in *X. ishibashii* was 40,537 bp long and encoded an NRPS of 12 modules and 38 domains that included a C_{start} domain and two terminal TE domains. Induced expression of this operon resulted in detection of Ishipeptide A (**1**) Ishipeptide

B(**2**) and **2-Me** (Figure 12) of masses (*m/z* 1342.8 [M+H] +), (*m/z* 1360.8 [M+H] +) and (687.9 $[M+2H]^{2+}$) respectively. **1** and **2** both had 13 *N* and 65 *C* and differed in mass by a single water molecule (Figure 13) indicating **2** as the linearised version of **1**. Confirmed amino acids building blocks by feeding experiments were: 4x Leu, 2x Val, Glu, Gln, Tyr (Figures 14&15). Ser, Leu and Ala were confirmed by $HRMS^2$ spectra analysis. Using $HRMS^2$ spectra analysis of **2**, the differences in masses of successive ions of B_3 to B_{12} were 99.1, 71.02, 128.06, 163.05, 113.08, 99.07, 87.02, 113.09, 113.07, 147.05 (Figure 16) and were established as monoisotopic masses for Val, Ala, Gln, Tyr, Leu, Val, Ser, Leu, Leu, Glu.H_2O respectively. The B_2 ion of 227 was established as the protonated monoisotopic mass for Leu-Leu. This enabled the amino acid sequence of **2** to be determined as: Leu-Leu-Val-Ala-Gln-Tyr-Leu-Val-Ser-Leu-Leu-Glu.H_2O. The following linear peptide was synthesised –LlVaQyLvsLIE– by first binding E to the resin and condensing amino acids in the reverse order (from E to L) using an iterative solid phase synthesis route (Figure 17). This synthetic **2** was of exactly the same mass, $HRMS^2$ spectra and retention times were compared to natural **2** (Figure 18) confirming the structure and absolute configuration of natural **2**. Macrocyclisation of **2** was proposed to create **1**. A structure macrolactonated between *β*-hydroxy of D-Ser^9 and terminal carboxy had hypothetical spectral fragments that matched those of **1** with the highest probability (P1:71.2; P2:31.9) when calculated using the iSNAP algorithm. Thus, this structure was synthesised as follows: D-Leu^{11} was bound to the resin overnight. Then, the following amino acids which all had protected side chains, apart from D-Ser^9 were condensed onto a growing peptide chain in the following order using iterative sold phase peptide synthesis: L-Leu^{10} D-Ser^9 D-Val^8 L-Leu^7 D-Tyr^6 L-Gln^5 D-Ala^4 L-Val^3 D-Leu^2 L-Leu^1 with L-Leu^1 having its *C*-terminal protected. The unprotected carboxy of an L-Glu^{12} with a protected side chain, was then esterified to the unprotected *β*-hydroxy of D-Ser^9. D-Leu^{11} was cleaved from resin and a late stage macrolactamation between its carboxy and the unprotected amine group of L-Glu^{12} resulted in macrocyclisation (Figure 19). All side chains were then deprotected resulting in synthetic **1**. Masses, $HRMS^2$ spectra and retention times of synthetic **1** and natural **1** were exactly the same confirming the structure and absolute configuration (Figure 20). Neutral losses for amino acid building blocks of **1** were identified in its $HRMS^2$ spectra

(Figure 21). Eluting samples with Acetonitrile instead of methanol eliminated **2-Me** (Figure 23). All amino acid building blocks of **2** were identified in $HRMS^2$ spectra of **2-Me** (Figure 24). The amino acid of position 143 of IshA protein was mutated from Q to H. This did not result in generation of any new derivatives (Figure S5). Biosynthesis of Ishipeptides was proposed (Figure 22). The following were revealed as new Stachelhaus codes for Gln, Glu and Leu respectively: DasnIGEVGK, DAvDLGVVDK and DAWlLGaVcK (Table S5). Photoditritide(**3**) from *P. temperata,* **1** and **2** were all non-inhibitory to *C. lusitaniae*. Peptides **3** and **1** inhibited *M. luteus* at 37 µM and 3 µM respectively while Ampicillin inhibited it at 4.2 µM. **3** was further inhibitory to *E. coli* at 24 µM (Table 7). Proposed helical wheel of Ishipeptides demonstrated a large hydrophobic face (Figure S6).

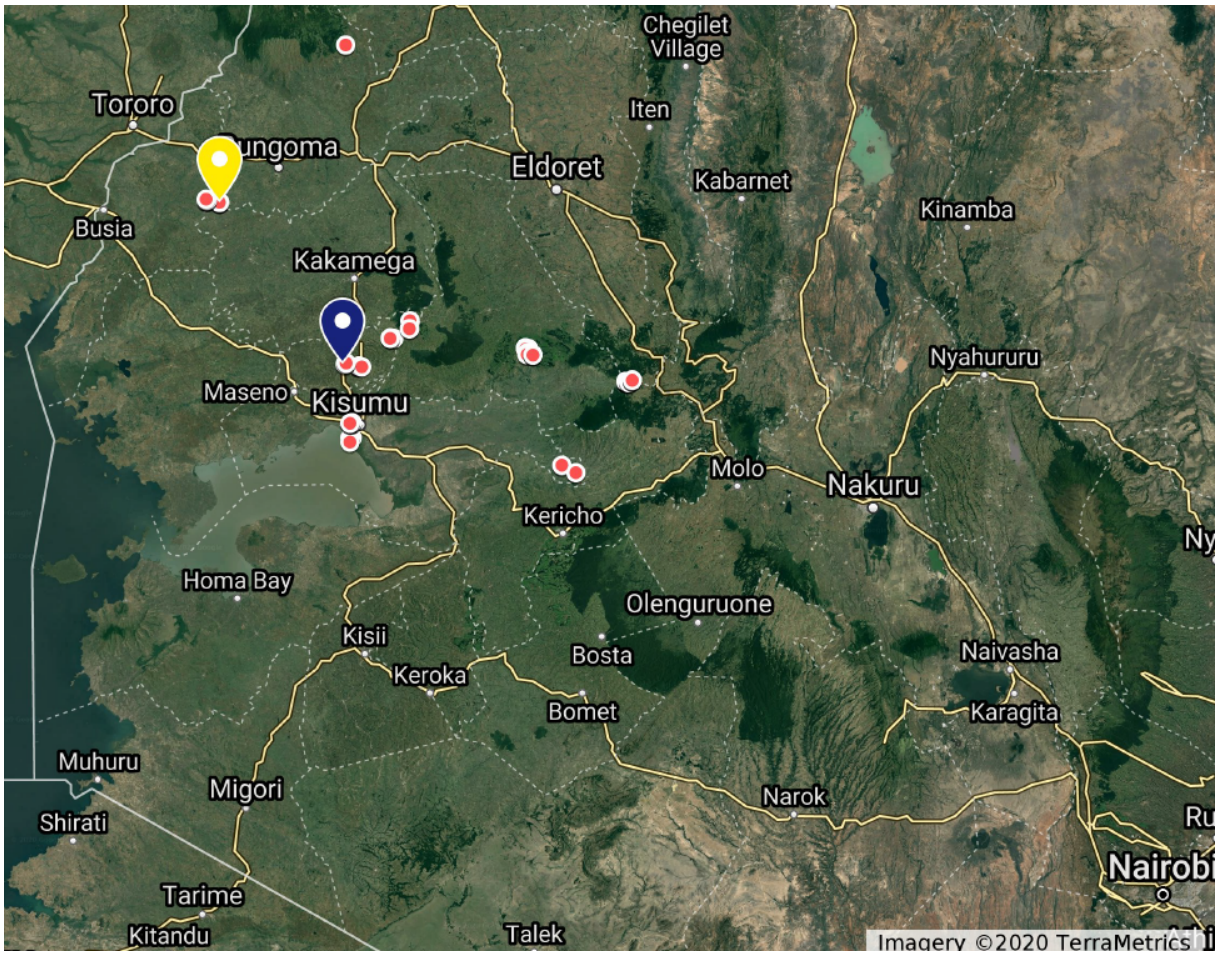


Figure 4. Map of geographical regions of soil collection points used to isolate entomopathogenic nematodes. Orange circles represent an area with collection points. Soil was collected from a total of 76 points. The two soil collection points from which nematodes and their respective *Xenorhabdus* bacteria were successfully isolated are in yellow (strain BG5) and blue (strain VH1). Map was created with Google MyMaps.

Table 2. Quality and characteristics of genome assemblies.

	Xenorhabdus sp. strain BG5	*X. griffiniae* XN45	*Xenorhabdus* sp. strain VH1	*Xenorhabdus* sp. strain BMMCB
Contigs	129	381	273	231
GC content	43.80	43.57	43.65	44.68
Contig L50	12	29	43	21
Contig N50	102,633 bp	45,296 bp	29,298 bp	57,901 bp
Genome length	3,933,551 bp	4,215,754 bp	4,224,998 bp	4,183,760 bp
Fine consistency	96.5%	95.9%	95.9%	95.7%
Coarse consistency	97.0%	96.7%	96.7%	96.7%
Contamination	0.7%	0%	0%	0%
Completeness	100%	100%	100%	100%

Table 3. Characteristics of RAST-k annotated genomes.

	Xenorhabdus sp. nov. BG5	*X. griffiniae* XN45	*X. griffiniae* VH1	*Xenorhabdus sp.* strain BMMCB
CDS	3,827	4,232	4,160	4318
Repeat Regions	127	66	69	70
tRNA	75	58	57	8
rRNA	8	5	10	7
Partial CDS	0	0	0	0
Hypothetical proteins	973	1,175	1,185	1,193
Protein with functional assignments	2,854	3,057	2,975	3,125

Table 4. Strains with the highest percentage identity scores to a query of ITS1-*5.8S rRNA*-ITS2 nucleotide sequence of *Steinernema* sp. scarpo(869 bp).

Strain name	Locality	Accession number	Percentage identity	E-value	total score
Steinernema sp. UH36	Kenya	AY230186.1	93.217	0	1308
Steinernema sp. 10 APR-2003	Sri Lanka	AY230184.1	90.263	0	1149
S. karii type strain	Kenya	AY230173.1	84.256	0	821
Steinernema sp. VP-2015a strain MW8A	Tanzania	KT358811.1	83.934	0	798
Steinernema sp. PT-2012 strain S3 18S	Tanzania	JQ687354.1	85.658	0	754
Steinernema sp. VK-2013 strain EPN15	India	KC252604.1	86.724	0	739
S. hermaphroditum isolate CS34	India	MF663703.1	86.705	0	737
S. hermaphroditum isolate S0905	India	MH802516.1	86.416	0	726
S. guangdongense strain GDc339	China	AY170341.1	84.735	0	726
S. taiwanensis strain T39	Taiwan	KX853101.1	87.048	0	719
Steinernema sp. WS9	South Africa	KP325086.1	85.073	0	710
Steinernema sp. PQ12	Vietnam	MF161467.1	86.617	0	702

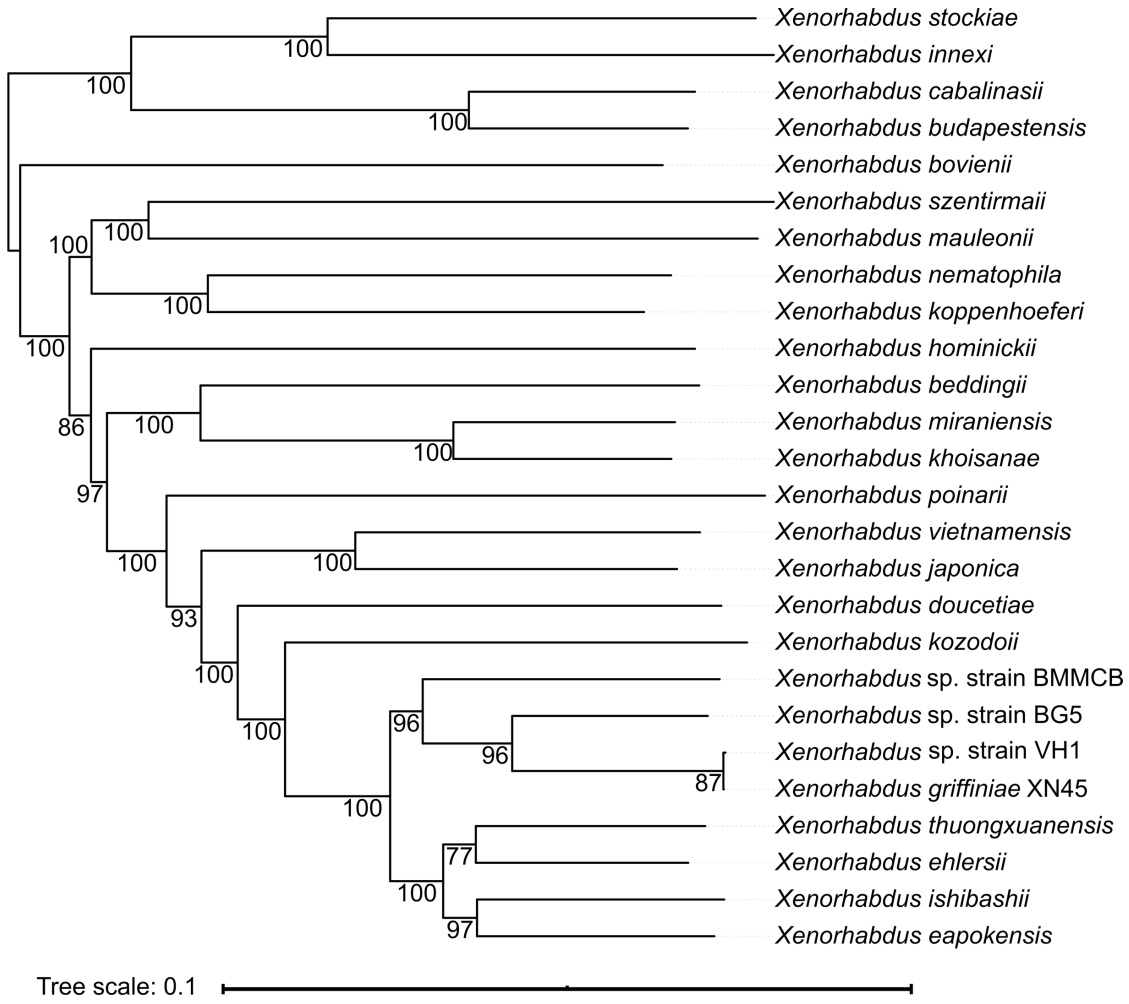


Figure 5. Phylogenomic reconstruction of 26 *Xenorhabdus* species using genome blast distance phylogeny approach (GBDP) distances calculated from genome sequences using the d_5 distance formula. Genome sequences of *Xenorhabdus* sp. strains VH1 & BG5 and *X. griffiniae* XN45 were obtained in this study. All other genomes are of type strains of respective species. There was no distance between *Xenorhabdus* sp. strain VH1 and *X. griffiniae* XN45, suggesting it as an *X. griffiniae* species. *Xenorhabdus* sp. strain BMMCB was previously classified as an *X. griffiniae* species. However, it did not cluster with *X. griffiniae* XN45 indicating that it is not an *X. griffiniae* species. *Xenorhabdus* sp. strain BG5 did not cluster together with any species suggesting it as a novel species. GBDP pseudo-bootstrap values above 60% are shown above branches.

Table 5. Orthologous average nucleotide identities(orthoANI), genome-to-genome distances (GGD) and digital DNA-DNA hybridization (in %) (dDDH) values for species most closely related to *Xenorhabdus* sp. strains VH1& BG5, *X. griffiniae* XN45 and *Xenorhabdus* sp. strain BMMCB. OrthoANI are in the top half, GGD in brackets and dDDH in the bottom half. Values that are within the threshold for two strains to be classified as one species are shaded in grey. The threshold values for species delineation are orthoANI values above 95.1%, dDDH values above 70% or GGD values below 0.0361. Based on these *Xenorhabdus* sp. strain VH1 was classified as *X. griffiniae* VH1, *Xenorhabdus* sp. strain BG5 as *Xenorhabdus* sp. nov. BG5 and *Xenorhabdus* sp. strain BMMCB as *Xenorhabdus* sp. nov. BMMCB. Type species are DSM 2270=*X. ishibashii,* DL20=*X. eapokensis*, DSM 16337=*X. ehlersii*, 30TX1= X. *thuongxuanensis.*

	DSM 22670	XN45	DL20	DSM 16337	VH1	BG5	30TX1	BM MCB
BM MCB	89.68 (0.103)	91.41 (0.088)	90.03 (0.010)	91.95 (0.082)	91.42 (0.088)	92.25 (0.080)	90.56 (0.094)	-
30TX1	92.55 (0.076)	91.02 (0.091)	93.42 (0.067)	93.76 (0.064)	91.01 (0.091)	91.72 (0.085)	-	49.6
BG5	90.44 (0.097)	94.24 (0.059)	90.94 (0.092)	93.08 (0.071)	94.22 (0.059)	-	57.3	57.1
VH1	89.80 (0.102)	99.99 (0.000)	90.32 (0.097)	91.95 (0.081)	-	67.4	51.1	50.5
DSM 16337	92.11 (0.080)	92.05 (0.081)	92.55 (0.077)	-	59.9	66.7	53.50	56.4
DL20	93.23 (0.069)	90.34 (0.097)	-	47.90	51.6	57.5	51.90	50.9
XN45	89.92 (0.102)	-	51.7	59.9	99.9	67.3	51.1	50.4
DSM 22670	-	50.3	51.2	46.4	50.3	55.0	48.3	48.9

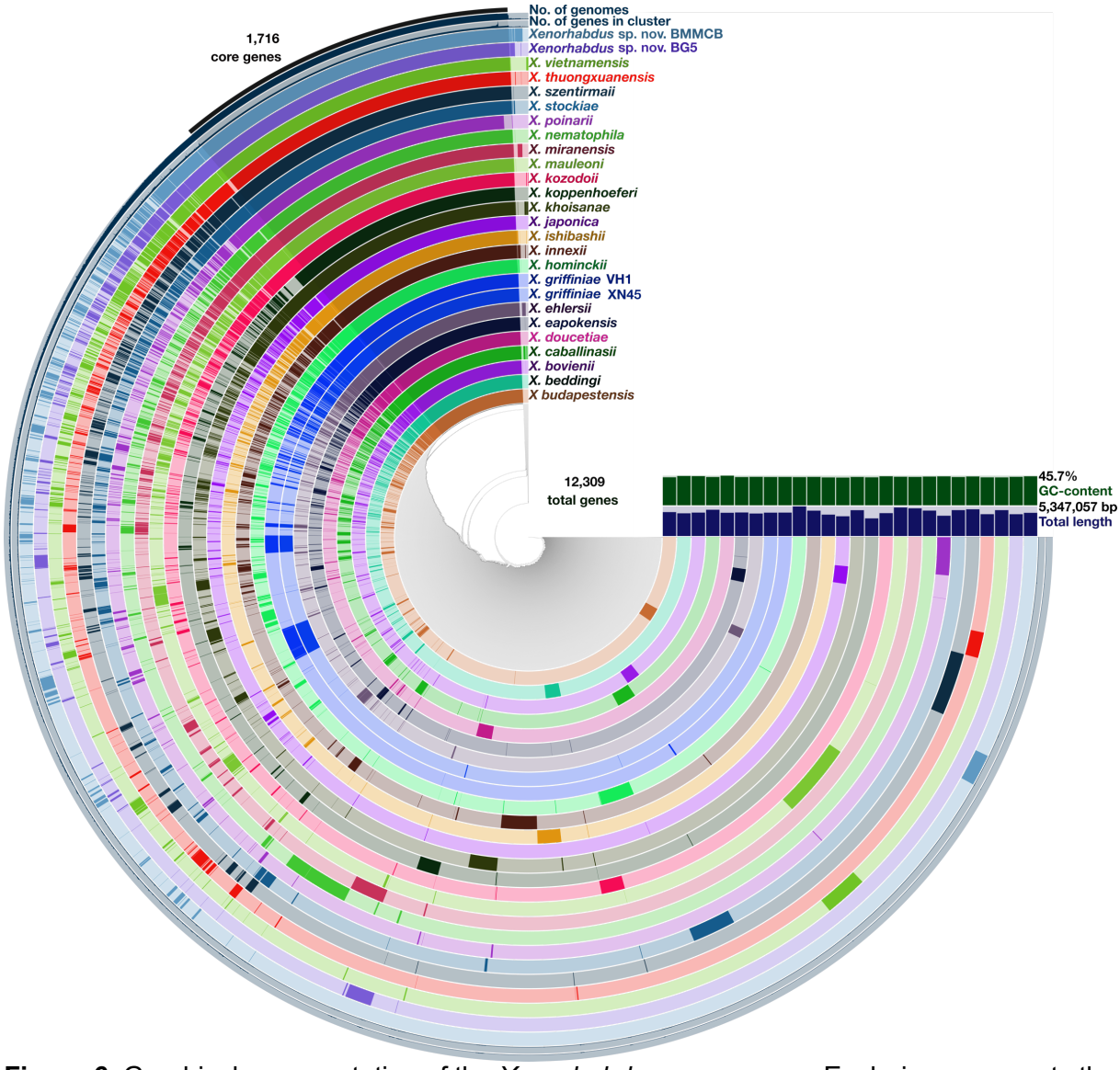


Figure 6. Graphical representation of the *Xenorhabdus* pangenome. Each ring represents the genome of a species and dark regions along a ring indicate a protein encoding gene sequence. Each radial position represents a cluster of orthologous protein sequences. Thus, the core genes are those shared by almost all species and regions towards the end of a ring represents singletons. A total of 1,716 protein encoding genes were found across all species representing 13.9% of the pangenome. GC content did not vary significantly while genome size did, with *X. hominickii* having the largest genome of 5,347,057 bp. Pangenome was created with the anvio workflow.

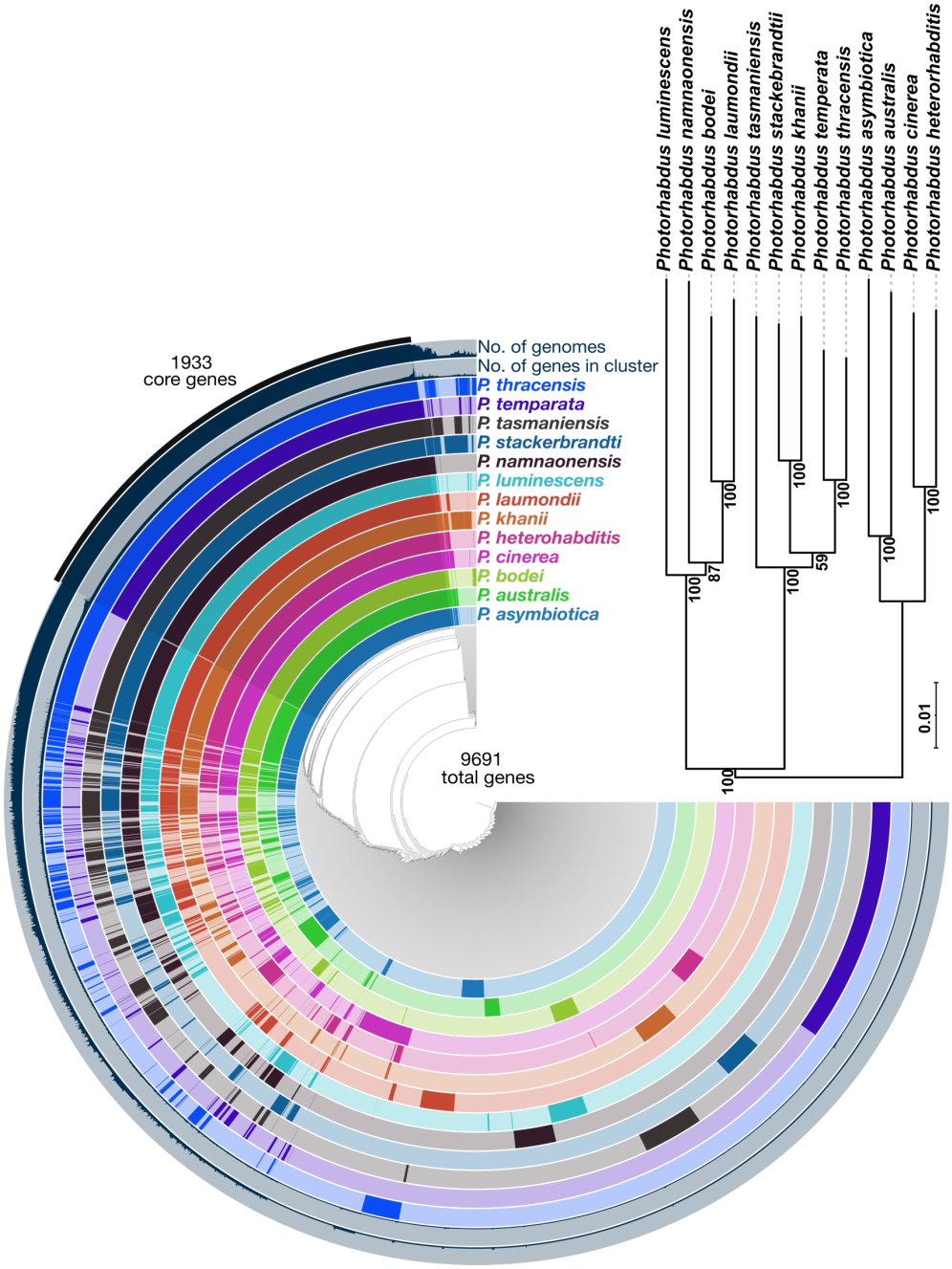


Figure 7. Graphical representation *Photorhabdus* pangenome and phylogenomic reconstruction. The core genome composed 19% (1933 protein encoding genes) of the pangenome (9691 protein encoding genes). Pangenome was created with anvio workflow and phylogenomic tree was reconstructed using genome blast distance phylogeny approach (GBDP) distances calculated from genome sequences using the d_5 distance formula. GBDP pseudo-bootstrap values above 50% are shown above branches.

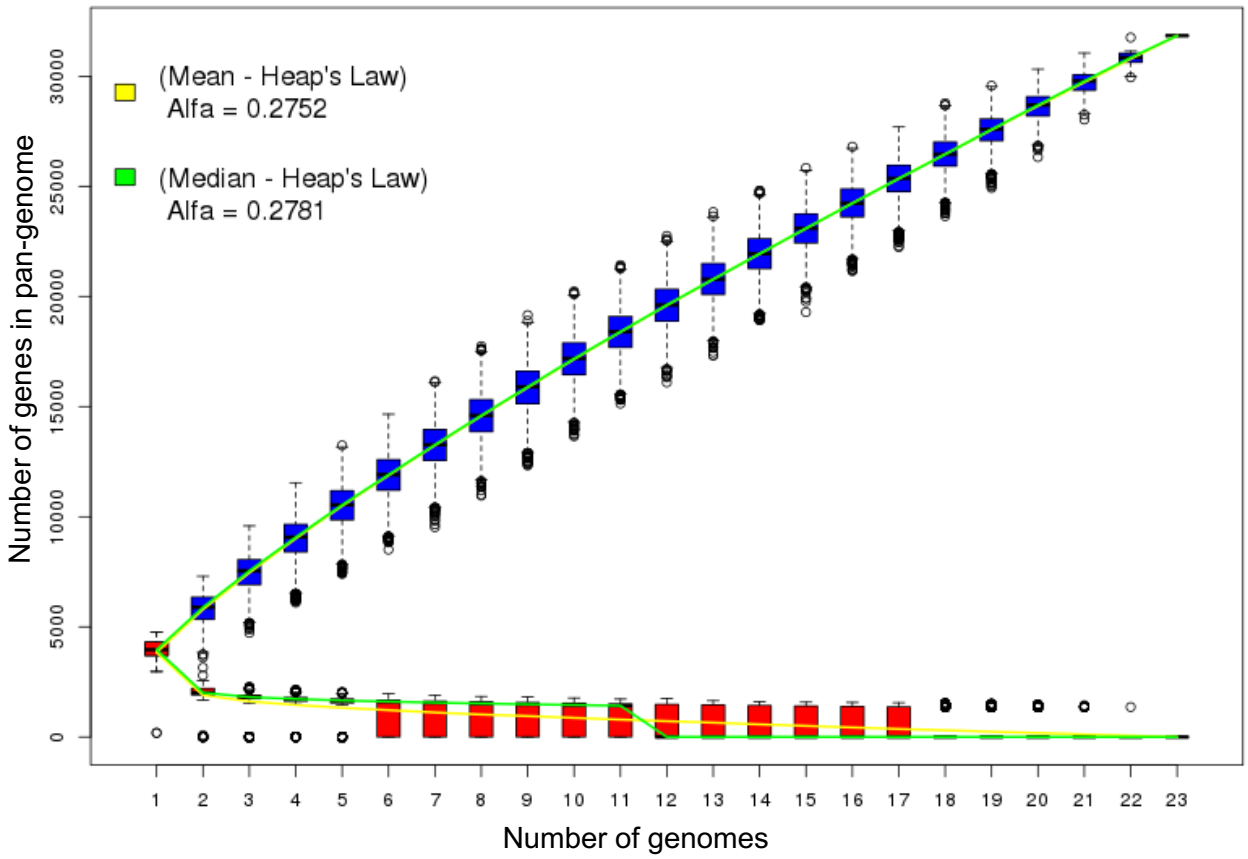


Figure 8. Box-plot graph depicting how α exponent of Heaps law was estimated for the *Xenorhabdus* genus. Red plots and lower line depict the size and trend of the core genomes as more genomes are added to the pangenome. Blue plots and upper line depict the size and trend of the accessory genome as more genomes are added to the pangenome. Pangenome was open because its mean α exponent of Heap's Law <1. This predicts that more species genomes need to be added to this genus so as to complete it. Pangenome analysis was determined by the Panweb workflow.

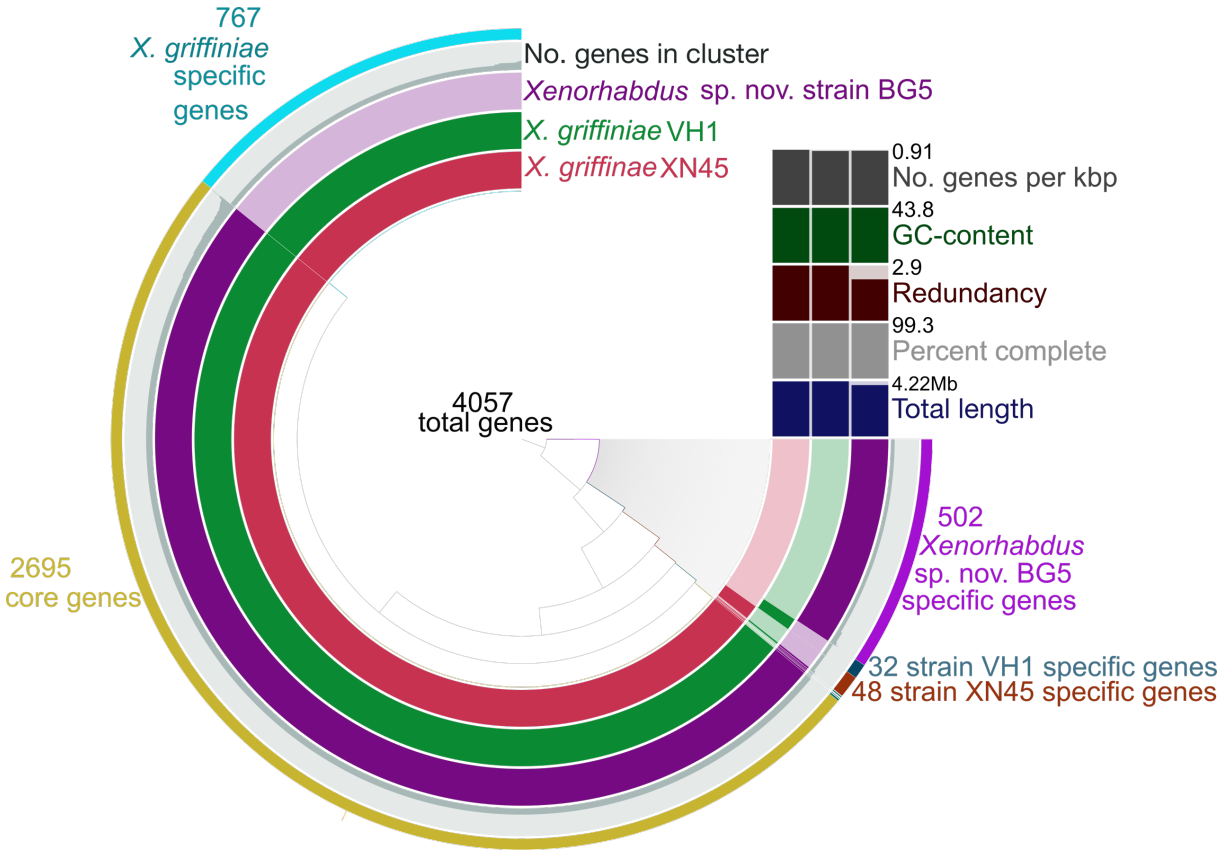


Figure 9. Graphical representation of the *X. griffiniae* clade pan-genome. Core genes were those shared *by X. griffiniae* and *Xenorhabdus* sp. nov. BG5 and were 66.4% of pan-genome. *X. griffiniae* species specific genes were 18.9% and *Xenorhabdus* sp. nov. BG5 species specific genes were 12.3% of the pangenome. This indicates that although they are most closely related, they are two distinct species. Pangenome was created with the anvio workflow.

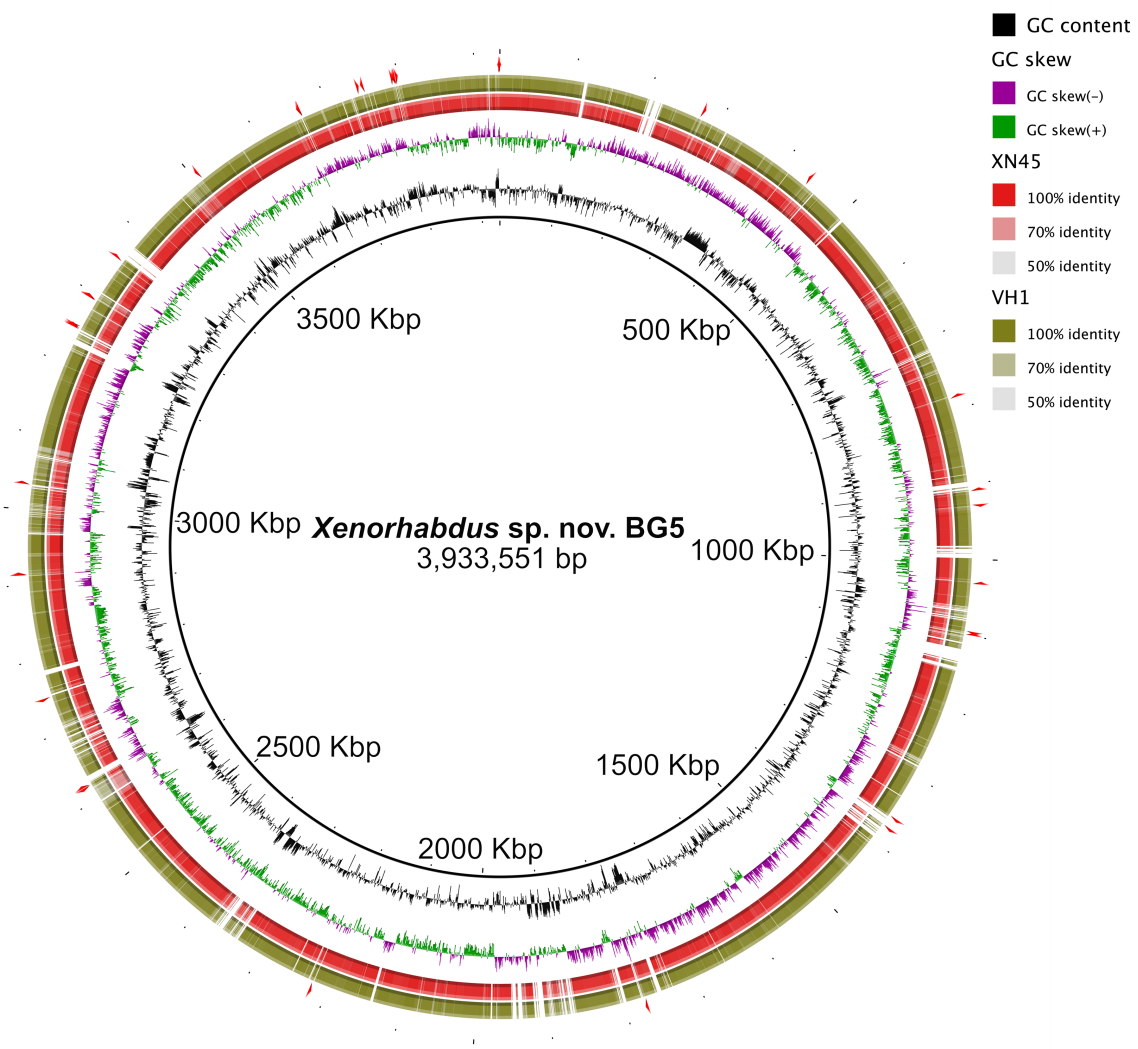


Figure 10. Genome visualisations of *Xenorhabdus* sp. nov. BG5 when compared to *X. griffiniae* XN45 (red) and *X. griffiniae* VH1 (green). Genomic islands of *Xenorhabdus* sp. nov. BG5 are denoted by white breaks in *X. griffiniae* genomes and these contain cognate nucleotide sequences that were less than 50% identical. Red arrows in outermost ring denote loci of IS family transposase genes on the BG5 genome. All transposase genes that were not on contig edges flanked genomic islands indicating that they played a role in acquisition of nucleotide sequences specific to *Xenorhabdus* sp. nov. BG5. Visualisation was created in Blast Ring Image Generator (BRIG).

