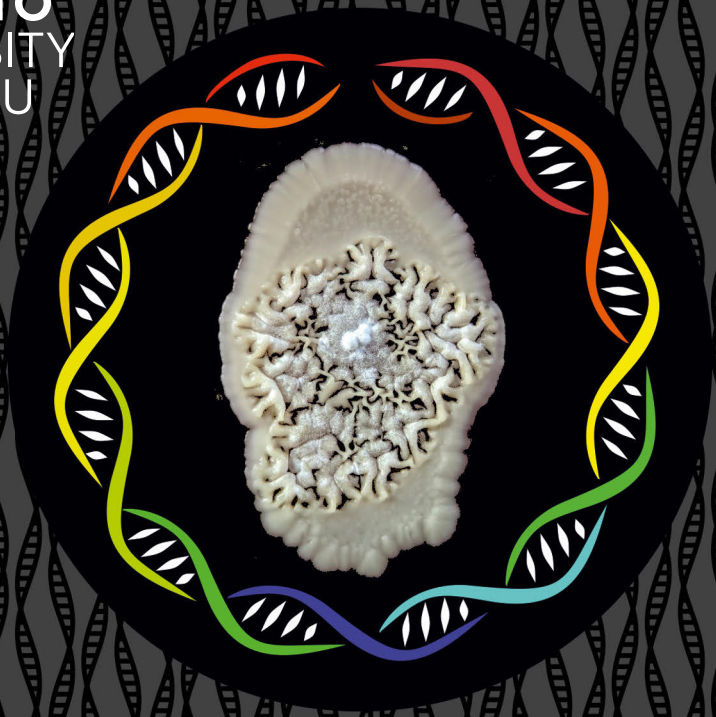
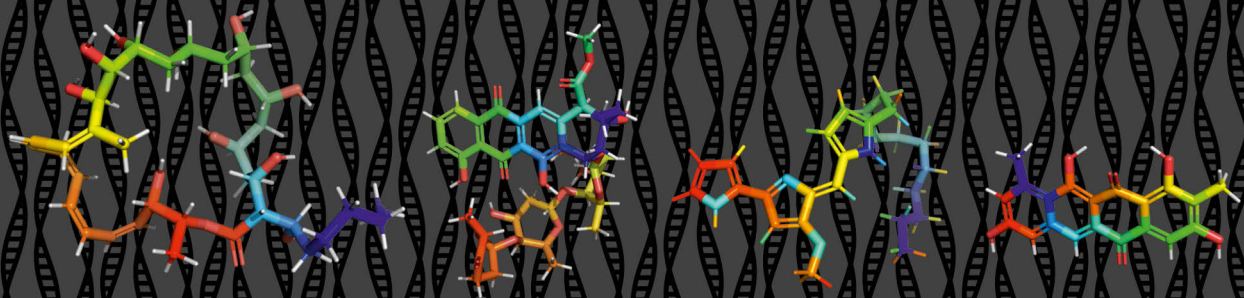
TURUN YLIOPISTO
UNIVERSITY OF TURKU



GENOME MINING ACTINOBACTERIA

Eliciting the Production of Natural Products

Keith Yamada



TURUN YLIOPISTON JULKAISUJA – ANNALES UNIVERSITATIS TURKUENSIS
SARJA – SER. AI OSA – TOM. 693 | ASTRONOMICA – CHEMICA – PHYSICA – MATHEMATICA | TURKU 2023

TURUN
YLIOPISTO
UNIVERSITY
OF TURKU

GENOME MINING ACTINOBACTERIA

Eliciting the Production of Natural Products

Keith Yamada

TURUN YLIOPISTON JULKAISUJA – ANNALES UNIVERSITATIS TURKUENSIS
SARJA – SER. AI OSA – TOM. 693 | ASTRONOMICA – CHEMICA – PHYSICA – MATHEMATICA | TURKU 2023

University of Turku

Faculty of Technology
Department of Life Technologies
Biochemistry
Doctoral Programme in Technology

Supervised by

Professor, Mikko Metsä-Ketelä
Department of Life Technologies
University of Turku
Turku, Finland

Reviewed by

Dr. Gerald Lackner
Leibniz Institute for Natural Product
Research and Infection Biology –
Hans Knöll Institute
Jena, Germany

Docent, Miia Mäkelä
Department of Microbiology
University of Helsinki
Helsinki, Finland

Opponent

Research Professor, Istvan Molnar
VTT Technical Research Centre of
Finland Ltd.
Espoo, Finland

The originality of this publication has been checked in accordance with the University of Turku quality assurance system using the Turnitin OriginalityCheck service.

Cover Image: Keith Yamada

ISBN 978-951-29-9287-4 (PRINT)
ISBN 978-951-29-9288-1 (PDF)
ISSN 0082-7002 (Print)
ISSN 2343-3175 (Online)
Painosalama, Turku, Finland 2023

A JACK OF ALL TRADES IS A MASTER OF NONE,
BUT OFTENTIMES BETTER THAN A MASTER OF ONE
-William Shakespeare

UNIVERSITY OF TURKU
Faculty of Technology
Department of Life Technologies
Biochemistry
KEITH YAMADA: Genome Mining Actinobacteria; Eliciting the Production of Natural Products
Doctoral Dissertation,181 pp.
Doctoral Programme in Technology
June 2023

ABSTRACT

Antimicrobial resistance is an imminent threat that is expected to kill 10 million people per year by 2050. Natural products have been a major source of antimicrobial compounds and encompass a chemical space far greater than synthetic chemistry can provide and have evolved over the ages to have biological activities. The natural products produced by *Streptomyces* have provided us with two-thirds of the antibiotics currently used, as well as chemotherapeutics, antifungals, and immunosuppressants. In recent years, the drug discovery pipeline from *Streptomyces* has run dry, largely because laboratory culture conditions lack their natural stimuli, resulting in the rediscovery of the same natural products. However, with the advent of modern genomics, we now realize that the genomes of *Streptomyces* have the have a greater capacity for natural product production than what we have observed, which providing us with the hope of new drug leads.

My doctoral research focused on two aims. First is the concept of eliciting *Streptomyces* to produce novel natural products; in other words, what triggers the production of natural products. Additional perspectives from ecology, evolution, and regulation evoked the idea that microbe-microbe interactions could be the key. Microscopic observations of the interactions between *Streptomyces* and yeast suggested that physical contact is essential for elicitation. Genomics, transcriptomics, and proteomics showed that 31% of the silent biosynthetic gene clusters are activated, notably an antifungal polyene cluster, as well as a suite of enzymes capable of digesting the cell wall of yeast. Arguably, *Streptomyces* can prey on yeast. The differential regulation of a homologous polyene gene clusters further suggested that natural product production is triggered by different ecological needs.

Second, genome mining provides insight into the genetic potential of Actinobacteria to produce natural products via the identification of their gene clusters and is a proven method that aids in drug discovery. Here, I sequenced the genome of a rare *Streptomonospora* isolate and identified the novel persiamycin gene cluster and its associated product. Moreover, genome mining was applied to publicly available *Streptomyces* genomes, which resulted in the identification of two new gene clusters that produce the antibiotic komodoquinone B.

KEYWORDS: Streptomyces, Actinobacteria, secondary metabolism, natural products, genome mining

TURUN YLIOPISTO
Teknillinen tiedekunta
Bioteknologian laitos
Biokemia
KEITH YAMADA: Genomin louhinta aktinobakteereissa; luonnontuotteiden tuotannon edistäminen
Väitöskirja, 181 s.
Teknologian tohtoriohjelma
Kesäkuu 2023

TIIVISTELMÄ

Mikrobilääkeresistenssi on välitön uhka, joka uhkaa tappaa vuosittain 10 miljoonaa ihmistä vuoteen 2050 mennessä. Luonnontuotteet ovat olleet merkittävä mikrobilääkkeiden lähde. Lisäksi luonnontuotteet kattavat paljon laajemman kemiallisen alueen kuin synteettinen kemia voi tarjota, ja ne ovat aikojen kuluessa kehittyneet niin, että niillä on biologisia vaikutuksia. Streptomykeetti-organismin tuottamat luonnontuotteet ovat tuottaneet kaksi kolmasosaa käytössä olevista antibiooteista sekä kemoterapeuttisia aineita, sienilääkkeitä, immunosuppressantteja ja matolääkkeitä. Viime vuosina Streptomykeetti-bakteerin lääkkeiden kehittämiskanava on kuitenkin kuivunut. Nykyaikaisen genomiikan myötä olemme kuitenkin nyt ymmärtäneet, että Streptomykeeteillä on potentiaalia tuottaa vielä löytämättömiä luonnontuotteita, jotka antavat meille toivoa uusista lääkkeistä, erityisesti antibiooteista.

Väitöstutkimuksellani oli kaksi tavoitetta. Ensimmäinen tarkoitus oli saada Streptomykeetit tuottamaan uusia luonnontuotteita; toisin sanoen tutkia sitä, mikä saa Streptomykeetin tuottamaan luonnontuotteita. Ekologian, evoluution ja säätelyn lisänäkökulmat herättivät ajatuksen, että mikrobien väliset vuorovaikutus voisivat olla avainasemassa. Mikroskooppiset havainnot Streptomykeetin ja hiivan välisestä vuorovaikutuksesta osoittivat, että fyysinen kontakti oli välttämätön yhdisteiden tuoton aktivoinnissa. Genomiikka, transkriptomiikka ja proteomiikka osoittivat, että jopa 31 prosenttia hiljaisista biosynteettisistä geeniryhmistä aktivoitui, samoin kuin joukko entsyymejä, jotka kykenevät pilkkomaan hiivan soluseinää. Tämä osoitti, että Streptomykeetit pystyvät saalistamaan hiivasoluja. Homologisten geeniryhmien erilainen säätely viittaa lisäksi siihen, että luonnontuotteiden tuotanto perustuu erilaisiin ekologisiin tarpeisiin.

Toisena tavoitteena oli käyttää genomien louhintaa selvittääksemme Streptomykeetin geneettisestä potentiaalista tuottaa luonnontuotteita. Tässä työssä selvitimme harvinaisen Streptomonaspora bakteerin genomin ja tunnistin uuden persiamysiini-yhdisteen biosynteesireitin. Tämän lisäksi löysimme genomin louhinnan avulla kaksi uutta biosynteettistä geeniryhmää julkisista tietokannoista, joiden osoitimme olevan vastuussa komodokinoni B antibiootin tuotannosta.

ASIASANAT: Streptomykeetit, Aktinobakteerit, sekundaarinen aineenvaihdunta, luonnonyhdisteet, genomin louhinta

Table of Contents

Abbreviations

A	adenylation
A + T	adenine + thymine
ACP	acyl carrier protein
AMR	antimicrobial resistance
ANI	average nucleotide identity
AT	acyltransferase
ATP	adenosine triphosphate
BGC	biosynthetic gene cluster
C	condensation
CAZyme	carbohydrate-active enzyme
CDS	coding sequence
CRC	colorectal cancer
DH	dehydratase
DNA	deoxyribonucleic acid
DSB	double strand break
E	epimerization
ER	enoylreductase
G + C	guanine + cytosine
GCF	gene cluster family
GFP	green fluorescent protein
GH	glycoside hydrolase
GlcNAc	*N*-acetylglucosamine
HGT	horizontal gene transfer
HMA	high-microbial-abundance
IBD	inflammatory bowel disease
KR	ketoreductase
KS	ketosynthase
KSα	ketosynthase alpha
KSβ	ketosynthase beta
LMA	low-microbial-abundance
MV	membrane vesicle

NP	natural product
NRPS	nonribosomal peptide synthetase
OMV	outer membrane vesicle
PCD	programmed cell death
PCP	peptidyl carrier protein
PCR	polymerase chain reaction
pHMM	profile hidden markov model
PKS	polyketide synthase
RFP	red fluorescent protein
RNA	ribonucleic acid
rRNA	ribosomal ribonucleic acid
SARP	*Streptomyces* antibiotic regulatory proteins
TE	thioesterase
VOC	volatile organic compound
WGS	whole genome sequencing

List of Original Publications

This dissertation is based on the following original publications, which are referred to in the text by their Roman numerals:

I **Yamada K**, Koroleva A, Laughlin M, Oksanen N, Akhgari A, Safronova V, Yakovleva E, Kolodyaznaya V, Buldakova T, Metsä-Ketelä M. Characterization and overproduction of cell-associated cholesterol oxidase ChoD from *Streptomyces lavendulae* YAKB-15. *Scientific Reports*, 2019; 9: 1-8.

II **Yamada K***, Koroleva A*, Tirkkonen H, Siitonen V, Laughlin M, Akhgari A, Mazurier G, Niemi J, Metsä-Ketelä M. Physical interactions trigger *Streptomyces* to prey on yeast using natural products and lytic enzymes. (Unpublished), 2023

III Baral B*, Siitonen V*, Laughlin M, **Yamada K**, Ilomäki M, Metsä-Ketelä M, Niemi J. Differential regulation of undecylprodigiosin biosynthesis in the yeast-scavenging *Streptomyces* strain MBK6. *FEMS Microbiology Letters*, 2021; 8: 1-9.

IV Matroodi S, Siitonen V, Baral B, **Yamada K**, Akhgari A, Metsä-Ketelä M. Genotyping-guided discovery of persiamycin A from sponge-associated halophilic *Streptomonospora* sp. PA3. *Frontiers in Microbiology*, 2020; 11: 1-15.

V Grocholski T, **Yamada K**, Sinkkonen J, Tirkkonen H, Niemi J, Metsä-Ketelä M. Evolutionary trajectories for the functional diversification of anthracycline methyltransferases. *ACS Chemical Biology*, 2019; 5: 850-856.

List of Contemporaneous Publications

I Rosenqvist P, Palmu K, Prajapati RK, **Yamada K**, Niemi J, Belogurov GA, Metsä-Ketelä M, Virta P. Characterization of C-nucleoside antimicrobials from *Streptomyces albus* DSM 40763: Streptouridin is pseudouridimycin. *Scientific Reports*, 2019; 9: 1-9.

II Biosynthetic aclacinomycin A invention (2021)

III Biosynthetic doxorubicin patent (2021)

1 Introduction

Antimicrobial resistance (AMR) is an increasing problem. AMR is when bacteria, viruses, fungi, and parasites (microbes) become resistant to medicines. In 2019 almost 5 million deaths were associated with bacterial AMR, including 1.27 million deaths attributed to bacterial AMR.[1] The Review on Antimicrobial Resistance, commissioned by the UK Government, estimates that AMR could kill 10 million people per year by 2050.[2] The World Bank estimates 28.3 million people could fall into extreme poverty and an economic loss of an estimated $100 trillion by 2050 because of AMR.[3]

The problem of AMR is recognized by many high-level global organizations. The World Health Organization (WHO) has repeatedly warned that the world is heading to a "post-antibiotic era" due to the rapid increase of AMR.[4] This post-antibiotic era would prevent routine surgery and chemotherapy because there would be no drugs to prevent or treat microbial infections[5] and harks back to the pre-antibiotic era. Other influential global organizations, including the United Nations General Assembly, the G7 and the G20, have also held high-level meetings about AMR.[6] Despite these major threats, the rate of emergence of AMR is increasing while the rate of new antimicrobials discovered is decreasing.[7–9]

The decrease in the rate of newly discovered antimicrobials can be attributed to the end of the Golden Age of antibiotic discovery, the source of most of the antibiotics used today. To fully understand and appreciate this Golden Age and its complex relationship with AMR, it is worthwhile to start from the beginning of the antibiotic revolution.