Table 6. IS family transposase-associated genomic islands of *Xenorhabdus* sp. nov. BG5

Name of transposase associated Genomic Islands	Length	GC content	Comparison to average genome GC content
GI_1	1,204	42.7	Significantly different from average GC content (p value= 1.21163E-06, T_{stat} = -4.738539315)
GI1.2	1,835	39.4	Significantly different from average GC content (p value= 4.0888E-74, T_{stat} =-19.05746349)
GI_2	15,414	43.9	Significantly different from average GC content (p value < 0.0001, T_{stat} =7.93847263)
GI_3	3,705	42.5	Significantly different from average GC content (p value= 1.59355E-06, T_{stat} = -4.665642783)
GI_4	4,230	31.76	Significantly different from average GC content (p value= 0.00E+00, T_{stat} = -57.92244733)
GI_6	10,052	42.81	Significantly different from average GC content (p value= 5.02543E-11, T_{stat} = -6.473117519)
GI_7	2,219	43.38	Not significantly different from average GC content (p value= 0.9881, T_{stat} = -0.015)
GI_8.1	13,626	43.2	Not significantly different from average GC content (p value= 0.056, T_{stat} = -1.913)
GI_8.2	3,416	41.2	Significantly different from average GC content (p value < 0.0001, T_{stat} = -17.574)
GI_8.3	1,086	29.2	Significantly different from average GC content (p value < 0.0001, T_{stat} = -63.871)
GI_9	749	37.2	Significantly different from average GC content (p value < 0.0001, T_{stat} = -18.466)
GI_10	2,418	33.2	Significantly different from average GC content (p value < 0.0001, T_{stat} = -51.345)
GI_11	2,004	43.1	Significantly different from average GC content (p value < 0.0001, T_{stat} = -51.345)
GI_12	935	41.6	Significantly different from average GC content (p value < 0.0001, T_{stat} = 7.938)
GI_14	4,246	44.6	Significantly different from average GC content

			(p value < 0.0001, T_{stat} = 10.177)
GI_15	4,872	37.3	Significantly different from average GC content (p value < 0.0001, T_{stat} = -49.604)
GI_16	1,764	38	Significantly different from average GC content (p value < 0.0001, T_{stat} = -18.131)
GI_18	1,982	41.4	Significantly different from average GC content (p value < 0.0001, T_{stat} = -11.437)

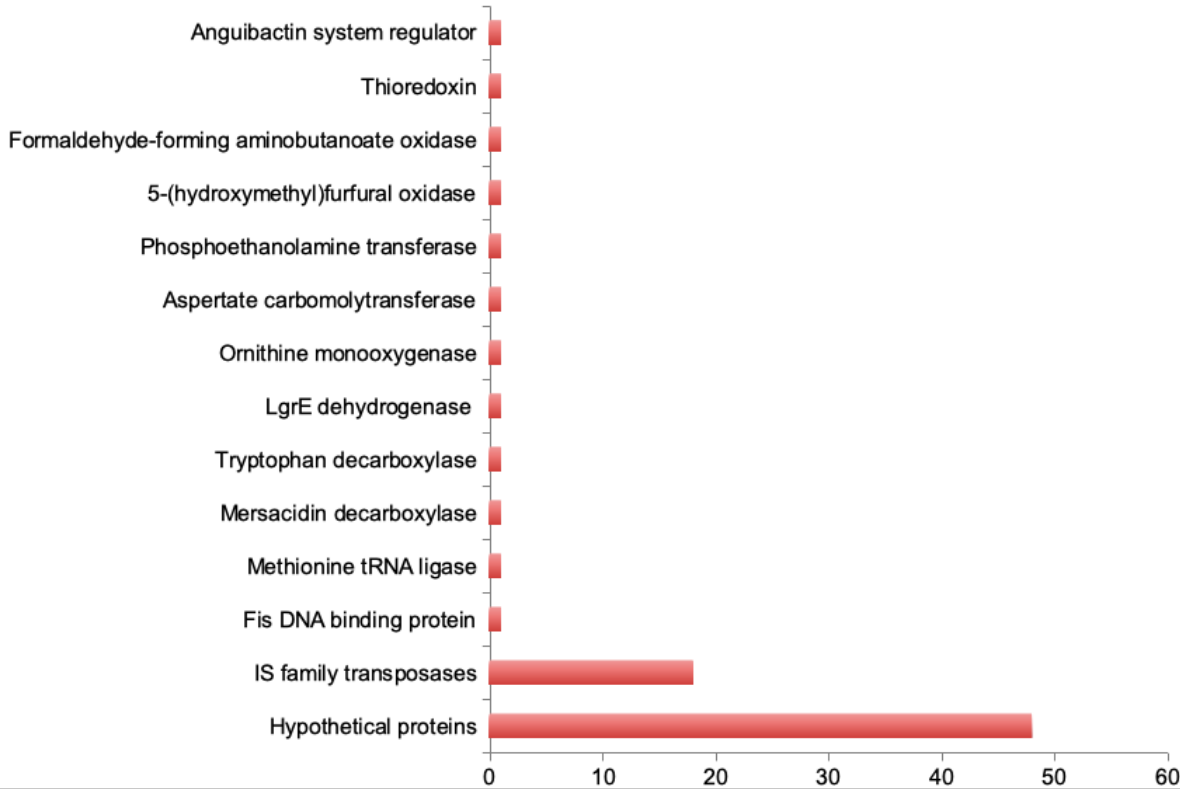


Figure 11. Bar chart of predicted products encoded by coding DNA sequences found in genomic islands associated with transposases genes in the *Xenorhabdus* sp. nov. BG5 genome. Islands were elucidated by genome comparisons with *X. griffiniae,* the most-closely related species to *Xenorhabdus* sp. nov. BG5. Genes encoding IS family transposases were 23% and those encoding unknown proteins were 62%. This indicated unknown protein products were the main group of proteins gained by putative transposase mediated horizontal gene transfer. BG5 gene prediction and annotation of genomic islands was done in PROKKA.

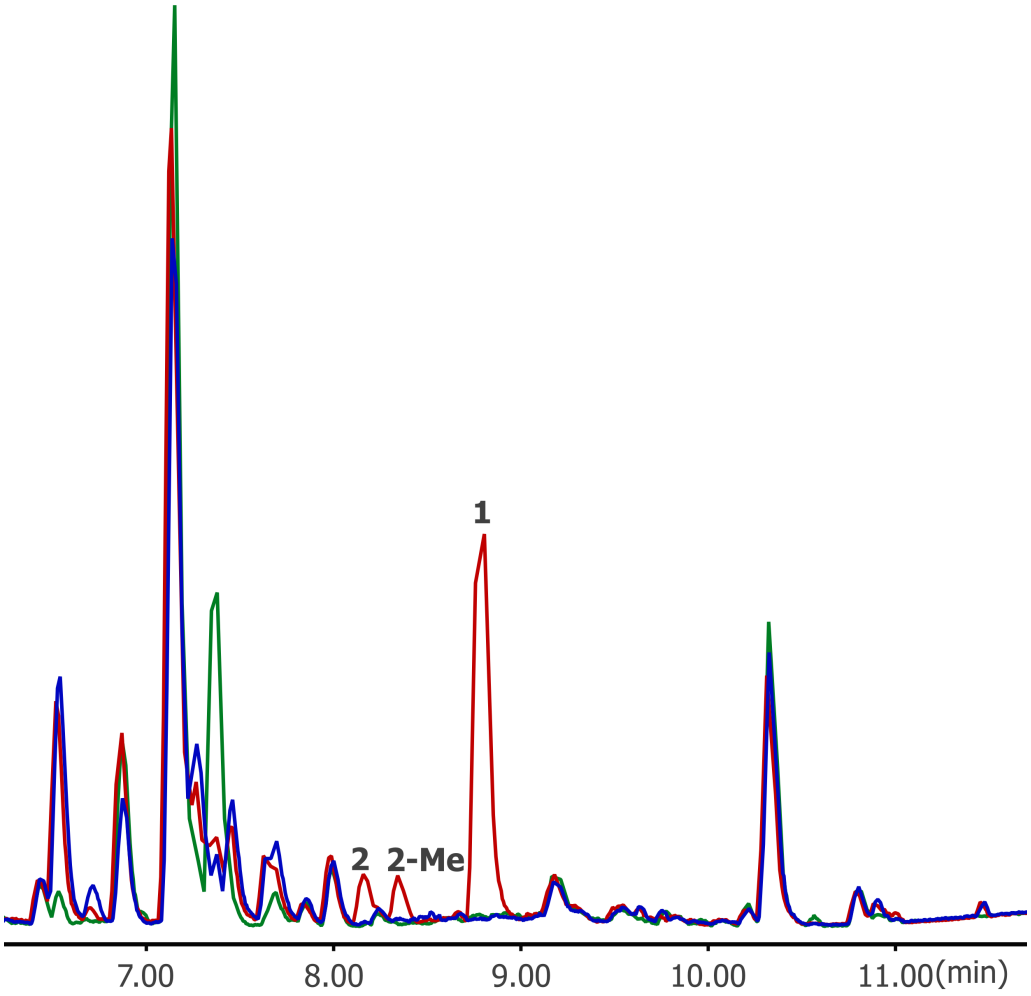


Figure 12. Base peak chromatograms (BPC) depicting two natural products **1** and **2** that resulted from induced expression of *ishAB* NRPS gene cluster. To determine the natural product, whose synthesis is encoded by the *ishA-B* operon in *X. ishibashii*, its native promoter was refactored by exchanging it with a P_{BAD} promoter. BPC were obtained from HPLC-ESI-HRMS analysis of methanol extracts of cell free supernatants from 72 h cultivated *X. ishibashii* pCEP_JC0094 mutant strain that had the native promoter for *ishA* gene refactored by exchanging it with a P_{BAD} promoter. To induce expression, the fermentation culture was supplemented with 0.2% L-arabinose (w/v) (red) resulting in detection of above products that were equally absent in extracts from *X. ishibashii* wild-type strain (green) and control strain (blue).

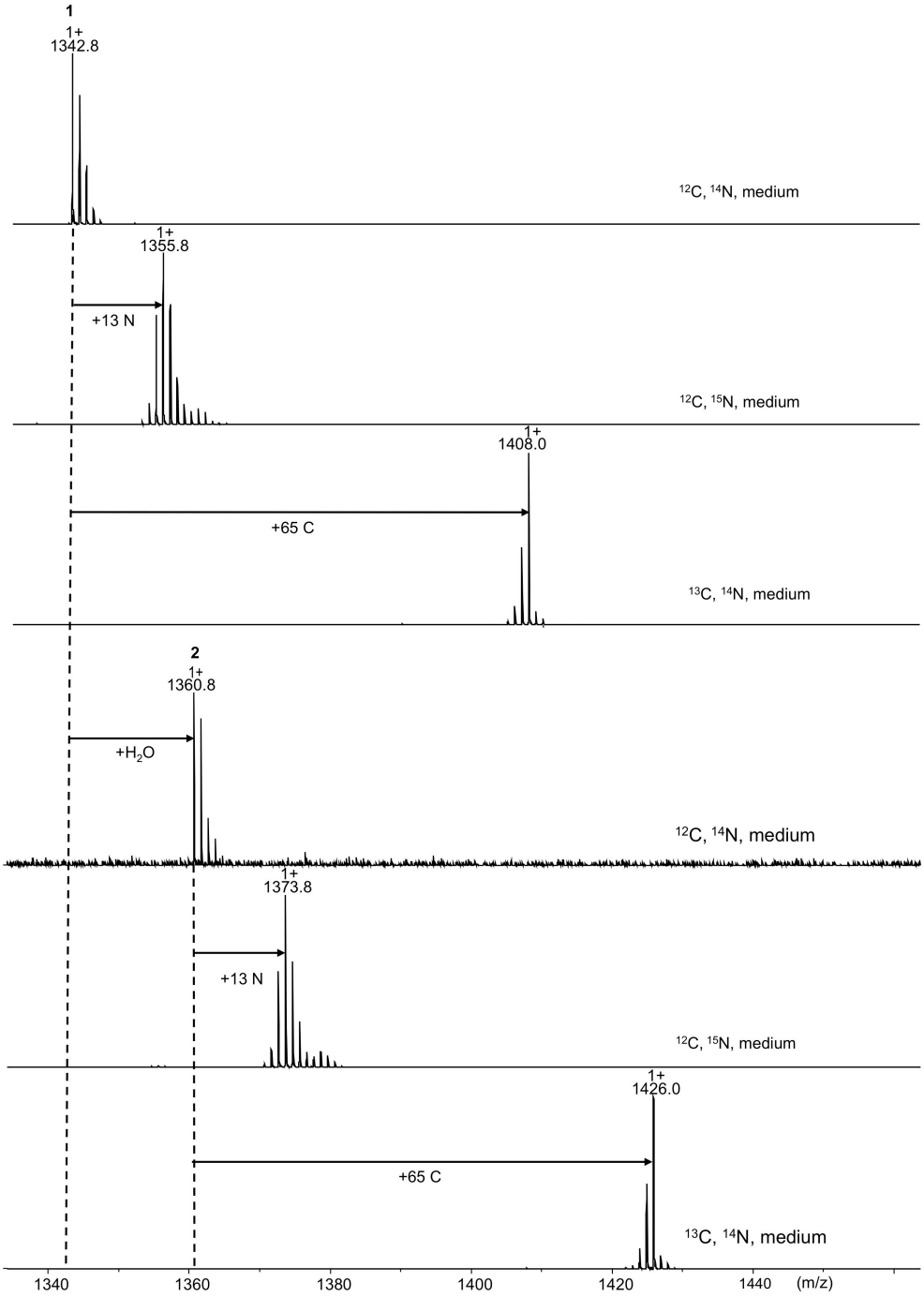


Figure 13. HPLC-ESI-HRMS obtained Base Peak Chromatograms (BPC) depicting Carbon and Nitrogen labelling of **1** (1342.8 $[M+H]^+$) and **2** (1360.8 $[M+H]^+$) to confirm their carbon and nitrogen atoms. Indicated mass shifts were determined by comparison of methanol extracts of **1** and **2** that were fermented from standard medium (^{12}C, ^{14}N) fully labelled ^{15}N medium (^{12}C, ^{15}N) and ^{13}C medium (^{13}C, ^{14}N). **1** and **2** differed only by a mass of 18 Da corresponding to a water molecule.

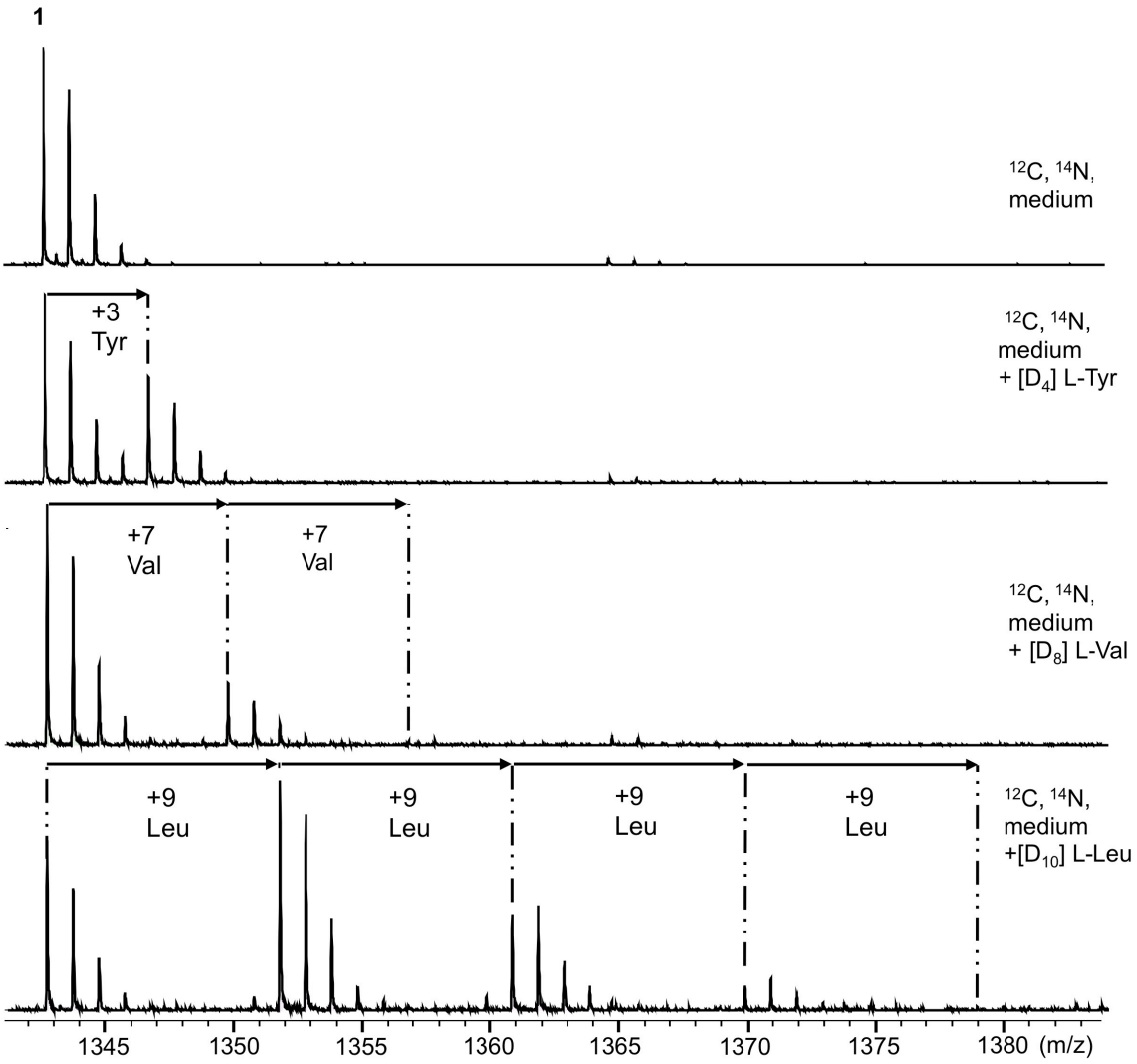


Figure 14. HPLC-ESI-HRMS obtained base peak chromatograms (BPC) depicting how amino acid building blocks of **1** were determined via labelling experiments with deuterated amino acids. Mass shifts were obtained by comparison of BPCs of extracts with **1** when fermented from sf-900 media with when fermented from the same media supplemented with 1mM of a deuterated amino acid ([D10] L-Leucine, [D4] L-Tyrosine, [D8] L-Valine)

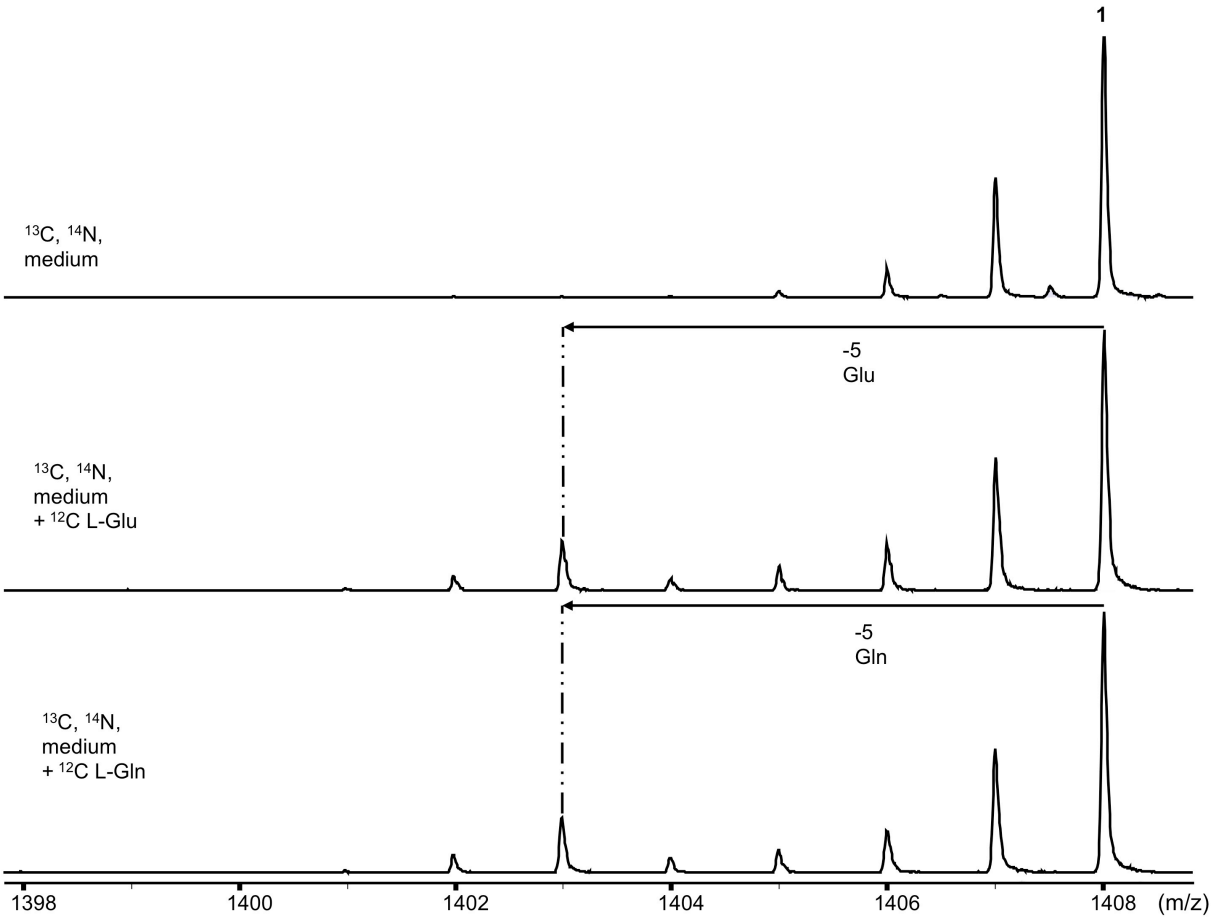


Figure 15. HPLC-ESI-HRMS obtained Base peak chromatograms (BPC) depicting how amino acid building blocks of **1** were determined via inverse labelling experiments. Mass shifts were obtained by comparison of BPCs of extracts with **1** when fermented from fully labelled ^{13}C medium to when fermented from fully labelled ^{13}C medium supplemented with 1mM ^{12}C L - amino acid (L-Glu, L-Gln).

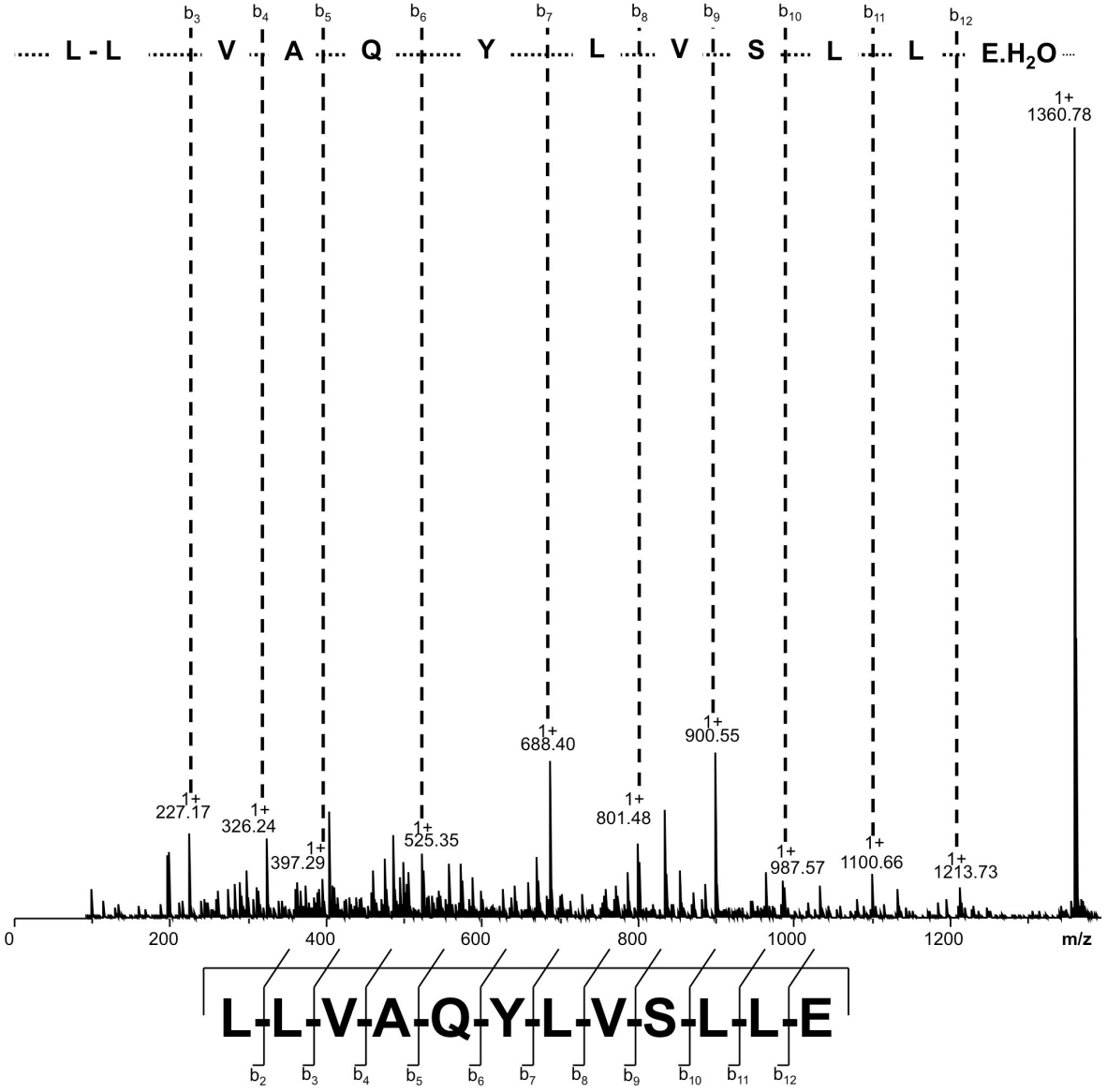


Figure 16. HRMS^2 spectra of **2** (1360.8 [M+H]$^+$) depicting its peptide sequencing and assignment b_2-b_{12} ions. Differences between successive B ions from B_3 to B_{12}, were 99.1, 71.02, 128.06, 163.05, 113.08, 99.07, 87.02, 113.09, 113.07, 147.05 and these corresponded to monoisotopic masses for Val, Ala, Gln, Tyr, Leu, Val, Ser, Leu, Leu, Glu.H_2O. The B_3 ion corresponded to the protonated monoisotopic mass of 2x Leu.

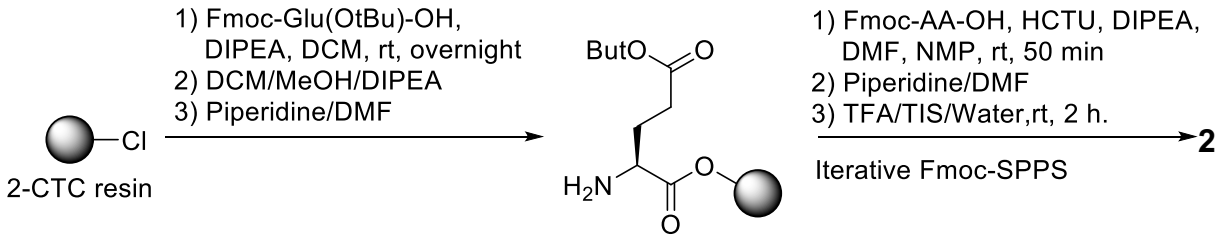


Figure 17. Chemosynthesis route for **2** via Fmoc/Boc chemistry. Linear peptide was synthesised by iterative solid phase peptide synthesis.

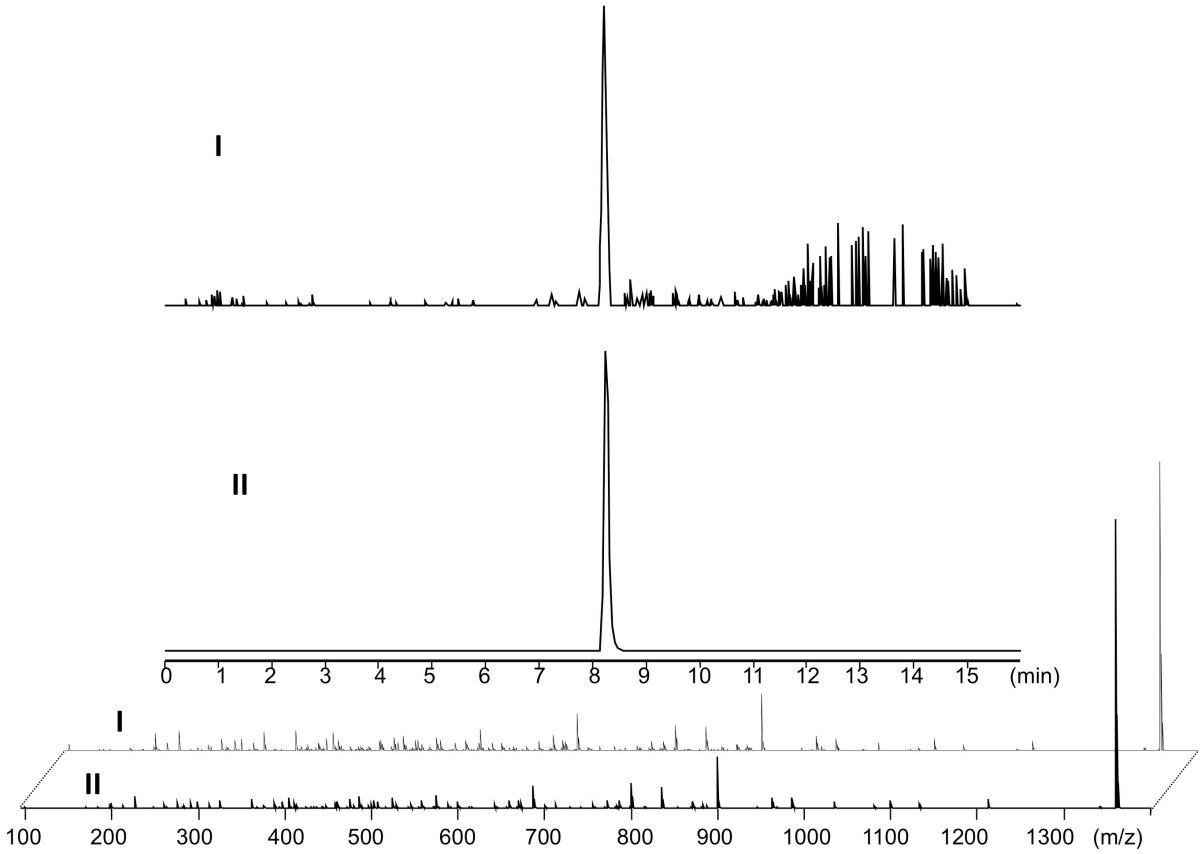


Figure 18. Extracted Ion Chromatogram from HPLC-ESI-HRMS analysis of extracts containing **2** (*m/z* 1360.8 [M+H] +) with concomitant MS^2 data that was used for structure elucidation. Natural (I) and synthetic (II) compounds were identical as seen from the above similar retention times, m/z values and fragmentation patterns. This confirmed the absolute configuration of **2** and amino acid residues of Ishipeptides A and B.

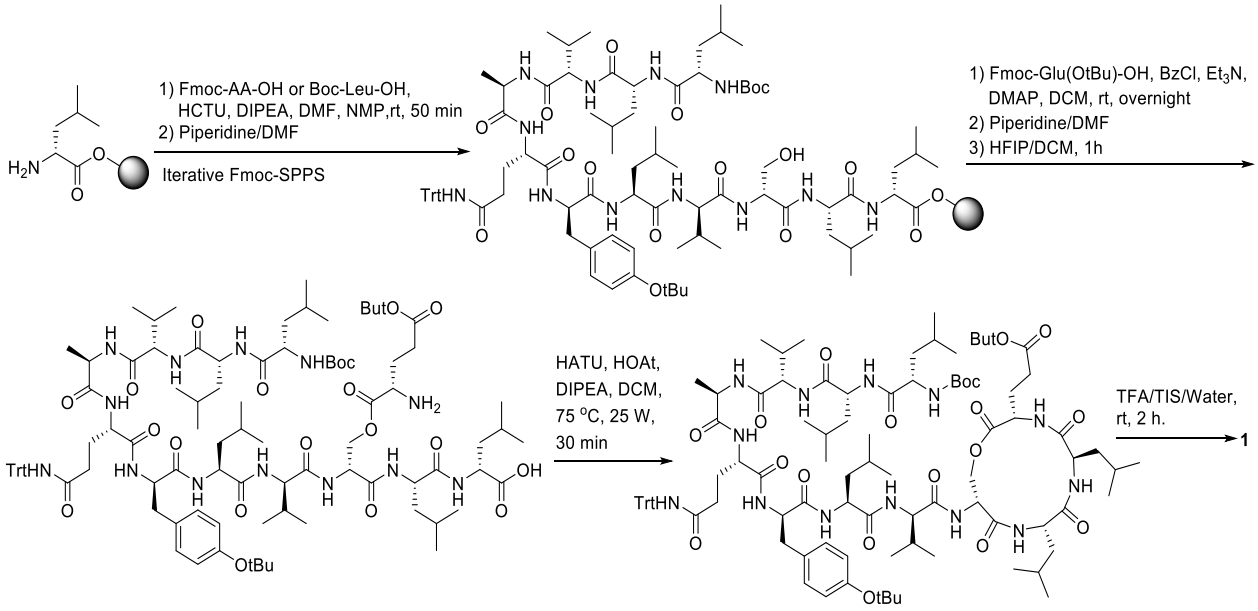


Figure 19. Chemosynthesis route for **1** via Fmoc/Boc chemistry. Linear undecapeptide was first synthesised by iterative solid phase peptide synthesis (SPPS). This followed by microwave-assisted esterification between the building blocks L-glutamic acid and D-serine. After cleavage from the resin, late stage macrolactamisation was then carried out between L-glutamic acid and the first loaded D-leucine.

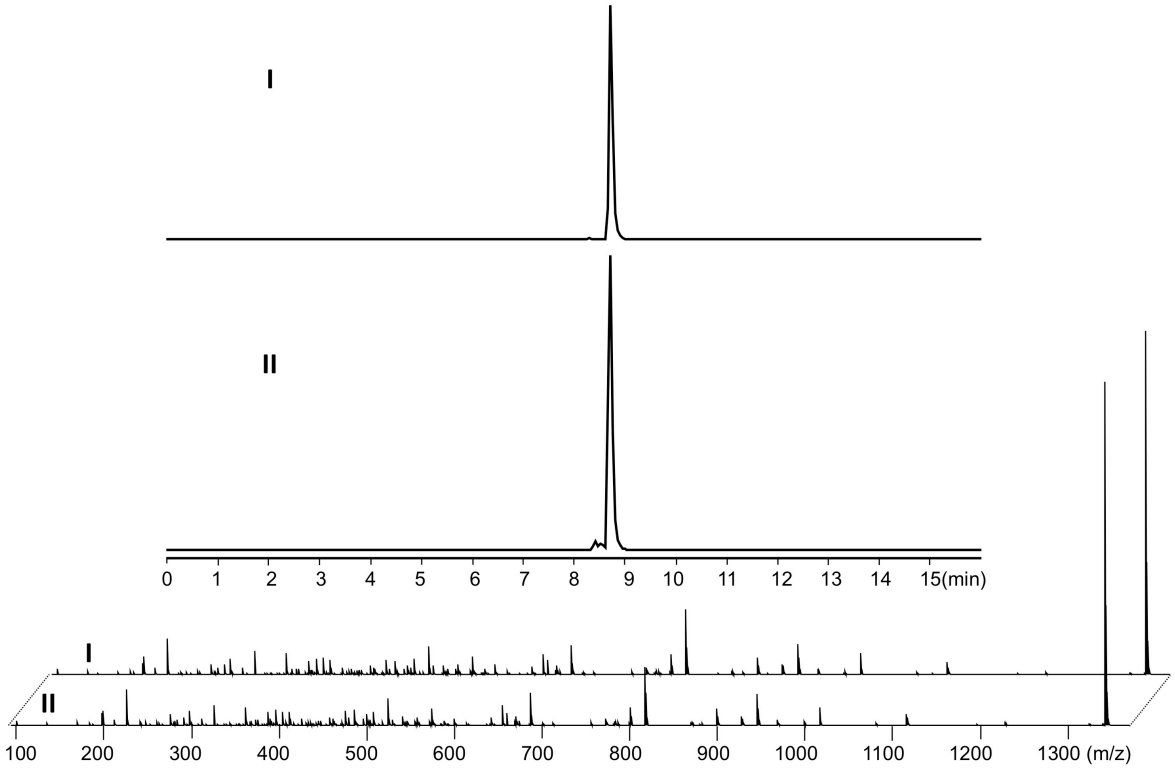


Figure 20. Extracted Ion Chromatogram (EIC) from HPLC-ESI-HRMS analysis of extracts containing **1** (*m/z* 1342.8 [M+H] +) with concomitant MS^2 data that was used for structure elucidation. Natural (I) and synthetic (II) compounds were identical as seen from the above similar retention times, m/z values and fragmentation patterns. This also confirmed the absolute configuration of **1**.

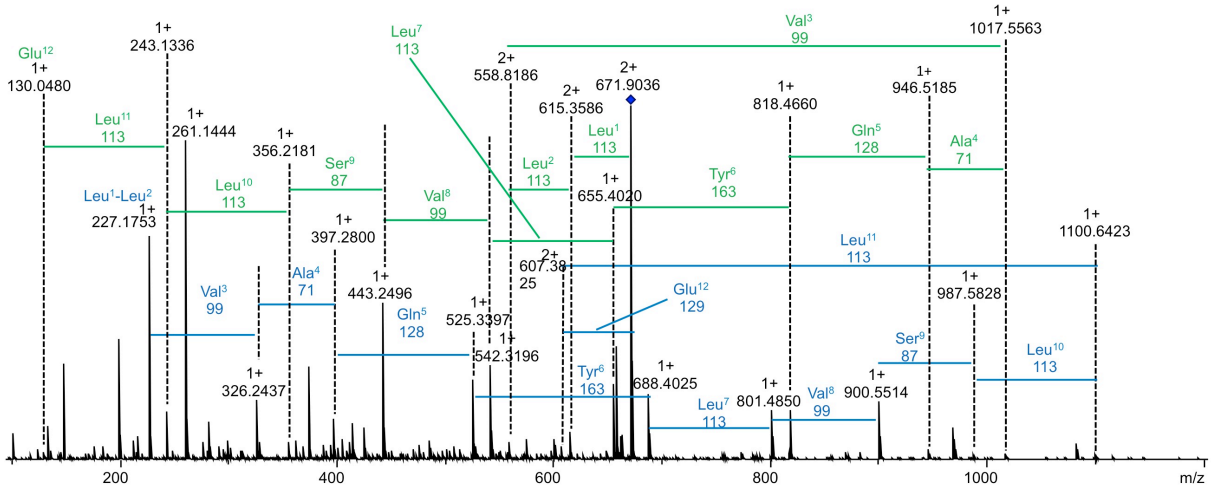


Figure 21. $HRMS^2$ spectra of **1** (671.9 $[M+2H]^{2+}$) labelled with neutral losses from its constituent amino acid building blocks. These were calculated from both b(blue) and y ions(green).

The proposed biosynthesis of **1** from IshA-B NRPS is as follows. Module one does not condense any starter unit to L-Leu[1] despite containing a C_{start} domain. Module two would condense L-Leu[1] to L-Leu[2]. Module three would epimerise L-Leu[2] to D-Leu[2] then condense it to L-Val[3]. Module four would condense L-Val[3] to L-Ala[4]. Module five would epimerise L-Ala[4] to D-Ala[4] then condense it L-Gln[5]. Module six would condense L-Gln[5] to L-Tyr[6]. Module seven would epimerise L-Tyr[6] to D-Tyr[6], then condense it to L-Leu[7]. Module eight would condense L-Leu[7] to L-Val[8]. Module nine would epimerise L-Val[8] to D-Val[8] then condense it to L-Ser[9]. Module ten would epimerise L-Ser[9] to D-Ser[9] then condense it to L-Leu[10]. Module 11 would condense L-Leu[10] to L-Leu[11]. Module 12 would epimerise L-Leu[11] to D-Leu[11] then condense it to L-Glu[12]. It would further catalyse would a nucleophilic attack of the *β*-hydroxy of D-Ser[9] on the terminal carboxy creating a branch cyclic depsidodecapeptide and simultaneously cleaving it from the NRPS.

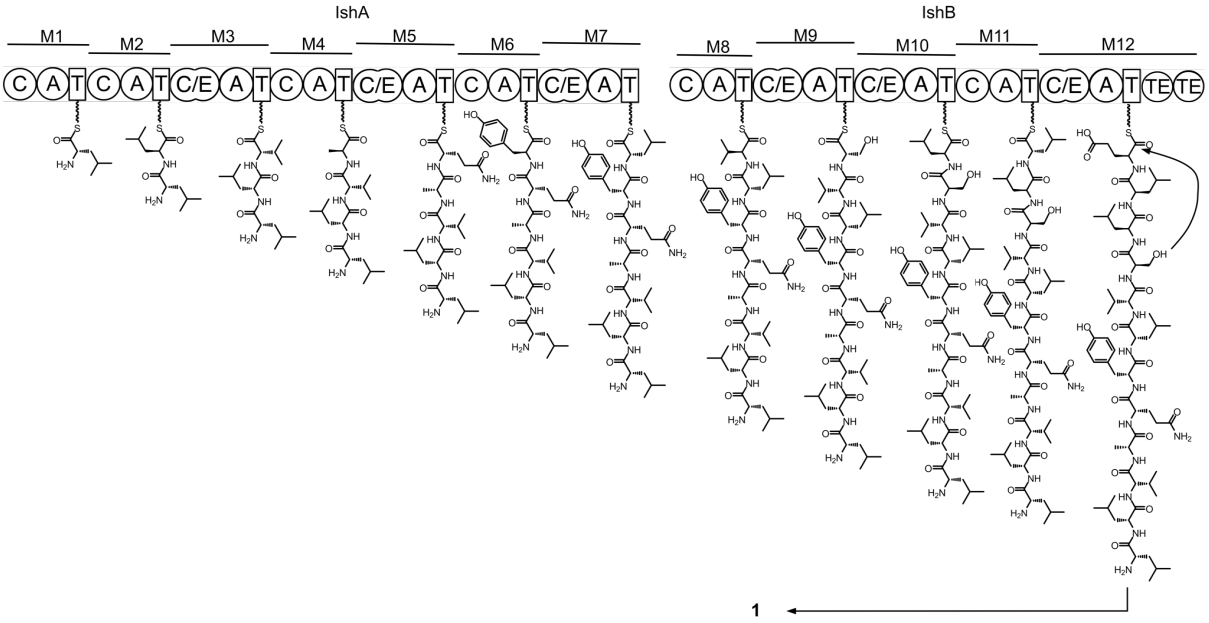


Figure 22. Proposed biosynthesis of ishipeptide A (**1**) from *X. ishibashii.* Domain organization of the 12 modules (M) of the two NRPSs -IshA & IshB- is represented by C (condensation domain), E (epimerisation domain), A (adenylation domain), T (thiolation domain) and TE (thioesterase domain).

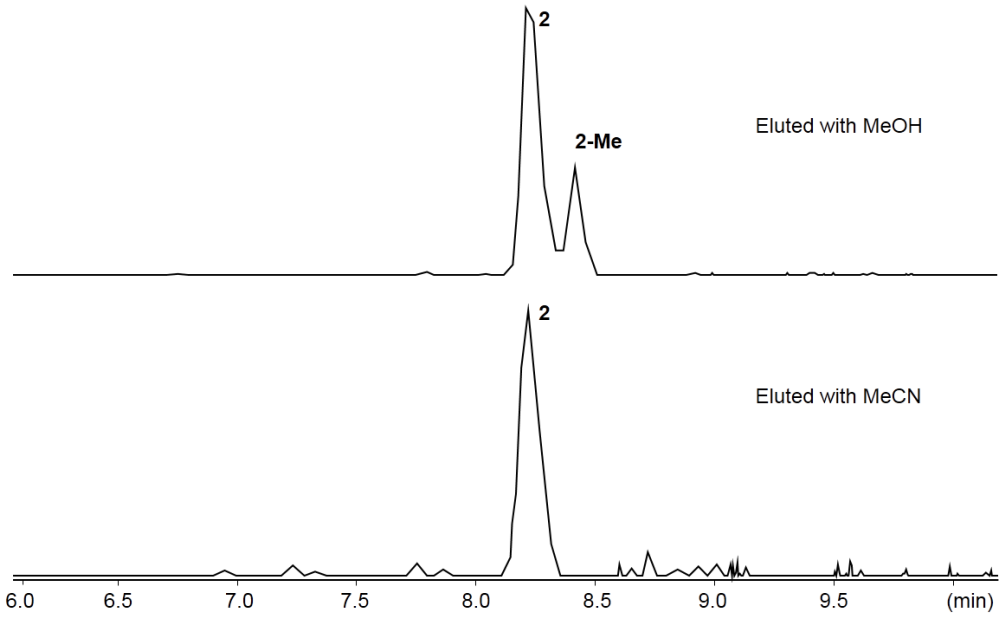


Figure 23. Extracted Ion Chromatogram from HPLC-ESI-HRMS analysis of extracts containing **2** (*m/z* 1360.8 [M+H] +). Sample preparation with MeOH resulted in **2-Me** which was absent when sample preparation without MeOH was done, confirming it as the methylated product of **2.**

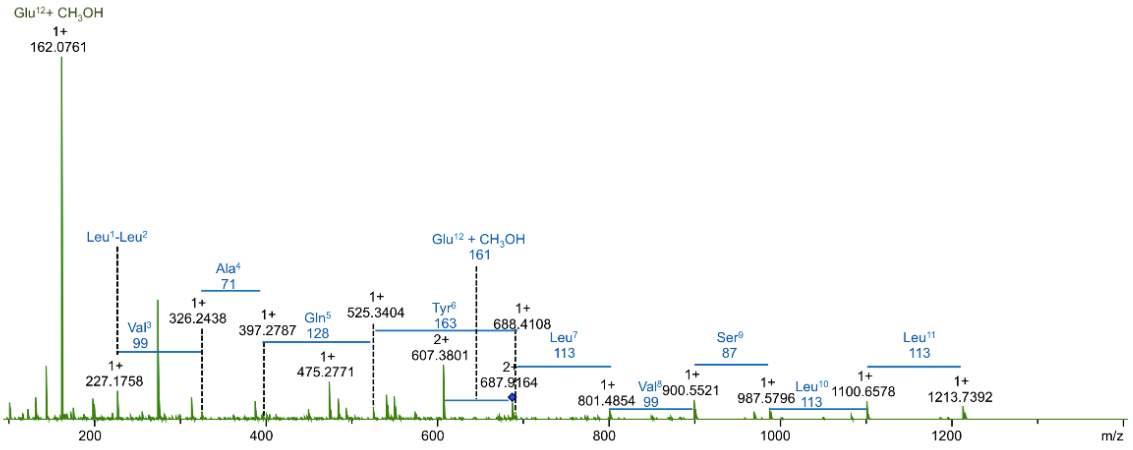


Figure 24. HRMS^2 spectra of **2-Me** (687.9 $[M+2H]^{2+}$) labelled with neutral losses from its constituent amino acid building blocks. These were calculated from both b(blue) and y ions(green).

We previously[106] obtained mammalian cell toxicity for **3**. As **3** was strongly inhibitory to Gram-positive bacteria, it was further tested against Gram-negative bacteria. Compound **2** was linear peptide and displayed no detectable antimicrobial activity while its macrocyclic derivative **1** was inhibitory to Gram-positive bacteria indicating that macrocyclisation affected antimicrobial activity.

Table 7. Antimicrobial activity of macrocyclic non-ribosomal peptides **1,3** and linearised **2.** Minimum inhibitory concentration (MIC) and toxicity to mammalian cells (IC_{50}) are in μM. Ampicillin, chloramphenicol, Amphotericin B and podophyllotoxin were used as controls for *M. luteus*, *E. coli, C. lusitaniae* and mammalian L6 cells respectively.

Test species	**1**	**3**	**2**	Control
M. luteus	37	3.0	>73.5	4.2
C. lusitaniae	>74.4	>96	>73.5	0.6
E. coli	N/A	24	N/A	0.42
Mammalian L6 cells	N/A	83	N/A	0.007

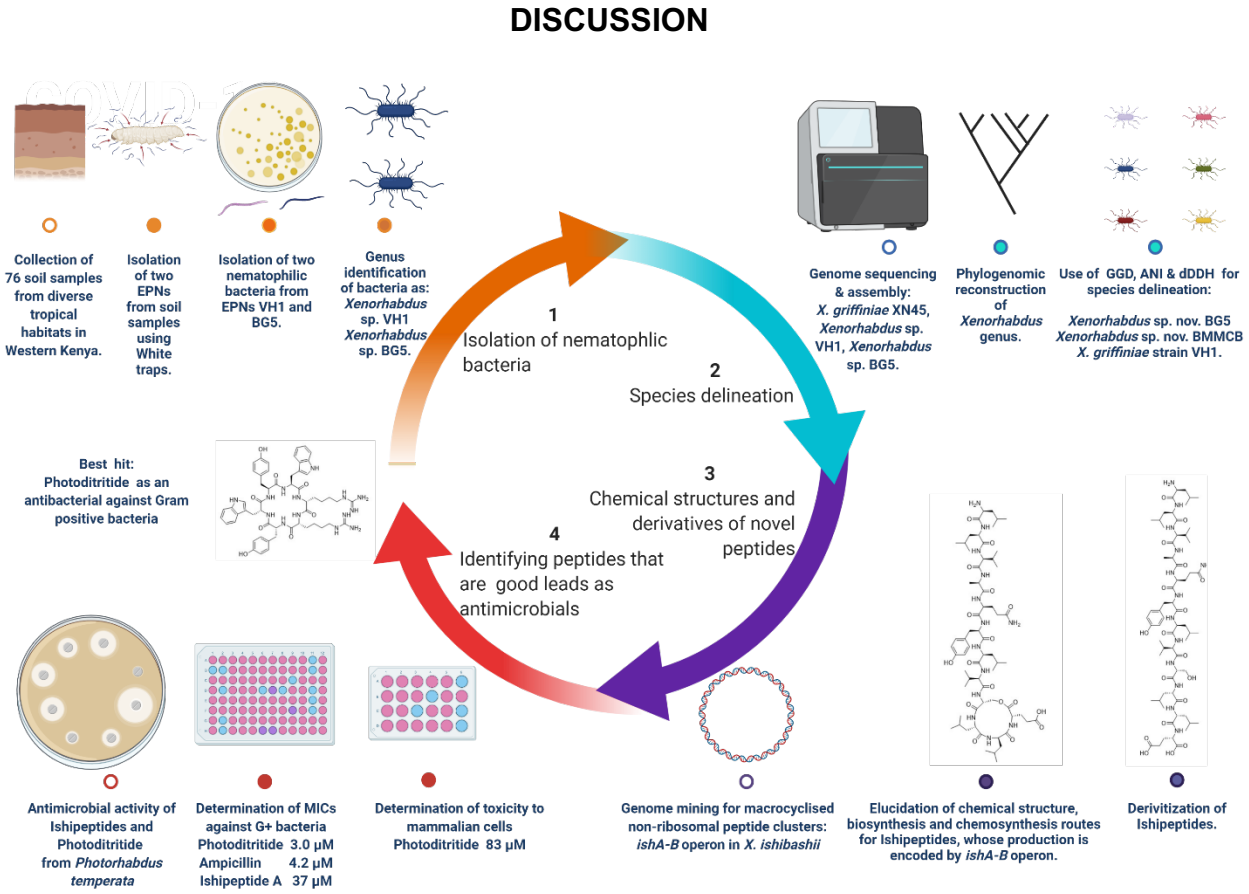


Figure 25. Flowchart of the process of how to discover antimicrobial peptides from nematophilic bacteria as demonstrated in this dissertation. Created with Biorender.com

Soils in Western Kenya contain entomopathogenic nematodes

Seventy-six soils samples collected from geographical points in Western Kenya (Table S1) were investigated for the presence of EPNs. This was done by baiting EPNs with *G. mellonella*, as recommended[180]. After seven days, viable larvae were present in most samples and these provided a default control for natural mortality. A few soil samples (Table S1) had dead larvae. These were assessed for characteristics that typify an EPN-caused death: limp body, red/tan colour, and no smell of putrefaction. Limp body is due to the bioconversion of tissues by enzymes and the lack of putrefaction is due to the production of antimicrobials and both these biomolecules are produced by the bacterial EPN endosymbionts. Such cadavers were placed in modified White traps[180] as recommended, to draw out nematodes. This happens

because IJ3 naturally migrate away from the cadaver once nutrients are depleted and into the distilled water of the White trap[211]. Traps of cadavers obtained from soil samples of points VH1 and BG5 (Figure S1, S2) were confirmed to have nematodes by light microscopy.

In order to culture EPN endosymbionts from these cadavers, cadavers were surface sterilised, aseptically dissected and haemolymph was obtained and inoculated onto nutrient agar for cultivation. Though endosymbionts can be obtained by hanging drop or crushing nematodes techniques[168], this indirect isolation via haemolymph is fast and accurate. The type species of *X. khoisanae*[33] was obtained using this method. This method is based on the natural lifecycle of EPNs where, once within an insect haemocoel they release their *Xenorhabdus* or *Photorhabdus* endosymbionts that then cause quick insect death and proliferate within the cadavers to high concentrations. Previously[168] it was assumed that *Xenorhabdus* and *Photorhabdus* create a monoxenic environment due to their production of antimicrobials. However, for *Steinernema* EPNs this paradigm was recently shifted to *Xenorhabdus* as the dominant bacteria in the cadaver and other bacteria of the following genera present at low levels: *Pseudomonas* sp., *Alcaligenes* sp., *Ochrobactrum* sp., *Stenotrophomonas* sp*., Achromobacter* sp., and *Brevundimonas* sp.,[16].

After four days, NBTA agar plate cultures were obtained and the following morphological characteristics were used to distinguish putative primary variants of *Photorhabdus/Xenorhabdus* from other bacteria: pigmented, umbonate, irregular margin and visible swarming patterns. Only *Xenorhabdus* and *Photorhabdus* possess all these four characteristics among bacteria genera possibly found in cadavers of EPN-killed insects supporting the use of this criteria for their selection.

Primary *Photorhabdus/Xenorhabdus* variants adsorb dyes, swarm, and produce high levels of secondary metabolites while secondary variants do not[168]. A genetic basis of this variation is the *hexA* gene[212] and it was demonstrated that its deletion resulted in a permanent primary variant of *P. luminescens* because HexA represses genes associated with primary variant traits[213]. Additionally, HexA interacts with Hfq which regulates secondary metabolite production, and Δ*Hfq* mutants produce almost no secondary metabolites[204]. HexA also interacts with the global regulator Lrp which

regulates swarming motility in *Xenorhabdus*[214]. Pure cultures from VH1 & BG5 isolates were catalase negative indicating putative *Xenorhabdus* isolates. As *Xenorhabdus* species have only been naturally isolated from *Steinernema* nematodes their respective nematode hosts were named *Steinernema* sp. VH1 and *Steinernema* sp. BG5.

Soils from VH1 collection point were red volcanic, found on land cultivated with cabbages (Figure S1) and those from BG5 were clay (Figure S2) found on riverine land, adjacent to cultivated land (Table S1). Previous studies correlated natural occurrence of EPNs to clay soils in riverine areas in Fiji[215], cultivated land in Central Kenya[216] and Colimo, Mexico[217] and land on crop edges[218] in Northern Spain. The soil and habitat from which EPNs were found in Western Kenya corroborate these previous studies.

Genome assemblies of isolated bacterial strains were of good quality for overall genome related index analysis

To delineate the species of putative *Xenorhabdus* isolated from *Steinernema* sp. VH1 and *Steinernema* sp. BG5, genomes were sequenced and assembled, including that of *X. griffiniae* XN45 that was previously isolated from *Steinernema* sp. scarpo.

Coverage was 100x for *X. griffiniae* XN45 and 70x, *Xenorhabdus* sp. strains BG5 & VH1. Assembled genomes for *Xenorhabdus* sp. strain BG5, *X. griffiniae* XN45 and *Xenorhabdus* sp. strain VH1 had GC contents of 43.8%, 43.57%, and 43.65% respectively. *Xenorhabdus* sp. strain BG5 had the smallest genome at 3,933,551 bp with those of *X. griffiniae* XN45 and *Xenorhabdus* sp. strain VH1 of sizes 4,215,754 bp and 4,224,998 bp respectively (Table 2). N50 values for strains BG5, XN45 and VH1 were 102,633bp, 45,296bp, 29,298 bp, respectively. Strain XN45 had a larger N50 than strain VH1 despite having more contigs and this can be attributed to the shorter insert size(100bp) used in sequencing. The contig N50 represents the length of the shortest contig among the fewest number of contigs in which half of the genome is contained. Low contig N50 values indicate genomes with low “draftness” i.e they are less fragmented. This is especially important for genome analysis that require long contigs such as the elucidation of BGC because of their size, which –only taking those found in *Streptomyces* species as an example– are between 11-108Kb[219]. For overall

genome related indices however, sequencing depth of >50x is the only explicitly outlined threshold[67], indicating that the assemblies were of good quality for such analysis.

All genome assemblies were 100% complete and only the BG5 genome was contaminated (0.7%) when assessed with the checkM algorithm[220]. With this algorithm, completeness was estimated by determining the presence of single-copy core genes from the most closely related complete genome –in this case this was *Xenorhabdus bovienii* SS-2004– and 100% indicated all were detected in the assessed genome. When any of these genes are detected more than once, contamination may be assumed. The estimated contamination of BG5 may have been due to occurrence of single-copy core genes on contig edges that makes them appear to be in two genomic loci. However, this is unlikely for the BG5 genome as partial CDS were not detected during annotation (Table 3). The most likely reason is due to the estimation of contamination by the occurrence of single copy core genes that are meant to be absent as per the reference genome. And this type of contamination was observed for genes encoding UDP-N-acetylmuramoyl-L-alanine--D-glutamate ligase, tRNA-(ms[2]io[6]A)-hydroxylase, S-adenosylmethionine decarboxylase proenzyme, RbsB, Oxygen-independent coproporphyrinogen-III oxidase-like protein, Phosphate ABC transporter, permease protein, Na+ dependent nucleoside transporter NupC, Membrane-associated zinc metalloprotease, Cell division protein ZapE, 2-keto-3-deoxy-D-arabino-heptulosonate-7-phosphate synthase II, Glutamate/aspartate ABC transporter and ATP-binding protein GltL among others. A simple reason for this latter misinterpretation of contamination is that BG5 is phylogenomically distantly-related from reference genome *X. bovienii* (Figure 5) and thus is likely to contain a high number of single copy core genes that are absent in *X. bovienii*. This supported the hypothesis it was of good quality for accurate downstream analysis.

The nematode host of *X. griffiniae* XN45*, Steinernema* sp. scarpo is a putative novel nematodes species.

Previously we[69] isolated *X. griffiniae* XN45 from the nematode *Steinernema* sp. scarpo that was originally isolated from soils in Muran'ga, Central Kenya. However, we did not do any molecular identification of this nematode. The nucleotide sequence of its

internal transcribed spacer region (ITS) was obtained (Peter Njenga Ng'ang'a, personal communication) and comparison of its percentage nucleotide identities to any known species yielded the highest value as 93.2% (Table 4). This was of *Steinernema* sp. UH36 that was isolated from the Aberdare Ranges above L. Elementaita[216] in Rift Valley Province, Kenya making it both geographically and phylogenetically the most closely related species to *Steinernema* sp. scarpo. The minimum percentage nucleotide identity of the ITS region for two *Steinernema* strains to be considered one species is 95%[221] indicating that the nematode host of *X. griffiniae* XN45, *Steinenema* sp. scarpo may be a novel *Steinernema* species.

Novel isolates *Xenorhabdus* sp. strain VH1 and BG5 are most closely related to *X. griffiniae* XN45

Genome distances between *X. griffiniae* XN45, *Xenorhabdus* sp. strains VH1 and BG5 and 22 *Xenorhabdus* type species were calculated with GGDC 2.1 formula within the type strain genome server workflow[184] and used to create a phylogenomic tree (Figure 5). Strains BG5 and VH1 did not cluster with any of the type strains thus excluding them as those species. VH1 clustered with *X. griffiniae* XN45 and there was no distance between the two (Figure 5), indicating that it is an *X. griffiniae* species. Additionally, the difference in GC content between the two was 0.08%, and this is well within 1% threshold that exists when two strains are of the same species[222]. Strain BG5 was most closely related to *X. griffiniae* XN45 and shared a common ancestor with its extinct primary ancestor. It was distantly related from any other species indicating that it may be a novel species. *Xenorhabdus* sp. strain BMMCB was identified an *X. griffiniae* strain[68]. Yet we[69] previously demonstrated that its *serC*, *recA* and *16s rRNA* genes nucleotide percentage identities to those of the type strain were below the same species thresholds of 97% for protein coding genes and 98.5% for *16s rRNA* genes. Its genome assembly was assessed as of good quality for overall genome relatedness index analysis (Table 2) and thus included in the phylogenomic reconstruction (Figure 5) in order to conclusively delineate its species. Indeed, it did not cluster in the *X. griffiniae* clade and was most closely-related to *Xenorhabdus* sp. strain BG5, sharing a common ancestor with its extinct primary ancestor. Additionally, the difference in GC content between it and *X. griffiniae* XN45 or strain VH1 was at

least 1.03%, and this is above the 1% threshold for two strains to belong to the same species[222]. Its distant relatedness to other *Xenorhabdus* taxon (Figure 5) indicated that it may be a novel species. These supported the hypothesis that *Xenorhabdus* sp. strain BMMCB is not an *X. griffiniae* species.

X. griffiniae stain VH1, _Xenorhabdus_ sp. nov. BG5, _Xenorhabdus_ sp. nov BMMCB are a novel _Xenorhabdus_ strain and two novel bacterium species.

To conclusively delineate species, strains of the 6 closest species to VH1, BMMCB and BG5 from the phylogenomic reconstruction were compared via orthoANI, GGD and dDDH. These six strains were: *X. ishibashii*, *X. ehlersii*, *X. eapokensis*, *X. thuoxuangnensis* and *X. griffiniae* XN45 (Table 5). *X. griffiniae* XN45 and *Xenorhabdus* VH1 had 0.000 GGD, 99% dDDH and 99% orthoANI to each other and these were all above the thresholds for same species delineation of < 0.0361, > 95.1% and >70% respectively. This demonstrated that VH1 is an *X. griffiniae* species. Strain BG5 was most closely related to *X. griffiniae* XN45 with GGD, orthoANI and dDDH of 0.059, 94.24% and 67.3% respectively. This was below all three thresholds for same species delineation demonstrating *Xenorhabus* sp. strain BG5 as a novel bacteria species. Additionally, the pangenome of the clade containing strains XN45, VH1 and BG5(Figure 5) elucidated 66% as core genes, 12.3% as *Xenorhabdus* sp. strain BG5 specific genes and 18.9% as *X. griffiniae* specific genes (Figure 9). *X. griffiniae* XN45 and *X. griffiniae* VH1 differed from each other by a total of only 80 genes. These results supported the hypothesis that VH1 is an *X. griffiniae* species and *Xenorhabdus* sp. strain BG5 is specifically distinct from *X. griffiniae* species. *Xenorhabdus* sp. strain BMMCB was mostly closely related to *Xenorhabdus* sp. strain BG5 with GGD, orthoANI, and dDDH of 0.08, 92.25%, 57.1% respectively. This was also below all three same species thresholds demonstrating it as a novel bacteria species. Thus, the strains were emended as follows: *X. griffiniae* VH1, *Xenorhabdus* sp. nov. BG5 and *Xenorhabdus* sp. nov. BMMCB.

The _Xenorhabdus_ and _Photorhabdus_ genera have open pangenomes.

The α exponents of Heaps Law[223] for genera *Xenorhabdus* and *Photorhabdus* pangenomes were 0.2792 and 0.3967, respectively. The core protein encoding genes

consisted of 13.9% and 19% of their pangenomes, respectively. Addition of new genomes results in an increase in pangenome size due to more accessory genes as the core genome decreases until a constant size (Figure 8, S3). This increase in pangenome size occurs at a decreasing rate and when a pangenome is complete, this rate would theoretically be 0[66]. The α exponent of Heaps Law is used to estimate how close to completion a pangenome is with >1 indicating a closed/ complete pangenome and values <1 indicating an open pangenome[66]. Both pangenomes were open as they had values <1. The pangenome of *Xenorhabdus* had a larger percentage of protein encoding genes (86.1%) compared to *Photorhabdus* (81%) that were either shared by two or more species or were species specific. This indicated that it was a more diverse genus corroborating with phylogenomic trees of the two (Figures 4, 7). The challenges in delineating species in *Photorhabdus*[38,48,224] when compared to *Xenorhabdus,* can be attributed to this lower phylogenetic diversity. For both species, open pangenomes means that novel genomes need to be added to each genus so that they can be complete. As *Xenorhabdus* is a more open pangenome one can predict novel species from this genus shall be discovered at a faster rate than those of *Photorhabdus*.

Loci of genes encoding transposases are adjacent to genomic islands in *Xenorhabdus* sp. nov. BG5

To gain insights on the genomic basis of speciation in *Xenorhabdus* sp. nov. BG5, differences in genomes were highlighted by identification of genomic islands of *Xenorhabdus* sp. nov. BG5 when compared to *X. griffiniae* species (Figure 10). A total of 18 genomic islands were identified and 16 of these had GC content that significantly differed from the genome average (Table 6). Additionally, these were flanked by genes that encode insertion sequence (IS) family transposases. In fact, all genes encoding IS family transposases flanked genomic islands. A number of genomic islands were not flanked by IS elements. One of this, prophage 7_GI3, was attributed to a prophage (Table S3) and further analysis elucidated seven prophage loci, all of which had significantly differed GC content (Table S3).

IS family transposase facilitate the movement of mobile genetic elements within and between genomes and thus contribute to genome diversification[225] through events like horizontal gene transfer (HGT). Thus, acquisition of genomic islands can be facilitated

by IS family transposase mediated-HGT and such islands would typically have different GC content from the rest of the genome[226]. This supports the hypothesis that BG5 has genomic islands that resulted from of IS family transposase mediated HGT. Within these islands, genes predicted to encode unknown proteins were the major CDS accounting for 66% followed by those encoding transposases at 23% (Figure 11). This indicated that IS family transposase mediated-HGT contributed towards the gain of unique genetic traits in *Xenorhabdus* sp. nov. BG5.

1 and 2 are related compounds whose synthesis is encoded by the *ishA-B* operon

Upon induced expression of *ishA-B* operon, **1** and **2** were detected (Figures 12, S4) and their HRMS-determined molecular formulae were experimentally verified as $C_{65}H_{108}N_{13}O_{17}$ (**1** *m*/*z* 1342.7731 $[M + H]^+$, calcd for $C_{65}H_{108}N_{13}O_{17}$, Δppm 19.288) and $C_{65}H_{110}N_{13}O_{18}$ (**2** *m*/*z* 1360.7869 $[M + H]^+$, calcd for $C_{65}H_{110}N_{13}O_{18}$, Δppm 16.2403) with the mass difference attributed to a H_2O molecule (Figure 13) suggesting a hydrolysis reaction. Detailed $HRMS^2$ analysis (Figures 16, 24) coupled with sample preparation with different solvents (Figure 23) revealed that elution of extracts with methanol resulted in **2-Me.** An example of NRPS-synthesised compound that undergoes non-enzymatic hydrolysis upon release from the TE domain is obafluorin (*m*/*z* 359.7 $[M + H]^+$) that is hydrolysed to obafluorin-COOH (*m*/*z* 377.7 $[M + H]^+$)[227]. This supported the hypothesis that these are related compounds with **2** being the hydrolysed version of **1.**

Amino acid residues for 1 and 2 are 5x Leu, 2x Val, Ala, Ser, Tyr, Gln, Glu.

Bioinformatic analysis of A-domain Stachelhaus codes and positions of E-domains predicted the NRPS encoded by *ishA-B* operon to synthesize a peptide of the following sequence: LIVaQyLvsLIQ (Table S5). Eight of these were experimentally confirmed as building blocks of **1** via both labelling experiments and $HRMS^2$ spectra analysis (Figures 14, 15, 21). Serine and alanine were confirmed by $HRMS^2$ spectra analysis only as serine and alanine are readily catabolized[228] by bacterial enzymes such as serine aminotransferase, serine hydroxymethyltransferase, serine dehydratase, aspartate decarboxylase and muramyl ligase –their putative genes were all identified in an annotated record of NJAK00000000.1– thus obscuring labelling experiments. An

isotope of **1** with all five labelled leucines was probably produced albeit below detection levels (Figure 14), thus the fifth leucine building block was confirmed by $HRMS^2$ spectra analysis only. Isotope labelling experiments coupled with $HRMS^2$ spectra and domain organization analysis for amino acid building block determination is an established method in NRP structure elucidation[229] supporting the hypothesis that the building blocks for **1** and **2** are 5x Leu, 2x Val, Ala, Ser, Tyr, Gln, Glu.

The main product, whose synthesis is encoded by the *ishA-B* operon, is a branch cyclic peptide and 2 is its linearised derivative

The peptide sequence was determined by $HRMS^2$ spectra analysis of **2** (Figure 16) that was hypothesised as the linearised peptide thus suitable for *de novo* peptide sequencing. To confirm the sequence LlVaQyLvsLlE and elucidate the structure of **2**, the peptide was synthesized, using the epimerization domain positions to determine amino acid configurations, via iterative solid phase synthesis (Figure 17). Identical values for HPLC-ESI-HRMS data and $HRMS^2$ spectral fragments for synthetic and natural **2** (Figure 18) elucidated the structure and absolute configuration of **2** as shown**.** This supported the hypothesis that **2** is the linearised peptide of **1.**

To determine which cyclisation catalysed by observed double TE domain resulted in **1**, a head to tail macrolactamisation was hypothesised creating a cyclic peptide or macrolactonisation creating a branch cyclic depsipeptide. Dissimilar retention times for natural **1** and a synthetic head-to-tail cyclised **2** revealed this as an inaccurate cyclisation. *In silico* comparisons of hypothetical spectral fragments of a branch cyclic depsipeptide with an ester bond between side *β*-hydroxy of D-Ser^9 and terminal carboxyl of L-Glu and another with an ester bond between side chain hydroxy of D-Tyr^6 and terminal carboxy of L-Glu with experimental MS^2 spectral fragments of **1** using the iSNAP algorithm[230] demonstrated the former to have a more probable peptide-spectral-match (P1:71.2; P2:31.9). This structure was synthesised (Figure 19) via a late-stage macrolactamisation route so as to increase ester bond formation[231]. Identical values for HPLC-ESI-HRMS data and $HRMS^2$ spectral fragments for synthetic and natural **1** (Figure 20) elucidated the structure and absolute configuration of as shown**.** This supported the hypothesis that **1** is the main NRPS product, macrocyclised from **2.**

C_{start} domain of IshA does not *N*-terminally acylate peptides

X. ishibashii biosynthesised Ishipeptides A (**1**) from IshA&IshB synthetases that composed of 12 modules (Figure 22). Ishipeptide B (**2**) is a dodecapeptide with the primary structure: L-Leu[1] D-Leu[2] L -Val[3] D -Ala[4] L -Gln[5] D -Tyr[6] L -Leu[7] D -Val[8] D -Ser[9] L -Leu[10] D -Leu[11] L -Glu[12]. Ishipeptide A is formed by a double thioesterase-catalysed ester bond formation between *β*-hydroxy of D -Ser[9] and the terminal carboxy of L -Glu[12] creating a cyclic tetrapeptide linked to a linear octapeptide that is consequently 18 Da less in mass than Ishipeptide B.

Both peptides were not *N*-terminally acylated, despite being biosynthesised from a C_{start} domain bearing synthetase. A mutation at the histidine active site HH $(X)_3D(X)_{14}$ from HH to HP of C_{start} domains was present in the non *N*-terminally acylating C_{start} domain that produced the non-acylated ambactin[117]. A probable single base mutation at position 429 of *ishA* from T/C to the current G resulted in a similar change from the histidine active site HH to HQ. A site directed mutagenesis approach to correct this mutation did not result in any production of acylated derivatives (Figure S5). This approach also did not result in production of acylated ambactin derivatives indicating that in addition to the histidine active site HH $(X)_3D(X)_{14}$, other sites are required for C_{start} domains to *N*-terminally acylate peptides. This is a crucial insight for NRPS-reengineering of C_{start} domains.

Macrocyclisation contributed to antimicrobial activity in Ishipeptides

The only inhibition by Ishipeptides was by Ishipeptide A against Gram-positive *M. luteus* at an MIC of 37 μM demonstrating that the macrocyclic peptide was inhibitory while the linear analogue was not. This indicated macrolactonation had an enhancing effect on antimicrobial activity as observed with the taxalllaids[171]. A reason for this may be change in hydrophobicity. Ishipeptide A had increased hydrophobicity as compared to linear Ishipeptide B as it was more retained in RP-HPLC columns (Figure S4), and this change in physiochemical properties due to macrocyclisation was also seen in the lipodepsipeptide Xenematide A which was more hydrophobic than its linear derivative[175]. Apart from the reduced conformation, this increase in hydrophobicity may be due to reduced hydrophilicity, because the side chain of serine is now involved in an ester bond with the terminal carboxyl. Hydrophobicity of an AMP determines how

'soluble' it is in the non-polar lipid matrix of phospholipid bilayers of cell membranes as apart from the polar phospholipid heads the rest of the bilayer is highly hydrophobic. Thus a degree of hydrophobicity is required for antimicrobials to insert into cell membranes with increasing hydrophobicity – up to an upper threshold beyond which there is loss of selectivity and resultant lysis of mammalian cells– among certain AMP analogues resulting in increasing antimicrobial activity[232,233]. This supported the hypothesis that increased hydrophobicity due to macrocyclisation enhanced antimicrobial activity in Ishipeptides.

The macrocyclic structure of Photoditritide may have contributed to its high antibacterial activity.

Photoditritide is a potent antimicrobial inhibitory to *M. luteus* at 3.0 μM, lower than the clinical beta-lactam antibiotic Ampicillin (4.2 μM). It was selective, with no inhibitory seen against fungi and mammalian cells at concentrations >96 μM and 83 μM respectively. *P. temperata* biosynthesised from Pdt synthetase, Photoditritide that is a head-to-tail cyclised hexapeptide of primary structure: D-Har[1] D-Har[2] D-Tyr[3] D-Trp[4] L-Tyr[5] L-Trp[6] and cyclised via a thioesterase-catalysed nucleophilic attack of the *α*-amino group of D-Har[1] on terminal carboxy of L-Trp[6]. The rare residue homoarginine, is an arginine analogue that differs by having its side chain elongated by a single carbon atom. Photoditritide is an arginine and tryptophan rich cyclic hexapeptide and for these, cyclisation improved antimicrobial activity as compared to linear analogues by creating conformational rigidity which resulted in an increased charge density[169]. An increase in charge density increases the probability[150] that peptide electrons will participate in electrostatic interactions with the bacterial surface. Additionally, the conformational rigidity of cyclic peptides increases affinity to bind to their target[234]. Photoditritide has a net charge is of (+2) due to its two homoarginine residues. This makes it highly electrostatically attracted to the cell surface of *M. luteus* that is conversely, negatively charged due to the teichoic and lipoteichoic acids and the negative carboxyl groups of cross-linking amino acids of the peptidoglycan cell wall[232]. This high negative charge as compared to animal, fungi and protozoal cell surfaces makes cationic AMPs most attracted to bacteria, creating a selective activity as was seen in Photoditritide that was

not inhibitory fungi and mammalian cells at high concentrations (Table 7). This selective activity can also be because most AMPs exhibit their bioactivity by interacting with cell membranes of target cells[235] exemplified by 5 out 7 of clinically-approved AMPs being membrane active peptides[132].The difference in cell membrane composition confer different proprieties that result in the selectivity of AMPs. For example, bacterial cell membranes lack any sterol, fungi have ergesterol, protozoa have lanesterol and mammals have cholesterol that makes it harder to reorganise the lipid bilayer, creating strong cohesive forces of cell membrane that are most difficult to overcome while those of bacterial cell membranes are conversely the least difficult[233]. Additionally, one third of bacterial proteins are associated with the cell membrane making them easy targets for AMPs as compared to proteins in mammalian cells. These characteristics make bacteria, especially Gram-positive that lack the outer protective lipopolysaccharide capsule, susceptible to even weak antimicrobials.

The net charge (+2) plus increased charged density due to cyclisation could create high adsorption and accumulation of peptides on bacterial surface[169] in a carpet like manner that then destabilises the membrane when enough molecules accumulate, finally causing membrane permeabilisation[132]. Additionally, the two hydrophobic tryptophan residues facilitate insertion of the peptide into the lipid matrix, as tryptophan has an affinity for cell membranes surfaces[176]. Photoditritide is macrocyclised via a lactam bond between D-Har^{1} and L-Trp^{6} which creates a cation-π interaction[236] between the aromatic ring of tryptophan and the cationic moieties of homoarginine, which results in shielding of the hydrophilic side chain facilitating longer interactions in the hydrophobic lipid matrix. Although this shielding effect would be less when compared to Arg residues because the cationic moieties of the longer Har side chain are more exposed[237], it is sufficient to enhance peptide insertion into the matrix, contributing either directly to inhibition by membrane disruption, or indirectly by facilitating passage of peptide to its cytoplasmic target. These support the hypothesis that macrocyclisation in Photoditritide may have contributed to its antibacterial activity.