The antibiotic revolution started in 1907.[10] Nobel laureate Paul Ehrlich postulated that a "magic bullet" could kill specific disease causing microbes without harming the body itself, which led to the development of Salvarsan, a portmanteau for "saving arsenic"[11]. This marked the end of the pre-antibiotic era, which is characterized by infectious diseases, such as smallpox, diphtheria, and tuberculosis, as the leading causes of death.[12]

Then in 1928, inspired by the magic bullet, Nobel laureate Sir Alexander Fleming discovered penicillin, the first modern antibiotic.[13] His discovery was from serendipitous contamination of a staphylococci plate culture with a fungus. He

noticed that the staphylococci colonies immediately surrounding the fungus had been destroyed, while the colonies further away were not affected. Fleming's screening method, using inhibition zones, was cheap and effective and became the standard for the screening of antibiotic-producing microbes.[14] However, it was not until 1943 that penicillin would be widely available due to lack of interest and problems in production.[15]

The first broadly effective commercially available antibiotic, the sulfonamide prontosil, was chemically synthesized and shown to be effective against bacterial infections by Nobel laureate Gerhard Domagk in 1932.[16] This discovery marked a new era in medicine, as seen by the transition from the use of disinfectants and topical antiseptics to modern antimicrobial drugs.[17] The success of Prontosil created the sulfa craze, which led to the current sulfonylureas and thiazide diuretics used today. However, the success of sulfa drugs was compromised by the broad dissemination of AMR, that is, resistance to sulfa drugs. This resistance, at least in part, prompted researchers to continually modify derivatives so that sulfa drugs can still be used today, exemplifying the arms race between antimicrobials and AMR.[10]

Later, in 1943, Nobel laureate Selman Waksman, who recognized Actinobacteria as important contributors to the decomposition of plant biomass,[18] discovered streptomycin and coined the term 'antibiotic' in the medical sense, which subsequently started the Golden Age of antibiotic discovery.[19]

The Golden Era of antibiotic discovery began with the Waksman Platform, which was based on the Fleming screening method but was more systematic in nature.[20] The Waksman Platform was based on the observation of ecological interactions between microbes, where one microbe is able to inhibit or kill another by secreting specialized metabolites known as natural products (NPs). Between the 1940s and 1970s more than 160 new NP antibiotics and semi-synthetic derivatives were commercialized, including tetracyclines, cephalosporins, aminoglycosides, and vancomycin, which are still in use today.[7] The success of the Golden Age has made infectious diseases falling to the fourth leading cause of death today.[21]

Although the Golden Age was quite successful in reducing deaths, AMR continued to compromise every new discovery.[22] For example, *Staphylococcus aureus* infections were initially controlled with penicillin, but resistance developed shortly after. Methicillin, the first designer anti-resistance antibiotic, was then developed in 1959 and within just 3 years methicillin-resistant *S. aureus* (MRSA) emerged, which inexorably led to other multidrug-resistant strains, i.e., 'superbugs'. This shows how quickly AMR outpaces our current drug discovery and development pipeline, 3 years and 15 years, respectively.[23]

Nevertheless, the arms race between antibiotics, particularly NPs, and AMR began long before the Golden Age of antibiotic discovery. Indeed, AMR is a natural phenomenon, with resistance genes having evolved more than two billion years

ago.[10] It is no surprise that Actinobacteria, the most prolific producer of NPs, have their origins dating back over 2.7 billion years ago[24] because without the threat of an attack by a NP there would be no selective pressure to evolve resistance. Actinobacteria themselves concomitantly confer AMR to the NP they produce to protect themselves against their own weapons, demonstrating co-evolution with NPs and AMR within producing organisms. In this way, the arms race is never ending, since for every new NP a new resistance mechanism will evolve and for every new resistance mechanism a new NP will evolve.

Humanity has only recently joined this race. We discovered the extensive utility of NPs that are valuable as antimicrobials and additionally as anticancer agents (doxorubicin and aclacinomycin), food preservatives and antifungals (natamycin), anthelmintic drugs and insecticides (avermectin and derivatives such as ivermectin and abamectin), immunosuppressants and anticancers (rapamycin), and anti-diabetics (acarbose). Furthermore, we have found that NPs have the advantage of greater scaffold diversity and structural complexity compared to typical synthetic small molecules[25] and are still an important source of drug leads.[26] We also discovered the hidden genetic potential of Actinobacteria to be able to produce more NPs that have been observed in laboratory conditions.[27] We know that NPs are produced to inhibit or kill competing microbes; however, we have yet to fully uncover other reasons for their production, thus limiting our ability to reveal the extent of their chemical space and biological activities.[28]

Our lack of knowledge about why NPs are produced likely limits our ability to elicit them in the laboratory.[29] NPs are energetically expensive to produce, and therefore they are likely to be produced when needed. This need is often an environmental signal that elicits production, such as a specific molecule and/or nutrient depletion. Consequently, our blindness to elicitors was likely a contributor to the decline of the Golden Age, where the rediscovery of already known NPs became rampant due to standard axenic laboratory conditions not being conducive towards production.[25] When the producers are cultured axenically they are unlikely to produce NPs, since NPs are thought to be important for survival by mediating interactions with other organisms. Recently, ecological interactions, especially symbiosis and predation, have been proven to provide information not only on the natural role of NPs but also on enzymes, which in turn can provide promising leads toward elicitation.[30–37] Therefore, we investigated the elicitation of NP and enzyme production by exploring ecological interactions of Actinobacteria and discuss how production is influenced by their life cycle, genomes, genetic regulation, and evolution.

1.1 The Life Cycle of *Streptomyces*

Actinobacteria are a large and highly biodiverse phylum and the most prolific source of NPs.[38] Actinobacteria also exhibit an enormous diversity in terms of morphology, physiology, and metabolic capabilities. Ecologically, its members have evolved to be pathogens (*Corynebacterium*, *Mycobacterium*, *Nocardia*, and *Tropheryma*), plant commensals (*Leifsonia*), gastrointestinal commensals (*Bifidobacterium*), marine (*Streptomonospora*) and soil inhabitants (*Streptomyces*). They also range in lifestyle from anaerobic (*Arachnia*, *Rothia*, and *Bifidobacterium*) to aerobic (e.g., *Nocardia*, *Rhodococcus*, and *Streptomyces*), unicellular (e.g., *Micrococcus* and *Mycobacterium*) to filamentous (e.g., *Amycolatopsis*, *Frankia* and *Streptomyces*), and most are spore forming and saprophytic. Their broad interactions with various organisms, tissues, and cells are likely origins for their large variability of metabolites and activities.[39]

Streptomyces are the largest, most well-known, and most studied genus of Actinobacteria. Historically, *Streptomyces* were considered fungi due to similar lifestyles[40,41] or a transitional form between bacteria and fungi[42] and the name itself means 'twisted fungi' in ancient Greek. Today, *Streptomyces* are known as soil-dwelling Gram-positive bacteria and are the most fruitful source of NPs, especially secondary metabolites and enzymes.[43] Secondary metabolites of *Streptomyces* account for two-thirds of antibiotics and one-third of chemotherapeutics in use, as well as antifungals, immunosuppressants, and anthelmintic drugs, to name but a few.[7] Moreover, *Streptomyces* produce enzymes that catalyze chemo-, regio-, enantio-, and stereo-selective reactions more efficiently than synthetic chemistry.[44] The life cycle of *Streptomyces* is complex and is classically thought of in three stages: spores, vegetative mycelium, and aerial hyphae (**Figure 1**).

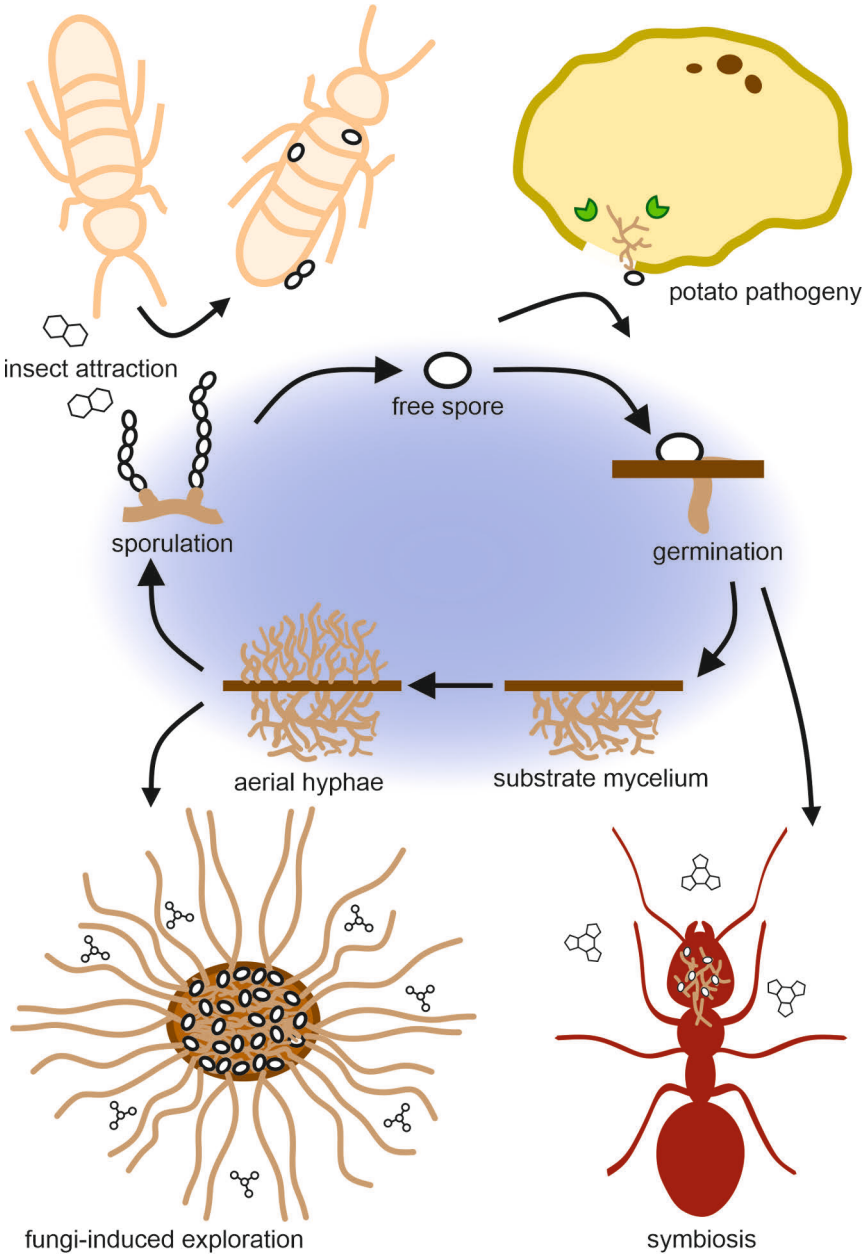


Figure 1. ***Streptomyces*** **generalized lifecycle.** The classical free-living lifecycle is highlighted in blue and recent additions are on the top and bottom. Top-left: *Streptomyces* produces geosmin to attract an arthropod that consumes mycelia and spreads the spores.[65] Top-right: Pathogenic *Streptomyces* infecting a potato.[63] Bottom-left: *Streptomyces* growing on yeast and producing trimethylamine, which can induce the exploration phenotype.[67] Bottom-right: Symbiotic *Pseudonodocardia* growing on an ant and producing antifungal NPs.[64]

Abiotic factors in the *Streptomyces* ecosystem have been shown to be elicitors of NPs with nutrient starvation being the most studied. During nutrient starvation *N*-acetylglucosamine (GlcNAc) acts as an elicitor of NP production. GlcNAc is a success story of an elicitor that is now used in drug discovery screening regimes.[29] This phenomenon is best described through the interconnected features of the *Streptomyces* life cycle, i.e., its large genome size, multicellularity, and morphological differentiation.

Morphological differentiation (sporulation) is temporally correlated with chemical differentiation, as shown in *Streptomyces coelicolor*.[45,46] First, a spore will germinate and then create a complex vegetative mycelial network.[47] Nutrients are obtained by the release of exoenzymes, which breakdown natural polymers, such as cellulose, glucans, mannans, and chitins. Upon nutrient depletion, *Streptomyces* is triggered to initiate the development of reproductive aerial hyphae by *bld* genes and the autolysis of its vegetative hyphae, i.e., programmed cell death (PCD), to provide the necessary nutrients.[48] The PCD lysate is largely composed of cell wall peptidoglycan, which are recycled to release the amino sugar GlcNAc. GlcNAc is internalized and metabolized until it becomes the ligand for the global regulator DasR.[49] DasR is consequently inactivated and metabolic pathways for NP production are activated.[50] In *S. coelicolor*, the production and secretion of NPs actinorhodin, calcium-dependent antibiotic, and coelimycin P1 are likely to defend PCD-released nutrients against motile microbes and to lyse competitors to gain more nutrients.[51,52] Meanwhile, prodiginine production but not secretion appears to apparently facilitate PCD via DNA damage.[53]

Multicellularity features include intraspecies communication, morphological differentiation, and PCD.[51] A multicellular population of bacteria needs to be able to communicate between subpopulations so they can coordinate differentiation, e.g., the coordination of vegetative mycelia PCD and aerial hyphae sporulation as described above. Intraspecies communication is also observed through the division of labor for antibiotic production, where a subpopulation differentiates into diverse antibiotic overproducers, by large deletions of the unstable genome, at the cost of reduced fitness.[54] Another example is the spatiotemporal patterning of gene expression determined by regulatory networks and physiological gradients.[55,56] Multicellularity is a way to simultaneously resolve trade-offs between mutually incompatible traits, e.g., morphological and chemical differentiation.[57,58]

The large genome of *Streptomyces* has been correlated with biosynthetic diversity.[59,60] In general, bacteria with larger genomes are generalists with expanded metabolism that allows them to use multiple nutrient sources through the production of NPs and hydrolytic exoenzymes, as observed in the most important producers, i.e., Actinobacteria and Myxobacteria. These generalist bacteria have a complex morphology, e.g., multicellularity and morphological differentiation, allowing them

to adapt to diverse environments.[51] On the contrary, bacteria with smaller genomes tend to be specialists with a dedicated life style in a specific environment and therefore require less diverse metabolism, e.g., *Mycoplasma*.[61,62]

Examples of branches to the *Streptomyces* life cycle are often based on ecological interactions and include, infecting potatoes[63], protecting ants[64], and attracting arthropods with geosmin[65] (**Figure 1**). Geosmin has also been shown to function as a deterrent of the bacteriophagous nematode *Caenorhabditis elegans*.[66]

Recently, a new stage in the *Streptomyces* life cycle was discovered called 'exploration' (**Figure 1**) and is characterized by a physically associated fungal trigger, pH-induced morphological switches, and volatile organic compound (VOC)-mediated communication.[67] Further, this study showed that the VOC elicitor trimethylamine raises the pH around the colony and specifically reduces iron availability, in turn perturbing the surrounding microbes and enhancing the exploration of other *Streptomyces* (interspecies interaction). During exploratory growth, *Streptomyces venezuelae* produces a diverse array of siderophore NPs highlighting the complexity of NP regulation, because it is an exception to the morphological and chemical differentiation link.[68] Finally, exploratory growth strengthens the correlation between the life cycle of *Streptomyces* and ecological interactions.

1.2 Ecological Interactions and Natural Products

The microbial world is a universe unto itself. Earth is home to an estimated 10^{12} microbial species, more than the number of humans that have ever lived.[69,70] There are 100 million times more bacteria in the oceans (13×10^{28}) than there are stars in the known universe. *Clostridium botulinum* produces a toxin so potent that 400 grams would kill everyone on the planet and the fraction of microbial diversity that we have sampled is effectively zero.[71]

Historically, Actinobacteria were thought of as endemic soil bacteria, but are now recognized as cosmopolitan; they are found in virtually all ecosystems. "Everything is everywhere: but the environment selects" is a fundamental assumption that was generally promulgated by Martinus Wilhelm Beijerinck in the early twentieth century and specifically articulated in 1934 by Lourens G.M. Bass Becking.[72] The Baas Becking hypothesis is still being debated.[73] Recent work has shown that the structure of the environment does not affect the microbial diversity, but significantly affect the secondary metabolites produced.[74] This hypothesis has also been investigated at the gene level with the conclusion that gene pools are shaped by their broad ecological niche (such as sea water, fresh water, host, and airborne) and that certain antibiotic resistance genes deviate from this general trend by exhibiting a high degree of cross-habitat mobility.[75] Finally, strong evidence has

been found to support the hypothesis that every metabolite is everywhere, but the environment selects.[76] However, more work needs to be done to clarify the links between microbial taxonomic profiles, metabolism, and functional diversification, i.e., enzymes on a planetary scale.

The vast range of Actinobacteria applications are a reflection of their ecological niche.[77] Examples include bioindicators[78] and bioremediation of contaminated soil,[79–81] probiotics in aquaculture,[82] bioweathering of volcanic rock,[83] and plant growth-promoting,[84,85] biofertilizers,[86] biopesticides,[87] and biocontrol agents in association with plants [86,88,89]. In these examples, Actinobacteria are able to survive and thrive in a breadth of different environments by mediation with a number of enzymes and NPs. Naturally, studies have traditionally focused on their ability to produce useful NPs and enzymes because of their commercial benefits, while neglecting their ecological context, i.e., why are they produced. Therefore, by understanding the ecological role of NPs and enzymes, we might be better able to elicit their production and identify niches that are enriched for specific types of NPs and enzymes.

In particular, the thermophilic ecological niche has proven to be a rich source of Actinobacteria capable of producing valuable enzymes and potentially NPs.[90] These thermophilic Actinobacteria produce highly active and thermostable enzymes such as cellulases (optima: 50 ˚C, pH 6.5),[91] α-amylases (optima: 60 ˚C, pH 7),[92] proteases (optima: 70 ˚C, pH 11),[93] and xylanases (optima: 80 ˚C, pH 8)[94]. These enzymes are commonly utilized in many industries, e.g., pulp and paper, leather, textile, food and beverage, animal feed, and detergent. Amylases alone account for ~25% of the global enzyme market.[95] Thermophilic Actinobacteria also produce potentially valuable pharmaceutical NPs, e.g., topoisomerase inhibitors, isoaurostatin[96] and topostatin[97]. However, thermophilic Actinobacteria have a reduced genome size and are depleted in genes involved in secondary metabolism, making them an unlikely source for novel NPs.[90]

Finally, stable ecosystems likely house a diverse array of Actinobacteria. The diversity-stability relationships of soil microbial communities parallel those of plant communities, i.e., greater species diversity is required to support greater ecosystem stability.[98] Stable ecosystems contain diverse microbes, many of which are potential competitors with Actinobacteria. Actinobacteria are likely to have co-evolved in this environment to have broader metabolic capacities (NPs and enzymes) to compete with these competitors.[99] Therefore, stable ecosystems are a target for the discovery of diverse Actinobacteria with broad metabolic capabilities, such as the symbiotic ecosystem.

1.2.1 Symbiosis and Defensive Natural Products

The vast majority of multicellular animals, as well as many plants and fungi, engage in mutualistic relationships with microorganisms that are often essential for successful growth and reproduction of the host (Margulis & Fester, 1991; Moran, 2006). Actinobacteria have been found to have symbiotic associations with a number of eukaryotic hosts. For example, Actinobacteria are associated with animals ranging from insects to mammals[100] and plants[32,101,102]. These associations include aiding hosts with nutrient acquisition, especially in insects that consume detritus,[103] serving as defensive mutualists,[102] and virulent pathogens[101].

NP production has also been shown to be maintained in symbiont microbial communities, but is lost as alone-living symbionts, e.g., Burkholderiales, Pseudomonadales, Rhizobiales, and Xanthomonadales are localized to the hindgut of *Cephalotes* and have BGCs, while Opitutales are found alone in the midgut and do not have BGC.[104] The role of *Streptomyces* as symbionts is gaining more attention and is increasingly more important to identify the elicitors of novel antibiotic production.[35]

The most well-known example of Actinobacteria symbiosis is with plants, specifically in the rhizosphere, the area around the roots of plants, and the endosphere, the area inside the plant tissue. In the endosphere, Actinobacteria are the second largest representatives (20%), just behind Proteobacteria.[105] Antibiotic NPs have been isolated from medicinal plants with their idea that some of the healing properties may be due to the products of endophytic microbes. Munumbicins, wide-spectrum antibiotics, have been isolated from endophytic *Streptomyces* NRRL 3052 that was isolated within the medicinal plant snakevine (*Kennedia nigriscans*).[106] Kakadumycins produced by endophytic *Streptomyces* NRRL 30566 isolated from the golden tree (*Grevillea pteridifolia*) were shown to be more bioactive than their structural relative, echinomycin.[107] Coronamycins isolated from *Streptomyces* sp. MSU-2100, an endophyte of a *Monstera* sp. from the upper Amazon of Peru, showed antifungal activities.[108] Recently, 40% of endophytic Actinobacteria from coastal marsh plants in Jiangsu Province, China were shown to produce fibrinolytic enzymes, which are used in the treatment of thrombosis.[109] These included a streptokinase, which is used as a thrombolytic drug to treat blood clots and is on the WHO List of Essential Medicines. Interestingly, the homologous streptokinase enzyme of *Streptococcus* is a virulence factor that is produced to break blood clots and spread from the initial site of infection.[110,111] Endophytic *Streptomyces* have been shown to be an untapped source of NPs and enzymes, but their exact ecological functions have not yet been elucidated, which could provide more insight into why they are produced.

In the rhizosphere, Actinobacteria generally promote plant growth and suppress disease by producing a number of extracellular enzymes and a wide variety of

NPs.[112] In particular, *Streptomyces* have been shown to be able to reassemble and optimize the rhizosphere microbiome by significantly increasing bacterial diversity, while simultaneously decreasing fungal diversity.[113] Rhizospheric Actinobacteria enhance plant growth by a variety of mechanisms such as enzyme production (1-aminocyclopropane-1-carboxylate deaminase, inorganic phosphate solubilization, and biological nitrogen fixation) and NP production (phytohormones, siderophores, antifungals, and VOCs).[84,114] VOCs have attracted attention for their ecological role as suppressants of harmful plant microbes[115–117] and as attractants of arthropods to disperse their spores[65,118]. Recently, VOCs have been utilized as a tool to activate NP production.[119]

In addition to promoting plant growth, Actinobacteria are also a major driver of the global carbon cycle, e.g., decomposition of plant detritus, either as free-living in the soil or as symbionts of eukaryotic herbivores. Although carbohydrate-active enzymes (CAZymes) are widespread in the genus, key enzyme families of high cellulolytic capabilities are mainly enriched in two clades of host-associated *Streptomyces* strains. This suggests that symbiotic *Streptomyces* with high cellulose degrading activities are selected for, in the host-associated niche, because they aid the host in nutrient acquisition, making them a priority in the screening efforts for CAZymes.[120] For example, SirexAA-E is an insect-associated (sirex wood wasp) *Streptomyces* strain that produces highly active cellulolytic enzymes. Studies have provided information on the regulation and elicitation of these enzymes in SirexAA-E, that is, CebR, a cellobiose-responsive repressor,[121] together with newly discovered ManR are both able to sense mannose and mannooligosaccharides[122].

Insect-associated Actinobacteria also provide us with the most complete *in situ* context of the nature of their ecological interactions. The best studied example is the bacteria *Pseudonodocardia*, the leafcutter ant *Acromyrmex*, the fungus *Leucoagaricus*, and the fungal parasite *Escovopsis* as a coevolving quadripartite symbiosis (**Figure 2**). Briefly, the ant cultivates the fungus, the sole food source for its larvae.[123] The ant feeds leaf cuttings to the fungus and grooms the fungus to remove spores of the fungal parasite.[124] Concurrently, the fungus is sterilized with antifungal-parasite compounds produced by the bacteria kept on the ant's cuticle, which feed the bacteria.[125] In turn, the fungal parasite fights back by producing virulence factors that kill bacteria and adversely affect the behavior of the ants, ultimately killing the ant as well.[64] Although no molecular elicitors have been identified yet, a number of novel *Pseudonodocardia* NPs have been identified. Examples include pseudonocardones,[126] dentigerumycins,[127] and the atypical antifungal polyene selvamicin[128]. The *Pseudonodocardia* symbionts of fungus-growing ants now serve as a model system to understand the evolution of NPs and have provided a significant source of novel antibiotics.[129]

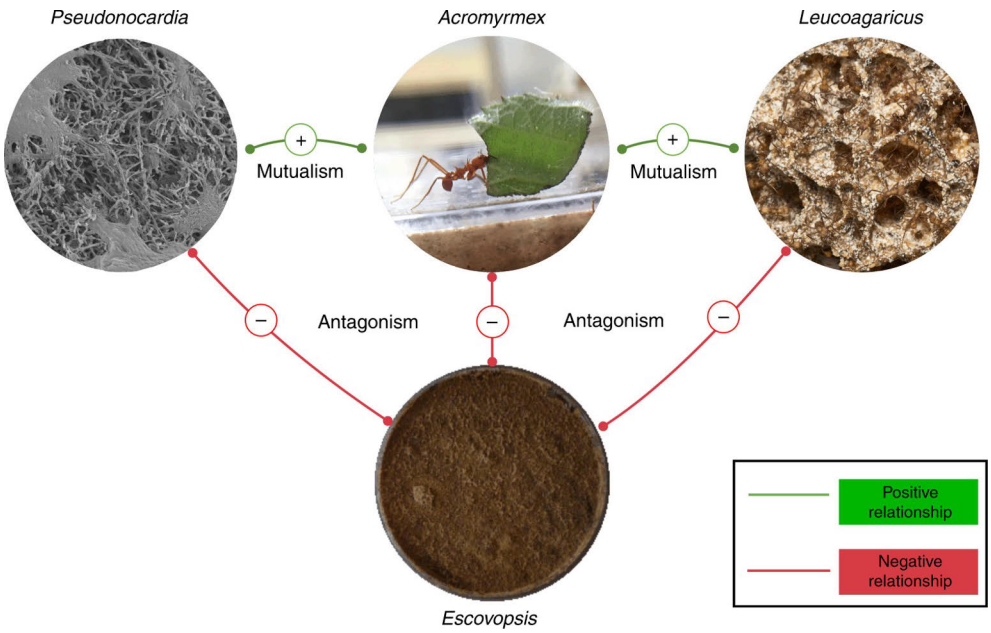


Figure 2. The *Pseudonodocardia*, the leafcutter ant *Acromyrmex*, the fungus *Leucoagaricus*, and the fungal parasite *Escovopsis* as a coevolving quadripartite symbiosis. Pairwise relationships are shown. Source: [64].

As pathogens, *Streptomyces*, especially *S. somaliensis* and *S. sudanensis*, can cause mycetomas, i.e., subcutaneous opportunistic infections, in humans.[38] Little is known about the molecular mechanism of pathogenicity, although recent work has shown that 2,5-diketopiperazines appear to be responsible.[130] Respiratory infections, for example, farmer's lung disease that has recently become widely spread in California due to the prevailing megadrought, have been associated with the inhalation of actinomycete spores.[131] Recently, a new strain of pathogenic *Streptomyces* was isolated and studied from human lungs.[39] However, most *Streptomyces* are actually beneficial because they are present in healthy skin,[132] the gut,[133] the respiratory tract,[134] and even the uterus[135]. Additionally, the human gut has a relatively low abundance of *Streptomyces* compared to other animals[133] and this may correlate with the high incidence of inflammatory bowel disease (IBD) and colorectal cancer (CRC) in humans[136].

The marine environment requires that marine metabolites be physicochemically adapted to their surroundings and have evolved to encompass a chemical space distinct from their terrestrial relatives. As a result of this diversity, marine NPs are increasingly being used in clinical trials of various diseases. These NPs are often derived from the dense microbiome of marine sponges and tunicates.[137] In sponges from the Great Barrier Reef, Actinobacteria are one of the dominant bacterial taxa,

especially members of the *Acidimicrobiaceae* family, representing a large sponge-specific lineage.[138,139] Sponge metagenomic studies have revealed a ubiquitous, evolutionarily distinct, and highly sponge-specific group of NPs affiliated with the candidate phylum Poribacteria,[140,141] providing evidence of a niche-specific NPs. It was shown that the sponge host can be categorized as high-microbial-abundance (HMA) and low-microbial-abundance (LMA) regions, which are composed of dense tissue and a more porous body matrix, respectively.[142] Actinobacteria tend to be associated with HMA sponges, further narrowing the search space for these NP producers, however, isolation still remains a challenge.[143] Finally, the core microbiome of *Porifera* sponges has also been shown to be distinguished by amensalism and commensalism rather than competition and mutualism,[144] therefore, amensalistic interacting microbes could be a promising source of NPs and enzymes.

1.2.2 Antagonism and Offensive Natural Products

Amensalism can generally be defined as a unilateral interaction where one organism harms another organism without any cost or benefit and is one of the least described ecological interactions.[145,146] For example, amensalism was the interaction observed during the discovery of penicillin, where the Ascomycete fungus *Penicillium* kills bacteria by producing penicillin[14] and is well documented in *Streptomyces*. However, viewing all *Streptomyces* antibiotic-mediated interactions as amensalistic is likely to be too general; a more nuanced interpretation is more appropriate[147] as seen with a few reports of this interaction as being predatory[148–151]. Moreover, the difference between ammensals, e.g., *Streptomyces*, and predators, e.g., *Myxococcus*, is minute and only differs in the antibiotic-producers ability to consume what they kill.[152] Yet, the perspective that *Streptomyces* can be predators has not been elucidated on a molecular level and is not widely accepted. Still, by recognizing *Streptomyces* predatory abilities, a link can be made to established predatory bacterial research, thus providing a more holistic understanding of microbial ecosystems in general.

The classical definitions of ecological relations are becoming blurred as new methodologies provide more evidence for exceptions, reversions, and transitions, thus creating a continuum spectrum of interactions. The interactions between species are more complex and dynamic than previously thought and whether an association is seen as positive or negative often depends on the environmental circumstances.[153] For example, the definitions of endophytes have recently come under scrutiny.[105] Endophytes were originally only considered to be fungi that live within plants; however, we now know that plant tissues can be colonized by bacteria as well.[154] Conceptually, plant pathogens are not considered endophytes because of their virulence, but what about when they lose their virulence?[155] What about when endophytes become harmful under specific conditions?[156] Our current understanding

of endophytes is built on a small set of controlled and optimized experimental conditions, that seldom replicate natural conditions thus biasing our understanding of these interactions.[105]

Bacterial predators drive diversity and thus contribute to overall stability of the community. A recent study has shown that unilateral interactions, e.g., amensalism or commensalism, are more stabilizing to the community than symmetrical interactions, e.g., mutualism or competition, but asymmetric antagonisms, e.g., predation or parasitism, are the most stabilizing.[157] Predatory bacteria have also been shown to be a major driver of diversity in the bacterial community[158–160] and commonly produce secondary metabolites[161]. Specifically, predators may mediate coexistence between superior and inferior competitors[162] and exert top-down control of microbial food webs resulting in higher productivity of microbial communities[163].

Bacterial predation is understudied although it is one of the earliest forms of predation[164] and may be the origin of eukaryotic cells[165,166] as well as multicellularity[51]. If *Streptomyces* are predators, this could explain, in part, why they are found in stable microbial communities in a wide range of environments.

Most studies of predatory prokaryotes have focused on the Gram-negative deltaproteobacteria: *Myxobacteria*, *Bdellovibrio* and like organisms (BALOs), and recently discovered *Bradymonas*. Predators have recently been proposed to be categorized into a three-group framework: (i) obligate, (ii) facultative, and (iii) opportunistic predators. This framework adds more granularity to the spectrum of predatory groups. Specifically, it reflects the level of auxotrophies of each group, with obligate predators having more, opportunistic predators having less, and facultative predators in-between. In this new spectrum of predation, *Bdellovibrio* are obligate predators, *Myxobacteria* are opportunistic predators, and *Bradymonas* are facultative predators.[167]

Bradymonas are novel bacterial predators with a versatile survival strategy and have a genome size of ~5 Mb and a G + C content of ~60%.[168] Bradymonas is likely to kill using both physical contact and by secreting outer membrane vesicles (OMVs) and NPs as supported by transciptomics data showing up-regulation of T4b pilins and type III secretion system, and OMVs during predation. However, no antibacterial activity was observed from the supernatant. *Bradymonas* is also able to store energy in the form of polyphosphates to survive during periods of starvation.