Ishipeptides can form an α-helix secondary structure

A major group of AMPs are those that form α-helices upon interaction with cell membranes of target cells[232]. Formation of such increases antimicrobial activity due to

increased interaction with the cell membrane. A determinant of α-helix formation is peptide length with at least 7 amino acids required and higher propensities with increasing length are observed among analogues[232]. Ishipeptide length of 12 residues make it sufficient to form an α-helix. Another determinant is composite residues of the peptide with Ala, Leu, Val having the highest propensity for α-helix formation and Glu and Gln having a high propensity[179]. Ten/twelve composite residues of Ishipeptides are the aforementioned giving it a high propensity for α-helix formation. The arrangement of 1-2 hydrophobic residues for every 2-3 residues is indicative that a peptide forms α-helix[233] and this is seen in Ishipeptides with 4 hydrophobic (2x Leu, Val, Ala) for the first 6 residues, then 2(Leu, Val) for the next three and 2(Leu, Leu) for the last three.

At least 5 hydrophobic residues indicates a peptide may form an α-helix AMP as this increases the chance of formation of a hydrophobic face[170]. Ishipeptides have 8/12 hydrophobic residues making it likely to form a hydrophobic face of an α-helix as seen in proposed helical wheel (Figure S6). Perfect amphipathicity in an *α*-helix peptide would mean all hydrophobic and hydrophilic residues on their own faces. This would enable the hydrophobic face to interact with the non-polar lipid matrix and the hydrophilic face with the polar phospholipid heads and heterogeneous extracellular aqueous environment. The hydrophobic face is thus less tolerant of hydrophilic residues as it interacts only with lipids while the hydrophilic face can tolerant even 2-3 hydrophobic residues[232]. The incorporation of D amino acids among L amino acids reduces the helicity & stability of α-helices[233] and 2x D-Leu, D -Ala, D -Val may contribute to this. However, α-helix stabilisation as well as increased helicity is observed through cyclisation of side chains[238]. For example, GRF analogues with a side chain lactamation between Lys-Asp had higher local and general helicity than linear analogues[238]. In apolipoprotein E analogues[239], cyclisation between the i^{th} and (i+3th) residues, as is found in Ishipeptide A, stabilised the α-helix. These support the hypothesis that Ishipeptides can form an α-helix secondary structure.

CONCLUSION

In view of exacerbating antimicrobial resistance, nematophilic bacteria are a source of novel antimicrobials, in particular macrocyclic non-ribosomal peptides. Yet there are few studies on a complete discovery process –from nematophilic bacteria isolation to demonstrated antimicrobial peptides. Thus, four parts of this process were demonstrated by investigating 6 different nematophilic bacteria.

The first two parts of isolation of bacteria and species delineation was done on strains isolated from natural soil-dwelling nematodes of Western Kenya. *Steinernema* sp. VH1 and BG5 were isolated from red volcanic soil on cultivated land and clay soil on riverine land respectively. Their bacteria endosymbionts *Xenorhabdus* sp. strains VH1 and BG5 were delineated as *X. griffiniae* VH1 and *Xenorhabdus* sp. nov. BG5. Using *X. griffiniae* XN45 genome data, *X. griffiniae* BMMCB was emended to *Xenorhabdus* sp. nov. BMMCB. The nematode host of *X. griffiniae* XN45, *Steinernema* sp. scarpo was highlighted as a new species.

The third part was genome mining, chemical structure elucidation, chemosynthesis and biosynthesis. This was done by investigations on *ishA-B* gene cluster of *X. ishibashii.* The IshA-B proteins biosynthesised Ishipeptide A that was a very hydrophobic branch cyclic dodecadepsipeptide, macrolactoned between serine and the terminal glutamate. Ishipeptide B was its linearised derivative. Chemosynthesis routes for both were elucidated. Insights into the functions of IshA C_{start} domain revealed that more than the HH active site is necessary for *N*-terminal acylation, aiding future derivitisation of these peptides.

The fourth part of the process was antimicrobial susceptibility testing. This was done for both Ishipeptides and the macrocyclic hexapeptide Photoditritide from *P. temperata*. Only the macrocyclic peptides were inhibitory, with Photoditritide being more potent than a current clinical compound Ampicillin in inhibiting Gram-positive bacteria. Ishipeptide A weakly potent against Gram-positive bacteria. Discussions into physiochemical characteristics revealed that macrocyclisation contributed to antimicrobial activity for both, and further suggested how net charge, hydrophobicity, amphipathicity and helicity may have a role in antimicrobial activity. Such insights aid

the design of derivatives with potent antimicrobial activity: needed entities in a world plagued by microbial pandemics.

RECOMMENDATIONS

1) Description of novel species: *Xenorhabdus* sp. nov BG5, *Xenorhabdus* sp. nov BMMCB and *Steinernema* sp. scarpo
2) Investigation of the exact composition of core protein encoding genes of the pangenomes of *Xenorhabdus* genera and *Photorhabdus* genera.
3) Swapping the C_{start} domain of IshA protein with one with the highest percentage similarity– *X. eapokensis* C_{start} domain found on contig MKGQ01000042 from position 17,130bp.
4) Constitutive expression the gene encoding the putative *X. eapokensis* NRPS that begins with C_{start} domain found on contig MKGQ01000042 from position 17,130bp to investigate production of ishipeptide derivatives.
5) Determining the cellular target of Photoditritide by comparison of genomes of resistant *M. luteus* strains to susceptible strains.

SUPPLEMENTARY FIGURES AND TABLES

Figure S1. Photograph of soil collection point VH1 from which nematodes were isolated. Soil was collected from this point located in Vihiga County for investigation of presence entomopathogenic nematodes (EPN). After successfully isolating EPNs from this soil, their gut bacteria were also successfully isolated. These were identified as *X. griffiniae* VH1 and their respective nematode was thus named *Steinernema* sp. strain VH1.

Figure S2. Photograph of soil collection point BG5 from which nematodes were isolated. Soil was collected from this point located in Bungoma County for investigation of presence entomopathogenic nematodes (EPN). After successfully isolating EPNs from this soil, their gut bacteria were also successfully isolated. These were identified as *Xenorhabdus* sp. nov. BG5 and their respective nematode was thus named *Steinernema* sp. strain BG5.

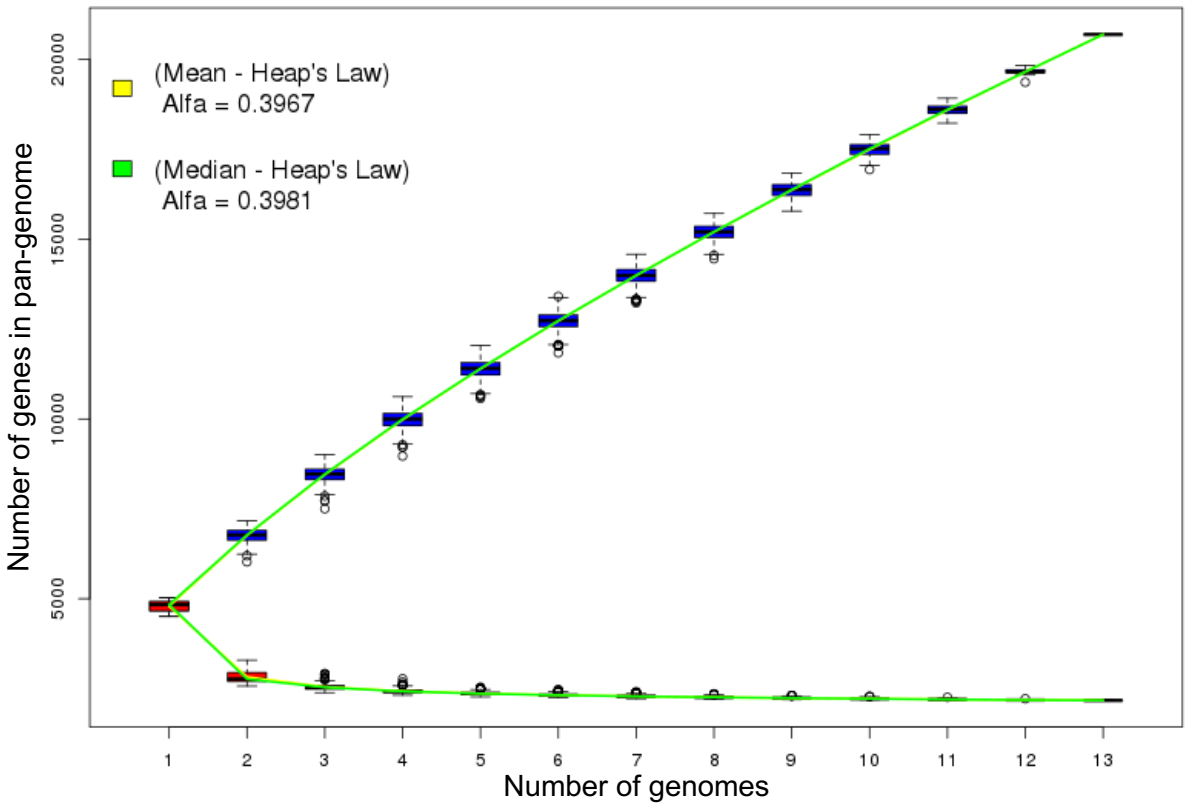


Figure S3. Box-plot graph depicting how α exponent of Heaps law was estimated for the *Photorhabdus* genus. Red plots and lower line depict the size and trend of the core genomes as more genomes are added to the pangenome. Blue plots and upper line depict the size and trend of the accessory genome as more genomes are added to the pan-genome. Pangenome was open because its mean α exponent of Heap's Law <1. This predicts that more species genomes need to be added to this genus so as to complete it. Pangenome analysis was determined by the Panweb workflow.

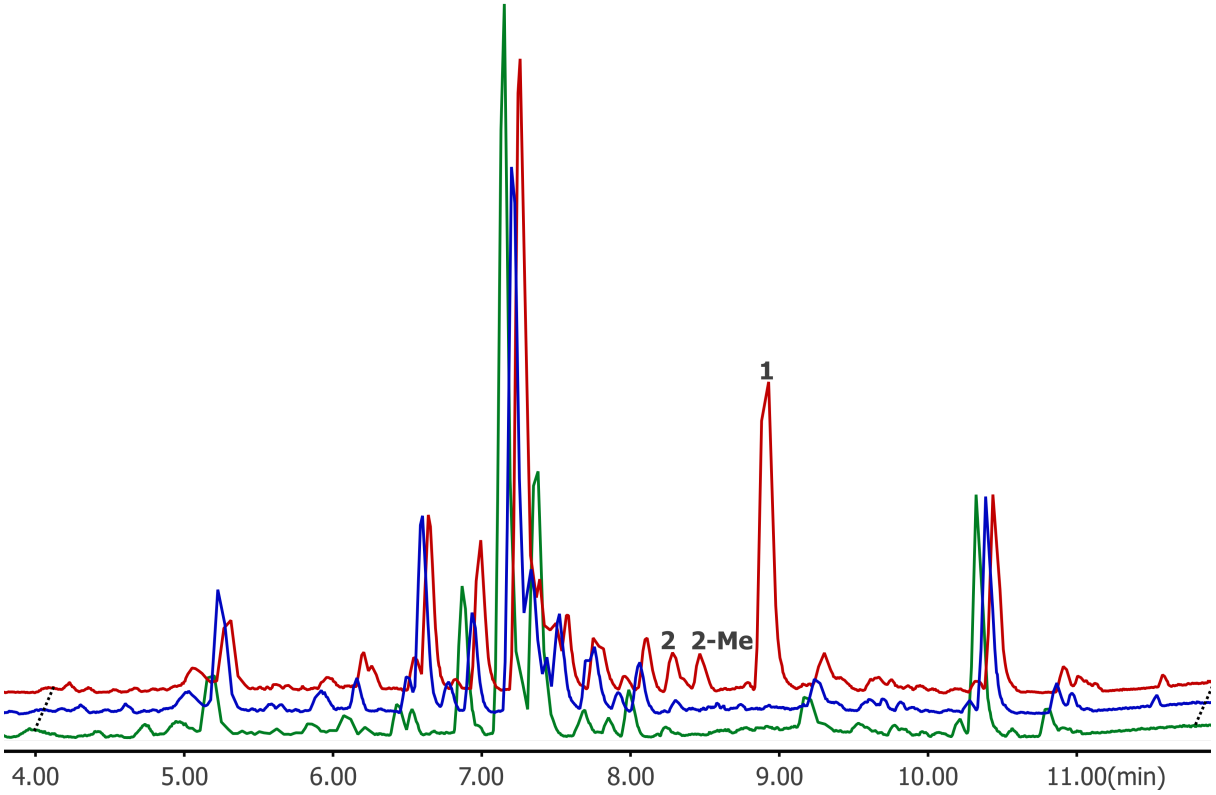


Figure S4. Extended base peak chromatograms (BPC) depicting two natural products **1** and **2** that resulted from induced expression of *ishAB* NRPS gene cluster. BPC were obtained from HPLC-ESI-HRMS analysis of methanol extracts of cell free supernatants from 72 h cultivated *X. ishibashii* pCEP_JC0094 mutant strain that had the native promoter for *ishA* gene refactored by exchanging it with a P_{BAD} promoter. To induce expression, the fermentation culture was supplemented with 0.2% L-arabinose (w/v) (red) resulting in detection of above products that were equally absent in extracts from *X. ishibashii* wild-type strain (green) and control strain (blue).

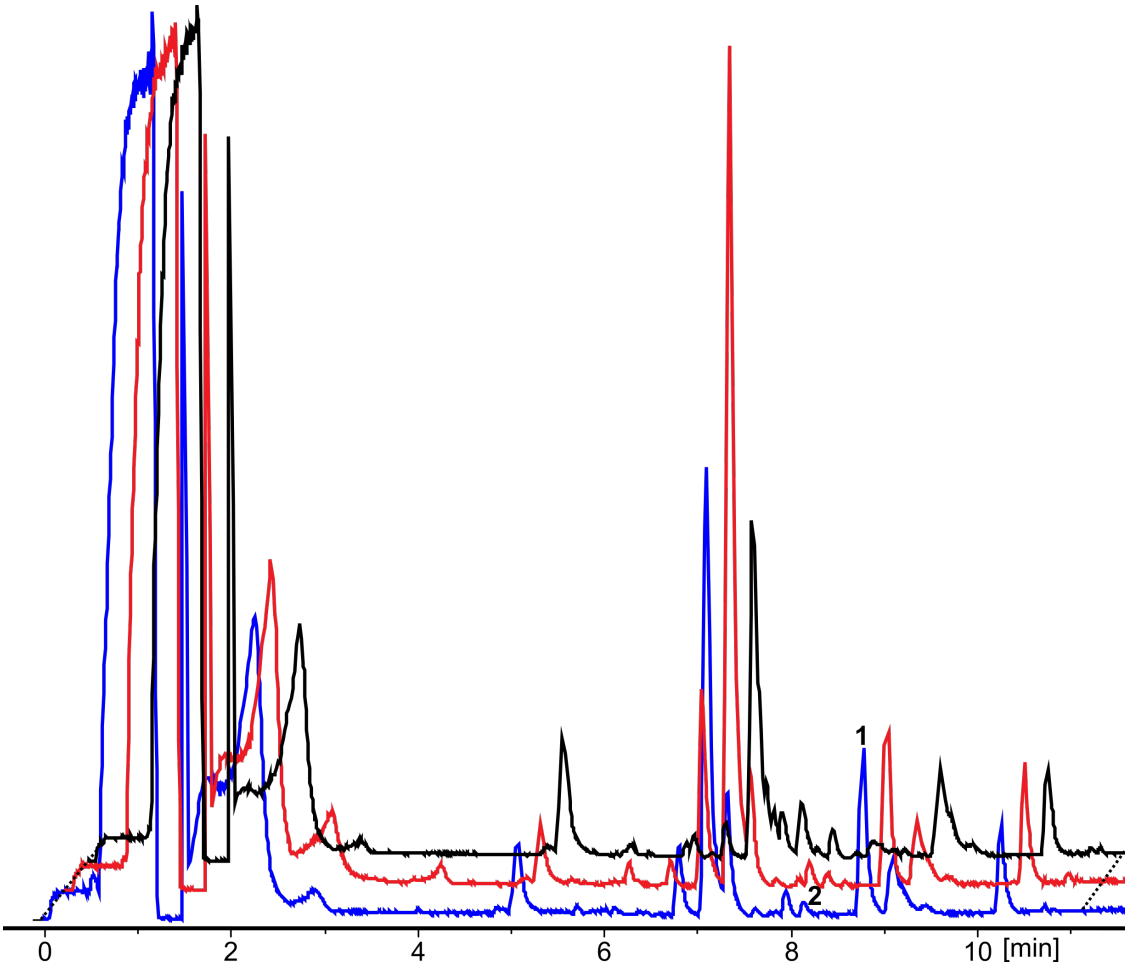


Figure S5. Base peak chromatograms (BPC) depicting two natural products **1** and **2** that resulted from induced expression of *ishAB* NRPS gene cluster. BPC were obtained from HPLC-ESI-HRMS analysis of methanol extracts of cell free supernatants from 72 h cultivated *X. ishibashii* pCEP_JC0094(red) and *X. ishibashii* pCEP_JCHIS143(blue) mutant strains that had the native promoter for *ishA* gene refactored by exchanging it with a P_{BAD} promoter. To induce expression, the fermentation culture was supplemented with 0.2% L-arabinose (w/v) (red&blue) resulting in detection of above products that were equally absent in extracts from *X. ishibashii* pCEP_JCHIS143 that were not induced(black). The *ishA* gene of *X. ishibashii* pCEP_JCHIS143 was mutated via site-directed mutagenesis to change the amino acid residue glutamine of position 143 to histidine so as to restore an active site within the C_{start} domain of the IshA protein, that is necessary for creating acylated derivatives. From the BPCs no new derivatives were produced by this mutant(blue) as compared to the one whose IshA protein was not altered at position 143. This indicates that an additional unknown mutation is required to restore the function of C_{start} domain of IshA protein.

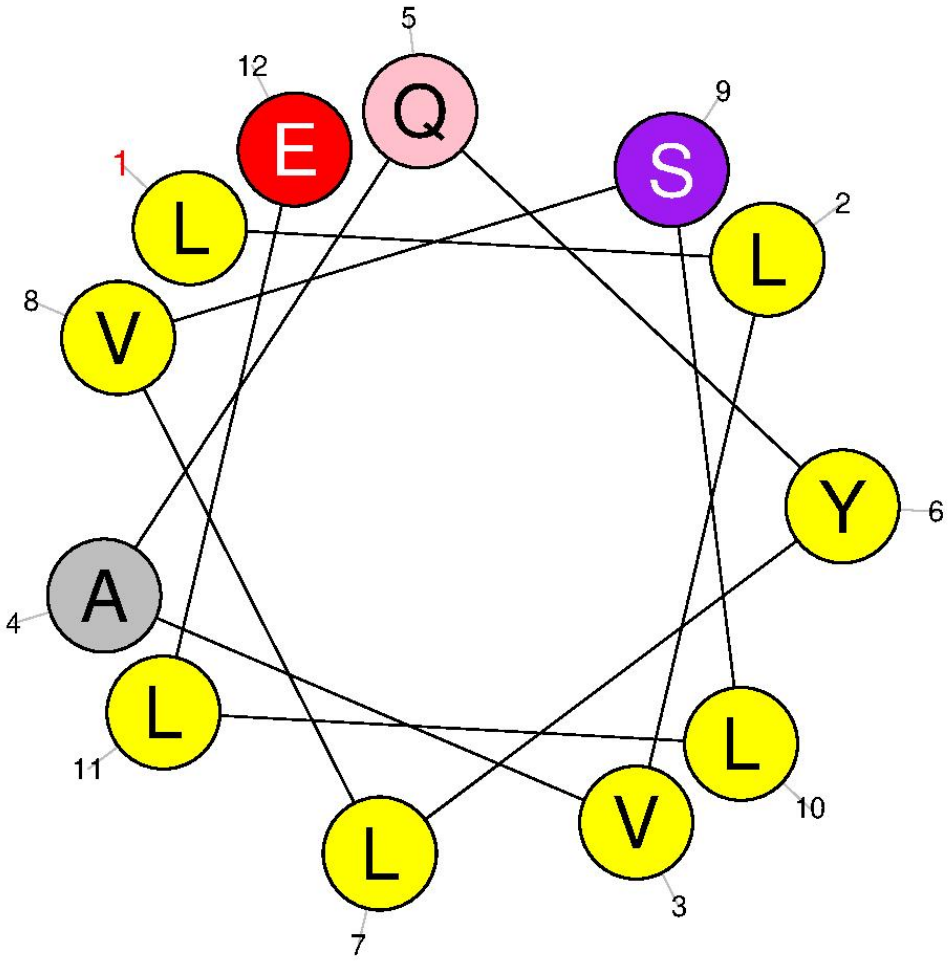


Figure S6. Proposed *α*-helix structure of Ishipeptides. Single letter amino acid notion is used. Yellow are hydrophobic residues while red, pink, and purple are hydrophilic residues. Proposed hydrophobic face is clockwise from position two to position eleven. Hydrophilic face is clockwise from position four to nine. This would create a weakly amphipathic peptide. Helical wheel was created with Heliquest.

Table S1: Soil samples names and their geographic regions of collection. Bacteria used in this dissertation were collected from embolden samples. Samples that had dead larvae after baiting are in red. The remaining samples that were investigated for entomopathogenic nematodes are in orange.

Sample name	Latitude	Longitude	Altitude(m)	Area description	Locality name if known
TN1	0.01971	35.44755	2485.7	Virgin forest	
TN3	0.01968	35.44765	2486.7	Virgin forest	
TN4	0.01968	35.44765	2486.7	Virgin forest	
TN5	0.01926	35.44772	2478.7	Virgin forest	
TN6	0.01908	35.44796	2478.7	Virgin forest	
TN7	0.01908	35.44804	2451.7	Virgin forest	
TN8	0.01673	35.4545	2595.7	Cultivated land with minimum disturbance	Coliby Forest
TN9	0.01662	35.45479	2613.7	Cultivated land with minimum disturbance	
TN10	0.01647	35.45539	2629.7	Cultivated land with minimum disturbance	
TN11	0.02258	35.46348	2773.6	Virgin forest	
TN12	0.02287	35.46344	2744.0	Virgin forest	
TN13	0.02223	35.46351	2694.0	Virgin forest	
TN14	0.02223	35.46341	2672.0	Virgin forest	Lorenge Forest
ND1	0.1053	35.19019	2045.0	Near tea cultivated land	
ND2	0.10535	35.19028	2062.0	Tea cultivated land	
ND3	0.10498	35.19019	2047.0	Near tea cultivated land	
ND4	0.10355	35.19836	2017.0	Eucalyptus cultivated land	Nandi Hills Tea Estate
ND5	0.10309	35.19858	1989.0	Eucalyptus cultivated land	Nandi Hills Tea Estate
ND6	0.10207	35.19084	2064.0	Tea cultivated land	Nandi Hills Tea Estate
ND7	0.10219	35.19069	2055.0	Tea cultivated land-1990 tea bush	Nandi Hills Tea Estate
ND8	0.0926	35.19466	1967.0	Eucalyptus cultivated land	Nandi Hills Tea Estate
ND9	0.0884	35.19389	1934.0	Riverine vegetation	Nandi Hills Tea Estate
ND10	0.08779	35.19403	1930.0	Riverine vegetation	Nandi Hills Tea Estate
ND11	0.08725	35.21023	2019.0	Tea cultivated land-1996 tea bush	Nandi Hills Tea Estate
ND12	0.08631	35.21016	2010.6	Indigenous trees cultivated land	Nandi Hills Tea Estate

ND13	0.0862	35.21019	1996.0	Indigenous trees cultivated land	Nandi Hills Tea Estate
ND14	0.08594	35.21032	1981.6	Riverine vegetation	Nandi Hills Tea Estate
KR1	-0.1957	35.28344	1422.7	Fallow land	
KR2	-0.19564	35.2839	1420.0	Fallow land	
KR3	-0.21374	35.31923	1429.4	Riverine vegetation	River Murgut
KR4	-0.21363	35.31924	1491.4	Riverine vegetation	River Murgut
KS1	-0.12548	34.7447	1143.0	River Wigia Inlet	Lake Victoria shores
KS2	-0.12547	34.7464	1143.0	River Wigia Inlet swamp area	Lake Victoria shores
KS4	-0.12565	34.74166	1117.0	Hippo Point	Lake Victoria shores
KS5	-0.12537	34.74238	1117.0	Hippo Point-River inlet swamp area	Lake Victoria shores
KS6	-0.12541	34.74238	1117.0	Hippo Point	Lake Victoria shores
KS7	-0.13592	34.74103	1133.0	Swamp Land	Lake Victoria shores
KS8	-0.08592	34.74627	1130.0	Nyanza Golf Club shoreline	Lake Victoria shores
KS9	-0.08596	34.74644	1130.0	Nyanza Golf Club shoreline	Lake Victoria shores
KS10	-0.08596	34.74654	1130.0	Nyanza Golf Club shoreline	Lake Victoria shores
KS11	-0.087818	34.74231	1139.0	Nyanza Golf Club shoreline	Lake Victoria shores
KS12	-0.08756	34.7432	1139.0	Nyanza Golf Club shoreline	Lake Victoria shores
KS13	-0.08761	34.74207	1139.0	Nyanza Golf Club shoreline	Lake Victoria shores
KS14	-0.08774	34.74140	1140.0	Nyanza Golf Club shoreline	Lake Victoria shores
GS1	0.0577	34.76411	1680.0	Cultivated land	
GS2	0.05824	34.76421	1680.0	Cultivated land-Maize	
GS3	0.05857	34.76486	1668.0	Eucalyptus cultivated land	
GS4	0.05895	34.76526	1634.6	Riverine vegetation	
GS5	0.05913	34.76493	1656.3	Man-made pond banks	
GS6	0.05623	34.77107	1719.6	Virgin land	Gisambai Rocks
VH1	**0.06293**	**34.72903**	**1624.0**	**Cultivated land-cabbage**	**Terry Amadi Farm**
VH2	0.0629	34.72911	1624.0	Cultivated land-cabbage	Terry Amadi Farm

VH3	0.06644	34.7316	1592.0	Virgin forest	River Idumu Forest
VH4	0.06637	34.73195	1592.0	Virgin forest	River Idumu Forest
ST1	0.12841	34.85427	1660.0	Virgin tropical rain forest	Kaimosi Tropical Rain forest
ST2	0.12972	34.8464	1657.0	Virgin tropical rain forest	Kaimosi Tropical Rain forest
ST3	0.12978	34.84423	1657.0	Virgin tropical rain forest	Kaimosi Tropical Rain forest
ST4	0.12968	34.84413	1610.0	Virgin tropical rain forest	Kakamega Forest Reserve
KG1	0.17356	34.89599	1627.0	Virgin tropical rain forest	Kakamega Forest Reserve
KG2	0.17382	34.8959	1627.0	Virgin tropical rain forest	Kakamega Forest Reserve
KG3	0.15393	34.8935	1618.0	Kakabuti Riverbed Sample	Kakamega Forest Reserve
KG4	0.15375	34.89375	1618.0	Virgin tropical rain forest	Kakamega Forest Reserve
BY1	0.48037	34.40839	1240.0	Virgin riverbank land	Lawrence Awori farm
BG5	**0.48044**	**34.40836**	**1239.0**	**Virgin riverbank land**	**Lawrence Awori farm**
BY2	0.48035	34.40849	1239.0	Virgin riverbank land	Lawrence Awori farm
BY3	0.48043	34.40851	1239.0	Virgin riverbank land	Lawrence Awori farm
BY4	0.48051	34.40845	1239.0	Virgin riverbank land	Lawrence Awori farm
BY5	0.48046	34.40828	1239.0	Virgin riverbank land	Lawrence Awori farm
BY6	0.48419	34.373222	1239.0	Cultivated land-Maize farm	
BY7	0.48466	34.37534	1230.0	Cultivated land-cassava	
BY8	0.4857	34.3763	1230.0	Fallow land	
BY9	0.48584	34.37654	1215.0	Riverine vegetation	
BY10	0.48575	34.37653	1229.0	Riverine vegetation	River Sianda
BY11	0.48937	34.37311	1229.0	Riverine vegetation	River Sianda
EL1	0.882906	34.731095	2149.0	Fallow land	
EL2	0.882906	34.731095	2149.0	Fallow land	

Table S2. Accessions numbers, strains and predicted gene counts of genomes used in phlyogenomic and pangenomic analyses

Genbank number of master-record for drafts/complete genome	Strain name	Number of predicted genes in Prodigal
This study	*X. griffiniae* XN45	3855
This study	*X. griffinae* VH1	3780
This study	*Xenorhabdus* sp. nov. BG5	3534
MUBK01000000	*X. beddingii*	3668
FN667741	*X. bovienii* ss-2004	3747
NIBS00000000.1	*X. budapestensis* DSM 16342	3756
NJGH00000000.1	*X. cabalinasii* JM26	4288
NZ_FO704550	*X. doucetiae* FRM16	3640
MKGQ00000000.1	*X. eapokenisis* DL20	3839
NIBT00000000.1	*X. ehlersii* DSM 16337	3727
NJAI00000000.1	*X. hominickii* DSM 17903	4937
NIBU00000000.1	*X. innexi* DSM 16336	4149
NJAK00000000.1	*X. ishibashii* DSM 22670	3556
FOVO00000000.1	*X. japonica* DSM 16522	3198
LFCV00000000.1	*X. khoisanae* MCB	4117
FPBJ00000000.1	*X. koppenhoeferi* DSM 18168	2770
NJCX00000000.1	*X. kozodoii* DSM 17907	3704
NITY00000000.1	*X. mauleonii* DSM 17908	4418
NITZ00000000.1	*X. miraniensis* DSM 17902	4313
LN681227	*X. nematophila* AN6/1	4123

FO704551	*X. poinarii* G6	3357
LDNM01000000.1	*Xenorhabdus* sp. nov. BMMCB	3951
NJAJ00000000.1	*X. stockiae* DSM 17904	4094
MKGR00000000.1	*X. thuogxuanensis* 30TX1	3533
MUBJ00000000.1	*X. vietnamensis* DSM 22392	3966

Photorhabdus genomes

LOIC00000000.1	*P. namnaonensis* PB45.5	4643
PUJV00000000.1	*P. stackerbranti* DSM 23271	4354
LJCS00000000.1	*P. heterohabditis* VMG	4317
FMWJ00000000.1	*P. luminescens* ATCC29999	4524
PUJW00000000.1	*P. cinerea* DSM19727	4542
PUJU00000000.1	*P. tasmaniensis* T327	4648
NSCI00000000.1	*P. laumondii* BOJ-47	4427
CP011104.1	*P thracensis* DSM15199	4545
JONO00000000.1	*P. australis* DSM17609	4247
RBLJ00000000.1	*P. asymbiotica* DSM15149	4389
AYSJ00000000.1	*P. khanii* NC-19	4548
NSCM00000000.1	*P. bodei* LJ24	4454

Table S3. Prophage associated genomic loci of *Xenorhabdus* sp. nov. BG5

Prophage loci name	Length	GC content	
Prophage 1	17,860	43.2	Significantly different from average GC content (p value= 0.005408171, T_{stat} = -2.548845258)
Prophage 2	86,072	43.12	Significantly different from average GC content (p value= 8.22695E-22, T_{stat} = -9.527857257)
Prophage 3	18,682	42.78	Significantly different from average GC content (p value=3.05511E-28, T_{stat} = -10.97542747)
Prophage 5	9,128	41.6	Significantly different from average GC content (p value= 9.1663E-102, T_{stat} = -21.65830959)
Prophage 4	11,829	47.25	Significantly different from average GC content (p value < 0.0001, T_{stat} = 64.31439891)
Prophage 6_GI3	29,967	43.52	Significantly different from average GC content (p value= 0.006, T_{stat} = 2.739173238)
Prophage 7_GI3	29,163	44.11	Significantly different from average GC content (p value< 0.0001, T_{stat} =14.74031146)

Table S4. Primers and plasmids used

Name	Sequence (5´- 3´)	Description
RMA87	GGCTAACAGGAGGC TAGCATATGATGCC TATGTCATGCAATAGTAG	Amplifies first 629bp of *ishA* gene with 23 bp overhang at 5' end (forward primer)
RMA88	CCGTTTAAACATTTAAA TCTGCAGGGAGGTAAG TTTTTGTAGCGAT	Amplifies first 629bp of *ishA* gene with 23 bp overhang at 5' end (Reverse primer)
RMA 88b	CCGTTTAAACATTT AAATCTGCAG CAATAAGATGA TGACAGCAAAACT	When used with RMA87 as the forward primer, it amplifies first 436bp of *ishA* gene with change of complementary base pair for

		position 429(G) to that of T. It has 23 bp overhang at 5' end (Reverse primer)
V_PCEP_FW	GCTATGCCATAGC ATTTTTATCCATAAG	Amplifies a 863 bp fragment from plasmid pCEP_JC0094 (forward primer)
PDS132_RV	ACATGTGGAATT GTGAGCGG	Amplifies a 863 bp fragment from plasmid pCEP_JC0094 (reverse primer)
ALPCEP1F	CCACTGGTGATA CCATTCGC	Amplifies a 1226 bp fragment only found in genomes of successfully modified *X. ishibashii* pCEP_JC0094 strains (forward primer)
RMAr032	ATAAGCTGGT TTCGGCACG	Amplifies a 1226 bp fragment only found in genomes of successfully modified *X. ishibashii* pCEP_JC0094 strains (reverse primer)
RMA_JC1F	TAGCCCAACCC ATGCCG	Amplifies locus 14026..14309 of *ishA* gene (284bp) (forward primer)
RMA_JC1R	CCGGATCGCTT GTTGTGGA	Amplifies locus 14026..14309 of *ishA* gene (284bp) (reverse primer)
RMA_JC2F	TGCAGGAAACC TATCAAAATCTG	Amplifies locus 20377..21792 of *ishA* gene (1416bp) (forward primer)

RMA_JC2R	TGGCAATCGGGCGACC	Amplifies locus 20377..21792 of *ishA* gene (1657bp) (reverse primer)
RMA_NJC2F	TGGATAACCCCGAACCGC	Amplifies locus 21196..153 of *ishA-ishB* genes (1657bp) (forward primer)
RMA_JC3R	TTGTTGGGCAA AAGAGAGTGG	Amplifies 1656bp of *ishA* gene with (reverse primer)
RMA_JC4F	TCATCAGGGCGCCAGATG	Amplifies locus 12073..12699 of *ishB* gene (627bp) (forward primer)
RMA_JC4R	ATTCGGGGTCAGCGGG	Amplifies locus 12073..12699 of *ishB* gene (627bp) (reverse primer)
pCEP plasmid	Integrative plasmid with a P_{BAD} promoter system and a kanamycin resistance cassette.	[204]
pCEP_JC0094 plasmid	pCEP plasmid that contains first 629bp of *ishA* gene, which is also the target locus for genomic integration, directly downstream of P_{BAD} promoter.	This study
pCEP_JCHIS143 plasmid	pCEP plasmid that contains first 436bp of *ishA* gene, which is also the target locus for genomic integration, directly downstream of P_{BAD}	This study

	promoter. It has a mutation at position 429 of the *IshA* gene, from G to T	

Table S5. Amino acid specificities of A domains of IshA-B

Position	Stachelhaus Code	Detected amino acid specificity	Predicted amino acid specificity
A domain 1	DAWlLGaVcK	Leu	Leu
A domain 2	DAWlLGaVcK	Leu	Leu
A domain 3	DAwWLGGTFK	Val	Val
A domain 4	DLYNNALTYK	Ala	Ala
A domain 5	DasnIGEVGK	Gln	Asn
A domain 6	DASTVAAVCK	Tyr	Tyr
A domain 7	DAWlLGaVcK	Leu	Leu
A domain 8	DAwWLGGTFK	Val	Val
A domain 9	DVWHLSLIDK	Ser	Ser
A domain 10	DAWlLGaVcK	Leu	Leu
A domain 11	DAWlLGaVcK	Leu	Leu
A domain 12	DAvDLGVVDK	Glu	Gln

REFERENCES

(1) World Health Organisation. *Antimicrobial Resistance: Global Report on Surveillance*; World Health Organization, **2014**.

(2) Nieuwlaat, R.; Mbuagbaw, L.; Mertz, D.; Burrows, L.; Bowdish, D. M.; Moja, L.; Wright, G. D.; Schünemann, H. J. COVID-19 and Antimicrobial Resistance: Parallel and Interacting Health Emergencies. *Clinical Infectious Diseases* **2020**.

(3) Murray, A. K. The Novel Coronavirus COVID-19 Outbreak: Global Implications for Antimicrobial Resistance. *Frontiers in Microbiology* **2020**, *11*, 1020. https://doi.org/10.3389/fmicb.2020.01020.

(4) Rawson, T. M.; Ming, D.; Ahmad, R.; Moore, L. S.; Holmes, A. H. Antimicrobial Use, Drug-Resistant Infections and COVID-19. *Nature Reviews Microbiology* **2020**, 1–2.

(5) Prevention, C. for D. C. and. *Antibiotic Resistance Threats in the United States, 2019. Atlanta, GA: US Department of Health and Human Services, Centers for Disease Control and Prevention; 2019*; **2019**.

(6) Collignon, P.; Beggs, J. J. CON: COVID-19 Will Not Result in Increased Antimicrobial Resistance Prevalence. *JAC-Antimicrobial Resistance* **2020**, *2* (dlaa051). https://doi.org/10.1093/jacamr/dlaa051.

(7) Lansbury, L.; Lim, B.; Baskaran, V.; Lim, W. S. Co-Infections in People with COVID-19: A Systematic Review and Meta-Analysis. *Journal of Infection* **2020**. https://doi.org/10.1016/j.jinf.2020.05.046.

(8) Spigaglia, P. COVID-19 and *Clostridioides difficile* Infection (CDI): Possible Implications for Elderly Patients. *Anaerobe* **2020**, *64*, 102233. https://doi.org/10.1016/j.anaerobe.2020.102233.

(9) Shinnick, T. M.; Starks, A. M.; Alexander, H. L.; Castro, K. G. Evaluation of the Cepheid Xpert MTB/RIF Assay. *Expert review of molecular diagnostics* **2015**, *15* (1), 9–22.

(10) Moran, A.; Beavis, K. G.; Matushek, S. M.; Ciaglia, C.; Francois, N.; Tesic, V.; Love, N. The Detection of SARS-CoV-2 Using the Cepheid Xpert Xpress SARS-CoV-2 and Roche Cobas SARS-CoV-2 Assays. *Journal of Clinical Microbiology* **2020**.

(11) O'Neill, J. Antimicrobial Resistance: Tackling a Crisis for the Health and Wealth of Nations. Review on Antimicrobial Resistance. *Review on Antimicrobial Resistance, London, United Kingdom: https://amr-review. org/sites/default/files/AMR% 20Review% 20Paper* **2014**.

(12) Clarke, D. J. *Photorhabdus:* A Tale of Contrasting Interactions. *Microbiology* **2020**, micro000907.

(13) Abebe, E.; Akele, F.-A.; Morrison, J.; Cooper, V.; Thomas, W. K. An Insect Pathogenic Symbiosis between a *Caenorhabditis* and *Serratia*. *Virulence* **2011**, *2* (2), 158–158.

(14) Bowen, D.; Blackburn, M.; Rocheleau, T. A.; Andreev, O.; Golubeva, E.; ffrench-Constant, R. H. Insecticidal Toxins from the Bacterium *Photorhabdus luminescens*: Gene Cloning and Toxin Histopathology. *Pesticide Science* **1999**, *55* (6), 666–668. https://doi.org/10.1002/(sici)1096-9063(199906)55:6<666::aid-ps5>3.0.co;2-g.

(15) Waterfield, N. An ABC Guide to the Bacterial Toxin Complexes. *Advances in applied microbiology* **2005**, *58*, 169–183.

(16) Ogier, J.-C.; Pagès, S.; Frayssinet, M.; Gaudriault, S. Entomopathogenic Nematode-Associated Microbiota: From Monoxenic Paradigm to Pathobiome. *Microbiome* **2020**, *8* (1), 25. https://doi.org/10.1186/s40168-020-00800-5.

(17) Shan, S.; Wang, W.; Song, C.; Wang, M.; Sun, B.; Li, Y.; Fu, Y.; Gu, X.; Ruan, W.; Rasmann, S. The Symbiotic Bacteria *Alcaligenes faecalis* of the Entomopathogenic Nematodes *Oscheius* Spp. Exhibit Potential Biocontrol of Plant- and Entomopathogenic Fungi. *Microbial Biotechnology* **2019**, *12* (3), 459–471. https://doi.org/10.1111/1751-7915.13365.

(18) Petersen, L. M.; Tisa, L. S. Friend or Foe? A Review of the Mechanisms That Drive *Serratia* towards Diverse Lifestyles. *Canadian Journal of Microbiology* **2013**, *59* (9), 627–640. https://doi.org/10.1139/cjm-2013-0343.

(19) Torres-Barragan, A.; Suazo, A.; Buhler, W. G.; Cardoza, Y. J. Studies on the Entomopathogenicity and Bacterial Associates of the Nematode *Oscheius carolinensis*. *Biological control* **2011**, *59* (2), 123–129.

(20) van den Hoogen, J.; Geisen, S.; Routh, D.; Ferris, H.; Traunspurger, W.; Wardle, D. A.; de Goede, R. G. M.; Adams, B. J.; Ahmad, W.; Andriuzzi, W. S.; Bardgett, R. D.;

Bonkowski, M.; Campos-Herrera, R.; Cares, J. E.; Caruso, T.; de Brito Caixeta, L.; Chen, X.; Costa, S. R.; Creamer, R.; Mauro da Cunha Castro, J.; Dam, M.; Djigal, D.; Escuer, M.; Griffiths, B. S.; Gutiérrez, C.; Hohberg, K.; Kalinkina, D.; Kardol, P.; Kergunteuil, A.; Korthals, G.; Krashevska, V.; Kudrin, A. A.; Li, Q.; Liang, W.; Magilton, M.; Marais, M.; Martín, J. A. R.; Matveeva, E.; Mayad, E. H.; Mulder, C.; Mullin, P.; Neilson, R.; Nguyen, T. A. D.; Nielsen, U. N.; Okada, H.; Rius, J. E. P.; Pan, K.; Peneva, V.; Pellissier, L.; Carlos Pereira da Silva, J.; Pitteloud, C.; Powers, T. O.; Powers, K.; Quist, C. W.; Rasmann, S.; Moreno, S. S.; Scheu, S.; Setälä, H.; Sushchuk, A.; Tiunov, A. V.; Trap, J.; van der Putten, W.; Vestergård, M.; Villenave, C.; Waeyenberge, L.; Wall, D. H.; Wilschut, R.; Wright, D. G.; Yang, J.; Crowther, T. W. Soil Nematode Abundance and Functional Group Composition at a Global Scale. *Nature* **2019**, *572* (7768), 194–198. https://doi.org/10.1038/s41586-019-1418-6.

(21) Plichta, K. L.; Joyce, S. A.; Clarke, D.; Waterfield, N.; Stock, S. P. *Heterorhabditis gerrardi* n. Sp. (Nematoda: *Heterorhabditidae*): The Hidden Host of *Photorhabdus asymbiotica* (Enterobacteriaceae: γ-Proteobacteria). *Journal of Helminthology* **2009**, *83* (4), 309–320.

(22) Adeolu, M.; Alnajar, S.; Naushad, S.; S. Gupta, R. Genome-Based Phylogeny and Taxonomy of the 'Enterobacteriales': Proposal for *Enterobacterales* ord. nov. Divided into the Families *Enterobacteriaceae*, *Erwiniaceae* fam. nov., *Pectobacteriaceae* fam. nov., *Yersiniaceae* fam. nov., *Hafniaceae* fam. nov., *Morganellaceae* fam. nov., and *Budviciaceae* fam. Nov. *International Journal of Systematic and Evolutionary Microbiology* **2016**, *66* (12), 5575–5599. https://doi.org/10.1099/ijsem.0.001485.

(23) Akhurst, R. J.; Boemare, N. E. A Numerical Taxonomic Study of the Genus *Xenorhabdus* (Enterobacteriaceae) and Proposed Elevation of the Subspecies of *X. nematophilus* to Species. *Microbiology* **1988**, *134* (7), 1835–1845.

(24) AKHURST, R. J. Taxonomic Study of *Xenorhabdus*, a Genus of Bacteria Symbiotically Associated with Insect Pathogenic Nematodes. *International Journal of Systematic and Evolutionary Microbiology,* 1983, *33*, 38–45.

(25) Lengyel, K.; Lang, E.; Fodor, A.; Szállás, E.; Schumann, P.; Stackebrandt, E. Description of Four Novel Species of *Xenorhabdus,* Family Enterobacteriaceae:

Xenorhabdus budapestensis sp. nov., *Xenorhabdus ehlersii* sp. nov., *Xenorhabdus innexi* sp. nov., and *Xenorhabdus szentirmaii* sp. nov. *Systematic and applied microbiology* **2005**, *28* (2), 115–122.

(26) Tailliez, P.; Pages, S.; Ginibre, N.; Boemare, N. New Insight into Diversity in the Genus *Xenorhabdus,* Including the Description of Ten Novel Species. *International Journal of Systematic and Evolutionary Microbiology* **2006**, *56* (12), 2805–2818.

(27) Fischer-Le Saux, M.; Mauléon, H.; Constant, P.; Brunel, B.; Boemare, N. PCR-Ribotyping of *Xenorhabdus* and *Photorhabdus* Isolates from the Caribbean Region in Relation to the Taxonomy and Geographic Distribution of Their Nematode Hosts. *Appl Environ Microbiol* **1998**, *64* (11), 4246–4254.

(28) Waturu, C. N.; Hunt, D. J.; Reid, A. P. *Steinernema karii* Sp. n.(Nematoda: Steinernematidae), a New Entomopathogenic Nematode from Kenya. *International Journal of Nematology* **1997**, *7* (1), 68–75.

(29) Sudershan, G.; Singh, L. K. Steinernema Thermophilum Sp. n.(Rhabditida: Steinernematidae) from India. *International Journal of Nematology* **2000**, *10* (2), 183–191.

(30) Nishimura, Y.; Hagiwara, A.; Suzuki, T.; Yamanaka, S. *Xenorhabdus japonicus* sp. nov. Associated with the Nematode Steinernema Kushidai. *World Journal of Microbiology and Biotechnology* **1994**, *10* (2), 207–210.

(31) Kuwata, R.; Qiu, L.; Wang, W.; Harada, Y.; Yoshida, M.; Kondo, E.; Yoshiga, T. *Xenorhabdus ishibashii* sp. nov., Isolated from the Entomopathogenic Nematode *Steinernema aciari*. *International journal of systematic and evolutionary microbiology* **2013**, *63* (5), 1690–1695.

(32) Qiu, L.; Yan, X.; Zhou, Y.; Nguyen, K. B.; Pang, Y. *Steinernema aciari* Sp. n. (Nematoda: *Steinernematidae*), a New Entomopathogenic Nematode from Guangdong, China. *Journal of Invertebrate Pathology* **2005**, *88* (1), 58–69. https://doi.org/10.1016/j.jip.2004.09.004.

(33) Ferreira, T.; Van Reenen, C. A.; Endo, A.; Spröer, C.; Malan, A. P.; Dicks, L. M. T. Description of *Xenorhabdus khoisanae* sp. nov., the Symbiont of the Entomopathogenic Nematode *Steinernema khoisanae*. *International Journal of Systematic and Evolutionary Microbiology* **2013**, *63* (9), 3220–3224.

(34) Bedding, R. A.; Akhurst, R. J. A Simple Technique for the Detection of Insect Paristic Rhabditid Nematodes in Soil. *Nematologica* **1975**, *21* (1), 109–110.

(35) Tailliez, P.; Pagès, S.; Edgington, S.; Tymo, L. M.; Buddie, A. G. Description of Xenorhabdus Magdalenensis Sp. Nov., the Symbiotic Bacterium Associated with Steinernema Australe. *International journal of systematic and evolutionary microbiology* **2012**, *62* (8), 1761–1765.

(36) Poinar, G. O.; Weiser, J.; Hansen, E. L.; Yarwood, E. A. Clarification of the Status of the DD-136 Strain of *Neoaplectana carpocapsae* Weiser. *Nematologica* **1972**, *18* (2), 288–290.

(37) Poinar, G. O. Generation Polymorphism in *Neoaplectana glaseri* Steiner (Steinernematidae: Nematoda), Redescribed From *Strigoderma arboricola* (Fab.) Scarabaeidae: Coleoptera) in North Carolina. *Nematologica* **1978**, *24* (1), 105–114. https://doi.org/10.1163/187529278X00137.

(38) Tailliez, P.; Laroui, C.; Ginibre, N.; Paule, A.; Pagès, S.; Boemare, N. Phylogeny of *Photorhabdus* and *Xenorhabdus* Based on Universally Conserved Protein-Coding Sequences and Implications for the Taxonomy of These Two Genera. Proposal of New Taxa: *X. vietnamensis* sp. nov., *P. luminescens* subsp. Caribbeanensis subsp. nov., P. l. *International journal of systematic and evolutionary microbiology* **2010**, *60* (8), 1921–1937.

(39) Phan, K. L.; Nguyen, N. C.; Moens, M. *Steinernema sangi* Sp n.(Rabditida: Steinernematidae) from Vietnam. *Russian Journal of Nematology* **2001**, *9* (1), 1–7.

(40) Kämpfer, P.; Tobias, N. J.; Ke, L. P.; Bode, H. B.; Glaeser, S. P. *Xenorhabdus thuongxuanensis* sp. nov. and *Xenorhabdus eapokensis* sp. nov., Isolated from *Steinernema* Species. *International Journal of Systematic and Evolutionary Microbiology* **2017**, *67* (5), 1107–1114. https://doi.org/10.1099/ijsem.0.001770.

(41) Zhang, C.-X.; Yang, S.-Y.; Xu, M.-X.; Sun, J.; Liu, H.; Liu, J.-R.; Liu, H.; Kan, F.; Sun, J.; Lai, R. *Serratia nematodiphila* sp. nov., Associated Symbiotically with the Entomopathogenic Nematode *Heterorhabditidoides chongmingensis* (Rhabditida: Rhabditidae). *International journal of systematic and evolutionary microbiology* **2009**, *59* (7), 1603–1608.

(42) Zhang, C.; Liu, J.; Xu, M.; Sun, J.; Yang, S.; An, X.; Gao, G.; Lin, M.; Lai, R.; He, Z.; Wu, Y.; Zhang, K. *Heterorhabditidoides chongmingensis* gen. nov., sp. nov. (Rhabditida: Rhabditidae), a Novel Member of the Entomopathogenic Nematodes. *Journal of Invertebrate Pathology* **2008**, *98* (2), 153–168. https://doi.org/10.1016/j.jip.2008.02.011.

(43) Ye, W.; Torres-Barragan, A.; Cardoza, Y. J. *Oscheius carolinensis* n. Sp. (Nematoda: Rhabditidae), a Potential Entomopathogenic Nematode from Vermicompost. *Nematology* **2010**, *12* (1), 121–135. https://doi.org/10.1163/156854109X458464.

(44) Serepa-Dlamini, M. H.; Gray, V. M. A New Species of Entomopathogenic Nematode *Oscheius safricana* n. Sp. (Nematoda: Rhabditidae) from South Africa. *Archives of Phytopathology and Plant Protection* **2018**, *51* (5–6), 309–321. https://doi.org/10.1080/03235408.2018.1475281.

(45) Lephoto, T. E.; Featherston, J.; Gray, V. M. Draft Whole-Genome Sequence of *Serratia* sp. Strain TEL, Associated with *Oscheius* sp. TEL-2014 (Nematoda: Rhabditidae) Isolated from a Grassland in South Africa. *Genome Announc.* **2015**, *3* (4), e00747-15.

(46) Zhou, G.; Yang, H.; Wang, F.; Bao, H.; Wang, G.; Hou, X.; Lin, J.; Yedid, G.; Zhang, K. *Oscheius microvilli* n. Sp.(Nematoda: Rhabditidae): A Facultatively Pathogenic Nematode from Chongming Island, China. *Journal of nematology* **2017**, *49* (1), 33.

(47) Abebe, E.; Jumba, M.; Bonner, K.; Gray, V.; Morris, K.; Thomas, W. K. An Entomopathogenic *Caenorhabditis briggsae*. *Journal of Experimental Biology* **2010**, *213* (18), 3223–3229.

(48) Machado, R. A. R.; Wüthrich, D.; Kuhnert, P.; Arce, C. C. M.; Thönen, L.; Ruiz, C.; Zhang, X.; Robert, C. A. M.; Karimi, J.; Kamali, S.; Ma, J.; Bruggmann, R.; Erb, M. Whole-Genome-Based Revisit of *Photorhabdus* Phylogeny: Proposal for the Elevation of Most *Photorhabdus* Subspecies to the Species Level and Description of One Novel Species *Photorhabdus bodei* sp. nov., and One Novel Subspecies *Photorhabdus laumondii* subsp. c. *International Journal of Systematic and Evolutionary Microbiology* **2018**, *68* (8), 2664–2681. https://doi.org/10.1099/ijsem.0.002820.

(49) Akhurst, R. J. Use of Starch Gel Electrophoresis in the Taxonomy of the Genus *Heterorhabditis* (Nematoda: Heterorhabditidae). *Nematologica* **1987**, *33* (1), 1–9. https://doi.org/10.1163/187529287X00155.

(50) An, R.; Grewal, P. S. *Photorhabdus luminescens* subsp. Kleinii Subsp. Nov.(Enterobacteriales: Enterobacteriaceae). *Current microbiology* **2011**, *62* (2), 539–543.

(51) Glaeser, S. P.; Tobias, N. J.; Thanwisai, A.; Chantratita, N.; Bode, H. B.; Kämpfer, P. *Photorhabdus luminescens* Subsp. *namnaonensis* Subsp. Nov., Isolated from *Heterorhabditis baujardi* Nematodes. *International Journal of Systematic and Evolutionary Microbiology* **2017**, *67* (4), 1046–1051. https://doi.org/10.1099/ijsem.0.001761.

(52) Malan, A. P.; Knoetze, R.; Moore, S. D. Isolation and Identification of Entomopathogenic Nematodes from Citrus Orchards in South Africa and Their Biocontrol Potential against False Codling Moth. *Journal of Invertebrate Pathology* **2011**, *108* (2), 115–125. https://doi.org/10.1016/j.jip.2011.07.006.

(53) Hazir, S.; Keskin, N.; Stock, S. P.; Kaya, H. K.; Özcan, S. Diversity and Distribution of Entomopathogenic Nematodes (Rhabditida: Steinernematidae and Heterorhabditidae) in Turkey. *Biodiversity & Conservation* **2003**, *12* (2), 375–386. https://doi.org/10.1023/A:1021915903822.

(54) Grewal, P. S.; Grewal, S. K.; Malik, V. S.; Klein, M. G. Differences in Susceptibility of Introduced and Native White Grub Species to Entomopathogenic Nematodes from Various Geographic Localities. *Biological Control* **2002**, *24* (3), 230–237. https://doi.org/10.1016/S1049-9644(02)00025-7.

(55) Akhurst, R. J.; Boemare, N. E.; Janssen, P. H.; Peel, M. M.; Alfredson, D. A.; Beard, C. E. Taxonomy of Australian Clinical Isolates of the Genus Photorhabdus and Proposal of *Photorhabdus asymbiotica* subsp. *asymbiotica* subsp. nov. and *P. asymbiotica* subsp. *australis* subsp. nov. *International Journal of Systematic and Evolutionary Microbiology* **2004**, *54* (4), 1301–1310.

(56) Farmer, J. J.; Jorgensen, J. H.; Grimont, P. A.; Akhurst, R. J.; Poinar, G. O.; Ageron, E.; Pierce, G. V.; Smith, J. A.; Carter, G. P.; Wilson, K. L. *Xenorhabdus*

luminescens (DNA Hybridization Group 5) from Human Clinical Specimens. *J. Clin. Microbiol.* **1989**, *27* (7), 1594.

(57) Tóth, T.; Lakatos, T. *Photorhabdus temperata* subsp. *cinerea* subsp. nov., Isolated from *Heterorhabditis* Nematodes. *International journal of systematic and evolutionary microbiology* **2008**, *58* (11), 2579–2581.

(58) THOMAS, G. M.; POINAR JR, G. O. *Xenorhabdus* gen. nov., a Genus of Entomopathogenic, Nematophilic Bacteria of the Family *Enterobacteriaceae*. *International Journal of Systematic and Evolutionary Microbiology* **1979**, *29* (4), 352–360.

(59) Naidoo, S.; Mothupi, B.; Featherston, J.; Mpangase, P. T.; Gray, V. M. Draft Genome Sequence and Assembly of *Photorhabdus heterorhabditis* Strain VMG, a Bacterial Symbiont Associated with the Entomopathogenic Nematode *Heterorhabditis zealandica*. *Genome announcements* **2015**, *3* (5).

(60) Mothupi, B. Genomics of Entomopathogenic Bacterial Endosymbiont Species Associated with Desiccation Tolerant Entomopathogenic Nematode, 2016.

(61) Khan, A.; Brooks, W. M.; Hirschmann, H. *Chromonema heliothidis* n. gen., n. sp. (Steinernematidae, Nematoda), a Parasite of *Heliothis zea* (Noctuidae, Lepidoptera), and Other Insects. *J Nematol* **1976**, *8* (2), 159–168.

(62) Chaston, J. M.; Suen, G.; Tucker, S. L.; Andersen, A. W.; Bhasin, A.; Bode, E.; Bode, H. B.; Brachmann, A. O.; Cowles, C. E.; Cowles, K. N. The Entomopathogenic Bacterial Endosymbionts *Xenorhabdus* and *Photorhabdus*: Convergent Lifestyles from Divergent Genomes. *PloS one* **2011**, *6* (11).

(63) Sajnaga, E.; Kazimierczak, W. Evolution and Taxonomy of Nematode-Associated Entomopathogenic Bacteria of the Genera *Xenorhabdus* and *Photorhabdus*: An Overview. *Symbiosis* **2020**, *80* (1), 1–13. https://doi.org/10.1007/s13199-019-00660-0.

(64) Nguyen, K. B. Species of *Steinernema* http://entnemdept.ufl.edu/nguyen/morph/steinsp1.htm (accessed Jul 25, 2020).

(65) Pantoja, Y.; Pinheiro, K.; Veras, A.; Araújo, F.; Lopes de Sousa, A.; Guimarães, L. C.; Silva, A.; Ramos, R. T. J. PanWeb: A Web Interface for Pan-Genomic Analysis. *PLOS ONE* **2017**, *12* (5), e0178154–e0178154.

(66) Medini, D.; Donati, C.; Rappuoli, R.; Tettelin, H. The Pangenome: A Data-Driven Discovery in Biology. In *The Pangenome*; Springer, Cham, 2020; pp 3–20.

(67) Chun, J.; Oren, A.; Ventosa, A.; Christensen, H.; Arahal, D. R.; da Costa, M. S.; Rooney, A. P.; Yi, H.; Xu, X.-W.; De Meyer, S.; Trujillo, M. E. Proposed Minimal Standards for the Use of Genome Data for the Taxonomy of Prokaryotes. *International Journal of Systematic and Evolutionary Microbiology*, 2018, *68*, 461–466. https://doi.org/10.1099/ijsem.0.002516.

(68) Mothupi, B.; Featherston, J.; Gray, V. Draft Whole-Genome Sequence and Annotation of *Xenorhabdus griffiniae* Strain BMMCB Associated with the South African Entomopathogenic Nematode *Steinernema khoisanae* Strain BMMCB. *Genome Announc* **2015**, *3* (4). https://doi.org/10.1128/genomeA.00785-15.

(69) Awori, R. M.; Ng'ang'a, P. N.; Nyongesa, L. N.; Amugune, N. O.; Masiga, D. Mursamacin: A Novel Class of Antibiotics from Soil-Dwelling Roundworms of Central Kenya That Inhibits Methicillin-Resistant *Staphylococcus aureus*. *F1000Research* **2017**, *5*. https://doi.org/10.12688/f1000research.9652.2.

(70) Kim, M.; Oh, H.-S.; Park, S.-C.; Chun, J. Towards a Taxonomic Coherence between Average Nucleotide Identity and 16S RRNA Gene Sequence Similarity for Species Demarcation of Prokaryotes. *International Journal of Systematic and Evolutionary Microbiology,* 2014, *64*, 346–351.

(71) Dreyer, J.; Malan, A. P.; Dicks, L. M. T. First Report of a Symbiotic Relationship between *Xenorhabdus griffiniae* and an Unknown *Steinernema* from South Africa. *Archives of Microbiology* **2018**, *200* (2), 349–353. https://doi.org/10.1007/s00203-017-1452-4.

(72) Richter, M.; Rosselló-Móra, R. Shifting the Genomic Gold Standard for the Prokaryotic Species Definition. *Proc Natl Acad Sci U S A* **2009**, *106* (45), 19126–19131. https://doi.org/10.1073/pnas.0906412106.

(73) Bobay, L.-M. The Prokaryotic Species Concept and Challenges. In *The Pangenome*; Springer, Cham, 2020; pp 21–49.

(74) Bobay, L.-M.; Ochman, H. Biological Species Are Universal across Life's Domains. *Genome Biology and Evolution* **2017**, *9* (3), 491–501. https://doi.org/10.1093/gbe/evx026.

(75) Moldovan, M. A.; Gelfand, M. S. Pangenomic Definition of Prokaryotic Species and the Phylogenetic Structure of *Prochlorococcus* Spp. *Frontiers in Microbiology* **2018**, *9* (MAR), 1–11. https://doi.org/10.3389/fmicb.2018.00428.

(76) Poinar Jr, G. O.; Leutenegger, R. Anatomy of the Infective and Normal Third-Stage Juveniles of Neoaplectana Carpocapsae Weiser (Steinernematidae: Nematoda). *The Journal of parasitology* **1968**, 340–350.

(77) Bedding, R. A.; Molyneux, A. S. Penetration of Insect Cuticle by Infective Juveniles of *Heterorhabditis* Spp.(Heterorhabditidae: Nematoda). *Nematologica* **1982**, *28* (3), 354–359.

(78) Poinar, G. O. J.; Grewal, P. S. History of Entomopathogenic Nematology. *J Nematol* **2012**, *44* (2), 153–161.

(79) Ciche, T. The Biology and Genome of *Heterorhabditis bacteriophora*. In *WormBook: The Online Review of C. elegans Biology [Internet]*; WormBook, **2007**.

(80) Heryanto, C.; Eleftherianos, I. Nematode Endosymbiont Competition: Fortune Favors the Fittest. *Molecular and Biochemical Parasitology* **2020**, *238*, 111298. https://doi.org/10.1016/j.molbiopara.2020.111298.

(81) Stilwell, M. D.; Cao, M.; Goodrich-Blair, H.; Weibel, D. B. Studying the Symbiotic Bacterium *Xenorhabdus nematophila* in Individual, Living *Steinernema carpocapsae* Nematodes Using Microfluidic Systems. *Msphere* **2018**, *3* (1).

(82) Waterfield, N. R.; Ciche, T.; Clarke, D. *Photorhabdus* and a Host of Hosts. *Annual review of microbiology* **2009**, *63*, 557–574.

(83) Crawford, J. M.; Portmann, C.; Zhang, X.; Roeffaers, M. B. J.; Clardy, J. Small Molecule Perimeter Defense in Entomopathogenic Bacteria. *Proc Natl Acad Sci USA* **2012**, *109* (27), 10821. https://doi.org/10.1073/pnas.1201160109.

(84) Fodor, A.; Fodor, A.; Forst, S. Comparative Analysis of Antibacterial Activities of *Xenorhabdus* Species on Related and Non-Related Bacteria in Vivo. *Journal of Microbiology and Antimicrobials* **2010**, *2* (4), 36–46.

(85) Ullah, I.; Khan, A. L.; Ali, L.; Khan, A. R.; Waqas, M.; Lee, I.-J.; Shin, J.-H. An Insecticidal Compound Produced by an Insect-Pathogenic Bacterium Suppresses Host Defenses through Phenoloxidase Inhibition. *Molecules* **2014**, *19* (12). https://doi.org/10.3390/molecules191220913.

(86) Song, C. J.; Seo, S.; Shrestha, S.; Kim, Y. Bacterial Metabolites of an Entomopathogenic Bacterium, *Xenorhabdus nematophila*, Inhibit a Catalytic Activity of Phenoloxidase of the Diamondback Moth, *Plutella xylostella*. *J Microbiol Biotechnol* **2011**, *21* (3), 317–322.

(87) Eugenia Nuñez-Valdez, M.; Lanois, A.; Pagès, S.; Duvic, B.; Gaudriault, S. Inhibition of *Spodoptera frugiperda* Phenoloxidase Activity by the Products of the *Xenorhabdus* Rhabduscin Gene Cluster. *PLOS ONE* **2019**, *14* (2), e0212809. https://doi.org/10.1371/journal.pone.0212809.

(88) Eleftherianos, I.; Boundy, S.; Joyce, S. A.; Aslam, S.; Marshall, J. W.; Cox, R. J.; Simpson, T. J.; Clarke, D. J.; ffrench-Constant, R. H.; Reynolds, S. E. An Antibiotic Produced by an Insect-Pathogenic Bacterium Suppresses Host Defenses through Phenoloxidase Inhibition. *Proc Natl Acad Sci USA* **2007**, *104* (7), 2419. https://doi.org/10.1073/pnas.0610525104.

(89) Sergeant, M.; Jarrett, P.; Ousley, M.; Morgan, J. A. W. Interactions of Insecticidal Toxin Gene Products from *Xenorhabdus nematophilus* PMFI296. *Appl. Environ. Microbiol.* **2003**, *69* (6), 3344. https://doi.org/10.1128/AEM.69.6.3344-3349.2003.

(90) Sheets, J. J.; Hey, T. D.; Fencil, K. J.; Burton, S. L.; Ni, W.; Lang, A. E.; Benz, R.; Aktories, K. Insecticidal Toxin Complex Proteins from *Xenorhabdus nematophilus*: STRUCTURE AND PORE FORMATION. *Journal of Biological Chemistry* **2011**, *286* (26), 22742–22749. https://doi.org/10.1074/jbc. M111.227009.

(91) Blackburn, M.; Golubeva, E.; Bowen, D.; Ffrench-constant, R. H. TCa of *Photorhabdus luminescens*. **1998**, *64* (8), 1–6.

(92) Vigneux, F.; Zumbihl, R.; Jubelin, G.; Ribeiro, C.; Poncet, J.; Baghdiguian, S.; Givaudan, A.; Brehélin, M. The XaxAB Genes Encoding a New Apoptotic Toxin from the Insect Pathogen *Xenorhabdus nematophila* Are Present in Plant and Human Pathogens. *Journal of Biological Chemistry* **2007**, *282* (13), 9571–9580. https://doi.org/10.1074/jbc.M604301200.

(93) Schubert, E.; Vetter, I. R.; Prumbaum, D.; Penczek, P. A.; Raunser, S. Membrane Insertion of α-Xenorhabdolysin in near-Atomic Detail. *eLife* **2018**, *7*, e38017. https://doi.org/10.7554/eLife.38017.

(94) Dowling, A. J.; Daborn, P. J.; Waterfield, N. R.; Wang, P.; Streuli, C. H.; Ffrench-Constant, R. H. The Insecticidal Toxin Makes Caterpillars Floppy (Mcf) Promotes Apoptosis in Mammalian Cells. *Cellular microbiology* **2004**, *6* (4), 345–353.

(95) Jank, T.; Lang, A. E.; Aktories, K. Rho-Modifying Bacterial Protein Toxins from *Photorhabdus* Species. *Toxicon* **2016**, *116*, 17–22. https://doi.org/10.1016/j.toxicon.2015.05.017.

(96) Yang, Q.; Zhang, J.; Li, T.; Liu, S.; Song, P.; Nangong, Z.; Wang, Q. PirAB Protein from *Xenorhabdus nematophila* HB310 Exhibits a Binary Toxin with Insecticidal Activity and Cytotoxicity in *Galleria mellonella*. *Journal of invertebrate pathology* **2017**, *148*, 43–50.

(97) Böszörményi, E.; Érsek, T.; Fodor, A.; Fodor, A. M.; Földes, L. S.; Hevesi, M.; Hogan, J. S.; Katona, Z.; Klein, M. G.; Kormány, A. Isolation and Activity of *Xenorhabdus* Antimicrobial Compounds against the Plant Pathogens *Erwinia amylovora* and *Phytophthora nicotianae*. *Journal of Applied Microbiology* **2009**, *107* (3), 746–759.

(98) Houard, J.; Aumelas, A.; Noël, T.; Pages, S.; Givaudan, A.; Fitton-Ouhabi, V.; Villain-Guillot, P.; Gualtieri, M. Cabanillasin, a New Antifungal Metabolite, Produced by Entomopathogenic *Xenorhabdus cabanillasii* JM26. *The Journal of antibiotics* **2013**, *66* (10), 617–620.

(99) Xiao, Y.; Meng, F.; Qiu, D.; Yang, X. Two Novel Antimicrobial Peptides Purified from the Symbiotic Bacteria *Xenorhabdus budapestensis* NMC-10. *Peptides* **2012**, *35* (2), 253–260. https://doi.org/10.1016/j.peptides.2012.03.027.

(100) Yang, X.; Qiu, D.; Yang, H.; Liu, Z.; Zeng, H.; Yuan, J. Antifungal Activity of Xenocoumacin 1 from *Xenorhabdus nematophilus* Var. Pekingensis against *Phytophthora infestans*. *World Journal of Microbiology and Biotechnology* **2011**, *27* (3), 523–528.

(101) Zhou, T.; Yang, X.; Qiu, D.; Zeng, H. Inhibitory Effects of Xenocoumacin 1 on the Different Stages of *Phytophthora capsici* and Its Control Effect on *Phytophthora* Blight of Pepper. *BioControl* **2017**, *62* (2), 151–160.

(102) Imai, Y.; Meyer, K. J.; Iinishi, A.; Favre-Godal, Q.; Green, R.; Manuse, S.; Caboni, M.; Mori, M.; Niles, S.; Ghiglieri, M. A New Antibiotic Selectively Kills Gram-Negative Pathogens. *Nature* **2019**, *576* (7787), 459–464.

(103) Xi, X.; Lu, X.; Zhang, X.; Bi, Y.; Li, X.; Yu, Z. Two Novel Cyclic Depsipeptides Xenematides F and G from the Entomopathogenic Bacterium *Xenorhabdus budapestensis*. *Journal of Antibiotics* **2019**, *72* (10), 736–743. https://doi.org/10.1038/s41429-019-0203-y.

(104) Lang, G.; Kalvelage, T.; Peters, A.; Wiese, J.; Imhoff, J. F. Linear and Cyclic Peptides from the Entomopathogenic Bacterium *Xenorhabdus nematophilus*. *Journal of natural products* **2008**, *71* (6), 1074–1077.

(105) Reimer, D.; Pos, K. M.; Thines, M.; Grün, P.; Bode, H. B. A Natural Prodrug Activation Mechanism in Non-ribosomal Peptide Synthesis. *Nature chemical biology* **2011**, *7* (12), 888–890.

(106) Zhao, L.; Awori, R. M.; Kaiser, M.; Groß, J.; Opatz, T.; Bode, H. B. Structure, Biosynthesis, and Bioactivity of Photoditritide from *Photorhabdus temperata* Meg1. *Journal of Natural Products* **2019**, *82* (12), 3499–3503. https://doi.org/10.1021/acs.jnatprod.9b00932.

(107) Givaudan, A.; Gualtieri, M.; Villain-Guillot, P.; Pages, S. Nemaucin, an antibiotic produced by entomopathogenic *Xenorhabdus cabanillasii* **2012**, WIPO PCT PATENT WO2012085177A1

(108) Grundmann, F.; Kaiser, M.; Kurz, M.; Schiell, M.; Batzer, A.; Bode, H. B. Structure Determination of the Bioactive Depsipeptide Xenobactin from *Xenorhabdus* Sp. PB30.3. *RSC Advances* **2013**, *3* (44), 22072–22077. https://doi.org/10.1039/c3ra44721a.

(109) Pantel, L.; Florin, T.; Dobosz-Bartoszek, M.; Racine, E.; Sarciaux, M.; Serri, M.; Houard, J.; Campagne, J. M.; de Figueiredo, R. M.; Midrier, C.; Gaudriault, S.; Givaudan, A.; Lanois, A.; Forst, S.; Aumelas, A.; Cotteaux-Lautard, C.; Bolla, J. M.; Vingsbo Lundberg, C.; Huseby, D. L.; Hughes, D.; Villain-Guillot, P.; Mankin, A. S.; Polikanov, Y. S.; Gualtieri, M. Odilorhabdins, Antibacterial Agents That Cause Miscoding by Binding at a New Ribosomal Site. *Molecular Cell* **2018**, *70* (1), 83-94.e7. https://doi.org/10.1016/j.molcel.2018.03.001.

(110) McInerney, B. V.; Taylor, W. C.; Lacey, M. J.; Akhurst, R. J.; Gregson, R. P. Biologically Active Metabolites from *Xenorhabdus* Spp., Part 2. Benzopyran-1-One Derivatives with Gastro-protective Activity. *Journal of Natural Products* **1991**, *54* (3), 785–795.

(111) Rhodes, S. H.; Lyons, G. R.; Gregson, R. P.; Akhurst, R. J.; Lacey, M. J. *Process of Making Xenorhabdin Antibiotics*; Google Patents, **1987**.

(112) Gualtieri, M.; Aumelas, A.; Thaler, J.-O. Identification of a New Antimicrobial Lysine-Rich Cyclolipopeptide Family from *Xenorhabdus nematophila*. *The Journal of antibiotics* **2009**, *62* (6), 295–302.

(113) Fuchs, S. W.; Proschak, A.; Jaskolla, T. W.; Karas, M.; Bode, H. B. Structure Elucidation and Biosynthesis of Lysine-Rich Cyclic Peptides in *Xenorhabdus nematophila*. *Organic and Biomolecular Chemistry* **2011**, *9* (9), 3130–3132. https://doi.org/10.1039/c1ob05097d.

(114) Zhao, L.; Vo, T. D.; Kaiser, M.; Bode, H. B. Phototemtide A, a Cyclic Lipopeptide Heterologously Expressed from *Photorhabdus temperata* Meg1, Shows Selective Antiprotozoal Activity. *ChemBioChem* **2020**.

(115) Zhou, Q.; Grundmann, F.; Kaiser, M.; Schiell, M.; Gaudriault, S.; Batzer, A.; Kurz, M.; Bode, H. B. Structure and Biosynthesis of Xenoamicins from Entomopathogenic Xenorhabdus. *Chemistry - A European Journal* **2013**, *19* (49), 16772–16779. https://doi.org/10.1002/chem.201302481.

(116) Ohlendorf, B.; Simon, S.; Wiese, J.; Imhoff, J. F. Szentiamide, an N-Formylated Cyclic Depsipeptide from *Xenorhabdus szentirmaii* DSM 16338 T. *Natural Product Communications* **2011**, *6* (9), 1247–1250. https://doi.org/10.1177/1934578x1100600909.

(117) Schimming, O.; Fleischhacker, F.; Nollmann, F. I.; Bode, H. B. Yeast Homologous Recombination Cloning Leading to the Novel Peptides Ambactin and Xenolindicin. *ChemBioChem* **2014**, *15* (9), 1290–1294. https://doi.org/10.1002/cbic.201402065.

(118) Grundmann, F.; Kaiser, M.; Schiell, M.; Batzer, A.; Kurz, M.; Thanwisai, A.; Chantratita, N.; Bode, H. B. Antiparasitic Chaiyaphumines from Entomopathogenic

Xenorhabdus Sp. PB61.4. *Journal of Natural Products* **2014**, *77* (4), 779–783. https://doi.org/10.1021/np4007525.

(119) Zhao, L.; Kaiser, M.; Bode, H. B. Rhabdopeptide/Xenortide-like Peptides from *Xenorhabdus innexi* with Terminal Amines Showing Potent Antiprotozoal Activity. *Organic Letters* **2018**, *20* (17), 5116–5120. https://doi.org/10.1021/acs.orglett.8b01975.

(120) Reimer, D.; Nollmann, F. I.; Schultz, K.; Kaiser, M.; Bode, H. B. Xenortide Biosynthesis by Entomopathogenic *Xenorhabdus nematophila*. *Journal of Natural Products* **2014**, *77* (8), 1976–1980. https://doi.org/10.1021/np500390b.

(121) Wenski, S. L.; Cimen, H.; Berghaus, N.; Fuchs, S. W.; Hazir, S.; Bode, H. B. Fabclavine Diversity in *Xenorhabdus* Bacteria. *Beilstein Journal of Organic Chemistry* **2020**, *16* (1), 956–965.

(122) Bi, Y.; Gao, C.; Yu, Z. Rhabdopeptides from *Xenorhabdus budapestensis* SN84 and Their Nematicidal Activities against Meloidogyne Incognita. *Journal of agricultural and food chemistry* **2018**, *66* (15), 3833–3839.

(123) Reimer, D.; Cowles, K. N.; Proschak, A.; Nollmann, F. I.; Dowling, A. J.; Kaiser, M.; Goodrich-Blair, H.; Bode, H. B. Rhabdopeptides as Insect-specific Virulence Factors from Entomopathogenic Bacteria. *ChemBioChem* **2013**, *14* (15), 1991–1997.