Bdellovibrio are aquatic, highly motile, and extremely small obligate predators with a circular chromosome of ~3.7 Mb and a G + C content of ~45% containing a wide range of lytic enzymes with the potential to be used as antimicrobial agents. However, *Bdellovibrio* lack the ability to produce antibiotics.[169] Consistent with being an obligate symbiont, the genomes feature elevated sequence evolution, gene loss, and a shift towards reduction in G + C content.[170,171] They are able to prey on Gram-negative cells by entering the periplasmic space (physical contact) of the prey

and then reseal the entry point to prevent leakage of their food and gain protection from other microbes.[172] Furthermore, *Bdellovibrio* can also store energy as polyphosphates, which can play an important role in maintaining and regulating a high activity of glycolysis.[173]

Bdellovibrio are of special interest for their application as living antibiotics. The Pathogen Predators program funded by the Defense Advanced Research Projects Agency (DARPA) is based on previous *in vitro* studies that have shown that *Bdellovibrio bacteriovorus* prey upon more than 100 different human pathogens, including several that are multi-drug resistant. These results suggest that it is possible to develop a predator-based therapeutic with efficacy against a broad spectrum of Gram-negative pathogens, including those resistant to antibiotics.[174] Importantly, *Bdellovibrio* produce a number of prey-destructive enzymes with potential genetic redundancies suggesting resistance is unlikely to occur[175,176] and are thus a strong candidate for the fight against AMR. Specifically, *Bdellovibrio* works alongside host immune cells to treat shigella infection in zebrafish larvae,[177] are effective in reducing the burden of *Klebsiella pneumoniae* in rat lungs,[178] and are a part of a healthy gut microbiota and could be utilized as a probiotic[179]. However, the industrial production of *Bdellovibrio* and the viability of the preparations will be a challenge.[180]

Myxobacteria are similar to Actinobacteria. Both have large genomes and a high G + C content, both can morphologically differentiate and form spores, and both can produce a broad range of structurally diverse NPs.[181] Although they are similar in many ways, *Myxobacteria* are distinguished by their vegetative cells being rod-shaped (**Figure 3**), rather than mycelial, and have two motility systems (**Figure 3**), while Actinobacteria are immotile. *Myxobacteria* can be divided into two ecological groups based on their nutritional requirements and phylogeny, i.e., predators and cellulose decomposers,[182,183] making them opportunistic predators. In contrast, while Actinobacteria are well-known cellulose decomposers, their predatory abilities are not well described.

The model species *Myxococcus xanthus* has been shown to have complex feeding behaviors such as rippling, where individual cells organize their movement and glide together in a swarm (or wolfpack) in order to make direct contact with large macromolecules or prey (**Figure 3**).[184] This complex behavior is similar to the newly observed exploration behavior in *Streptomyces*, where explorer cells quickly traverse solid surfaces and release VOC to inhibit competitors and coax other *Streptomyces* to start exploring as well,[67] although this behavior has not been studied in the context of predation.

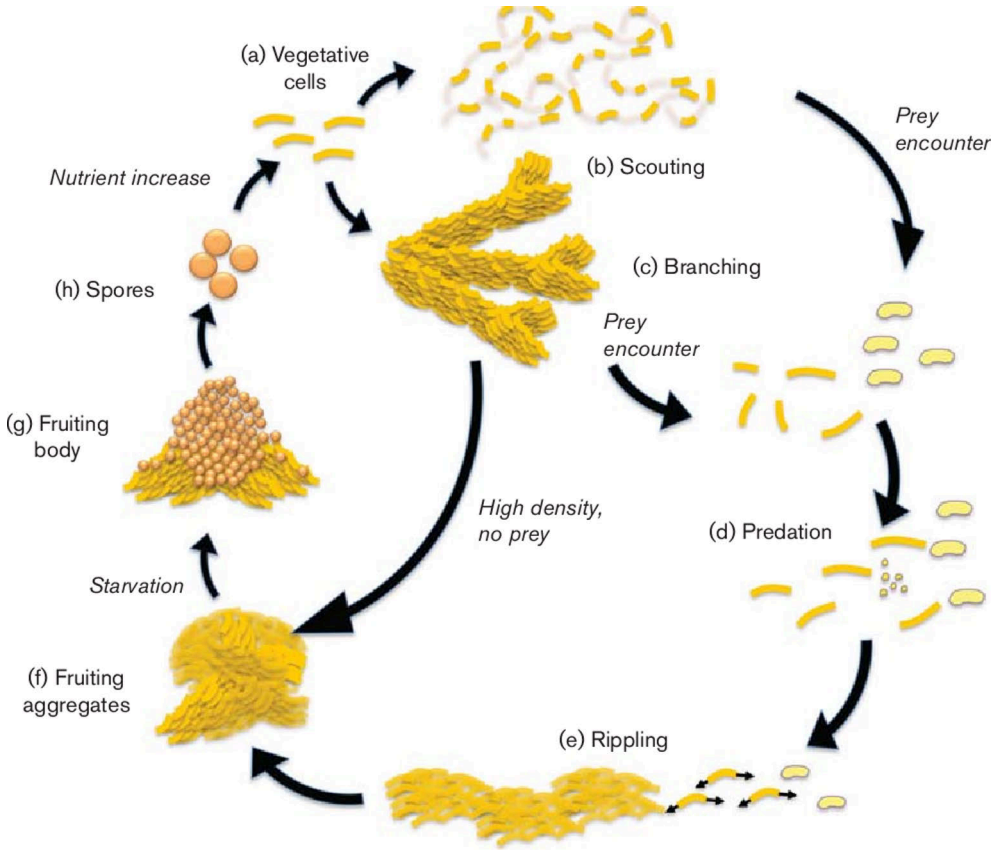


Figure 3. The *Myxococcus* predatory *life cycle.* The different life stages are referred to by letters. Source: [410].

Although predation in *Myxobacteria* is well established, little is known about its molecular mechanism.[185] However, it is known that *M. xanthus* can kill by secreting extracellular bacteriolytic enzymes and NPs to inhibit and weaken its prey.[186,187] *Streptomyces* has these same capabilities. The role of NPs in predation is understudied, with only two NPs, myxovirescin and myxoprincomide, having demonstrated a role in predation. Both NPs are shown to play a prey-dependent role.[188] Furthermore, while myxovirescin is known to target signal peptidase II (LspA),[189] the mode of action of myxoprincomide is still unknown[190]. Most (80%) of the known myxobacterial NPs have unknown modes of action,[181] making it more difficult to understand their role in predation. Additionally, myxovirescin production is transcriptionally controlled by alternative sigma factor σ54 in conjunction with specific enhancer binding proteins, which have been shown to regulate other NPs during development, thus linking NP production, development, and predation,[191] which is reminiscent of DasR regulation in *Streptomyces*.[49]

Similar to NPs, bacteriolytic proteins, such as amidases and glucosaminidases, are well documented, but their role in predation is not well understood.[187,188] In fact, the genome of *M. xanthus* shows an abundance of hydrolytic enzymes potentially capable of lysing bacterial cell walls, as well as a multitude of proteases and peptidases presumed to play a role in predatory feeding by degrading proteins from its prey.[188,192] *Streptomyces* have been shown to produce these same enzymes,[193–196] although again not in the context of predation. Recently, an outer membrane β-1,6-glucanase, GluM, of *Corallococcus* was shown to be required to lyse the cell wall of certain fungi,[197] providing evidence that a bacteriolytic protein plays a role in predation. However, the homologous protein in *M. xanthus*, Oar, in addition to being a glucanase, is also directly involved in secreting PopC, a signal-peptideless protease, which is essential for completion of the cell differentiation program.[198] This is reminiscent of the cholesterol oxidase, PimE, of *S. natalensis*, which was shown to have the dual role of being an enzyme and a putative signaling molecule for the production of the pimaricin antifungal compound.[199]

Physical interactions have been shown to be mediated by flagellum, nanotubes, and membrane-membrane contact to adhere, sense the environment, and exchange cytoplasmic constituents.[200] Direct physical contact also leads to lysis and may be essential for predation in *M. xanthus*.[201,202] This is supported by the genome of *M. xanthus*, which has half of its putative proteases predicted to be periplasmic or secreted to the cell surface and by the use of OMVs.[188,192] Physical contact mediated interactions of *Streptomyces* have been shown from the perspective of co-cultivations for the production of NPs, e.g., *Microcystis*,[149] *Aspergillus*,[203] *Bacillus*/*Staphylococcus*,[204] and *Tsukamurella* (and other bacteria containing mycolic acid),[205] however, direct observation of the nature of physical contact is challenging because it requires the use of electron microscopy[200].

Intimate physical-contact mediated predation is also supported from the perspective of public goods. Microbial public goods are analogous to economic public goods, e.g., public parks that are financed by tax payers but provide a benefit to all, in that they are compounds (goods) produced and secreted into the public milieu by some and enjoyed by all, e.g., enzymes, antibiotics, and leaky metabolites.[206] Physical contact would create spatial structure to limit diffusion and thereby prevent cheaters (scavengers) from benefiting from predators' creation of public goods, i.e., enzymes, antibiotics, and nutrients.[206,207] Evidence for this type of predatory mechanism has been observed in *M. xanthus*, where reduced half-life hydrolases are packed into OMVs to reduce the rate and area of diffusion, which are overproduced during starvation.[208] *Streptomyces* are also capable of creating membrane vesicles (MVs), which have been linked to antibiotic production.[209] In one case, the antifungal linearmycin was packaged into MVs and the abolishment of linearmycin production resulted in a reduction of MV production.[210] In another case,

S. albus S4, producer of the antifungals candicidin and antimycin, was shown to selectively package candicidin, but not antimycin, and the abolishment of one or both of the antifungals did not impact the biogenesis of MVs.[211]

M. xanthus uses polyphosphates for motility, development, and predation. Specifically, the polyphosphate metabolism genes *ppk1* and *pap* were shown to cause deficiencies in all of the above.[202] Inorganic polyphosphates have also been shown to be involved in virulence in bacteria[212–214] and development, sporulation, and predation in the amoeba *Dictyostelium discoideum*[215]. In Actinobacteria, polyphosphate research is focused on *Microlunatus phosphovorus* and its ability to recycle wasted phosphorus.[216] In *Streptomyces*, phosphorus deprivation is known to trigger antibiotic production during late growth. It is proposed that excess ATP is stored as polyphosphate by polyphosphate kinase (Ppk), which would later be degraded to inorganic phosphate and used as a source of energy during development and delay the expression of the Pho regulon, thus block antibiotic production.[217] Later it was shown that under inorganic phosphate limiting conditions, a *ppk* mutant could not regenerate ATP using polyphosphates and instead triggered the degradation of triacylglycerol to regenerate ATP, resulting in a massive increase in acetyl-CoA, while concomitantly reducing Krebs cycle activity, thus shifting metabolic flux towards the overproduction of actinorhodin.[218] The exact mechanisms connecting polyphosphates to predation and antibiotic production have not yet been elucidated, but current research provides a promising start.

Predation does not require motility. Fungal predators are able to attack nematodes and other microorganisms using a remarkable array of trapping devices to attract, capture, kill, and consume (**Figure 4**).[219] There are 5 types of traps: adhesive network, adhesive knob, non-constricting rings, adhesive column, and constricting rings.[220] Nematophagous fungi are opportunistic predators and can produce a cornucopia of NPs and enzymes that are used for predation.[221,222] Blumenol A attracts nematodes. Phomalactone, talathermophilins, and aurovertins, as well as extracellular enzymes, are nematicidal. Athrosporols and paganins play a role in the development and formation of trapping organs. Their capacity to feed on nematodes has been compared to carnivorous higher plants that complement photosynthetic energy with protein from captured insects[223,224] and is akin to ambush predation. Recently, G-protein signaling and the cAMP-PKA pathway have been shown to be necessary for prey detection in *Arthrobotrys oligospora*, however, more research is needed to fully uncover the molecular mechanisms behind their predatory abilities.[225,226] *Streptomyces* and fungi share the same mycelial form and both produce several NPs and enzymes, but *Streptomyces* has yet to be described to have the ability to trap other microbes.

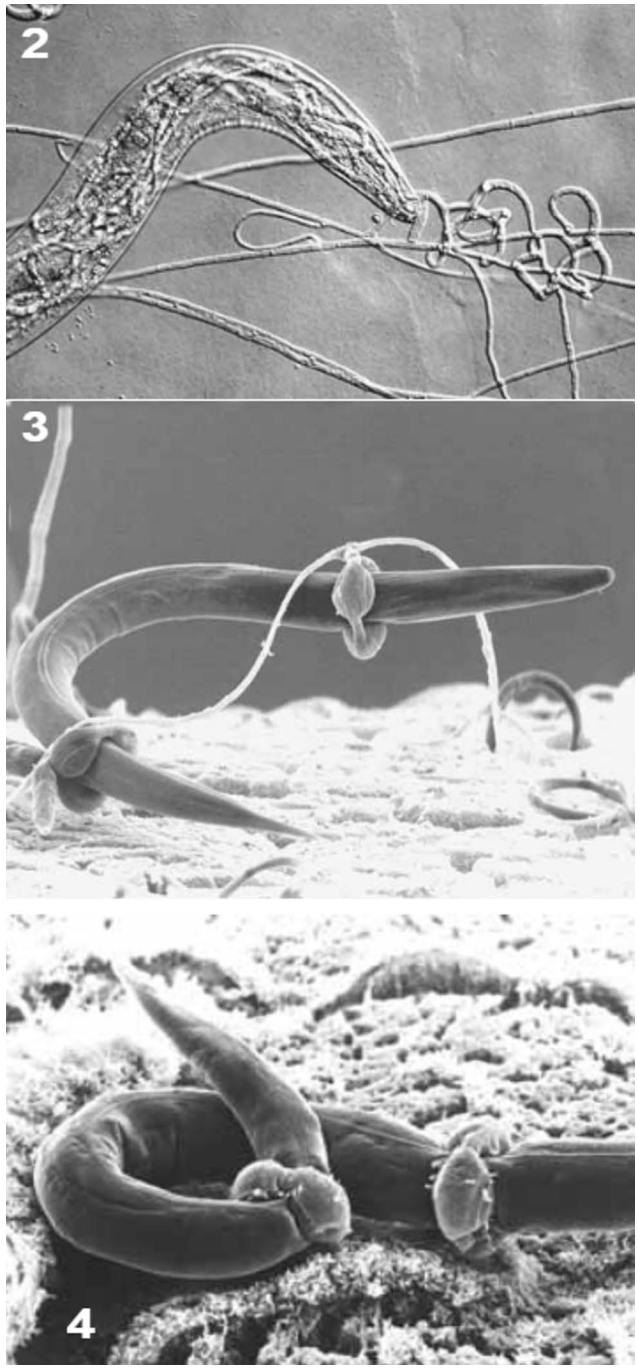


Figure 4. Nematophagous fungi trapping devices. A nematode caught in a sticky net trapping device, Top. A constricting ring trapping device before and after crushing the body of the nematode, Middle and Bottom, respectively. Source: [219].

Competition interactions have also been studied in *Streptomyces*, although this type of interaction is the least influential on community stability. A competition analysis revealed strong regulation of *Streptomyces* antibiotic production in response to interactions between social and resource environments.[227] However, coculture interactions only induced secondary metabolite production ~1/3 of the time and surprisingly suppressed production ~1/2 of the time.[227–230] *Streptomyces* tend to have a fight response when co-cultured with susceptible competitors and a flight response when co-cultured with inhibitory competitors.[227] This may be due to one-to-one encounters being typically costly for the provoker, since they can receive a strong reciprocal attack. However, provocation can be strongly beneficial in communities with more than two toxin-producers, since an aggression-provoking toxin producer can manipulate its competitors to kill each other, similar to a 'divide-and-conquer' strategy.[231]

1.3 Genomic Features of *Streptomyces*

Actinobacteria is one of the largest phyla in the domain *Bacteria* and is comprised of 5 subclasses and 14 suborders.[232] The general features of Actinobacteria genomes are a high guanine-plus-cytosine (G + C) content, a relatively large size, and a linear chromosome that is a rare characteristic in bacteria.[233] As I discussed previously, large genome size is correlated with multicellularity, morphological differentiation, and chemical differentiation. However, the content of G + C can range from less than 42%, i.e., *Gardnerella vaginalis* (the causative agent of vaginosis), to more than 70%, e.g., *Streptomyces*; the size of the genome can range from less than 1 Mb, i.e., the obligate pathogen *Tropheryma whipplei* (the causative agent of Whipple's disease), to more than 11 Mb, i.e., *Streptomyces bingchenggensis*; the chromosome can also be circular, e.g., *Mycobacterium tuberculosis* (the causative agent of tuberculosis) and *Corynebacterium diphtheriae* (the causative agent of diphtheria).