(124) Proschak, A.; Schultz, K.; Herrmann, J.; Dowling, A. J.; Brachmann, A. O.; ffrench-Constant, R.; Müller, R.; Bode, H. B. Cytotoxic Fatty Acid Amides from *Xenorhabdus*. *ChemBioChem* **2011**, *12* (13), 2011–2015.

(125) Gulcu, B.; Hazir, S.; Lewis, E. E.; Kaya, H. K. Evaluation of Responses of Different Ant Species (Formicidae) to the Scavenger Deterrent Factor Associated with the Entomopathogenic Nematode-Bacterium Complex. *European Journal of Entomology* **2018**, *115*.

(126) Gulcu, B.; Hazir, S.; Kaya, H. K. Scavenger Deterrent Factor (SDF) from Symbiotic Bacteria of Entomopathogenic Nematodes. *Journal of invertebrate pathology* **2012**, *110* (3), 326–333.

(127) Zhou, X.; Kaya, H. K.; Heungens, K.; Goodrich-Blair, H. Response of Ants to a Deterrent Factor (s) Produced by the Symbiotic Bacteria of Entomopathogenic Nematodes. *Applied and Environmental Microbiology* **2002**, *68* (12), 6202–6209.

(128) Heungens, K.; Cowles, C. E.; Goodrich-Blair, H. Identification of Xenorhabdus *nematophila* Genes Required for Mutualistic Colonization of *Steinernema carpocapsae* Nematodes. *Molecular Microbiology* **2002**, *45* (5), 1337–1353. https://doi.org/10.1046/j.1365-2958.2002.03100.x.

(129) Smith, S. H.; Jayawickreme, C.; Rickard, D. J.; Nicodeme, E.; Bui, T.; Simmons, C.; Coquery, C. M.; Neil, J.; Pryor, W. M.; Mayhew, D.; Rajpal, D. K.; Creech, K.; Furst, S.; Lee, J.; Wu, D.; Rastinejad, F.; Willson, T. M.; Viviani, F.; Morris, D. C.; Moore, J. T.; Cote-Sierra, J. Tapinarof Is a Natural AhR Agonist That Resolves Skin Inflammation in Mice and Humans. *Journal of Investigative Dermatology* **2017**, *137* (10), 2110–2119. https://doi.org/10.1016/j.jid.2017.05.004.

(130) Zhao, M.; Lepak, A. J.; Marchillo, K.; VanHecker, J.; Andes, D. R. In Vivo Pharmacodynamic Characterization of a Novel Odilorhabdin Antibiotic, NOSO-502, against *Escherichia coli* and *Klebsiella pneumoniae* in a Murine Thigh Infection Model. *Antimicrobial agents and chemotherapy* **2018**, *62* (9).

(131) Gold, L. S.; Bhatia, N.; Tallman, A. M.; Rubenstein, D. S. A Phase IIb, Randomized Clinical Trial of Tapinarof Cream for the Treatment of Plaque Psoriasis: Secondary Efficacy and Patient-Reported Outcomes. *Journal of the American Academy of Dermatology* **2020**. https://doi.org/10.1016/j.jaad.2020.04.181.

(132) Chen, C. H.; Lu, T. K. Development and Challenges of Antimicrobial Peptides for Therapeutic Applications. *Antibiotics (Basel)* **2020**, *9* (1). https://doi.org/10.3390/antibiotics9010024.

(133) Walsh, C. T.; O'Brien, R. V.; Khosla, C. Nonproteinogenic Amino Acid Building Blocks for Nonribosomal Peptide and Hybrid Polyketide Scaffolds. *Angew Chem Int Ed Engl* **2013**, *52* (28), 7098–7124. https://doi.org/10.1002/anie.201208344.

(134) Rausch, C.; Hoof, I.; Weber, T.; Wohlleben, W.; Huson, D. H. Phylogenetic Analysis of Condensation Domains in NRPS Sheds Light on Their Functional Evolution. *BMC Evolutionary Biology* **2007**, *7*, 1–15. https://doi.org/10.1186/1471-2148-7-78.

(135) Finking, R.; Marahiel, M. A. Biosynthesis of Nonribosomal Peptides. *Annual Review of Microbiology* **2004**, *58* (1), 453–488. https://doi.org/10.1146/annurev.micro.58.030603.123615.

(136) Li, Y. F.; Tsai, K. J. S.; Harvey, C. J. B.; Li, J. J.; Ary, B. E.; Berlew, E. E.; Boehman, B. L.; Findley, D. M.; Friant, A. G.; Gardner, C. A.; Gould, M. P.; Ha, J. H.; Lilley, B. K.; McKinstry, E. L.; Nawal, S.; Parry, R. C.; Rothchild, K. W.; Silbert, S. D.; Tentilucci, M. D.; Thurston, A. M.; Wai, R. B.; Yoon, Y.; Aiyar, R. S.; Medema, M. H.; Hillenmeyer, M. E.; Charkoudian, L. K. Comprehensive Curation and Analysis of Fungal Biosynthetic Gene Clusters of Published Natural Products. *Fungal Genetics and Biology* **2016**, *89*, 18–28. https://doi.org/10.1016/j.fgb.2016.01.012.

(137) Ziemert, N.; Podell, S.; Penn, K.; Badger, J. H.; Allen, E.; Jensen, P. R. The Natural Product Domain Seeker NaPDoS: A Phylogeny Based Bioinformatic Tool to Classify Secondary Metabolite Gene Diversity. *PLOS ONE* **2012**, *7* (3), e34064. https://doi.org/10.1371/journal.pone.0034064.

(138) Johnson, L. S.; Eddy, S. R.; Portugaly, E. Hidden Markov Model Speed Heuristic and Iterative HMM Search Procedure. *BMC Bioinformatics* **2010**, *11* (1), 431. https://doi.org/10.1186/1471-2105-11-431.

(139) Medema, M. H.; Blin, K.; Cimermancic, P.; de Jager, V.; Zakrzewski, P.; Fischbach, M. A.; Weber, T.; Takano, E.; Breitling, R. AntiSMASH: Rapid Identification, Annotation and Analysis of Secondary Metabolite Biosynthesis Gene Clusters in Bacterial and Fungal Genome Sequences. *Nucleic Acids Research* **2011**, *39* (suppl_2), W339–W346. https://doi.org/10.1093/nar/gkr466.

(140) Skinnider, M. A.; Dejong, C. A.; Rees, P. N.; Johnston, C. W.; Li, H.; Webster, A. L. H.; Wyatt, M. A.; Magarvey, N. A. Genomes to Natural Products PRediction Informatics for Secondary Metabolomes (PRISM). *Nucleic Acids Res* **2015**, *43* (20), 9645–9662. https://doi.org/10.1093/nar/gkv1012.

(141) Prieto, C.; García-Estrada, C.; Lorenzana, D.; Martín, J. F. NRPSsp: Non-Ribosomal Peptide Synthase Substrate Predictor. *Bioinformatics* **2011**, *28* (3), 426–427. https://doi.org/10.1093/bioinformatics/btr659.

(142) Blin, K.; Shaw, S.; Steinke, K.; Villebro, R.; Ziemert, N.; Lee, S. Y.; Medema, M. H.; Weber, T. AntiSMASH 5.0: Updates to the Secondary Metabolite Genome Mining Pipeline. *Nucleic Acids Research* **2019**, *47* (W1), W81–W87. https://doi.org/10.1093/nar/gkz310.

(143) Marahiel, M. A.; Essen, L.-O. Chapter 13 Nonribosomal Peptide Synthetases: Mechanistic and Structural Aspects of Essential Domains. In *Complex Enzymes in Microbial Natural Product Biosynthesis, Part A: Overview Articles and Peptides*; Methods in Enzymology; Academic Press, 2009; Vol. 458, pp 337–351. https://doi.org/10.1016/S0076-6879(09)04813-7.

(144) Stachelhaus, T.; Mootz, H. D.; Marahiel, M. A. The Specificity-Conferring Code of Adenylation Domains in Nonribosomal Peptide Synthetases. *Chemistry and Biology* **1999**, *6* (8), 493–505. https://doi.org/10.1016/S1074-5521(99)80082-9.

(145) Döhren, H. von; Dieckmann, R.; Pavela-Vrancic, M. The Nonribosomal Code. *Chemistry & Biology* **1999**, *6* (10), R273–R279. https://doi.org/10.1016/S1074-5521(00)80014-9.

(146) McErlean, M.; Overbay, J.; Van Lanen, S. Refining and Expanding Nonribosomal Peptide Synthetase Function and Mechanism. *Journal of Industrial Microbiology and Biotechnology* **2019**, *46* (3–4), 493–513. https://doi.org/10.1007/s10295-018-02130-w.

(147) Donmez Ozkan, H.; Cimen, H.; Ulug, D.; Wenski, S.; Yigit Ozer, S.; Telli, M.; Aydin, N.; Bode, H. B.; Hazir, S. Nematode-Associated Bacteria: Production of Antimicrobial Agent as a Presumptive Nominee for Curing Endodontic Infections Caused by *Enterococcus faecalis*. *Frontiers in Microbiology* **2019**, *10*, 2672. https://doi.org/10.3389/fmicb.2019.02672.

(148) Winn, M.; Fyans, J. K.; Zhuo, Y.; Micklefield, J. Recent Advances in Engineering Non-ribosomal Peptide Assembly Lines. *Nat Prod Rep* **2016**, *33* (2), 317–347. https://doi.org/10.1039/c5np00099h.

(149) Belshaw, P. J.; Walsh, C. T.; Stachelhaus, T. Aminoacyl-CoAs as Probes of Condensation Domain Selectivity in Nonribosomal Peptide Synthesis. *Science* **1999**, *284* (5413), 486. https://doi.org/10.1126/science.284.5413.486.

(150) P. Muller. Glossary of Terms Used in Physical Organic Chemistry (IUPAC Recommendations 1994). *Pure and Applied Chemistry* **1994**, *66* (5), 1077–1184. https://doi.org/10.1351/pac199466051077.

(151) Gao, L.; Liu, H.; Ma, Z.; Han, J.; Lu, Z.; Dai, C.; Lv, F.; Bie, X. Translocation of the Thioesterase Domain for the Redesign of Plipastatin Synthetase. *Scientific Reports* **2016**, *6* (1), 38467. https://doi.org/10.1038/srep38467.

(152) Miao, V.; Coëffet-LeGal, M.-F.; Brian, P.; Brost, R.; Penn, J.; Whiting, A.; Martin, S.; Ford, R.; Parr, I.; Bouchard, M.; Silva, C. J.; Wrigley, S. K.; Baltz, R. H. Daptomycin Biosynthesis in Streptomyces Roseosporus: Cloning and Analysis of the Gene Cluster and Revision of Peptide Stereochemistry. *Microbiology* **2005**, *151* (Pt 5), 1507–1523. https://doi.org/10.1099/mic.0.27757-0.

(153) Choi, S.-K.; Park, S.-Y.; Kim, R.; Kim, S.-B.; Lee, C.-H.; Kim, J. F.; Park, S.-H. Identification of a Polymyxin Synthetase Gene Cluster of *Paenibacillus polymyxa* and Heterologous Expression of the Gene in *Bacillus subtilis*. *J. Bacteriol.* **2009**, *191* (10), 3350. https://doi.org/10.1128/JB.01728-08.

(154) Tambadou, F.; Caradec, T.; Gagez, A.-L.; Bonnet, A.; Sopéna, V.; Bridiau, N.; Thiéry, V.; Didelot, S.; Barthélémy, C.; Chevrot, R. Characterization of the Colistin (Polymyxin E1 and E2) Biosynthetic Gene Cluster. *Arch Microbiol* **2015**, *197* (4), 521–532. https://doi.org/10.1007/s00203-015-1084-5.

(155) Hwang, S.; Lee, N.; Cho, S.; Palsson, B.; Cho, B.-K. Repurposing Modular Polyketide Synthases and Non-Ribosomal Peptide Synthetases for Novel Chemical Biosynthesis. *Frontiers in Molecular Biosciences* **2020**, *7*, 87. https://doi.org/10.3389/fmolb.2020.00087.

(156) Bozhüyük, K. A. J.; Fleischhacker, F.; Linck, A.; Wesche, F.; Tietze, A.; Niesert, C. P.; Bode, H. B. De Novo Design and Engineering of Non-Ribosomal Peptide Synthetases. *Nature Chemistry* **2018**, *10* (3), 275–281. https://doi.org/10.1038/NCHEM.2890.

(157) Heckman, K. L.; Pease, L. R. Gene Splicing and Mutagenesis by PCR-Driven Overlap Extension. *Nature protocols* **2007**, *2* (4), 924–932.

(158) Steenbergen, J. N.; Alder, J.; Thorne, G. M.; Tally, F. P. Daptomycin: A Lipopeptide Antibiotic for the Treatment of Serious Gram-Positive Infections. *J Antimicrob Chemother* **2005**, *55* (3), 283–288. https://doi.org/10.1093/jac/dkh546.

(159) JONES, T. S. G. The Chemical Basis for the Classification of the Polymyxins. *Biochem J* **1948**, *43* (2), xxvi.

(160) Biavasco, F.; Vignaroli, C.; Lupidi, R.; Manso, E.; Facinelli, B.; Varaldo, P. E. In Vitro Antibacterial Activity of LY333328, a New Semisynthetic Glycopeptide. *Antimicrob Agents Chemother* **1997**, *41* (10), 2165–2172. https://doi.org/10.1128/AAC.41.10.2165.

(161) Chen, A.; Zervos, M.; Vazquez, J. Dalbavancin: A Novel Antimicrobial. *International journal of clinical practice* **2007**, *61* (5), 853—863. https://doi.org/10.1111/j.1742-1241.2007.01318.x.

(162) Higgins, D. L.; Chang, R.; Debabov, D. V.; Leung, J.; Wu, T.; Krause, K. M.; Sandvik, E.; Hubbard, J. M.; Kaniga, K.; Schmidt, D. E. J.; Gao, Q.; Cass, R. T.; Karr, D. E.; Benton, B. M.; Humphrey, P. P. Telavancin, a Multifunctional Lipoglycopeptide, Disrupts Both Cell Wall Synthesis and Cell Membrane Integrity in Methicillin-Resistant *Staphylococcus aureus*. *Antimicrob Agents Chemother* **2005**, *49* (3), 1127–1134. https://doi.org/10.1128/AAC.49.3.1127-1134.2005.

(163) Leadbetter, M. R.; Adams, S. M.; Bazzini, B.; Fatheree, P. R.; Karr, D. E.; Krause, K. M.; Lam, B. M. T.; Linsell, M. S.; Nodwell, M. B.; Pace, J. L.; Quast, K.; Shaw, J.-P.; Soriano, E.; Trapp, S. G.; Villena, J. D.; Wu, T. X.; Christensen, B. G.; Judice, J. K. Hydrophobic Vancomycin Derivatives with Improved ADME Properties: Discovery of Telavancin (TD-6424). *J Antibiot (Tokyo)* **2004**, *57* (5), 326–336. https://doi.org/10.7164/antibiotics.57.326.

(164) Kussmann, M.; Obermueller, M.; Berndl, F.; Reischer, V.; Veletzky, L.; Burgmann, H.; Poeppl, W. Dalbavancin for Treatment of Implant-Related Methicillin-Resistant Staphylococcus Aureus Osteomyelitis in an Experimental Rat Model. *Scientific Reports* **2018**, *8* (1), 9661. https://doi.org/10.1038/s41598-018-28006-8.

(165) Sader, H. S.; Watters, A. A.; Fritsche, T. R.; Jones, R. N. Daptomycin Antimicrobial Activity Tested against Methicillin-Resistant *Staphylococci* and Vancomycin-Resistant *Enterococci* Isolated in European Medical Centers (2005). *BMC Infect Dis* **2007**, *7*, 29. https://doi.org/10.1186/1471-2334-7-29.

(166) Tobias, N. J.; Wolff, H.; Djahanschiri, B.; Grundmann, F.; Kronenwerth, M.; Shi, Y. M.; Simonyi, S.; Grün, P.; Shapiro-Ilan, D.; Pidot, S. J.; Stinear, T. P.; Ebersberger, I.; Bode, H. B. Natural Product Diversity Associated with the Nematode Symbionts *Photorhabdus* and *Xenorhabdus*. *Nature Microbiology* **2017**, *2* (12), 1676–1685. https://doi.org/10.1038/s41564-017-0039-9.

(167) Dreyer, J.; Malan, A. P.; Dicks, L. M. T. Bacteria of the Genus *Xenorhabdus*, a Novel Source of Bioactive Compounds. *Frontiers in Microbiology* **2018**, *9*, 3177. https://doi.org/10.3389/fmicb.2018.03177.

(168) Boemare, N.; Akhurst, R. The Genera *Photorhabdus* and *Xenorhabdus*. In *The Prokaryotes: Volume 6: Proteobacteria: Gamma Subclass*; Dworkin, M., Falkow, S., Rosenberg, E., Schleifer, K.-H., Stackebrandt, E., Eds.; Springer New York: New York, NY, 2006; pp 451–494. https://doi.org/10.1007/0-387-30746-X_16.

(169) Dathe, M.; Nikolenko, H.; Klose, J.; Bienert, M. Cyclization Increases the Antimicrobial Activity and Selectivity of Arginine- and Tryptophan-Containing Hexapeptides. *Biochemistry* **2004**, *43* (28), 9140–9150. https://doi.org/10.1021/bi035948v.

(170) Wang, G.; Li, X.; Wang, Z. APD3: The Antimicrobial Peptide Database as a Tool for Research and Education. *Nucleic Acids Res* **2016**, *44* (D1), D1087-1093. https://doi.org/10.1093/nar/gkv1278.

(171) Kronenwerth, M.; Bozhüyük, K. A. J.; Kahnt, A. S.; Steinhilber, D.; Gaudriault, S.; Kaiser, M.; Bode, H. B. Characterisation of Taxlllaids A-G; Natural Products from *Xenorhabdus indica*. *Chemistry - A European Journal* **2014**, *20* (52), 17478–17487. https://doi.org/10.1002/chem.201403979.

(172) Sweet, R. M.; Eisenberg, D. Correlation of Sequence Hydrophobicities Measures Similarity in Three-Dimensional Protein Structure. *Journal of Molecular Biology* **1983**, *171* (4), 479–488. https://doi.org/10.1016/0022-2836(83)90041-4.

(173) Nollmann, F. I.; Dowling, A.; Kaiser, M.; Deckmann, K.; Grösch, S.; Ffrench-Constant, R.; Bode, H. B. Synthesis of Szentiamide, a Depsipeptide from Entomopathogenic *Xenorhabdus szentirmaii* with Activity against *Plasmodium falciparum*. *Beilstein Journal of Organic Chemistry* **2012**, *8*, 528–533. https://doi.org/10.3762/bjoc.8.60.

(174) Lang, G.; Kalvelage, T.; Peters, A.; Wiese, J.; Imhoff, J. F. Linear and Cyclic Peptides from the Entomopathogenic Bacterium *Xenorhabdus nematophilus*. *Journal of natural products* **2008**, *71* (6), 1074–1077.

(175) Crawford, J. M.; Portmann, C.; Kontnik, R.; Walsh, C. T.; Clardy, J. NRPS Substrate Promiscuity Diversifies the Xenematides. *Organic Letters* **2011**, *13* (19), 5144–5147. https://doi.org/10.1021/ol2020237.

(176) Yau, W.-M.; Wimley, W. C.; Gawrisch, K.; White, S. H. The Preference of Tryptophan for Membrane Interfaces. *Biochemistry* **1998**, *37* (42), 14713–14718. https://doi.org/10.1021/bi980809c.

(177) Baltz, R. H. Function of MbtH Homologs in Non-ribosomal Peptide Biosynthesis and Applications in Secondary Metabolite Discovery. *J Ind Microbiol Biotechnol* **2011**, *38* (11), 1747–1760. https://doi.org/10.1007/s10295-011-1022-8.

(178) Linck, A. Mass Spectrometric Characterisation of Natural Products and Reprogramming of Non-Ribosomal Peptide Synthetases from Entomopathogenic Bacteria, Johann Wolfgang Goethe-Universität, Frankfurt am Main, 2018.

(179) Pace, C. N.; Scholtz, J. M. A Helix Propensity Scale Based on Experimental Studies of Peptides and Proteins. *Biophys J* **1998**, *75* (1), 422–427. https://doi.org/10.1016/s0006-3495(98)77529-0.

(180) Kaya, H. K.; Stock, S. P. Techniques in Insect Nematology. In *Manual of techniques in insect pathology*; Elsevier, 1997; pp 281–324.

(181) Bankevich, A.; Nurk, S.; Antipov, D.; Gurevich, A. A.; Dvorkin, M.; Kulikov, A. S.; Lesin, V. M.; Nikolenko, S. I.; Pham, S.; Prjibelski, A. D.; Pyshkin, A. V.; Sirotkin, A. V.; Vyahhi, N.; Tesler, G.; Alekseyev, M. A.; Pevzner, P. A. SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. *J Comput Biol* **2012**, *19* (5), 455–477. https://doi.org/10.1089/cmb.2012.0021.

(182) Brettin, T.; Davis, J. J.; Disz, T.; Edwards, R. A.; Gerdes, S.; Olsen, G. J.; Olson, R.; Overbeek, R.; Parrello, B.; Pusch, G. D.; Shukla, M.; Thomason, J. A. 3rd; Stevens, R.; Vonstein, V.; Wattam, A. R.; Xia, F. RASTtk: A Modular and Extensible Implementation of the RAST Algorithm for Building Custom Annotation Pipelines and Annotating Batches of Genomes. *Sci Rep* **2015**, *5*, 8365. https://doi.org/10.1038/srep08365.

(183) Wattam, A. R.; Brettin, T.; Davis, J. J.; Gerdes, S.; Kenyon, R.; Machi, D.; Mao, C.; Olson, R.; Overbeek, R.; Pusch, G. D.; Shukla, M. P.; Stevens, R.; Vonstein, V.; Warren, A.; Xia, F.; Yoo, H. Assembly, Annotation, and Comparative Genomics in

PATRIC, the All Bacterial Bioinformatics Resource Center. In *Comparative Genomics: Methods and Protocols*; Setubal, J. C., Stoye, J., Stadler, P. F., Eds.; Springer New York: New York, NY, 2018; pp 79–101. https://doi.org/10.1007/978-1-4939-7463-4_4.

(184) Meier-Kolthoff, J. P.; Göker, M. TYGS Is an Automated High-Throughput Platform for State-of-the-Art Genome-Based Taxonomy. *Nature Communications* **2019**, *10* (1). https://doi.org/10.1038/s41467-019-10210-3.

(185) Ondov, B. D.; Treangen, T. J.; Melsted, P.; Mallonee, A. B.; Bergman, N. H.; Koren, S.; Phillippy, A. M. Mash: Fast Genome and Metagenome Distance Estimation Using MinHash. *Genome Biology* **2016**, *17* (1), 132. https://doi.org/10.1186/s13059-016-0997-x.

(186) Meier-Kolthoff, J. P.; Auch, A. F.; Klenk, H.-P.; Göker, M. Genome Sequence-Based Species Delimitation with Confidence Intervals and Improved Distance Functions. *BMC Bioinformatics* **2013**, *14* (1), 60. https://doi.org/10.1186/1471-2105-14-60.

(187) Lefort, V.; Desper, R.; Gascuel, O. FastME 2.0: A Comprehensive, Accurate, and Fast Distance-Based Phylogeny Inference Program. *Mol Biol Evol* **2015**, *32* (10), 2798–2800. https://doi.org/10.1093/molbev/msv150.

(188) Farris, J. S. Estimating Phylogenetic Trees from Distance Matrices. *The American Naturalist* **1972**, *106* (951), 645–668. https://doi.org/10.1086/282802.

(189) Kreft, Ł.; Botzki, A.; Coppens, F.; Vandepoele, K.; Van Bel, M. PhyD3: A Phylogenetic Tree Viewer with Extended PhyloXML Support for Functional Genomics Data Visualization. *Bioinformatics* **2017**, *33* (18), 2946–2947. https://doi.org/10.1093/bioinformatics/btx324.

(190) Meier-Kolthoff, J. P.; Hahnke, R. L.; Petersen, J.; Scheuner, C.; Michael, V.; Fiebig, A.; Rohde, C.; Rohde, M.; Fartmann, B.; Goodwin, L. A.; Chertkov, O.; Reddy, T.; Pati, A.; Ivanova, N. N.; Markowitz, V.; Kyrpides, N. C.; Woyke, T.; Göker, M.; Klenk, H.-P. Complete Genome Sequence of DSM 30083(T), the Type Strain (U5/41(T)) of Escherichia Coli, and a Proposal for Delineating Subspecies in Microbial Taxonomy. *Stand Genomic Sci* **2014**, *9*, 2. https://doi.org/10.1186/1944-3277-9-2.

(191) Lee, I.; Ouk Kim, Y.; Park, S.-C.; Chun, J. OrthoANI: An Improved Algorithm and Software for Calculating Average Nucleotide Identity. *International Journal of Systematic and Evolutionary Microbiology,* **2016**, *66*, 1100–1103.

(192) Eren, A. M.; Esen, Ö. C.; Quince, C.; Vineis, J. H.; Morrison, H. G.; Sogin, M. L.; Delmont, T. O. Anvi'o: An Advanced Analysis and Visualization Platform for 'omics Data. *PeerJ* **2015**, *3*, e1319–e1319. https://doi.org/10.7717/peerj.1319.

(193) Hyatt, D.; Chen, G.-L.; Locascio, P. F.; Land, M. L.; Larimer, F. W.; Hauser, L. J. Prodigal: Prokaryotic Gene Recognition and Translation Initiation Site Identification. *BMC Bioinformatics* **2010**, *11*, 119. https://doi.org/10.1186/1471-2105-11-119.

(194) Altschul, S. F.; Gish, W.; Miller, W.; Myers, E. W.; Lipman, D. J. Basic Local Alignment Search Tool. *Journal of molecular biology* **1990**, *215* (3), 403–410.

(195) van Dongen, S.; Abreu-Goodger, C. Using MCL to Extract Clusters from Networks. *Methods Mol Biol* **2012**, *804*, 281–295. https://doi.org/10.1007/978-1-61779-361-5_15.

(196) Buchfink, B.; Xie, C.; Huson, D. H. Fast and Sensitive Protein Alignment Using DIAMOND. *Nature Methods* **2015**, *12* (1), 59–60. https://doi.org/10.1038/nmeth.3176.

(197) Alikhan, N.-F.; Petty, N. K.; Zakour, N. L. B.; Beatson, S. A. BLAST Ring Image Generator (BRIG): Simple Prokaryote Genome Comparisons. *BMC genomics* **2011**, *12* (1), 402.

(198) Bah, T. *Inkscape: Guide to a Vector Drawing Program*; Prentice Hall Press, **2007**.

(199) Darling, A. C. E.; Mau, B.; Blattner, F. R.; Perna, N. T. Mauve: Multiple Alignment of Conserved Genomic Sequence with Rearrangements. *Genome Res* **2004**, *14* (7), 1394–1403. https://doi.org/10.1101/gr.2289704.

(200) Lima-Mendez, G.; Van Helden, J.; Toussaint, A.; Leplae, R. Prophinder: A Computational Tool for Prophage Prediction in Prokaryotic Genomes. *Bioinformatics* **2008**, *24* (6), 863–865.

(201) GOSSET, W. The Probable Error of a Mean. *Biometrika* **1908**, *6*, 1–25.

(203) Thoma, S.; Schobert, M. An Improved Escherichia Coli Donor Strain for Diparental Mating. *FEMS Microbiology Letters* **2009**, *294* (2), 127–132. https://doi.org/10.1111/j.1574-6968.2009.01556.x.

(204) Bode, E.; Heinrich, A. K.; Hirschmann, M.; Abebew, D.; Shi, Y. N.; Vo, T. D.; Wesche, F.; Shi, Y. M.; Grün, P.; Simonyi, S.; Keller, N.; Engel, Y.; Wenski, S.; Bennet, R.; Beyer, S.; Bischoff, I.; Buaya, A.; Brandt, S.; Cakmak, I.; Çimen, H.; Eckstein, S.; Frank, D.; Fürst, R.; Gand, M.; Geisslinger, G.; Hazir, S.; Henke, M.; Heermann, R.; Lecaudey, V.; Schäfer, W.; Schiffmann, S.; Schüffler, A.; Schwenk, R.; Skaljac, M.; Thines, E.; Thines, M.; Ulshöfer, T.; Vilcinskas, A.; Wichelhaus, T. A.; Bode, H. B. Promoter Activation in Δhfq Mutants as an Efficient Tool for Specialized Metabolite Production Enabling Direct Bioactivity Testing. *Angewandte Chemie - International Edition* **2019**, *58* (52), 18957–18963. https://doi.org/10.1002/anie.201910563.

(205) Fu, C.; Donovan, W. P.; Shikapwashya-Hasser, O.; Ye, X.; Cole, R. H. Hot Fusion: An Efficient Method to Clone Multiple DNA Fragments as Well as Inverted Repeats without Ligase. *PloS one* **2014**, *9* (12), e115318–e115318. https://doi.org/10.1371/journal.pone.0115318.

(206) Pluskal, T.; Castillo, S.; Villar-Briones, A.; Orešič, M. MZmine 2: Modular Framework for Processing, Visualizing, and Analyzing Mass Spectrometry-Based Molecular Profile Data. *BMC Bioinformatics* **2010**, *11*. https://doi.org/10.1186/1471-2105-11-395.

(207) Bisch, G.; Ogier, J. C.; Médigue, C.; Rouy, Z.; Vincent, S.; Tailliez, P.; Givaudan, A.; Gaudriault, S. Comparative Genomics between Two *Xenorhabdus bovienii* Strains Highlights Differential Evolutionary Scenarios within an Entomopathogenic Bacterial Species. *Genome Biology and Evolution* **2016**, *8* (1), 148–160. https://doi.org/10.1093/gbe/evv248.

(208) CLSI. *Methods for Dilution Antimicrobial Susceptibility Tests for Bacteria That Grow Aerobically. 10 Th Ed. CLSI Standard M07.*; **2018;** p 188.

(209) Kumaraswamy, M.; Lin, L.; Olson, J.; Sun, C. F.; Nonejuie, P.; Corriden, R.; Döhrmann, S.; Ali, S. R.; Amaro, D.; Rohde, M.; Pogliano, J.; Sakoulas, G.; Nizet, V. Standard Susceptibility Testing Overlooks Potent Azithromycin Activity and Cationic Peptide Synergy against MDR *Stenotrophomonas maltophilia*. *Journal of Antimicrobial Chemotherapy* **2016**, *71* (5), 1264–1269. https://doi.org/10.1093/jac/dkv487.

(210) Subcommittee on Antifungal Susceptibility Testing (AFST) of the ESCMID European Committee for Antimicrobial Susceptibility Testing (EUCAST). EUCAST

Definitive Document EDef 7.1: Method for the Determination of Broth Dilution MICs of Antifungal Agents for Fermentative Yeasts. *Clinical Microbiology and Infection* **2008**, *14* (4), 398–405. https://doi.org/10.1111/j.1469-0691.2007.01935.x.

(211) White, G. F. A Method for Obtaining Infective Nematode Larvae from Cultures. *Science (Washington)* **1927**, *66* (1709).

(212) Joyce, S. A.; Clarke, D. J. A HexA Homologue from Photorhabdus Regulates Pathogenicity, Symbiosis and Phenotypic Variation. *Molecular Microbiology* **2003**, *47* (5), 1445–1457. https://doi.org/10.1046/j.1365-2958.2003.03389.x.

(213) Langer, A.; Moldovan, A.; Harmath, C.; Joyce, S. A.; Clarke, D. J.; Heermann, R. HexA Is a Versatile Regulator Involved in the Control of Phenotypic Heterogeneity of *Photorhabdus luminescens*. *PLoS One* **2017**, *12* (4), e0176535. https://doi.org/10.1371/journal.pone.0176535.

(214) Cowles, K. N.; Cowles, C. E.; Richards, G. R.; Martens, E. C.; Goodrich-Blair, H. The Global Regulator Lrp Contributes to Mutualism, Pathogenesis and Phenotypic Variation in the Bacterium *Xenorhabdus nematophila*. *Cellular Microbiology* **2007**, *9* (5), 1311–1323. https://doi.org/10.1111/j.1462-5822.2006.00873.x.

(215) Kour, S.; Khurma, U.; Brodie, G.; Hazir, S. Natural Occurrence and Distribution of Entomopathogenic Nematodes (*Steinernematidae, Heterorhabditidae*) in Viti Levu, Fiji Islands. *The Journal of Nematology* **2020**, *52*, 1–17.

(216) Mwaniki, S. W.; Nderitu, J. H.; Olubayo, F.; Kimenju, J. W.; Nguyen, K. Factors Influencing the Occurrence of Entomopathogenic Nematodes in the Central Rift Valley Region of Kenya. *African Journal of Ecology* **2008**, *46* (s1), 79–84. https://doi.org/10.1111/j.1365-2028.2008.00933.x.

(217) Zepeda-Jazo, I.; Molina-Ochoa, J.; Lezama-Gutiérrez, R.; Skoda, S. R.; Foster, J. E. Survey of Entomopathogenic Nematodes from the Families *Steinernematidae* and *Heterorhabditidae* (Nematoda: Rhabditida) in Colima, México. *International Journal of Tropical Insect Science* **2014**, *34* (1), 53–57.

(218) Campos-Herrera, R.; Escuer, M.; Labrador, S.; Robertson, L.; Barrios, L.; Gutiérrez, C. Distribution of the Entomopathogenic Nematodes from La Rioja (Northern Spain). *J Invertebr Pathol* **2007**, *95* (2), 125–139. https://doi.org/10.1016/j.jip.2007.02.003.

(219) Nah, H.-J.; Pyeon, H.-R.; Kang, S.-H.; Choi, S.-S.; Kim, E.-S. Cloning and Heterologous Expression of a Large-Sized Natural Product Biosynthetic Gene Cluster in *Streptomyces* Species. *Frontiers in Microbiology* **2017**, *8*, 394. https://doi.org/10.3389/fmicb.2017.00394.

(220) Parks, D. H.; Imelfort, M.; Skennerton, C. T.; Hugenholtz, P.; Tyson, G. W. CheckM: Assessing the Quality of Microbial Genomes Recovered from Isolates, Single Cells, and Metagenomes. *Genome research* **2015**, *25* (7), 1043–1055.

(221) Nguyen, K.; Hunt, D. *Entomopathogenic Nematodes: Systematics, Phylogeny and Bacterial Symbionts*; Brill, **2007**.

(222) Meier-Kolthoff, J. P.; Klenk, H.-P.; Göker, M. Taxonomic Use of DNA G+C Content and DNA–DNA Hybridization in the Genomic Age. *International Journal of Systematic and Evolutionary Microbiology*, **2014**, *64*, 352–356. https://doi.org/10.1099/ijs.0.056994-0.

(223) Herdan, G. *Type-Token Mathematics*; Mouton, **1960**; Vol. 4.

(224) Fischer-Le Saux, M.; Viallard, V.; Brunel, B.; Normand, P.; Boemare, N. E. Polyphasic Classification of the Genus *Photorhabdus* and Proposal of New Taxa: *P. luminescens* subsp*. luminescens* subsp. nov., *P. luminescens* subsp. *akhurstii* subsp. nov., *P. luminescens* subsp. *laumondii* subsp. nov., *P. temperata* sp. nov., *P. temperata* Subs. *International Journal of Systematic and Evolutionary Microbiology* **1999**, *49* (4), 1645–1656.

(225) Ooka, T.; Ogura, Y.; Asadulghani, M.; Ohnishi, M.; Nakayama, K.; Terajima, J.; Watanabe, H.; Hayashi, T. Inference of the Impact of Insertion Sequence (IS) Elements on Bacterial Genome Diversification through Analysis of Small-Size Structural Polymorphisms in Escherichia Coli O157 Genomes. *Genome research* **2009**, *19* (10), 1809–1816.

(226) Tu, Q.; Ding, D. Detecting Pathogenicity Islands and Anomalous Gene Clusters by Iterative Discriminant Analysis. *FEMS Microbiology Letters* **2003**, *221* (2), 269–275. https://doi.org/10.1016/S0378-1097(03)00204-0.

(227) Schaffer, J. E.; Reck, M. R.; Prasad, N. K.; Wencewicz, T. A. β-Lactone Formation during Product Release from a Non-ribosomal Peptide Synthetase. *Nature Chemical Biology* **2017**, *13* (7), 737–744. https://doi.org/10.1038/nchembio.2374.

(228) Brody, T. *Nutritional Biochemistry*; Elsevier, **1998**.

(229) Götze, S.; Stallforth, P. Structure Elucidation of Bacterial Non-ribosomal Lipopeptides. *Organic and Biomolecular Chemistry* **2020**, *18* (9), 1710–1727. https://doi.org/10.1039/c9ob02539a.

(230) Ibrahim, A.; Yang, L.; Johnston, C.; Liu, X.; Ma, B.; Magarvey, N. A. Dereplicating Non-ribosomal Peptides Using an Informatic Search Algorithm for Natural Products (ISNAP) Discovery. *Proceedings of the National Academy of Sciences* **2012**, *109* (47), 19196–19201.

(231) Hung, K. Y.; Harris, P. W. R.; Heapy, A. M.; Brimble, M. A. Synthesis and Assignment of Stereochemistry of the Antibacterial Cyclic Peptide Xenematide. *Organic and Biomolecular Chemistry* **2011**, *9* (1), 236–242. https://doi.org/10.1039/c0ob00315h.

(232) Tossi, A.; Sandri, L.; Giangaspero, A. Amphipathic, α-Helical Antimicrobial Peptides. *Peptide Science* **2000**, *55* (1), 4–30. https://doi.org/10.1002/1097-0282(2000)55:1<4::AID-BIP30>3.0.CO;2-M.

(233) Schmidtchen, A.; Pasupuleti, M.; Malmsten, M. Effect of Hydrophobic Modifications in Antimicrobial Peptides. *Advances in Colloid and Interface Science* **2014**, *205*, 265–274. https://doi.org/10.1016/j.cis.2013.06.009.

(234) Davies, J. S. The Cyclization of Peptides and Depsipeptides. *J Pept Sci* **2003**, *9* (8), 471–501. https://doi.org/10.1002/psc.491.

(235) Bahar, A. A.; Ren, D. Antimicrobial Peptides. *Pharmaceuticals (Basel)* **2013**, *6* (12), 1543–1575. https://doi.org/10.3390/ph6121543.

(236) Junkes, C.; Wessolowski, A.; Farnaud, S.; Evans, R. W.; Good, L.; Bienert, M.; Dathe, M. The Interaction of Arginine- and Tryptophan-Rich Cyclic Hexapeptides with *Escherichia coli* Membranes. *J Pept Sci* **2008**, *14* (4), 535–543. https://doi.org/10.1002/psc.940.

(237) Frederiksen, N.; Hansen, P. R.; Björkling, F.; Franzyk, H. Peptide/Peptoid Hybrid Oligomers: The Influence of Hydrophobicity and Relative Side-Chain Length on Antibacterial Activity and Cell Selectivity. *Molecules* **2019**, *24* (24). https://doi.org/10.3390/molecules24244429.

(238) Taylor, J. W. The Synthesis and Study of Side-Chain Lactam-Bridged Peptides. *Peptide Science* **2002**, *66* (1), 49–75. https://doi.org/10.1002/bip.10203.

(239) Luo, P.; Braddock, D. T.; Subramanian, R. M.; Meredith, S. C.; Lynn, D. G. Structural and Thermodynamic Characterization of a Bioactive Peptide Model of Apolipoprotein E: Side-Chain Lactam Bridges to Constrain the Conformation. *Biochemistry* **1994**, *33* (41), 12367–12377.

JOURNAL ARTICLES

.

1) The following journal article resulted from a part of this dissertation. It is reprinted with permission from Structure, Biosynthesis, and Bioactivity of Photoditritide from *Photorhabdus temperata* Meg1. Lei Zhao, Ryan Musumba Awori, Marcel Kaiser, Jonathan Groß, Till Opatz, and Helge B. Bode. *Journal of Natural Products* **2019** *82* (12), 3499-3503. DOI: 10.1021/acs.jnatprod.9b00932.

2) The following journal article is not a result of this dissertation. It was from a project used as the foundation for a transcriptome annotation training Jamboree in 2014 in Nairobi, Kenya. The project was concluded in 2020 with this publication. As it is only accepted for publication but not yet printed only the title page is given herein.

   Comprehensive transcriptome of the maize stalk borer, *Busseola fusca*, from multiple tissue types, developmental stages, and parasitoid wasp exposures. *Busseola fusca* Genomics Consortium. Accepted. Genome Biology and Evolution, manuscript ID:GBE-200645.R3.

Structure, Biosynthesis, and Bioactivity of Photoditritide from *Photorhabdus temperata* Meg1

Lei Zhao,[†,‡] Ryan Musumba Awori,[†] Marcel Kaiser,[§] Jonathan Groß,[] Till Opatz,[] and Helge B. Bode*[†,‖]

[†]Molecular Biotechnology, Department of Biosciences, Goethe University Frankfurt, 60438 Frankfurt am Main, Germany

[‡]Institute of Botany, Jiangsu Province and Chinese Academy of Sciences, 210014 Nanjing, China

[§]Swiss Tropical and Public Health Institute, 4051 Basel, Switzerland

[]Institute of Organic Chemistry, Johannes Gutenberg University Mainz, 55128 Mainz, Germany

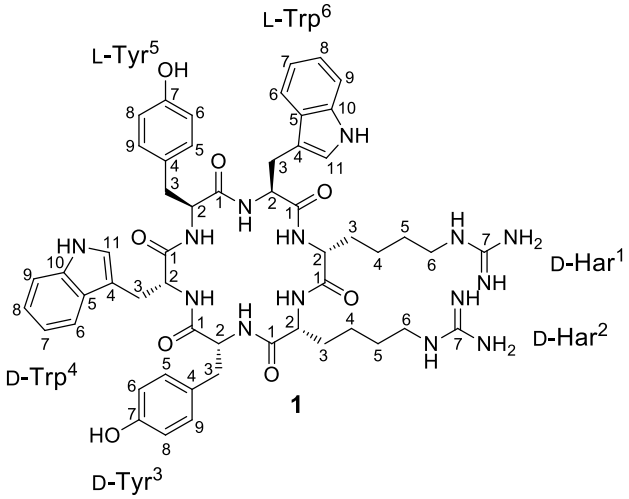
[‖]Buchmann Institute for Molecular Life Sciences (BMLS), Goethe University Frankfurt, 60438 Frankfurt am Main, Germany

Abstract

A new cyclic peptide photoditritide (**1**), containing two rare amino acid D-homoarginine residues, was isolated from *Photorhabdus temperata* Meg1 after the nonribosomal peptide synthetase encoding gene *pdtS* was activated *via* promoter exchange. The structure of **1** was elucidated by HR-MS and NMR experiments. The absolute configurations of amino acids were determined according to the advanced Marfey's method after hydrolysis of **1**. Bioactivity testing of **1** revealed potent antimicrobial activity against *Micrococcus luteus* with an MIC value of 3.0 *µ*M and weak antiprotozoal activity against *Trypanosoma brucei rhodesiense* with an IC_{50} value of 13 *µ*M. Additionally, the biosynthetic pathway of **1** was also proposed.

Photorhabdus is a genus of Gram-negative entomopathogenic bacteria that live in symbiosis with nematodes of the genus *Heterorhabditis*.[1,2] The nematode-bacterium complex is highly pathogenic for a broad range of insects and is well known as a model system for the investigation of mutualistic and pathogenic symbiosis.[2–4] The *Heterorhabditis* nematode releases *Photorhabdus* bacteria from its gut into the insect hemocoel after insect infection.[2] The bacteria replicate and produce a large number of protein toxins and secondary metabolites to kill the insect and protect the prey cadaver for nutrition against soil-living food competitors, including bacteria, fungi and protozoa.[5,6] The unique niche makes *Photorhabdus* a rich source of bioactive natural products,[6–8] and several compounds, such as isopropylstilbene,[9] phurealipids,[10] kolossins,[11] photolose,[12] lumizinones,[13] rhabduscins,[14] and photopyrones[15] with antibiotic,[16] insecticidal,[10] antiprotozoal,[11] immune regulation,[12] protease inhibitory,[13] virulence factor,[14] and signaling activity,[15] respectively, have been identified recently.

Natural products have played and will continue to play a highly significant role in the drug discovery and development process.[17] Our efforts to search for additional natural products from promising *Photorhabdus* strains resulted in the identification of photoditritide (**1**), a new cyclohexapeptide containing two uncommon amino acid homoarginine residues, after its biosynthetic gene *pdtS* was activated in *Photorhabdus temperata* Meg1. Here we report the discovery, structure elucidation, biosynthesis, and bioactivity of **1**.



1

In the genome of *P. temperata* Meg1, in total 25 biosynthesis gene clusters of secondary metabolites have been predicted by antiSMASH.[18] The *pdtS* (photoditritide synthetase) gene (*MEG1_RS04325*, 19.5 kbp, accession number NZ_JGVH01000009) was identified encoding an unknown nonribosomal peptide synthetase consisting of six modules with overall 18 domains (Figure 1). PdtS was hence expected to produce a hexapeptide. As no such peptide could be identified in *P. temperata* Meg1 when the strain was cultivated in the lab in different media such as lysogeny broth (LB) or Sf-900 II SFM medium, a promoter exchange approach was applied to activate *pdtS* as previously described,[19] thus resulting in the detection of one new natural product based on HPLC-MS analysis (Figure S1).

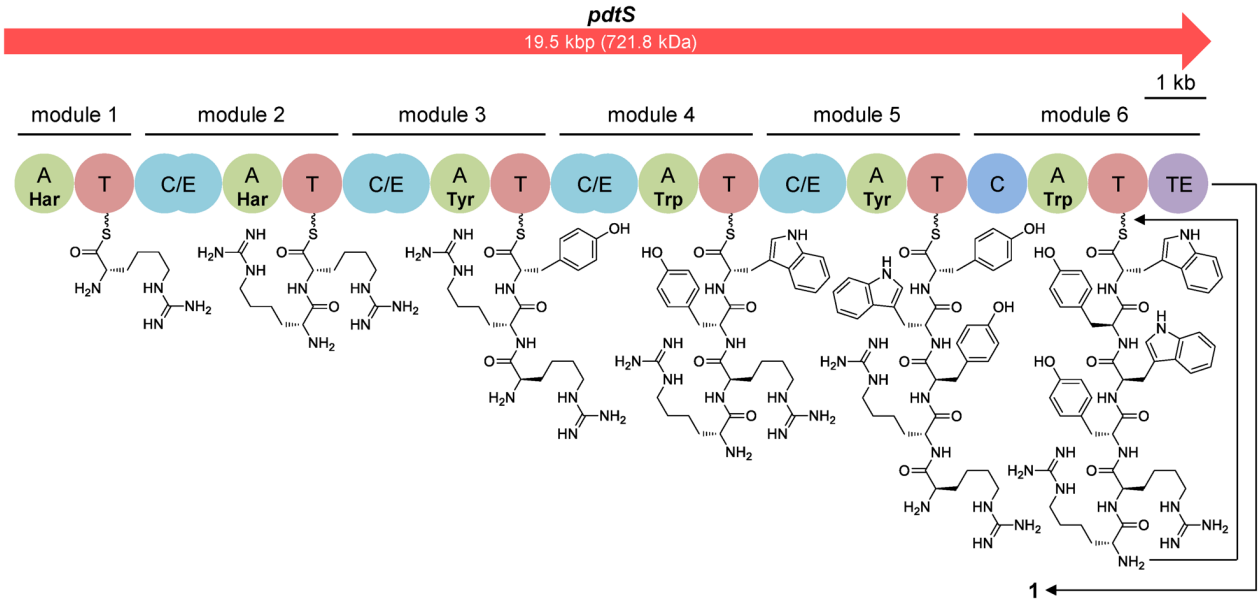


Figure 1. Domain organization of the PdtS and proposed biosynthesis of photoditritide (**1**). Domains: A: adenylation, T: thiolation, C/E: dual condensation/epimerization, C: condensation, TE: thioesterase.

To isolate **1**, the promoter exchange mutant *P. temperata* Meg1 pCEP-*pdtS* was grown in *Xenorhabdus* and *Photorhabdus* production medium (XPPM), an amino acid rich medium developed from M9 medium,[20] which gives a clear background for compound isolation from *Xenorhabdus* and *Photorhabdus* strains. From 4 L of cultures, 12 mg of **1** was isolated from the XAD-16 extracts by using Sephadex LH-20 chromatography, followed by semipreparative HPLC.

Compound **1** was obtained as a brown powder. The molecular formula of **1** was determined to be $C_{54}H_{66}N_{14}O_8$ by its HRESIMS data combined with isotopic labelling experiments as previously described (Figure S2),[21] indicating 29 degrees of unsaturation. Detailed analysis of 1D and 2D NMR data of **1** (Table 1, Figure 2) implied the presence of six amino acids: two homoarginines (Har), two tyrosines (Tyr), and two tryptophans (Trp), and therefore **1** was named as photoditritide. As nonproteinogenic amino acids like homoarginine are rarely found in natural products from entomopathogenic bacteria, a reverse labelling experiment with homoarginine was performed in ^{13}C medium verifying the incorporation of two homoarginines (Figure S3). Compound **1** was further identified to be a monocyclic peptide on the basis of the remaining one degree of unsaturation because the six amino acids accounted for 28 of the 29 degrees of unsaturation. The connectivity of the six amino acid building blocks was established by HMBC (Figure 2) and was confirmed by the assembly order of predicted amino acids from the adenylation (A) domain specificity conferring codes (Table S1). Thereby, compound **1** was unequivocally elucidated to be a cyclic hexapeptide.

Table 1. ^{1}H (500 MHz) and ^{13}C (125 MHz) NMR Data of 1 in DMSO-d_6 (δ in ppm)

subunit	position	δ_C, type	δ_H (J in Hz)	subunit	position	δ_C, type	δ_H (J in Hz)
Har[1]	1	170.3, C			3b		2.91, overlap
	2	52.2, CH	4.23, dd (13.4, 6.5)		4	108.8, C	
	2-NH		7.23, d (7.0)		5	127.9, C	
	3a	32.1, CH_2	1.52, m		6	118.5, CH	7.05, overlap
	3b		1.45, m		7	118.1, CH	6.83, t (7.4)
	4	22.0, CH_2	1.12, m		8	120.6, CH	6.99, overlap
	5	28.1, CH_2	1.40, m		9	111.0, CH	7.26, d (8.1)
	6	40.4, CH_2	2.97, overlap		10	135.8, C	
	6-NH		8.45, brs		10-NH		10.79, s
	7	157.2, C			11	123.8, CH	6.48, s
Har[2]	1	171.7, C		Tyr[5]	1	171.2, C	
	2	53.0, CH	3.99, m		2	57.4, CH	3.83, m
	2-NH		8.36, submerge		2-NH		8.27, d (4.5)
	3	29.9, CH_2	1.30, m		3a	36.0, CH_2	2.43, m
	4a	22.0, CH_2	0.96, m		3b		2.35, m
	4b		0.71, m		4	127.8, C	
	5	28.1, CH_2	1.30, m		5	130.0, CH	6.79, d (8.2)
	6	40.2, CH_2	2.91, overlap		6	114.9, CH	6.53, d (8.2)
	6-NH		8.51, brs		7	155.9, C	
	7	157.2, C			8	114.9, CH	6.53, d (8.2)
Tyr[3]	1	171.3, C			9	130.0, CH	6.79, d (8.2)
	2	56.3, CH	4.17, m	Trp[6]	1	170.4, C	

	2-NH		8.70, d (7.6)	2	53.3, CH	4.45, dd (12.6, 9.5)
	3a	36.5, CH_2	3.09, brd (12.4)	2-NH		8.03, d (7.7)
	3b		2.61, t (12.7)	3a	25.9, CH_2	3.34, overlap
	4	128.4, C		3b		2.97, overlap
	5	130.0, CH	7.04, overlap	4	110.8, C	
	6	115.0, CH	6.65, d (8.1)	5	127.3, C	
	7	156.0, C		6	118.2, CH	7.49, d (7.8)
	8	115.0, CH	6.65, d (8.1)	7	118.3, CH	6.99, overlap
	9	130.0, CH	7.04, overlap	8	120.9, CH	7.07, overlap
Trp^4	1	170.9, C		9	111.3, CH	7.31, overlap
	2	54.4, CH	4.57, m	10	136.1, C	
	2-NH		7.31, overlap	10-NH		10.85, s
	3a	27.3, CH_2	3.32, overlap	11	123.2, CH	7.07, overlap

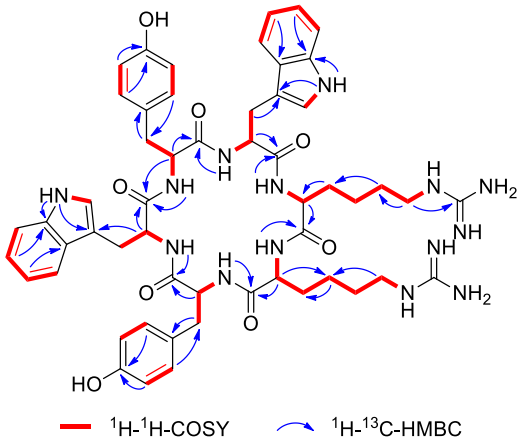


Figure 2. COSY and key HMBC correlations of **1**.

To address the absolute configurations of the amino acids in **1**, four dual condensation/epimerization (C/E) domains were identified in PdtS by antiSMASH (Figure 1).[22] C/E domains have been shown to transform the initially bound L-amino acids into their D-form,[21,23] suggesting **1** with the configuration sequence DDDDLL in its biosynthesis order. To confirm the configurations for homoarginine (D), tyrosine (D), tryptophan (D), tyrosine (L), and tryptophan (L), compound **1** was hydrolyzed with 6 M HCl supplemented with 2.5% triisopropylsilane to prevent degradation of tryptophan.[24] The hydrolysate was analyzed according to the advanced Marfey's method,[25] and the finding was in accordance with the results predicted above (Figure S4).

Finally, the antimicrobial and antiprotozoal activity of **1** was investigated in view of the ecological function of *Photorhabdus* (Table 2). Compound **1** exhibited potent activity against Gram-positive bacterium *Micrococcus luteus* with an MIC value of 3.0

*μ*M and weak activity against the causative agent of sleeping sickness *Trypanosoma brucei rhodesiense* with an IC_{50} value of 13 *μ*M, but no activity against other tested organisms and no cytotoxic activity against mammalian L6 cells.

In summary, a new cyclic peptide photoditritide was unequivocally identified from *P. temperata* Meg1 by using a combination of bioinformatics analysis, promoter exchange, isotopic labelling experiments, HPLC-MS analysis, and NMR. Although nonproteinogenic amino acid homoarginine containing peptides have been reported in several marine organisms such as a sponge,[26] cyanobacteria,[27] and marine-derived actinomycetes,[28] photoditritide so far is the only example of a peptide from entomopathogenic bacteria that contains homoarginine.

Table 2. Bioactivity of 1 Against Different Microorganisms (MIC in *μ*M), Protozoan Parasites (IC_{50} in *μ*M), and Mammalian L6 Cells (IC_{50} in *μ*M)

species	**1**	control[a]
M. luteus	3.0	4.2
Escherichia coli	24	0.60
Candida lusitaniae	>96	0.42
T. brucei rhodesiense	13	0.005
Trypanosoma cruzi	71	2.1
Leishmania donovani	>100	0.73
Plasmodium falciparum	27	0.009
mammalian L6 cells	83	0.007

[a]The positive control is different for each target organism: ampicillin for *M. luteus*, chloramphenicol for *E. coli*, amphotericin B for *C. lusitaniae*, melarsoprol for *T. brucei rhodesiense*, benznidazole for *T. cruzi*, miltefosine for *L. donovani*, chloroquine for *P. falciparum*, and podophyllotoxin for L6 cells.

EXPERIMENTAL SECTION

General Experimental Procedures. Optical rotation was measured on a PerkinElmer 241 polarimeter. ECD spectrum was obtained on a JASCO J-815 spectropolarimeter. ^{1}H, ^{13}C, COSY, HSQC, and HMBC NMR spectra were recorded on a Bruker AV 500 spectrometer at 500 MHz (^{1}H) and 125 MHz (^{13}C) using DMSO-d_6 as solvent; the ^{1}H and ^{13}C NMR chemical shifts were referenced to the solvent peaks at δ_H 2.50 and δ_C 39.52 for DMSO-d_6. HPLC-ESI-MS analysis was recorded on a Dionex UltiMate 3000 system coupled to a Bruker AmaZonX mass spectrometer; the MeOH

extracts were eluted on an ACQUITY UPLC BEH C_{18} column (130 Å, 2.1 mm × 100 mm, 1.7 *μ*m) using a gradient from 5% to 95% MeCN/H_2O solution containing 0.1% formic acid at a flow rate of 0.6 mL/min for 16 min. HPLC-HR-ESI-MS analysis was performed on a Dionex UltiMate 3000 system coupled to a Bruker Impact II QTOF mass spectrometer; the MeOH extracts were eluted on an ACQUITY UPLC BEH C_{18} column (130 Å, 2.1 mm × 50 mm, 1.7 *μ*m) using a gradient from 5% to 95% MeCN/H_2O solution containing 0.1% formic acid at a flow rate of 0.4 mL/min for 16 min. Positive mode with scan range from 100 to 1200 *m*/*z* for HPLC-ESI-MS and HPLC-HR-ESI-MS analysis was used to detect photoditritide (**1**). Semipreparative HPLC was conducted on an Agilent 1260 system with a ZORBAX Eclipse XDB-C_{18} column (9.4 mm × 250 mm, 5.0 *μ*m).

Strain Construction. The construction of the promoter exchange mutant *P. temperata* Meg1 pCEP-*pdtS* was carried out as described previously.[19] Briefly, the plasmid pLZ46 carrying the first 650 bp of *pdtS* from *P. temperata* Meg1 was constructed by using Gibson cloning and was transformed into *E. coli* ST18. For conjugation, *P. temperata* Meg1 and *E. coli* ST18 were grown in LB medium. Kanamycin (50 *μ*g/mL) and δ-aminolevulinic acid (ALA, 50 *μ*g/mL) were added to *E. coli* ST18. After OD_{600} 0.5–0.7, the cells were harvested and washed three times with fresh LB medium. Subsequently, *E. coli* ST18 and *P. temperata* Meg1 were mixed on a LB agar plate containing ALA in a ratio of 1:3 and were incubated at 37 °C for 3 hours followed by growth at 30 °C overnight. The next day, the bacterial cell layer was harvested and resuspended in fresh LB medium. Serial dilutions were spread out on selective LB agar plates with kanamycin and were incubated at 30 °C for 2–3 days. The genotype of individual clones was verified by PCR.

Production Cultivation. 100 *μ*L of overnight cultures of *P. temperata* Meg1 pCEP-*pdtS* were inoculated into 10 mL fresh liquid LB medium with kanamycin (50 *μ*g/mL), 0.1% L-arabinose (from a 25% stock solution), and 2% Amberlite XAD-16 resin added. The cultures were grown at 30 °C on a rotary shaker at 200 rpm and were harvested after 3 days. The XAD-16 beads were filtered and extracted with 10 mL MeOH for 1 h to give extracts for HPLC-MS analysis.

Labelling Experiments. Labelling experiments were performed in 5 mL fully ^{13}C labeled medium as described previously.[21] Culture conditions, extract preparation, and HPLC-MS analysis were carried out as described above. To confirm the incorporation of homoarginine, L-homoarginine was supplied to *P. temperata* Meg1 pCEP-*pdtS* in ^{13}C medium.

Compound Isolation. To isolate **1**, the XAD-16 beads from 4 L XPPM cultures of *P. temperata* Meg1 pCEP-*pdtS* were extracted with MeOH. The XAD-16 extract was fractionated by Sephadex LH-20 chromatography using MeOH as the eluent. The enriched fractions containing **1** were collected and further purified by semipreparative HPLC with a gradient 20% to 40% MeCN/H_2O solution containing 0.1% formic acid at a flow rate of 3 mL/min for 20 min to yield **1** (12 mg).

*Photoditritide (**1**):* brown, amorphous solid; $[\alpha]^{25}_D$ −1490.9 (*c* 0.022, MeOH); ECD (*c* 5.3 × 10^{-5} M, MeOH) λ_{max} ($\Delta\varepsilon$) 232 (35.25), 214 (7.42), and 202 (–11.60) nm; ^1H and ^{13}C NMR data, Table 1; HRESIMS *m*/*z* 520.2669 $[M + 2H]^{2+}$ (calcd for $C_{54}H_{66}N_{14}O_8$, 520.2667; Δppm 0.5).

Configuration Determination. The configurations of amino acids in **1** were determined by the advanced Marfey's method as described previously.[24,25] Approximately 1 mg of peptide was dissolved in 200 μL MeOH and was hydrolyzed with 800 μL HCl (6 M) in an Ace high-pressure tube at 110 ℃ for 1 h. To prevent the degradation of tryptophan, 2.5% triisopropylsilane was added. The hydrolysate was evaporated to dryness and was redissolved in 100 μL H_2O. To each half portion (45 μL) were added 10 μL 1M $NaHCO_3$ and 80 μL 1% Nα-(5-fluoro-2,4-dinitrophenyl)-L-leucinamide or Nα-(5-fluoro-2,4-dinitrophenyl)-D-leucinamide (L-FDLA or D-FDLA, dissolved in acetone). The brown reaction vials were incubated in a water bath at 40 ℃ for 1 h. After that the reactions were cooled to room temperature, quenched with 10 μL 1 M HCl and evaporated to dryness. The residue was dissolved in 400 μL MeOH and was analyzed by HPLC-MS.

Antimicrobial Activity Testing. Antimicrobial activity was tested against representative Gram-positive (*Micrococcus luteus*), Gram-negative (*Escherichia coli* OP50), and yeast (*Candida lusitaniae* DSM 70102) strains. For antibacterial activity, a microdilution broth assay modified from the Clinical and Laboratory Standards Institute

(CLSI) protocol[29] was carried out in round-bottom 96-well plates with a final inoculum concentration of 1.3×10^5 CFU/mL, incubation time of 20 h at 30 °C, and final culture volume of 100 *μ*L. A two-fold dilution series was created for both **1** and the positive control ampicillin or chloramphenicol with a concentration range between 100 *μ*g/mL and 0.02 *μ*g/mL. RPMI 1640 medium supplemented with 10% LB medium was used instead of Muller Hinton broth to ensure the detection of antibacterial activity.[30] For antifungal activity, a microdilution broth assay modified from the European Committee for Antimicrobial Susceptibility Testing [33] protocol[31] was performed in round-bottom 96-well plates with a final inoculum concentration 2.8×10^4 CFU/mL, incubation period of 49 h at 30°C in YPD medium, and final culture volume of 100 *μ*l. A two-fold dilution series was created for both **1** and the positive control amphotericin B with a concentration range between 100 *μ*g/mL and 0.02 *μ*g/mL. For both assays, minimum inhibitory concentration (MIC) was defined as the lowest antibiotic concentration that inhibits visible growth of the microorganisms.

Antiprotozoal Activity and Mammalian Cell Cytotoxicity Testing. Bioactivity testing of **1** against the parasites *Trypanosoma brucei rhodesiense* STIB900, *Trypanosoma cruzi* Tulahuen C4, *Leishmania donovani* MHOM-ET-67/L82, and *Plasmodium falciparum* NF54 was performed as described previously.[32] Cytotoxicity of **1** against rat skeletal myoblasts (L6 cells) was evaluated as described previously.[32] IC_{50} (50% inhibitory concentration) values against these cells were calculated.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI:

A domain specificity prediction of PdtS; bacterial strains, plasmids, and primers used; PdtS and its nearby proteins with predicted function; HPLC-MS analysis; ECD and NMR spectra for **1**

AUTHOR INFORMATION

Corresponding Author

*E-mail: h.bode@bio.uni-frankfurt.de

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported by the LOEWE Schwerpunkt MegaSyn funded by the State of Hessen. L.Z. holds a Ph.D. scholarship from the China Scholarship Council (CSC). We thank S. Li for part work on construction of the promoter exchange mutant *P. temperata* Meg1 pCEP-*pdtS*.

REFERENCES

(1) Thanwisai, A.; Tandhavanant, S.; Saiprom, N.; Waterfield, N. R.; Long, P. K.; Bode, H. B.; Peacock, S. J.; Chantratita, N. *PLoS One* **2012**, *7*, e43835.

(2) Forst, S.; Dowds, B.; Boemare, N.; Stackebrandt, E. *Annu. Rev. Microbiol.* **1997**, *51*, 47–72.

(3) Ehlers, R. U. *Commun. Agric. Appl. Biol. Sci.* **2003**, *68*, 3–16.

(4) Mulley, G.; Beeton, M. L.; Wilkinson, P.; Vlisidou, I.; Ockendon-Powell, N.; Hapeshi, A.; Tobias, N. J.; Nollmann, F. I.; Bode, H. B.; Van Den Elsen, J; Ffrench-Constant, R. H.; Waterfield, N. R. *PLoS One* **2015**, *10*, e0144937.

(5) Ffrench-Constant, R. H.; Dowling, A.; Waterfield, N. R. *Toxicon* **2007**, *49*, 436–451.

(6) Shi, Y.-M.; Bode, H. B. *Nat. Prod. Rep.* **2018**, *35*, 309–335.

(7) Bode, H. B. *Curr. Opin. Chem. Biol.* **2009**, *13*, 224–230.

(8) Vizcaino, M. I.; Guo, X.; Crawford, J. M. *J. Ind. Microbiol. Biotechnol.* **2014**, *41*, 285–299.

(9) Kronenwerth, M.; Brachmann, A. O.; Kaiser, M.; Bode, H. B. *ChemBioChem* **2014**, *15*, 2689–2691.

(10) Nollmann, F. I.; Heinrich, A. K.; Brachmann, A. O.; Morisseau, C.; Mukherjee, K.; Casanova-Torres, Á. M.; Strobl, F.; Kleinhans, D.; Kinski, S.; Schultz, K; Beeton, M. L.; Kaiser, M.; Chu, Y.-Y.; Ke, L. P.; Thanwisai, A.; Bozhüyük, K. A. J.; Chantratita, N.; Götz, F.; Waterfield, N. R.; Vilcinskas, A.; Stelzer, E. H. K.; Goodrich-Blair, H.; Hammock, B. D.; Bode, H. B. *ChemBioChem* **2015**, *16*, 766–771.

(11) Bode, H. B.; Brachmann, A. O.; Jadhav, K. B.; Seyfarth, L.; Dauth, C.; Fuchs, S. W.; Kaiser, M.; Waterfield, N. R.; Sack, H.; Heinemann, S. H.; Arndt, H.-D. *Angew. Chem. Int. Ed.* **2015**, *54*, 10352–10355.

(12) Perez, C. E.; Crawford, J. M. *Biochemistry* **2019**, *58*, 1131–1140.

(13) Park, H. B.; Crawford, J. M. *J. Antibiot.* **2016**, *69*, 616–621.

(14) Crawford, J. M.; Portmann, C.; Zhang, X.; Roeffaers, M. B. J.; Clardy, J. *Proc. Natl. Acad. Sci.* **2012**, *109*, 10821–10826.

(15) Brachmann, A. O.; Brameyer, S.; Kresovic, D.; Hitkova, I.; Kopp, Y.; Manske, C.;

Schubert, K.; Bode, H. B.; Heermann, R. *Nat. Chem. Biol.* **2013**, *9*, 573–578.

(16) Eleftherianos, I.; Boundy, S.; Joyce, S. A.; Aslam, S.; Marshall, J. W.; Cox, R. J.; Simpson, T. J.; Clarke, D. J.; Ffrench-Constant, R. H.; Reynolds, S. E. *Proc. Natl. Acad. Sci.* **2007**, *104*, 2419–2424.

(17) Newman, D. J.; Cragg, G. M. *J. Nat. Prod.* **2016**, *79*, 629–661.

(18) Blin, K.; Wolf, T.; Chevrette, M. G.; Lu, X.; Schwalen, C. J.; Kautsar, S. A.; Suarez Duran, H. G.; de Los Santos, E. L. C.; Kim, H. U.; Nave, M.; Dickschat, J. S.; Mitchell, D. A.; Shelest, E.; Breitling, R.; Takano, E.; Lee, S. Y.; Weber, T.; Medema, M. H. *Nucleic Acids Res.* **2017**, *45*, W36–W41.

(19) Bode, E.; Brachmann, A. O.; Kegler, C.; Simsek, R.; Dauth, C.; Zhou, Q.; Kaiser, M.; Klemmt, P.; Bode, H. B. *ChemBioChem* **2015**, *16*, 1115–1119.

(20) Bode, H. B. *Angew. Chem. Int. Ed.* **2019**, DOI: 10.1002/anie.201910563.

(21) Bode, H. B.; Reimer, D.; Fuchs, S. W.; Kirchner, F.; Dauth, C.; Kegler, C.; Lorenzen, W.; Brachmann, A. O.; Grün, P. *Chem. Eur. J.* **2012**, *18*, 2342–2348.

(22) Blin, K.; Shaw, S.; Steinke, K.; Villebro, R.; Ziemert, N.; Lee, S. Y.; Medema, M. H.; Weber, T. *Nucleic Acids Res.* **2019**, 47, W81–W87.

(23) Stachelhaus, T.; Walsh, C. T. *Biochemistry* **2000**, *39*, 5775–5787.

(24) Grundmann, F.; Kaiser, M.; Kurz, M.; Schiell, M.; Batzer, A.; Bode, H. B. *RSC Adv.* **2013**, *3*, 22072–22077.

(25) Fujii, K.; Ikai, Y.; Oka, H.; Suzuki, M.; Harada, K. *Anal. Chem.* **1997**, *69*, 5146–5151.

(26) Bonnington, L. S.; Tanaka, J.; Higa, T.; Kimura, J.; Yoshimura, Y.; Nakao, Y.; Yoshida, W. Y.; Scheuer, P. J. *J. Org. Chem.* **1997**, *62*, 7765–7767.

(27) Saito, K.; Konno, A.; Ishii, H.; Saito, H.; Nishida, F.; Abe, T.; Chen, C. *J. Nat. Prod.* **2001**, *64*, 139–141.

(28) Cha, J. W.; Park, J.-S.; Sim, T.; Nam, S.-J.; Kwon, H. C.; Del Valle, J. R.; Fenical, W. *J. Nat. Prod.* **2012**, *75*, 1648–1651.

(29) Patel, J. B.; Cockerill III, F. R.; Bradford, P. A.; Eliopoulos, G. M.; Hindler, J. A.; Jenkins, S. G.; Lewis II, J. S.; Limbago, B.; Miller, L. A.; Nicolau, D. P.; Powell, M.; Swenson, J. M.; Traczewski, M. M.; Turmidge, J. D.; Weinstein, M. P.; Zimmer, B. L. *Methods for Dilution Antimicrobial Susceptibility Tests for Bacteria That*

Grow Aerobically; Approved Standard M07–A10; Clinical and Laboratory Standards Institute: Wayne, PA, USA, 2015, Vol *35*, No.2.

(30) Kumaraswamy, M.; Lin, L.; Olson, J.; Sun, C.-F.; Nonejuie, P.; Corriden, R.; Döhrmann, S.; Ali, S. R.; Amaro, D.; Rohde, M. Pogliano, J.; Sakoulas, G.; Nizet, V. *J. Antimicrob. Chemother.* **2016**, *71*, 1264–1269.

(31) Subcommittee on Antifungal Susceptibility Testing (AFST) of the ESCMID European Committee for Antimicrobial Susceptibility Testing (EUCAST). *Clin. Microbiol. Infect.* **2008**, *14*, 398–405.

(32) Orhan, I.; Şener, B.; Kaiser, M.; Brun, R.; Tasdemir, D. *Mar. Drugs* **2010**, *8*, 47–58.

Supporting Information

Structure, Biosynthesis, and Bioactivity of Photoditritide from *Photorhabdus temperata* Meg1

Lei Zhao,[†,‡] Ryan Musumba Awori,[†] Marcel Kaiser,[§] Jonathan Groß,[□] Till Opatz,[□] and Helge B. Bode*[,†,∥]

[†]Molecular Biotechnology, Department of Biosciences, Goethe University Frankfurt, 60438 Frankfurt am Main, Germany

[‡]Institute of Botany, Jiangsu Province and Chinese Academy of Sciences, 210014 Nanjing, China

[§]Swiss Tropical and Public Health Institute, 4051 Basel, Switzerland

[□]Institute of Organic Chemistry, Johannes Gutenberg University Mainz, 55128 Mainz, Germany

[∥]Buchmann Institute for Molecular Life Sciences (BMLS), Goethe University Frankfurt, 60438 Frankfurt am Main, Germany

*Corresponding author: E-mail: h.bode@bio.uni-frankfurt.de

Supplementary Tables

Table S1. A domain specificity prediction of PdtS

A domain	Stachelhaus sequence	most likely amino acid predicted	amino acid detected
A1	DVESIGGVTK	Lys	Har
A2	DVESIGGVTK	Lys	Har
A3	DASFIADVCK	Tyr	Tyr
A4	DVQCIGDVCK	Tyr	Trp
A5	DASFIADVCK	Tyr	Tyr
A6	DASLVGDVCK	Tyr	Trp

Table S2. Bacterial strains used in this study

strain	genotype	reference
E. coli ST18	*E. coli* S17 λpir Δ*hemA*	[1]
P. temperata Meg1	wild type	[2]

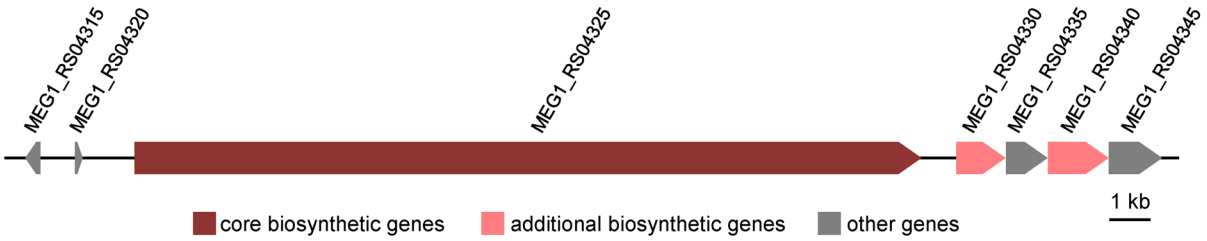
Table S3. Plasmids used in this study

plasmid	genotype/description	reference
pCEP	pDS132 based vector, R6Kγ ori, oriT, *ara*C, *ara*BAD promoter, Km^R	[3]
pLZ46	5443 bp, first 650 bp of *pdtS* gene from Meg1 genomic DNA assembled into pCEP, Km^R	this work

Table S4. Primers used in this study

primer	sequence (5'-3')	targeting DNA fragment	plasmid
LZ_112	TTTGGGCTAACA GGAGGCTAGCATATGAAAG ATAGCATTGCTGAAGC	first 650 bp of *pdtS* from Meg1 (650 bp)	pLZ46
LZ_113	TCTGCAGAGCTCGAGCATGCAC ATCACGGTGTGTCATCTTGTATCA		
pCEP_fw_gib	ATGTGCATGCTCGAGCTC	pCEP vector backbone (4841 bp)	
pCEP_rv_gib	ATGCTAGCCTCCTGTTAGC		
V_pCEP_fw	GCTATGCCATAGCATTTTTATCCATAAG	insert verification from pLZ46 (836 bp)	
V_pCEP_rv	ACATGTGGAATTGTGAGCGG		
V_pCEP_fw	GCTATGCCATAGCATTTTTATCCATAAG	conjugation verification from Meg1 pCEP-*pdtS* (1050 bp)	
LZ_114	ACTGAGCATCTGCTCTTGATAC		

Table S5. PdtS and its nearby proteins with predicted function



locus tag	predicted function
MEG1_RS04315	hypothetical protein
MEG1_RS04320	hypothetical protein
MEG1_RS04325 (PdtS)	non-ribosomal peptide synthase
MEG1_RS04330	aspartate aminotransferase family protein
MEG1_RS04335	arginine *N*-succinyltransferase
MEG1_RS04340	succinylglutamate-semialdehyde dehydrogenase
MEG1_RS04345	*N*-succinylarginine dihydrolase

Supplementary Figures

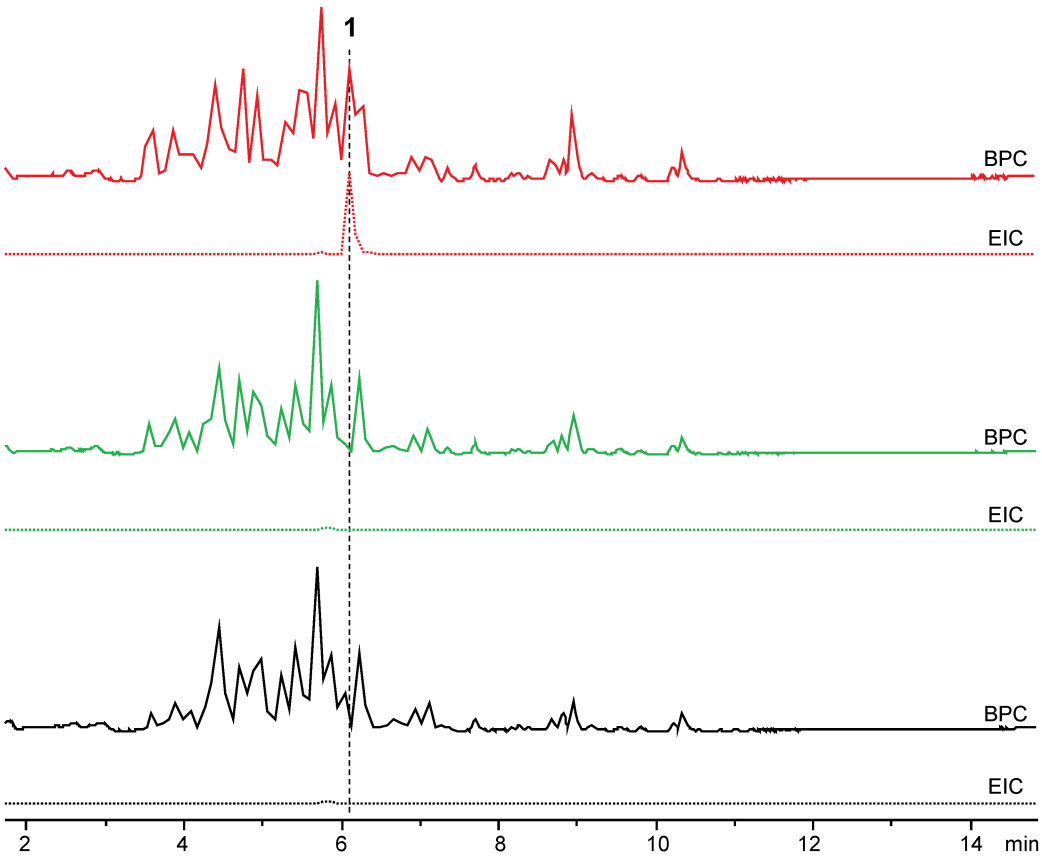


Figure S1. HPLC-MS analysis of XAD extracts from *P. temperata* Meg1 wild type (black line) and *P. temperata* Meg1 pCEP-*pdtS* without (green line) and with (red line) L-arabinose induction in LB medium. Base-peak chromatograms (BPCs) and extracted ion chromatograms (EICs) for 1 are shown, respectively.