The linearity of the *Streptomyces* chromosome contributes to the diversity and stability of BGCs. In general, the chromosome of *Streptomyces* is organized into a conserved central region (primary metabolism) flanked by variable arms (secondary metabolism) with telomeric ends and can be complemented by linear and circular plasmids (**Figure 5**). This organization may have originated in the linearization process, coinciding with the emergence of the *Streptomyces* genus.[234] Briefly, the acquisition of telomeres and chromosomal arms would have resulted from a single integration event of a linear replicon within the ancestral circular chromosome. As a result, the chromosomal arms would have included contingency genes (nonessential genes), since this early linearization event.[235]

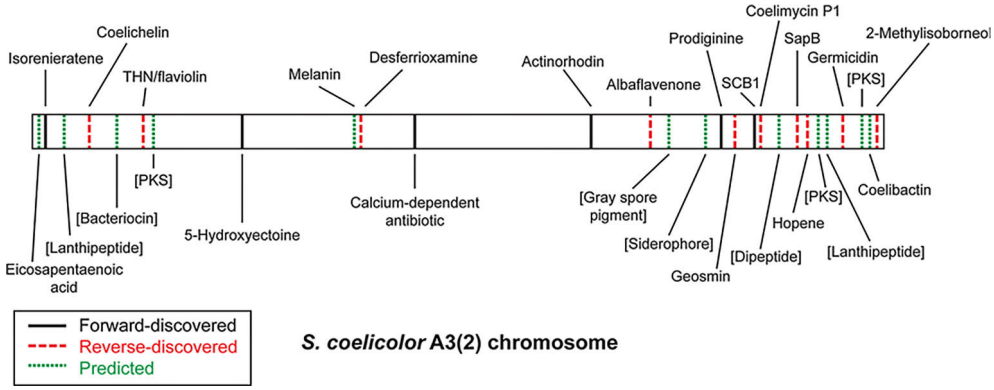


Figure 5. Schematic representation of the *Streptomyces coelicolor* A3(2) chromosome. All known and predicted BGCs are shown. Source: [411]

In *Streptomyces*, these contingency genes are often involved in secondary metabolism. NPs produced by secondary metabolism are typically biosynthesized by a series of biochemical reactions. These reactions are catalyzed by enzymes encoded in the same genomic locus forming a biosynthetic gene cluster (BGC). These BGCs are believed to aid *Streptomyces* in adapting to different environments. Similarly, in parasitic and pathogenic bacteria, these contingency genes are often associated within the subtelomeric regions and aid in their adaptation to different host environments because they are highly mutable.[236,237]

The subtelomeric chromosomal arms are enriched with diverse genes and BGC and evolve much faster than the central region. This diversity comes from instability and results in insertions, deletions, duplications, rearrangements, and horizontal gene transfer (HGT) from other chromosomes and plasmids that result in novel gene combinations.[235,238] Moreover, linearity has been suggested to facilitate the exchange of genetic information between DNA molecules via single recombination events.[239] For example, the exchange of ends between the linear plasmid, pPZG101, and the linear chromosome of *Streptomyces rimosus* led to the formation of a linear prime plasmid containing one chromosomal end and the chromosomal oxytetracycline BGC.[240] HGT is also thought to allow for rapid adaptation, which in turn may lead to diversification and speciation.[241]

Recently, gene expression was correlated with the folding of the linear chromosome of *Streptomyces ambofaciens*, which is the industrial producer of the antiparasitic macrolide spiramycin. Specifically, metabolic differentiation is accompanied by the chromosome changing its architecture from an ‘open’ to a ‘closed’ conformation, in which BGCs form new boundaries and provides insight into BGC regulation in relation to the overall dynamics of the genome.[242]

The location of BGCs on the chromosome correlates with the stability of their inheritance. Vertical inheritance of BGCs seems to be stabilized as they migrate

towards the core of the genome through selective pressure.[232,243] These 'core' BGCs tend to encode ectoines, siderophores, and terpenes, which are generally useful and likely essential for survival as seen in *Streptomyces*, *Amycolatopsis*, and *Salinospora*.[244,245] In *Streptomyces albus* J1074, a common host for heterologous expression, the expression of the aranciamycin BGC varied up to eight-fold depending on its position in the chromosome, with the highest levels being toward the center of the chromosomal shoulder, illustrating that chromosomal location also has an effect on expression levels.[246] In contrast, species-specific BGCs tend to be located further away in the chromosomal arms.

However, it is unclear how often HGT occurs in Actinobacteria; the current assumption is that HGT is rampant, but a recent study suggested that only one gene is stably acquired through lateral gene transfer every 100,000 years.[247] Yet, at large scale, the loss of a BGC by the acquisition of a similar BGC could mask a large portion of historical HGT events and obfuscate analysis of their evolution.[29,248] Nevertheless, HGT does play a role in BGC diversity as seen in the transfer of an actinomycin BGC from *Streptomyces padanus* to *Rhodococcus fascians* in a competitive coculture, resulting in rhodostreptomycins that are structurally unrelated to actinomycins.[249]

Recently, the pan-genome, i.e., the entire set of genes from all strains under study, of 121 *Streptomyces* strains revealed 273,372 clusters. The clusters were further classified according to their conservation level (from highest to lowest) among the analyzed genomes and consists of 633 core genes, 1,080 soft-core genes, 6,040 shell genes, and 137,709 cloud genes. Additionally, there were 81,568 unique clusters, i.e., they were unique to only one genome.[250] The small number of core genes is consistent with the 482 genes in *Mycoplasma genitalium*, which is a model of minimal genome research.[251] This underscores that the vast majority of genes can be considered part of secondary metabolism. Moreover, the pan-genome is also considered to be open, i.e., the full pan-genome is impossible to predict since many new genes will be added with each additional genome,[196,250,252] emphasizing the heterogeneity of the *Streptomyces* genome. This high level of heterogeneity is assumed to reflect their high level of biodiversity,[253] which can be correlated with NP diversity. Furthermore, no core group of intergenic regions was found suggesting that regulation is highly variable between strains.[250]

In terms of CAZymes and NPs, another pangenomic analysis showed that the number of CAZymes can range from 121 to 381 in a given species, with 2,096 proteins as core.[196] These large numbers of CAZymes are unsurprising, given that Actinobacteria is a major contributor to the global carbon cycle and highlights the large and diverse carbohydrate catabolizing capacity of *Streptomyces*. NPs capacities per strain range from 18 to 52 BGCs with a median of 31.[196] Although individual strains are capable of producing tens of NPs, historically only two to four were

observed under laboratory conditions, highlighting our lack of understanding of how to elicit their hidden potential.

The most common types of BGCs identified were nonribosomal peptide synthases (NRPSs), followed by terpenes and type 1 polyketide synthetases (PKSs).[196] NRPSs and PKSs need to be activated by incorporation of a phosphopantetheinyl side chain of coenzyme A into the peptidyl carrier domain or acyl carrier domain. This is achieved by phosphopantetheinyl transferases (PPTases).[254] Each strain analyzed contained 1-12 PPTases with a median of 4.[196] Strains, such as *S. collinus* Tu 365, that contain a diverse set of PPTases are likely to be capable of activating a wide variety of NRPSs and PKSs, making them promising hosts for heterologous expression of these types of BGCs. It is also worth noting that these analyses are limited by the majority of regulatory proteins of BGCs being poorly characterized and that the conservation of enzyme active sites and other residues involved in substrate recognition are more important than overall homology, which is the current basis of these types of analyses.[196]

1.4 Regulation of *Streptomyces* Natural Product Production

The regulation of BGCs in Actinobacteria is highly complex. Regulation can occur on a global, cluster-specific, and translational level. Moreover, regulation of BGCs and enzymes in Actinobacteria is highly responsive to external signals. Approximately 12% of the total chromosome is dedicated to regulatory genes with, most acting as activators or repressors and sometimes both.[255] Regulators can influence production directly, indirectly, via autoregulatory hormone-like signaling molecules, and cross-regulation.[46,256,257] The cognate products of BGCs can also be autoregulators of their own biosynthesis and for disparate BGCs.[258]

As an example, the complexity of environmental signal influence on the regulation of NPs can be seen through the signal GlcNAc and its influence on the regulation of the undecylprodigiosin BGC in *S. coelicolor*. GlcNAc is a monomer of chitin (from e.g., fungal cell walls) and a constituent of bacterial peptidoglycan and is common in many environments. It is also a preferred carbon and nitrogen source of Actinobacteria, as well as a signaling molecule in the DasR-mediated nutrient sensing system. GlcNAc is internalized as *N*-acetylglucosamine-6-phosphate (GlcNAc-6P) via the phosphoenolpyruvate-dependent sugar phosphotransferase system and catabolized by GlcNAc-6P deacetylase NagA resulting in glucosamine-6-phosphate (GlcN-6P), which can act as a ligand for DasR or be further catabolized by GlcN-6P deaminase/isomerase NagB.[48] GlcN-6P activates morphological development and NP production under famine conditions, by inhibiting DasR binding to DNA, and has the opposite effect under feast conditions. The DasR regulon includes the primary

metabolic *nag* genes and the secondary metabolic *red* genes.[259] The prodiginine BGC is directly controlled by the pathway-specific regulator *redZ*, which is directly repressed by DasR. When *redZ* is derepressed, the BGC is activated and produces undecylprodigiosin. RedZ lacks the conserved Asp and Lys residues of the phosphorylation pocket and therefore uses a phosphorylation-independent mechanism of regulation, i.e., end-product-mediated regulation, where the cognate product of the cluster is able to regulate its own expression.[260] In summary, GlcNAc is an environmental signal that can activate or repress the production of prodigiosins, via the pleiotropic regulator DasR and, subsequently, the pathway-specific regulator RedZ, which requires ecological context to determine its function. Furthermore, RedZ is also regulated by end-product-mediated regulation.

The *bldA* gene is an example of a global regulator of morphological differentiation and NP production that can exert its regulatory control at the translational level. *BldA* encodes a tRNA that is capable of translating the rare TTA codon into the amino acid leucine.[261,262] Although the TTA codon is rare in general, due to the high overall G + C content in *Streptomyces*, the codon seems to be enriched in BGC genes and especially regulatory genes of BGCs.[263] The AdpA-BldA feedback loop is able to control morphological differentiation and NP production through the auto-regulatory mechanisms of AdpA as shown in *S. griseus* (**Figure 6**).[264] In a *S. coelicolor bldA* mutant, replacing the TTA codon of *redZ*, *actII-ORF4*, *mmyB* and *mmfL* with one of the other five leucine codons restored the production of undecylprodigiosin, actinorhodin, and methylenomycin, respectively.[265]

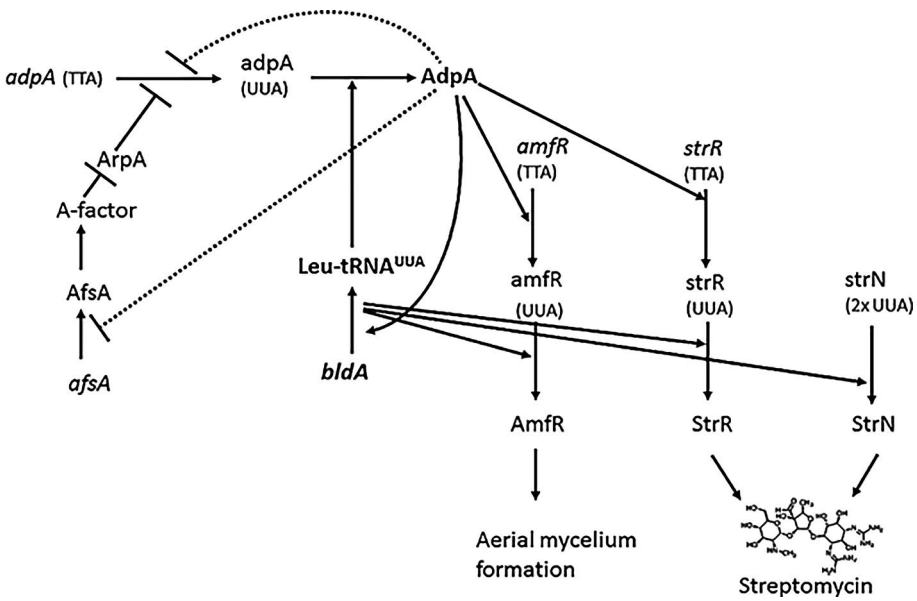


Figure 6. The AdpA-BldA feedback loop in *S. griseus*. The auto-regulatory mechanism of AdpA is shown. Source: [412].

Enzymatic proteins have also been shown to act as a signal. *Streptomyces natalensis*, a producer of the antifungal polyene macrolide pimaricin, has a cholesterol oxidase gene, *pimE*, within the pimaricin BGC. Adding any cholesterol oxidase, not only from *S. natalensis*, to the culture medium triggered pimaricin production in the *pimE* knockout mutant.[199] Furthermore, this phenomenon was also observed in *Streptomyces gilvosporeus*, where an additional copy of *pimE* was inserted into the genome, resulting in a 153% improvement in production.[266] However, it remains to be elucidated how cholesterol oxidase is capable of signaling or regulating pimaricin still needs to be elucidated.

The global regulator of CAZyme production, CebR has been shown to be the master regulator of cellulose catabolism and affects the morphological development in *Streptomyces griseus*.[267] CebR, in conjunction with TxtR, is the gatekeeper of *Streptomyces scabies* pathogenicity; this pathogenicity can also be thought of as herbivorous predation. Cellobiose and cellotriose are signals that inhibit the DNA-binding of CebR, leading to the production of thaxtomin A, a phytotoxin that inhibits cellulose biosynthesis and provides hypervirulence.[268]

1.5 Genome Mining and Types of Clusters

Streptomyces coelicolor A3(2) was studied for some 40 years and was only known to produce four SMs: actinorhodin,[269] undecylprodigiosin,[270] calcium-dependent antibiotic,[271] and methylenomycin[272]. The principal pioneer in the area of genetic techniques for the discovery of NP biosynthetic enzymes was Sir David Hopwood, who used actinorhodin as his test system.[273] However, with the advent of whole genome sequencing and advanced computational tools (bioinformatics), the genome of *S. coelicolor* A3(2) was sequenced and assembled in 2002 and revealed 18 additional BGCs.[27] This exciting revelation reignited interest in Actinobacteria antibiotic discovery and marks the beginning of the Genomic Age of antibiotic discovery.[274]

BGCs consist of genes for regulation, modification, transport (often resistance), and core biosynthesis of NPs. Many of these genes share homology and can be used to predict the presence of BGCs in genomes using bioinformatic tools. Genome mining for BGCs has become an integral part of NP discovery. Through genome mining, BGCs can be compared, allowing the identification of putatively novel NPs.

The explosion of genomic data can be seen through the various databases that attempt to organize it all. There are currently 4,390 genomes and 41,894 assemblies in the National Center for Biotechnology Information (NCBI) database for Actinobacteria.[275] Recent estimates show that the average genome size of *Streptomyces* is 8.5 Mb and 33 BGCs per genome with a linear relationship between

them.[276] The total number of predicted BGCs can be estimated from the antibiotics and secondary metabolite analysis shell (antiSMASH) database with 147,517 putative BGCs, the most common being traditional (multi)modular NRPSs with 37,320 BGCs, from 25,802 unique species/strains; Actinobacteria are represented by 3,215 species/strains and *Streptomyces* by 544 species/strains.[277] However, only a small fraction have been characterized as seen through the Minimum Information about a Biosynthetic Gene cluster (MIBiG) repository, which contains 2,502 clusters and the largest phylum represented is Actinobacteria with 1,042 entries.[278]

With this massive amount of data, new breakthrough computational methods are needed to mine these genomes and turn data into knowledge. One major problem is linking BGCs to NPs. This can be done either through a forward approach, i.e., starting from a genome and predicting the cognate NP, or a reverse approach, i.e., starting from a NP and predicting the BGC it came from. The forward approach has become more popular, with many tools having been developed, such as Prediction informatics for secondary metabolism (PRISM),[279] secondary metabolite unknown region finder (SMURF),[280] and antiSMASH (**Figure 7**)[281,282]. Only a limited number of reverse approaches are available such as, generalized retro-biosynthetic assembly prediction engine (GRAPE) and global alignment for natural products chemoinformatics (GARLIC).[283,284] It should be noted that all of these approaches are homology-based and are limited to known classes of NPs, e.g., polyketides and nonribosomal peptides.

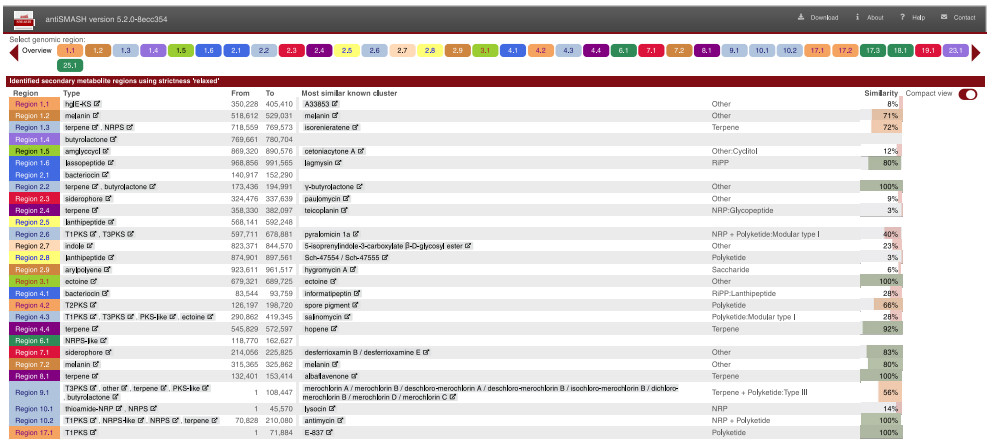


Figure 7. Genome mining of *S. longwoodensis* by antiSMASH. The overview page shows the predicted BGCs, along with their location, type, and similarity to the most similar known cluster. Retrieved from the antiSMASH database [277] on 3.3.2023.

Many of these genome mining tools are starting to form their own ecosystems by adding other types of omics data or adding more comprehensive functionalities.

For example, PRISM is a genome miner with a focus on predicting chemical structures and is closely connected to the genome-to-natural products (GNP) platform,[285] which incorporates MS/MS data, and to GRAPE/GARLIC, which compares the retro-biosynthetic predictions to the forward ones. Another example is the biosynthetic gene similarity clustering and prospecting engine (BiG-SCAPE), which incorporates antiSMASH and MIBiG to facilitate a global analysis of the relationships between BGCs across large numbers of genomes by grouping them into gene cluster families (GCFs) and aid in dereplication, along with the core analysis of synthetic orthologs to prioritize natural product biosynthetic gene clusters (CORASON), which guides in the exploration of gene cluster diversity linked to enzyme phylogenies. Recently, BiG-SLiCE[286] and the new clust-o-matic algorithm were used to estimate that only 3% of NPs in bacteria have been characterized, with *Streptomyces* holding the most potential.[287] GCF families can also be easily explored online through the biosynthetic gene cluster families database (BiG-FAM) database, which contains 1,225,071 BGCs forming 29,955 gene cluster families.[288]

Metagenomics has given insight into the unseen world of microbial dark matter, i.e., unculturable microbes. This microbial dark matter is likely to contain novel chemistry and biosynthetic machinery.[289] Recently, two phylotypes of the candidate genus *Entotheonella*, symbionts of the marine sponge *Theonella swinheoi*, were described to have ~9 Mb genomes that contain multiple distinct, BGCs and they are proposed to belong to the candidate phylum *Tectomicrobia*. Another recent report describes MetaBGC, that is, a modular probabilistic strategy to uncover the metamycin meta-BGCs from human microbiomes. Metamycins exert strong inhibitory activities against members of the human microbiome that occupy the same niche as their producer, implying an ecological role in microbial interactions.[290] Despite the numerous successes of genome mining, there is still a lack of a widely applicable pipeline that can match that of the Golden Age of antibiotic discovery.

EvoMining is an extension of genome mining that uses functional phylogenomics for the identification of expanded, re-purposed enzyme families, that have the potential to catalyze new conversion in NP biosynthesis; this method does not rely solely on homology, and thus can identify new potential pathways. Additionally, when complemented with CORASON, EvoMining can aid in the prioritization of BGCs that may involve novel chemistry.[291] The applicability of this method was illustrated by the discovery of arseno-organic metabolites in *Streptomyces coelicolor* and *Streptomyces lividans*.[292]

NPs can be classified according to their chemical scaffolds and signature biosynthetic enzymes that synthesize their core structures. The most common classes are polyketides and nonribosomal peptides, which are made by signature biosynthetic enzymes PKSs and NRPSs, respectively. The most popular genome mining tool antiSMASH uses manually curated gene cluster rules to identify core

(signature) and additional biosynthetic genes via profile Hidden Markov Models (pHMMs). This tool can additionally predict the core chemical structure by using the substrate specificity predictions of protein domains of PKS and NRPS modules. The common biosynthetic logic of these NPs facilitates their identification from genomes.

1.5.1 Type I Polyketide Synthases

Type I PKSs are large, multifunctional proteins with multiple modules containing several domains that are responsible for the non-iterative catalysis of one cycle of polyketide chain elongation. Examples of type I polyketides include erythromycin (**Figure 8**) and natamycin. They contain, at a minimum, an acyltransferase (AT) domain, a ketosynthase (KS) domain, and an acyl carrier protein (ACP) domain. The AT domain is responsible for choosing the monomer to be added during elongation, the KS domain catalyzes the decarboxylative Claisen condensation to bond the acyl moiety to the extender unit, and the ACP domain shuttles acyl intermediates between the catalytic domains. Modules may also contain optional domains, e.g., a ketoreductase (KR) domain, a dehydratase (DH) domain, and an enoylreductase (ER) domain. All optional domains successively reduce the β-keto group to a hydroxy (KR), a hydroxy to a double bond (DH), and a double bond to methylene (ER). Finally, a thioesterase (TE) at the C-terminal end of the PKSs releases the polyketide chain by cyclization or direct hydrolysis.[293–295]

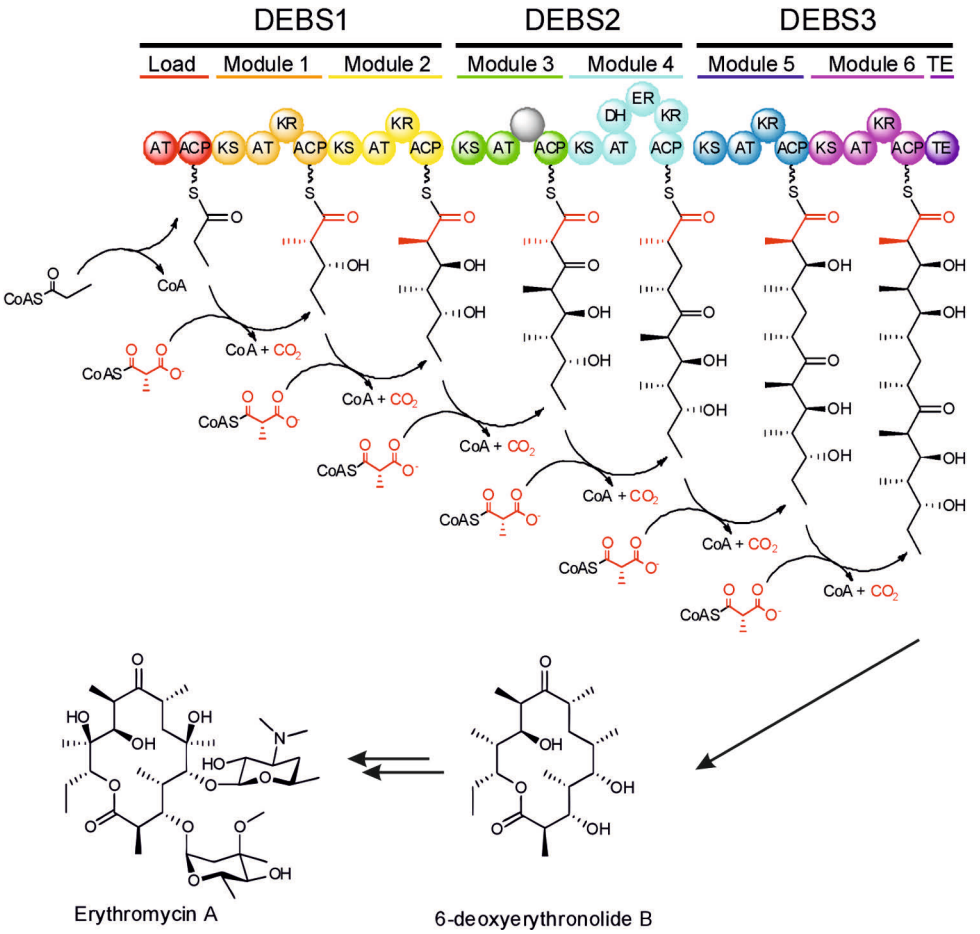


Figure 8. Erythromycin A biosynthesis. Three type I polyketide synthases containing six modules form the aglycone of erythromycin. Subsequently, the aglycone is converted into erythromycin A. Modified from [413].

1.5.2 Type II Polyketide Synthases

Type II PKSs are very similar to type I PKSs, with the key difference being that instead of one large megaenzyme, multiple monofunctional enzymes are used with invariable PKS cores.[296] At a minimum, the PKS consist of an ACP, and an α/β heterodimeric ketosynthase (KSα-KSβ).[297] Together, these enzymes function similarly to the AT, KS, and ACP domains in the type I PKS system, although they often lack an AT enzyme.[298] Type II PKSs often work iteratively, where multiple chain elongation steps are performed by the same enzyme.[299] Subsequently, additional enzymes of the BGC, e.g., KRs, aromatases, cyclases, and lyases, convert the nascent poly-β-keto intermediate into an aromatic scaffold.[300] Finally, tailoring enzymes such as oxygenases, methyltransferases, and glycosyltransferases modify the scaffold, diversifying the bioactivity of the respective compounds.[301] Importantly, glycosylation is how most NPs derive their bioactivities. The biosynthetic origins of these sugars are mainly derived from glycolytic intermediates and eventually convert to sugar-1-phosphate, which can then be activated by the appropriate nucleotidyltransferase.[302] Examples include aclacinomycins, doxorubicin (**Figure 9**), and tetracycline.

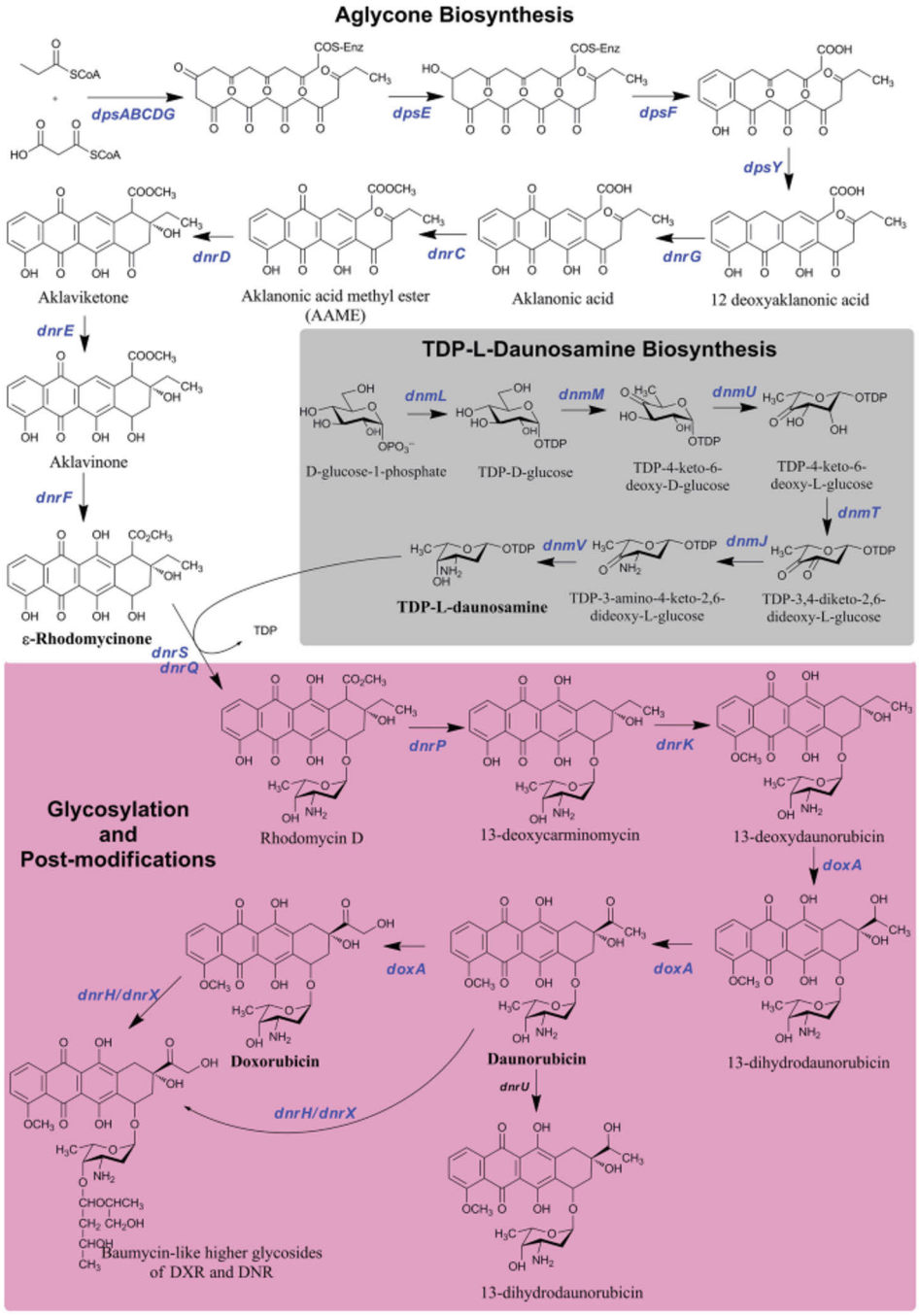


Figure 9. Doxorubicin biosynthesis. The aglycone is formed by type II polyketide synthases from propionyl-CoA and malonyl-CoA (white area). Thymidine diphosphate-L-daunosamine is formed from D-glucose-1-phosphate (grey area), which is then attached to the aglycone, and the intermediates are further modified by tailoring enzymes, finally, resulting in doxorubicin. Source: [414].

1.5.3 Nonribosomal Peptide Synthetases

NRPSs are enzymes that are involved in the synthesis of nonribosomal peptides. These enzymes function by linking individual amino acids to form a peptide without the use of ribosomes. Nonribosomal peptides include the important antibiotics penicillin, cephalosporins, and daptomycin. NRPSs are similar to type I PKSs in that they are large multifunctional enzymes. Furthermore, these enzymatic assembly lines consist of distinct modules responsible for the initiation, elongation, and termination of the synthesized intermediates. Like type I PKSs, these modules contain catalytic domains. The essential domains are adenylation (A), peptidyl carrier protein (PCP), condensation (C), and TE. The A domain activates amino acids, which are then post-translationally attached to the PCP domain. The C domain catalyzes the peptide bond formation between modules and the TE domain releases the fully assembled peptide.[303] In addition, numerous optional domains can be present, including epimerization (E), N-methylation, and cyclization that are responsible for diverse tailoring reactions.[304] The clinically important antibiotic daptomycin is shown in **Figure 10**, as an example of nonribosomal peptide biosynthesis.