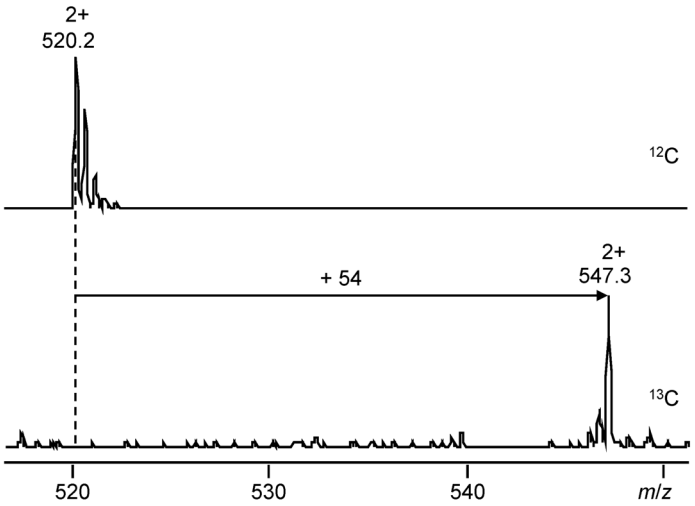


Figure S2. Determination of the number of carbon atoms for 1 by growing *P. temperata* Meg1 pCEP-*pdtS* in standard ^{12}C medium and fully ^{13}C labelled medium, respectively.

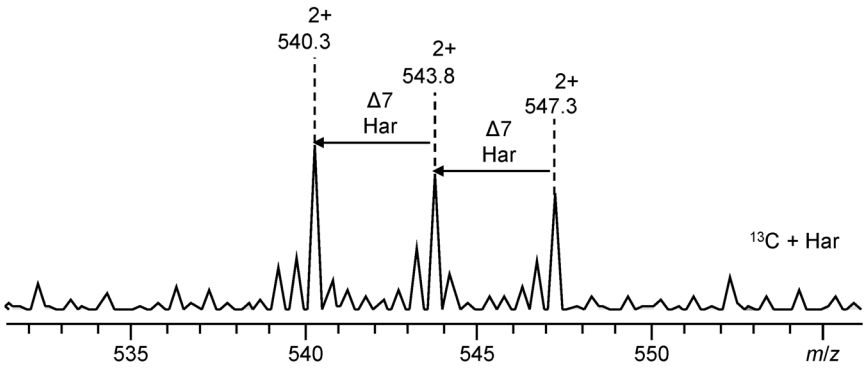


Figure S3. HPLC-MS analysis of 1 from reverse labelling experiment with L-homoarginine fed to *P. temperata* Meg1 pCEP-*pdtS* in fully ^{13}C labelled medium.

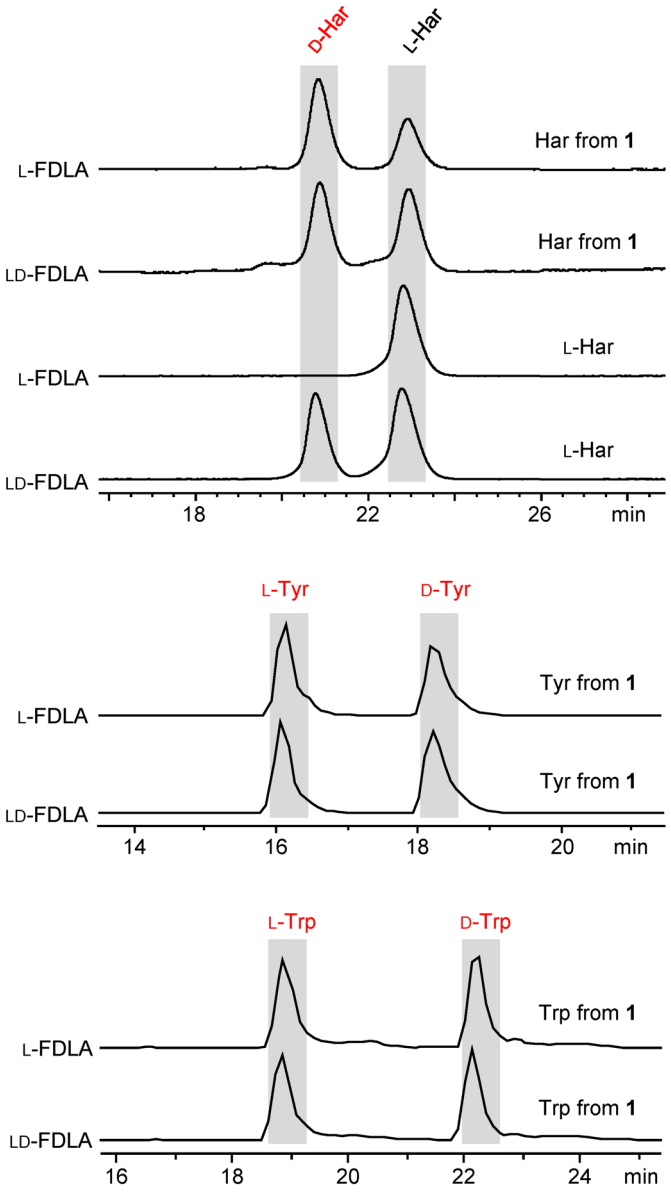


Figure S4. Configuration determination of amino acids in 1 using the advanced Marfey's method. HPLC-MS analysis of hydrolyzed **1** and derivatized with L-FDLA and LD-FDLA. Depicted are extracted ion chromatogram (EIC) traces for homoarginine (Har, *m*/*z* 483 $[M + H]^+$), tyrosine (Tyr, *m*/*z* 476 $[M + H]^+$), and tryptophan (Trp, *m*/*z* 499 $[M + H]^+$). The configurations are determined by the elution order: L-FDLA derivatized tyrosine and tryptophan elute prior to its D-enantiomer, while D-FDLA derivatized homoarginine elutes prior to its L-enantiomer. The small amount of L-Har might be from racemization of D-Har during acid hydrolysis of **1**.

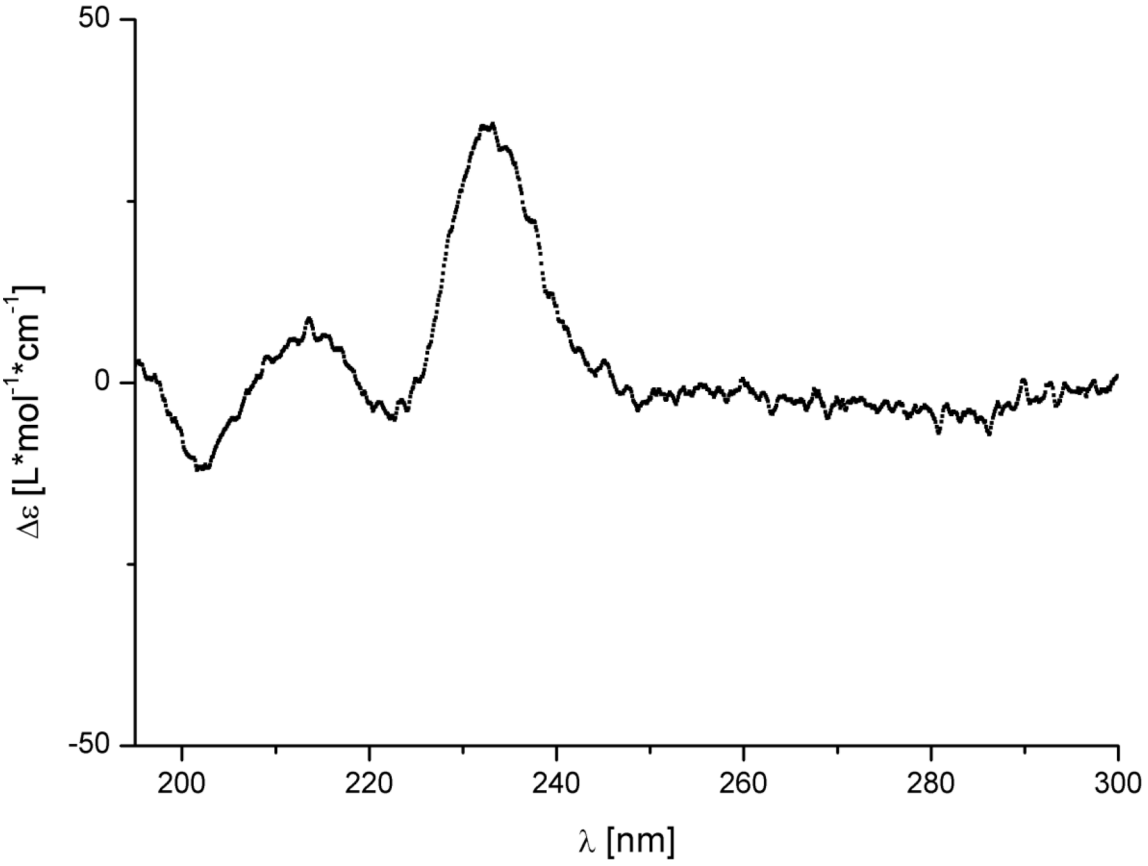


Figure S5. ECD spectrum of 1.

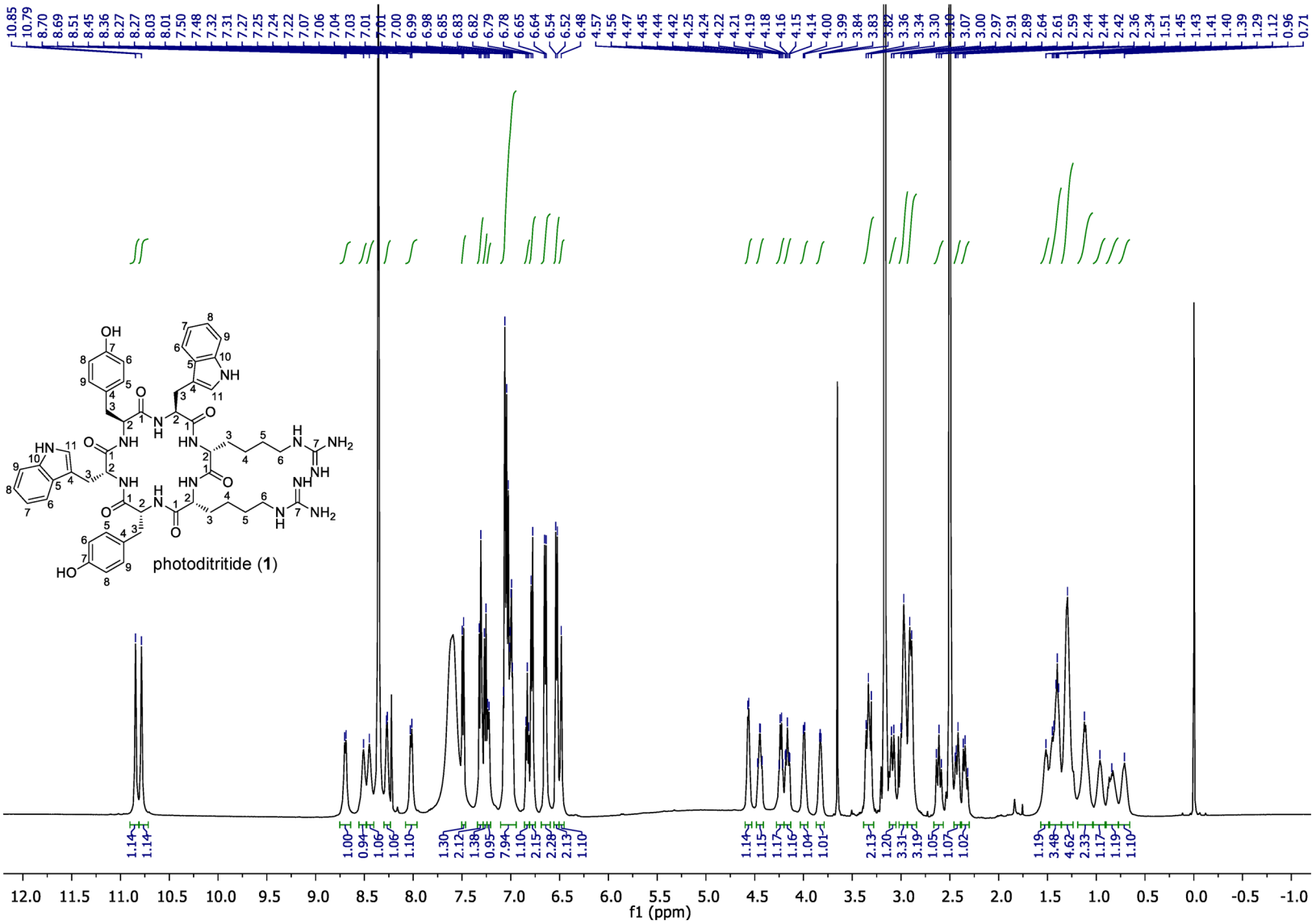


Figure S6. ^{1}H NMR (500 MHz, DMSO-d_6) spectrum of 1.

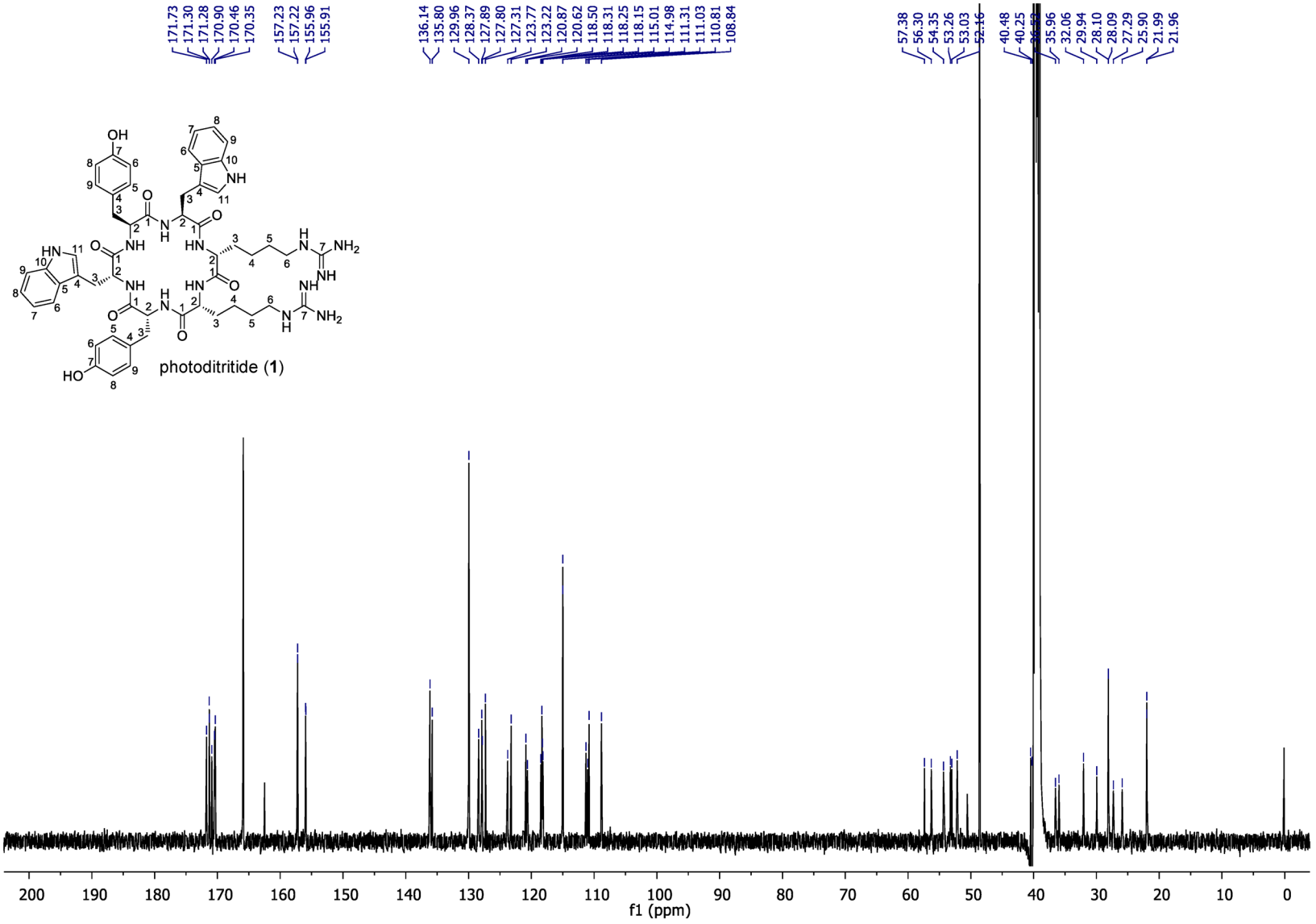


Figure S7. ^{13}C NMR (125 MHz, DMSO-d_6) spectrum of **1**.

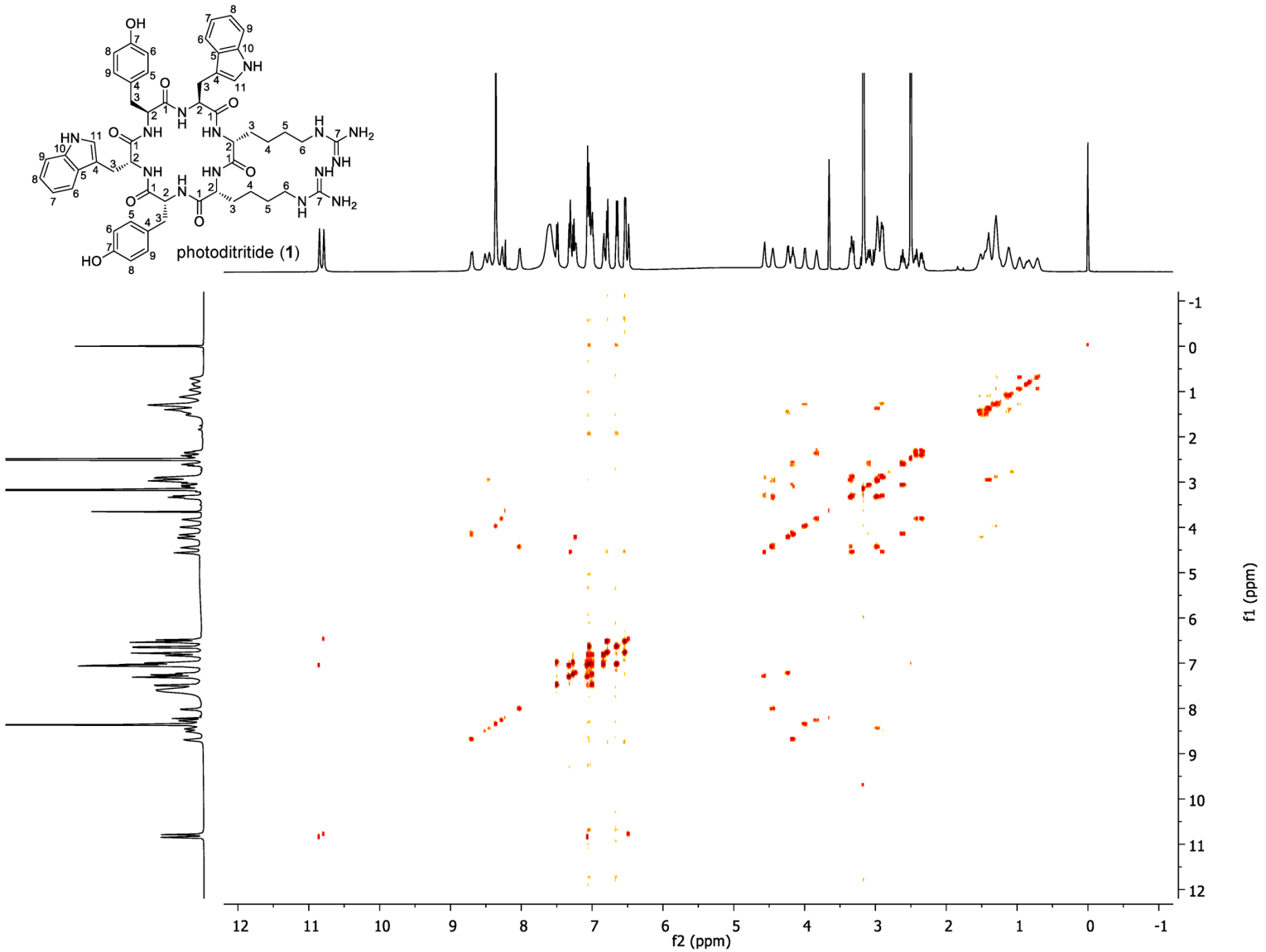


Figure S8. COSY (DMSO-d_6) spectrum of **1**.

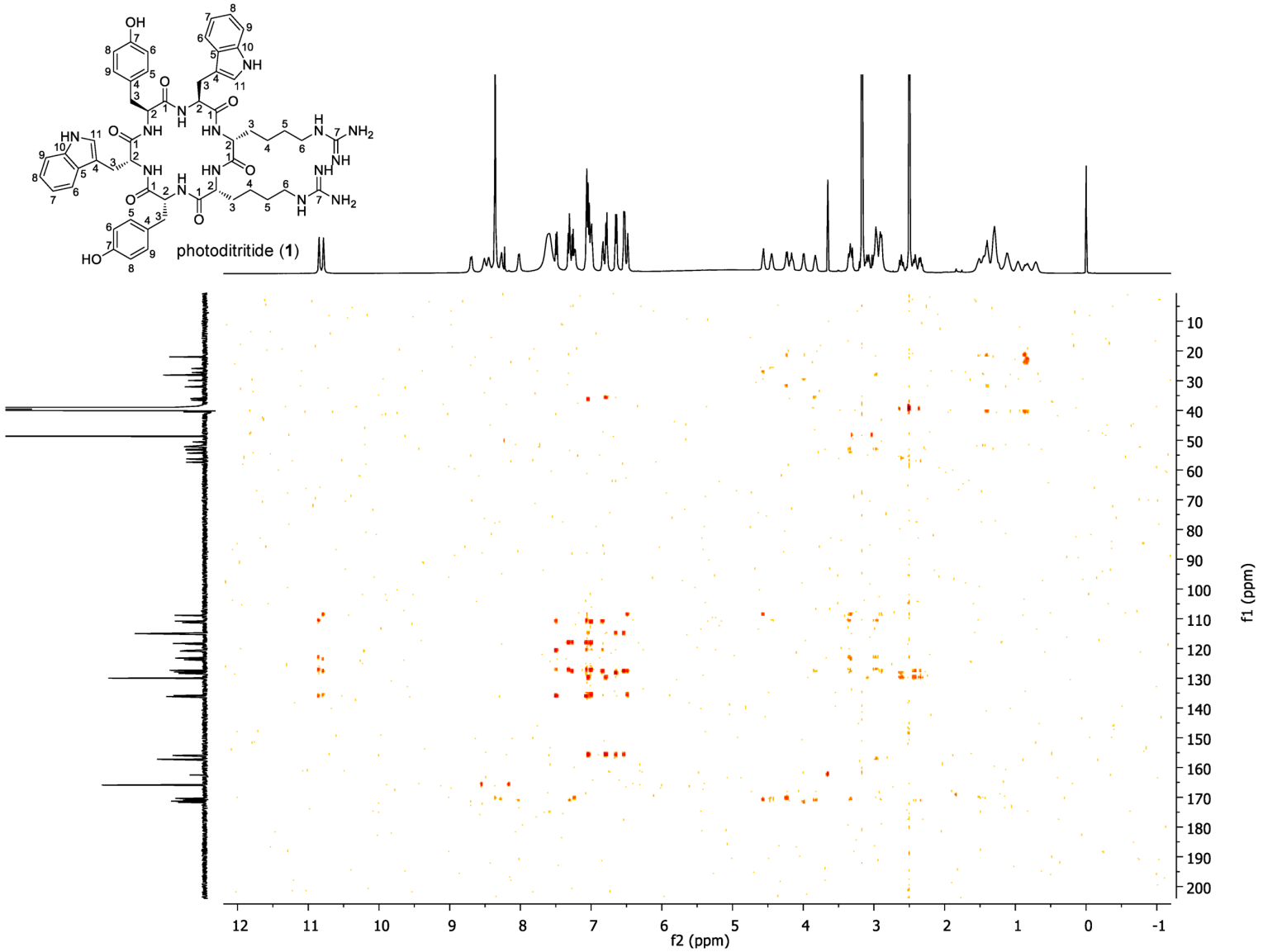


Figure S9. HMBC (DMSO-d_6) spectrum of **1**.

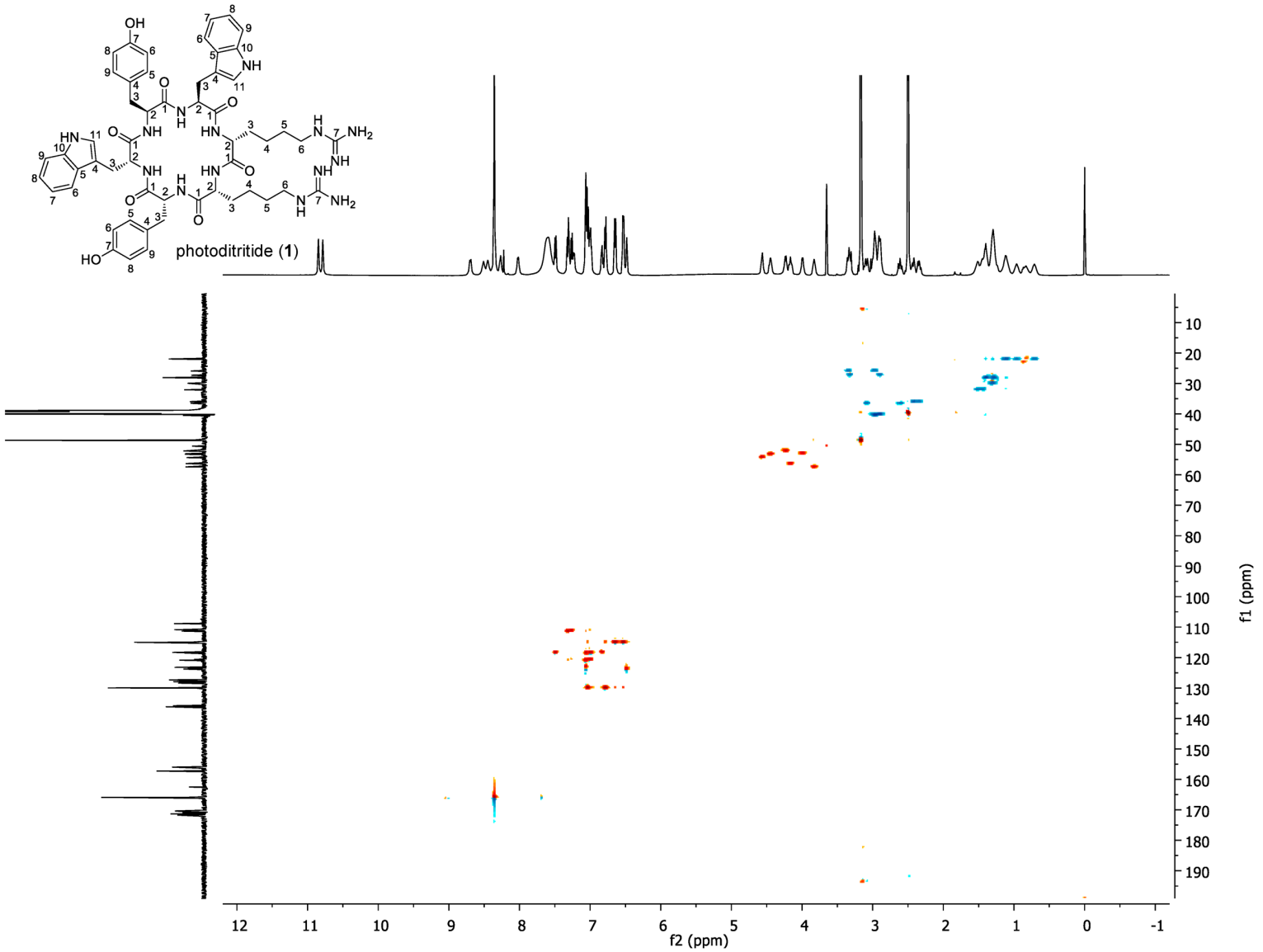


Figure S10. HSQC (DMSO-d_6) spectrum of **1**.

References

(1) Thoma, S.; Schobert, M. *FEMS Microbiol. Lett.* **2009**, *294*, 127–132.

(2) Hurst, S. G.; Ghazal, S.; Morris, K.; Abebe-Akele, F.; Thomas, W. K.; Badr, U. M.; Hussein, M. A.; AbouZaied, M. A.; Khalil, K. M.; Tisa, L. S. *Genome Announc.* **2014**, *2*, e01273-14.

(3) Bode, E.; Brachmann, A. O.; Kegler, C.; Simsek, R.; Dauth, C.; Zhou, Q.; Kaiser, M.; Klemmt, P.; Bode, H. B. *ChemBioChem* **2015**, *16*, 1115–1119.

Comprehensive transcriptome of the maize stalk borer, *Busseola fusca*, from multiple tissue types, developmental stages, and parasitoid wasp exposures

Kayla M. Hardwick[1,2,3] , Gladys Bosibori Bichang'a[3] , Andnet Bayleyegn Abtew[3], Ryan Musumba Awori[3], Leah C. S. Cepko[3], Lorna Jemosop Chebon-Bore[3], Alistair Darby[3], Jon D. deVries[1,3], Jonathan Filée[3], Muna Fuad[3], Grace Gachara[3], Dedan K. Githae[3], Patrick Gunga[3], Madison Held[1,3,] Hellen Wambui Kariuki[3], Evans Sioma Kataka[3], Susan Diana Kerfua[3], Kelvin Muteru Kimenyi[3], Bruno Le Ru[3], Ernest Mukweyi Lutomia[3], Elizabeth Ajema Luvai[3], Maureen K. Luvanda[3], Beatus Modest Lyimo[3], Eunice Magoma Machuka[3], Daniel G. Maeda[3], Solomon Maina[3], Peterson Gitonga Mathenge[3], Damaris Matoke-Muhia[3], Caitlin H. Miller[3], Silviane Aoko Miruka[3], Alfred Mitema[3], Antoinette Aluoch Miyunga[3], Sarah Akwabi Mukolwe[3], Edward Kirwa Muge[3], Simon Ngao Mule[3], Mary Kanyiri Murithi[3], Jean Pierre Musabyimana[3], Saphan Muzoora[3], Esther Wangui Mwangi[3], Harrison Ndung'u Mwangi[3], Nduta Mwangi[3], Ann Njeri Mwaura[3], Nelly Ndungu[3], Bidii Stephen Ngalah[3], Peter Njenga Ng'ang'a[3], Harun N. Njoroge[3], Brenda Nyaboke Nyandika[3], Johnson Ounya Nyasani[3], Faith Akinyi Obange[3], Shem Joabh Ochieng[3], Willingtone Otieno Odhiambo[3], Hellen Adhiambo Ogot[3], Maureiq Awino Edith Ojwang'[3], Fred Odhiambo Osowo[3], Billy Omboki Ratemo[3], Tolbert Sonda[3], Alexander Ssamula[3], Sharon Towett-Kirui[3], Augustin Penda Twizerimana[3],Kelvin Mwangi Wachiuri[3], Alex Paul Wacoo[3], Kevin Wamae[3], Mark Kilongosi Webale[3], Reuben Mangi Yaa[3], Dawn E. Gunderson-Rindal[3], Francesca Stomeo[3] Appolinaire Djikeng[3], Paul-Andre Calatayud[3], Sarah Schaack*[1,3]

Author affiliations:

[1]Department of Biology, Reed College, Portland, OR 9720228

[2]current address: Phylos Bioscience, Portland, OR 9721429

[3]Busseola fusca Genomics Consortium (for a complete list of all author addresses and affiliations, please see Appendix A)

*Corresponding Author, email: schaack@reed.edu

Keywords: agricultural pest, Lepidoptera, insect genomics

CURRICULUM VITAE

PERSONAL INFORMATION

Name: Ryan Musumba Awori **Nationality:** Kenya

Date of birth: 09/01/1989 **Place of Birth:** Nakuru, Kenya

ACADEMIC EDUCATION AND TRAINING

15/09/2019 – 17/09/2019 – Jena

International VAAM Workshop on Biology of Microorganisms Producing Natural Products – VAAM

Presented a Session VI talk titled "**Refactoring a silent biosynthetic gene cluster in *Xenorhabdus ishibashii* yields a novel dodecapeptide**"

01/07/2016 – 05/07/2016 – Nairobi, Kenya

Certificate in Science Paper Writing – International Centre of Insect Physiology and Ecology

---

09/2012 – 05/2015 – Nairobi, Kenya

Master of Science in Microbiology – University of Nairobi

The MSc thesis was titled **"Phylogeny and antibiotic activity of the bacteria *Xenorhabdus* isolated from nematode symbionts found in Kenya"**

Part of the dissertation was published as:

1. Awori RM, Ng'ang'a PN, Nyongesa LN et al. Mursamacin: a novel class of antibiotics from soil-dwelling roundworms of Central Kenya that inhibits methicillin-resistant *Staphylococcus aureus* [version 2; peer review: 2 approved, 1 approved with reservations]. F1000Research 2017, 5:2431 (https://doi.org/10.12688/f1000research.9652.2)
2. Antimicrobial agents produced by *Xenorhabdus griffiniae* XN45. WIPO PATENT NUMBER: WO2015173778. EPO Patent number EP3142676.KE PATENT NUMBER: KE 839

Academic Supervisors were:

1. Prof. Francis B. Mwaura,
   Thematic Head, Microbiology, Parasitology and Immunology,

School of Biological Sciences, University of Nairobi, Kenya
2. Dr. Daniel K. Masiga
   Head, Molecular biology and bioinformatics Department,
   International Centre for Insect Physiology and Ecology, Nairobi, Kenya
3. Dr. Charles N. Waturu
   Centre Director, Horticultural Research Institute, Thika, Kenya

---

09/04/2014 – 11/04/2014 – Nairobi, Kenya

East Africa *Busseola fusca* Genomics Consortium Annotation Jamboree and Workshop-Biosciences East and Central Africa Hub, Intentional Livestock Research Institute.

10/02/2014 – 14/02/2014 – Kampala, Uganda

2014 Workshop on Bioinformatics and Genomics – Department of Zoology, Makerere University.

22/04/2013 – 28/04/2013 – Nairobi, Kenya

Advanced Genomics and Bioinformatics Workshop: Next Generation Sequencing and its application to infectious diseases – Biosciences East and Central Africa Hub, International Livestock Research Institute**.**

09/2008 – 08/2012 – Nairobi, Kenya

Bachelor of Science in Microbiology and Biotechnology (Second Class Honours-Upper Division) – University of Nairobi
The BSc thesis was titled **"Isolation of *Xenorhabdus* bacteria from *Steinernema***

nematodes found in Kenya."

Academic Supervisors were:
1. Prof. Francis B. Mwaura, Thematic Head, Microbiology, Parasitology and Immunology, School of Biological Sciences, University of Nairobi, Kenya.
2. Dr. Nelson Amugune, School of Biological Sciences, University of Nairobi, Kenya.

05/2007 – 07/2010 – Nairobi, Kenya

Diploma in Business Information Technology (Merit) – Strathmore University

---

01/2003 – 11/2006 – Thika, Kenya

Kenya Certificate of Secondary Education (mean grade A) – Mang'u High School

---

ACADEMIC HONOURS AND AWARDS

1. First Place Award at the 2016 DAAD Falling Walls Lab Nairobi for an innovation titled "Breaking the Walls of Antimicrobial Resistance". DAAD-German Academic Exchange Service. Nairobi, Kenya. (13/09/2016)
2. Innovation of the Year-first runners-up Award. National Commission of Science Technology and Innovation. Nairobi, Kenya. (20/05/2016)
3. Best Poster Presentation-first runners-up Award. Africa International Biotechnology and Biomedical Conference. Nairobi, Kenya. (18/09/2015)
4. Carl Storm International Diversity Fellowship worth EUR 1131. Gordon Research Conferences, West Kingston, Ireland. (23/04/2015)
5. AuthorAid Travel Grant Award worth EUR 1378. INSAP. Oxford, United Kingdom (06/03/2015)
6. MSc Scholarship Award worth EUR 1869. World Federation of Scientists. Geneva, Switzerland. (01/06/2014)
7. ST&I Innovation grant Award worth EUR 8056. National Commission for Science Technology and Innovation. Nairobi, Kenya. (16/10/2011)