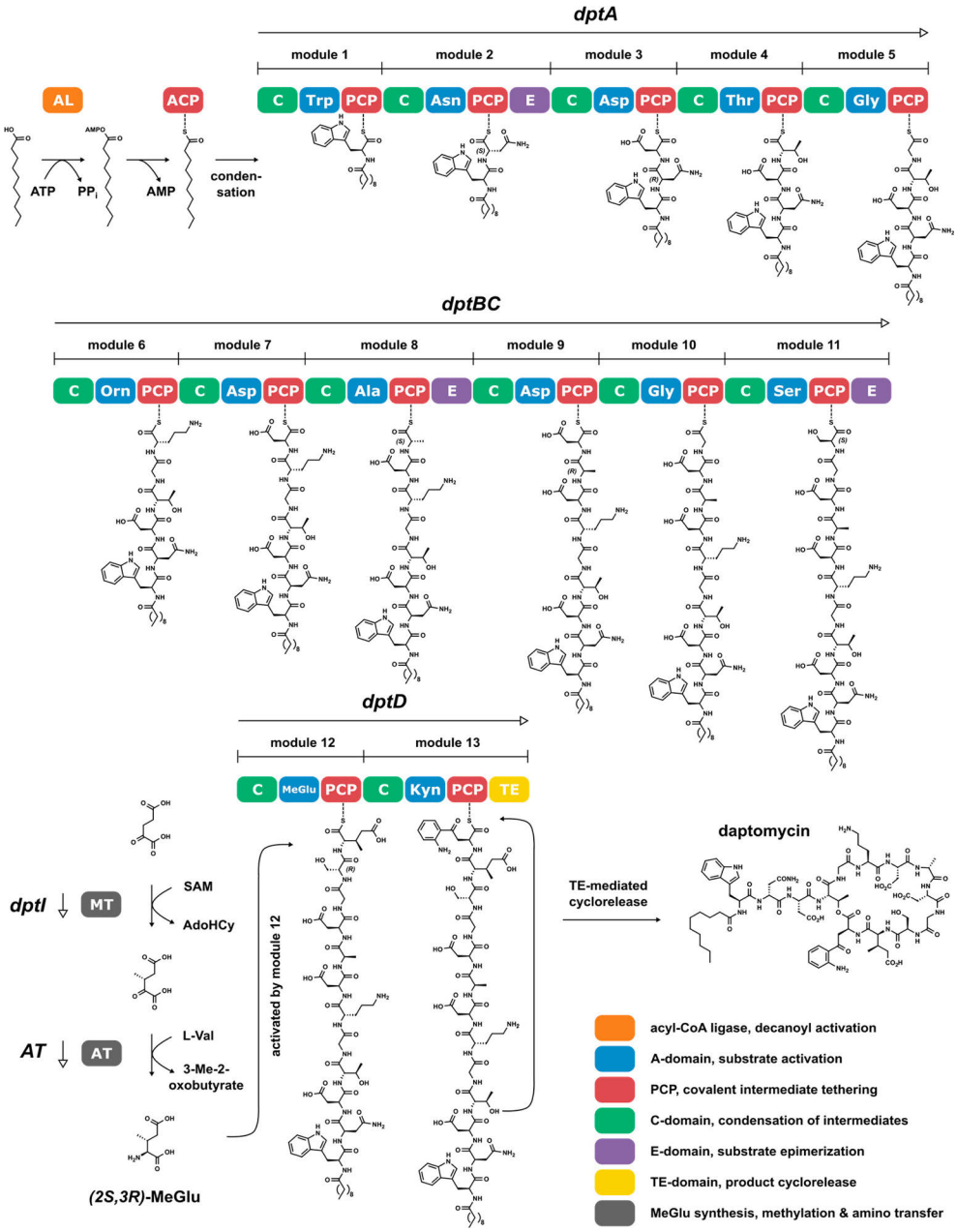


Figure 10. Nonribosomal biosynthesis of daptomycin. DptE and DptF initiate daptomycin biosynthesis, followed by the formation of the peptide core by the nonribosomal peptide synthetases, DptA, DptBC, and DptD. Source: [415].

1.6 Natural Product Diversification by Molecular Evolution

Molecular evolution, especially enzyme evolution, aids in explaining the arms race between organisms, i.e., how organisms are able to continuously innovate new NPs.[305] Primary metabolism and secondary metabolism are intimately related. BGCs are likely to have originated from primary metabolism, with HGT being an important factor in their evolution.[306] This is exemplified by fatty acid synthetases being the ancestor of PKSs. Specifically, AT and KS domains have revealed multiple gene duplications, gene losses, and HGTs have contributed to the evolution of type I PKSs.[307] Conversely, primary metabolism gene clusters are likely to have evolved independently and by different mechanisms, as demonstrated in multiple GAL pathway gene clusters in fungi.[308] A new evolutionary hypothesis has also been proposed, in which a non-essential secondary metabolic enzyme acts as an 'evolutionary seed' to generate an essential primary metabolic enzyme.[309] Metabolic pathways are likely to evolve by recruitment, that is, enzyme families are drawn to perform novel chemistries or use different substrates.[310] By understanding the diversification of enzymes and enzyme families, it is possible to engineer enzymes with novel functions and substrate specificities.[311]

Chemical variation is a hallmark of secondary metabolism.[7,312] It has been envisaged that a two-step evolutionary model can explain much of the chemical diversity of microbial NPs.[313] In this model, the first step is the slow appearance of ancestral biologically active metabolites, as reflected in the relatively few classes of NPs. The second step is the rapid diversification of pre-existing secondary metabolic pathways, resulting in the huge NP chemical space seen today.

Examples of the second step can be seen through simple evolutionary events such as single point mutations and HGT resulting in novel NPs.[313] Angucyclines are a large group of aromatic polyketides, e.g., gaudimycins, urdamycins, jadomycins, and landomycins.[314] The homologs JadH (jadomycin pathway) and PgaE (gaudimycin pathway) use the last common intermediate, prejadomycin (**Figure 11A**), as a substrate and show their catalytic promiscuity by performing 4a,12b-dehydration and 12b-hydroxylation, respectively.[315–317] Substrate promiscuity can be seen in the homologs UrdMred (urdamycin pathway) and LanV (landomycin pathway) which both catalyze 6-ketoredution (**Figure 11B**) but use the double hydroxylated intermediate or the earlier 12-hydroxylated intermediate as respective substrates.[318,319] Both substrate and catalytic promiscuity are key driving forces for the diversification of NP biosynthetic enzymes and thus NPs themselves,[320] and can emerge from single point mutation as in the examples above.

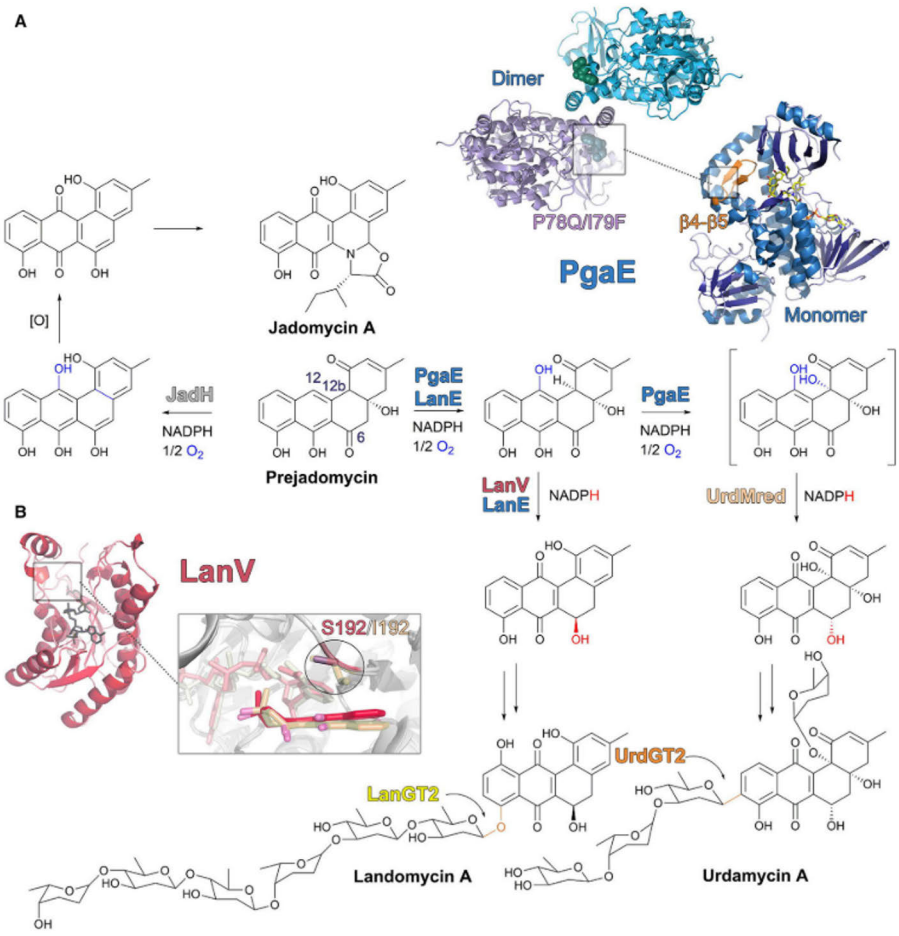


Figure 11. Diversification of the angucycline biosynthetic pathway from enzyme promiscuity. (A) Prejadomycin is the substrate for JadH, PgaE, and LanE, but they catalyze different reactions. (B) Enzymatic activity toward UrdMred-type chemistry was achieved by a single point mutation in LanV. Source: [313].

Another example is, the HGT of promiscuous enzymes can also lead to diversification of NPs, as shown by the ABBA prenyltransferases superfamily (**Figure 12**). ABBA prenyltransferases catalyze the aromatic prenylation of an unusually wide variety of NPs in bacteria and fungi (**Figure 12B**).[321–324] These enzymes share low sequence homology and instead are defined by a conserved tertiary structure (**Figure 12C**).[325] They typically exhibit strict specificity for their C5, C10, or C15 prenyl donor groups, but they can accept diverse aromatic acceptors.[323,326] ABBA prenyltransferases have been recruited in a wide variety of NP classes, including indole alkaloids, peptides, flavonoids, coumarins, terpenoids, and phenazines,[321–325,327] and their broad substrate specificity is likely to facilitate

their spread and integration[313]. Indole alkaloid ABBA prenyltransferases have been shown to catalyze the regioselective normal or reverse prenylation of a variety of NPs.[325,328–331] Furthermore, this broad substrate specificity has been applied to improve the clinically used antibiotic, daptomycin, with the permissive prenyltransferases PriB, FgaPT2, and CdpNPT (**Figure 12**).[332] The active exchange of ABBA prenyltransferases between different classes of NPs increases structural diversity, as suggested by phylogenomic analysis.[333]

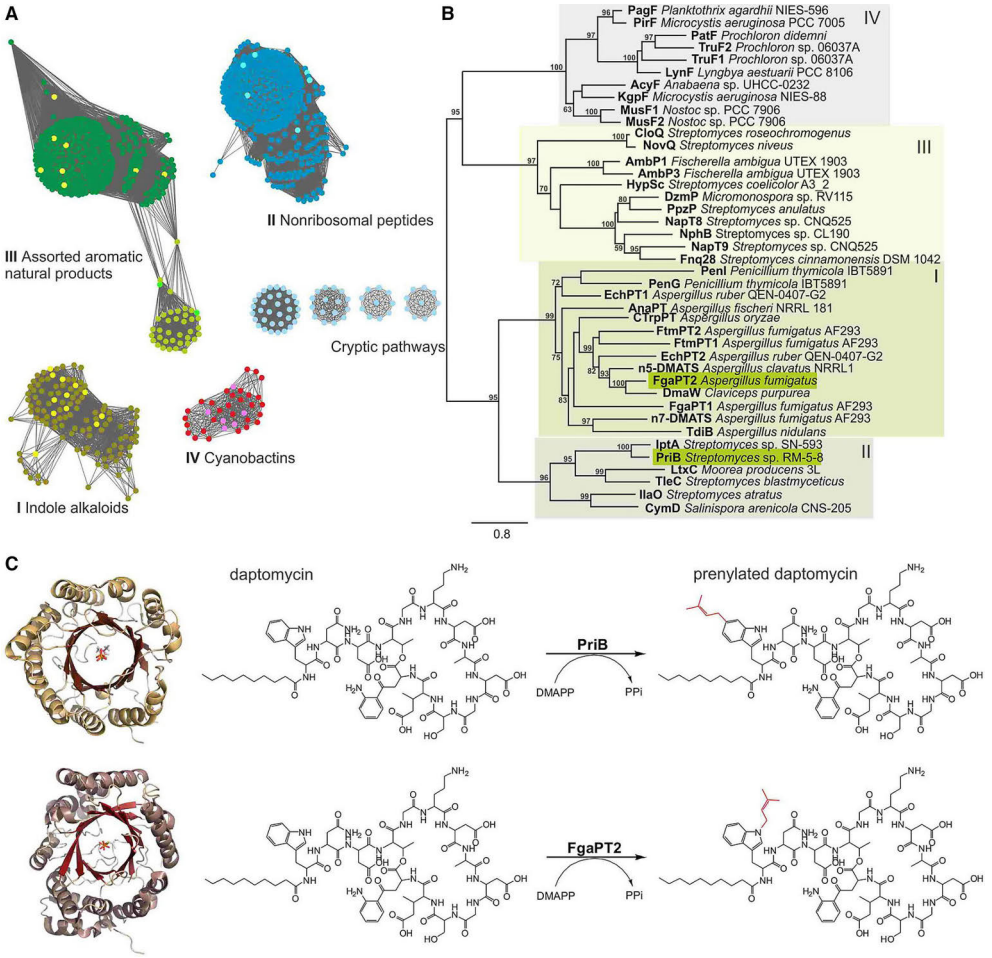


Figure 12. Gain of function in natural product biosynthetic gene clusters by enzyme promiscuity. (A) ABBA prenyltransferases catalyze the isoprenylation of various aromatic natural products. (B) They can be grouped into several families and are widely distributed in natural product biosynthetic gene clusters of a variety of bacteria and fungi. (C) ABBA prenyltransferases have a distinctive barrel-shaped fold, lined with β-sheets. *In vitro*, daptomycin can be prenylated by the unrelated permissive PriB and FgaPT2 enzymes. Source: [313]

2 Aims

This work explores the relationship between genomics, ecology, regulation and production of natural products of Actinobacteria through genome mining.

Specifically, the aims are the following:

I To use microbe-microbe interactions to elicit production of natural products and enzymes with the aid of genome mining (Articles I, II, & III).

II To use genome mining for discovery of aromatic polyketide natural products (Articles IV & V).

3 Materials and Methods

An overview of the materials and methods is given here; experimental details are found in the original publications (I-V).

3.1 Molecular biology and Microbiology

Standard molecular biology and microbiology techniques were used in all articles.[334]

3.1.1 Media, Strains, and Culture Conditions

Actinobacteria media used included, M1 medium (10 g/L starch, 4 g/L yeast extract, 2 g/L peptone, ±18 g/L agar in artificial seawater), P medium (1 g/L peptone, 4.55 g/L glucose anhydrase, 0.4 g/L $MgSO_4 \cdot 7\ H_2O$, 0.4 g/L K_2HPO_4, 100 g/L potato juice), Y medium (9.1 g/L glucose anhydrase, 2 g/L NH_4NO_3, 2 g/L $CaCO_3$, 26 g/L bakery yeast), YE medium (the same as Y medium but with yeast extract instead of bakery yeast), Y2 medium (20 g/L glycerol, 2.5 g/L bakery yeast, 2.5 g/L yeast extract, 1 g/L K_2HPO_4, 1 mL/L trace salts solution [$FeSO_4 \cdot 7\ H_2O$ 1 g/L, $MnCl_2 \cdot 4\ H_2O$ 1g/L, $ZnSO_4 \cdot 7\ H_2O$ 1 g/L]), YE2 medium without bakery yeast, ISP4 (BD Difco™), TSB medium (17 g/L tryptone, 3 g/L soy, 5 g/L NaCl, 2.5 g/L glucose), and GYM medium (4 g/L glucose, 4 g/L yeast extract, 10 g/L malt extract).The *E. coli* media used included: 2 x TY (16 g/L tryptone, 10 g/L yeast extract, 5 g/L NaCl), TB (20 g/L tryptone, 24 g/L yeast extract, 4 mL/L glycerol, 100 mL/L phosphate buffer [0.17 M KH_2PO_4 and 0.72 M K_2HPO_4]) and nutrient broth (BD Difco™). Standard and appropriate concentrations of antibiotics were added to cultures as needed.

The general culture conditions were, rotary sharking at 200 – 300 rpm, 30 ˚C for 3 – 10 days for Actinobacteria. For *E. coli* cultures, rotary shaking was at 200 – 250 rpm, 16 – 37 ˚C for 16 – 24 hours. The strains used are presented in **Table 1**.

Table 1. Microbial strains used in this thesis.

GENUS	SPECIES	STRAIN	NOTES
Escherichia	*coli*	TOP10	Cloning
Escherichia	*coli*	ET12567	Conjugation
Saccharomyces	*cerevisiae*	BY25610	Prey / elicitor
Streptomonospora	sp.	PA3	Sponge symbiont bioprospect
Streptomyces	*albus*	J1074	Expression host
Streptomyces	*candidus*	NRRL 3601	Predator
Streptomyces	*erythrochromogenes*	NRRL B-2112	Enzyme diversification
Streptomyces	*galilaeus*	ATCC 31615	Predator, enzyme diversification
Streptomyces	*kanamyceticus*	DSM 40500	Predator
Streptomyces	*lavendulae*	YAKB-15	ChoD, predator
Streptomyces	*lividans*	TK24	Expression host
Streptomyces	*peucetius*	ATCC 27952	Predator
Streptomyces	*platensis*	NRRL 8035	Predator
Streptomyces	*showdoensis*	ATCC 15227	Predator
Streptomyces	sp.	MBK6	Soil isolate, predator
Streptomyces	sp.	PGA64	Predator
Streptomyces	sp.	S-378	Enzyme diversification

3.1.2 Plasmids

The plasmids used were pSET152 for overexpression via conjugal transfer from *E. coli* to *Streptomyces* with the help of plasmid pUZ8002 (Article I & II). Plasmids pIJ486 or pIJE486 were used for overexpression via protoplast transformation of *Streptomyces* (Articles I & III). A pWHM3-based plasmid was used for gene inactivation in *Streptomyces* (Article II). Plasmids pBADHisBΔ and pBHBΔ were used for protein production in *E. coli*, and pUC18 for cloning in *E. coli* (Articles I, II, & III). Plasmid construction, primers, inserts, and restriction sites are described in their respective articles.

3.1.3 Nucleic Acid Isolation

Actinobacteria were cultured in a medium supplemented with 0.5% + glycine for three to four days using a previously developed protocol[335] with slight modifications

to extract genomic DNA (Articles I, III, & IV). Briefly, cells were centrifuged and washed twice with a 10% sucrose solution. Lysozyme, 10% SDS, and proteinase K were added to lyse cells and degrade proteins. Chloroform and 5M NaCl were added to remove everything except DNA. The DNA was precipitated with isopropanol, washed with 70% ethanol, and stored at 4 ˚C.

Biological quadruplicate samples were pooled for total RNA isolation by adding ice cold STOP solution (5% phenol in ethanol) to stop the degradation of the RNA and then flash frozen in liquid nitrogen. Total RNA was extracted from liquid nitrogen frozen samples, ground using a mortar and pestle, and then isolated using a RNeasy kit (Qiagen) with on-column DNase treatment (Article II).

3.1.4 Next Generation Sequencing

Genomic DNA was sent to Eurofins Genomics (Ebersberg, Germany) or the Finnish Functional Genomics Centre (Turku, Finland) for PCR-free shotgun library preparation (Illumina) and sequenced using MiSeq v3 2 x 300 bp paired-end reads (Illumina) (Articles I, III, & IV).

Total RNA was sent to Novogene (Cambridge, UK) for quality control (Agilent 2100), rRNA depletion (Ribo-Zero kit), library preparation (Illumina), and sequencing using NovaSeq 6000 (Illumina) to produce 2 x 150 bp reads (Article II).

3.2 Bioinformatics

The bioinformatics analyses were carried out on the Taito or Puhti super-cluster of CSC – IT Center for Science (Espoo, Finland).

3.2.1 Genome Assembly and Annotation

The quality of the reads was manually checked before and after trimming and error correction using FASTQC[336]. The reads were then trimmed, filtered, error corrected, and assembled using A5-miseq[337]. The contigs were contiguated with ABACAS[338], using *Streptomyces albus* NK660 (CP007574.1) as the reference, and the gaps were filled using IMAGE[339]. The final assembly was annotated using RAST[340], evaluated for completeness using BUSCO[341] and deposited in the National Center for Biotechnology Information (NCBI) database. The accession numbers for the genomes are VTZW00000000, JACERG000000000, and SMSN00000000 for *Streptomonospora* sp. PA3 (Article IV), *Streptomyces* sp. MBK6 (Article III), and *Streptomyces lavendulae* YAKB-15 (Article I), respectively. The genomes were annotated using RAST[340].

3.2.2 Genome Mining

The genomes were mined for putative BGCs using antiSMASH[281]. The characterized BGCs were deposited in MIBiG under the accession numbers BGC0002045 for the persiamycin A BGC of *Streptomonospora* sp. PA3 (Article IV), BGC0002090 for the metacycloprodigiosin / undecylprodigiosin BGC of *Streptomyces* sp. MBK6 (Article III), BGC0002789 for the pentamycin/filipin BGC of *Streptomyces lavendulae* YAKB-15 (Article II), and BGC0002788 for the 14-hydroxyisochainin BGC of *Streptomyces peucetius* ATCC 27952 (Article II).

3.2.3 Phylogenetics

Multiple sequence alignments (MSA) were performed using Jalview[342] and ClustalO[343]. The phylogenetic trees were constructed using FastTree[344], and the trees were visualized with Dendroscope[345]. The MSA was also used to correlate sequence similarities and secondary structure data using ESPript[346] (Article V).

3.2.4 Transcriptomics

The reads were manually checked using FASTQC and trimmed using TRIMMOMATIC[347]. The trimmed reads were aligned to the genome using Bowtie2[348]. The aligned reads were counted using HTSeq[349]. Differential expression was performed using edgeR[350]. All analyses were performed using Chipster[351]. Active BGC were determined on the basis of a previously developed method[352]. The raw and processed data were deposited to the GEO database under the accession number GSE228628 (Article II).

4 Results and Discussion

This section puts an emphasis on the work I did within the context of the larger story.

4.1 Microbe-Microbe Interactions for Eliciting Natural Product Production

4.1.1 Polyenes and Cholesterol Oxidases

In Article I, we noted that *Streptomyces lavendulae* YAKB-15 only produced cholesterol oxidase (ChoD) in a medium containing whole yeast cells and observed the disappearance of yeast cells. A literature search revealed that cholesterol oxidase can also oxidize other sterols, including ergosterol, which is a constituent of the yeast cell membrane. Moreover, a cholesterol oxidase gene had been identified in the polyene antifungal pimaricin (also known as natamycin) BGC of *S. natalensis*.[199] In that study, they showed that the cholesterol oxidase gene produced a functional cholesterol oxidase, and gene-inactivation and complementation experiments showed that it was involved in pimaricin biosynthesis. In a later study, overexpression of the cholesterol oxidase of the pimaricin BGC in *S. gilvosporeus* resulted in an impressive yield of 7 g/L of pimaricin, a 153% improvement.[266] With this knowledge we hypothesized that the cholesterol oxidase we studied may also be involved in the production of an antifungal NP, which could link the production of our cholesterol oxidase to the disappearance of yeast cells.

In Article II, we continued our investigation by analyzing extracts from *S. lavendulae* YAKB-15 co-cultures, which revealed the production of macrolide polyene-type compounds that were elucidated as pentamycin and filipin III by NMR. With this information, I was able to identify a pentamycin-like BGC from the genome as the cluster likely responsible for biosynthesis. Further, I compared the predicted pentamycin-like BGC to the characterized cluster of *Streptomyces* sp. S816.[353] The average nucleotide identity (ANI) analysis showed that the clusters were 90.55% identical and they had nearly identical synteny (**Figure 13**). The pentamycin-like BGC was located at the end of a contig, and four genes (two regulators, a cholesterol oxidase, and a thioesterase) putatively part of the cluster

were located at the end of another contig. It is likely that these contigs should be joined, based on the conserved synteny of the four genes in both *S. lavendulae* YAKB-15 and *Streptomyces* sp. S816. However, there was a transposase gene located at the very end of the contigs, which could suggest that these four genes have been transposed to another locus.

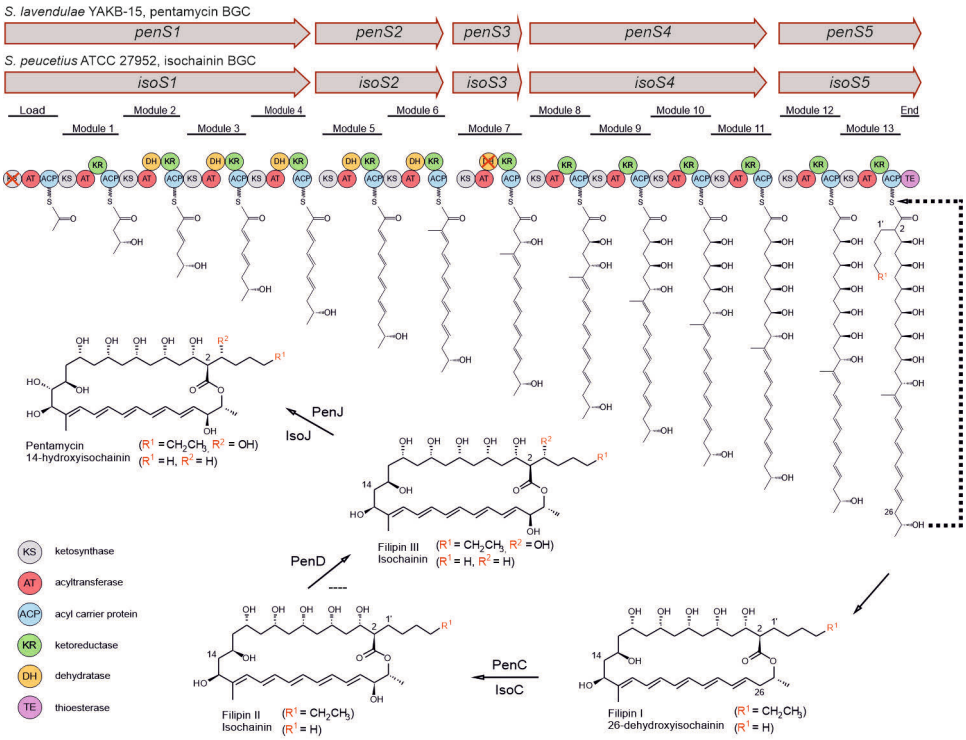


Figure 13. Comparison of pentamycin and isochainin biosynthesis of *S lavendulae* YAKB-15 and *S. peucetius* ATCC 27952, respectively. Notably, the homolog of *penD* in the *iso* cluster is absent, resulting in the absence of C-1' hydroxylation.

Next, we tried to determine whether cholesterol oxidase and / or antifungal polyenes were involved in the disappearance of yeast cells. We attempted to inactivate the cholesterol oxidase gene and one of the core PKS genes of the pentamycin BGC in *S. lavendulae* YAKB-15, but without success. We then identified another strain, *S. peucetius* ATCC 27952, with a putative pentamycin-like BGC, which also contained a cholesterol oxidase. This strain has been well studied for its ability to produce anticancer compounds, daunorubicin and doxorubicin, and is known to be genetically tractable.[354,355] We confirm that *S. peucetius* ATCC 27952 produced functional cholesterol oxidase, observed the disappearance of yeast cells,

and continued with this strain as a substitute for the genetically intractable strain of *S. lavendulae* YAKB-15.

The pentamycin-like BGC of *S. peucetius* ATCC 27952 was analyzed and I discovered that the ANI value was lower, 84.38%, compared to *Streptomyces* sp. S816. Further analysis showed that one cytochrome P450 monooxygenase gene (*penD* homolog) was absent from the pentamycin-like BGC of *S. peucetius* ATCC 27952 (**Figure 13**). Structural elucidation of the NP revealed that it was 14-hydroxyisochainin, a known antifungal polyene compound.[356,357] Recently, 14-hydroxyisochainin was shown to have a broader spectrum of antibiosis than pentamycin and could inhibit the take-all fungus, *Gaeumannomyces graminis* var. *tritici*, and conferred the producer, *Streptomyces* sp. N2, with growth-promoting benefits for *Arabidopsis thaliana*.[356]

There are two differences between pentamycin and 14-hydroxyisochainin. The first difference was found in the length of the side chain (**Figure 13**). This may have been a result of the final AT domain having a different specificity for different length extender units.[358] The amino acid sequences of the final AT domains were aligned using three pentamycin/filipin III producers (*Streptomyces* sp. S816, *S. lavendulae* YAKB-15, *S. avermitilis* MA-4680) and two 14-hydroxyisochainin producers (*Streptomyces* sp. N2, *S. peucetius* ATCC 27952). Unfortunately, no common mutation could be identified between the two groups. In addition, *Streptomyces* sp. N2 actually co-produced pentamycin and 14-hydroxyisochainin,[356] so it is likely the last AT domain in this strain can utilize both extender units. Finally, no sequences were available for the other 14-hydroxyisochainin producer, *S. cellulosae*,[357] which could have aided in the analysis. Nevertheless, the differences in the pre-cursor pool of each strain could also have directed the biosynthesis.[357]

The second difference was found at position C-1'. CYP105D6, *penD* homolog, and CYP105P1, *penC* homolog, of *S. avermitilis* have been shown to catalyze the hydroxylation of C-1' and C-26, respectively.[359,360] The *penD* homolog was absent in the 14-hydroxyisochainin producer *S. peucetius* ATCC 27952 and the structure of 14-hydroxyisochainin lacked C-1' hydroxylation (**Figure 13**). Interestingly, *Streptomyces* sp. N2, the co-producer of pentamycin and 14-hydroxyisochainin, encodes a *penD* homolog, indicating that the longer six-carbon fatty acid side chain may be required for the hydroxylation of C-1'.

After the analysis of the compounds and BGCs, we proceeded to knockout the cholesterol oxidase gene (*isoG*) and a core PKS gene (*isoS1*) of *S. peucetius* ATCC 27952. The mutant knockout strains *S. peucetius* ATCC 27952 Δ*isoG* and *S. peucetius* ATCC 27952 Δ*isoS1* were confirmed for their lack of cholesterol oxidase and polyenes production, respectively. Inactivation of the PKS in *S. peucetius* ATCC 27952 Δ*isoS1* did not block the production of the cholesterol oxidase, IsoG. A previous report has shown that the cholesterol oxidase (*pimE*) within the pimaricin

BGC is involved in the regulation of the polyene metabolic pathway.[199] However, 14-hydroxyisochainin production was not blocked in *S. peucetius* ATCC 27952 Δ*isoG*. I envisaged that this may be due to the location of *pimE*, at the center of the pimaricin BGC, which is in contrast to *isoG* located at the edge of its BGC. PimE may have had its functionality expanded to be an antifungal sensor / regulator and was recruited by the pimaricin BGC, according to the expansion-and-recruitment evolutionary principle.[291] Finally, pentamycin/isochainin and pimaricin were elicited differently. A medium containing yeast extract was sufficient to elicit pimaricin production, while pentamycin/isochainin production was not elicited by yeast extract, but only whole yeast cells. This suggested that the BGCs were regulated differently and provides an explanation why *S. peucetius* ATCC 27952 Δ*isoG* still produces polyenes.

The mutants were then tested for their ability to disappear yeast cells; both mutants retained this ability. These results showed that the predatory abilities of *Streptomyces* do not depend on cholesterol oxidase or the polyenes alone. Indeed, cholesterol oxidase has been known to exhibit potent insecticidal activity, by attacking insect sterols,[361] and polyenes were well known antifungals.[362] However, we found it interesting that cholesterol oxidase and polyenes were both elicited by whole yeast cells but did not have an apparent connection to each other. A redundant multipronged attack model by *Streptomyces* may connect the two, and is akin to how bacterial pathogens, such as *Legionella*, employ redundant virulence mechanisms.[363]

4.1.2 Predation as the Elicitor of Natural Product Production

Microscopic observation of *Streptomyces* interacting with yeast cells revealed that *Streptomyces* sequestered yeast cells within its mycelia. The sequestered yeast cells displayed obvious morphological changes compared to untouched yeast cells, i.e., the yeast cells shriveled (**Figure 14a**). Based on our initial idea that cholesterol oxidase was regulating the polyene BGC, we created a green fluorescent protein (GFP)-promoter probe of the cholesterol oxidase gene in *S. lavendulae* YAKB-15 to monitor the expression of both cholesterol oxidase and polyene biosynthesis. Fortunately, even though cholesterol oxidase was later found to not regulate polyene biosynthesis, they were still regulated in a similar manner. I then co-cultured *S. lavendulae* YAKB-15, with the GFP-promoter probe construct, with *Saccharomyces cerevisiae* BY25610, a strain that constitutively expresses mCherry, a red fluorescent protein (RFP) (**Figure 14b**). Images were taken using fluorescence microscopy in a time series to create a time-lapse video. I found that when *Streptomyces* co-localized with yeast cells, GFP was overexpressed and that the intensity of GFP was much lower when the cells did not co-localize (**Figure 14c**). This experiment suggested that the interaction was mediated by physical contact rather than diffusible signals.

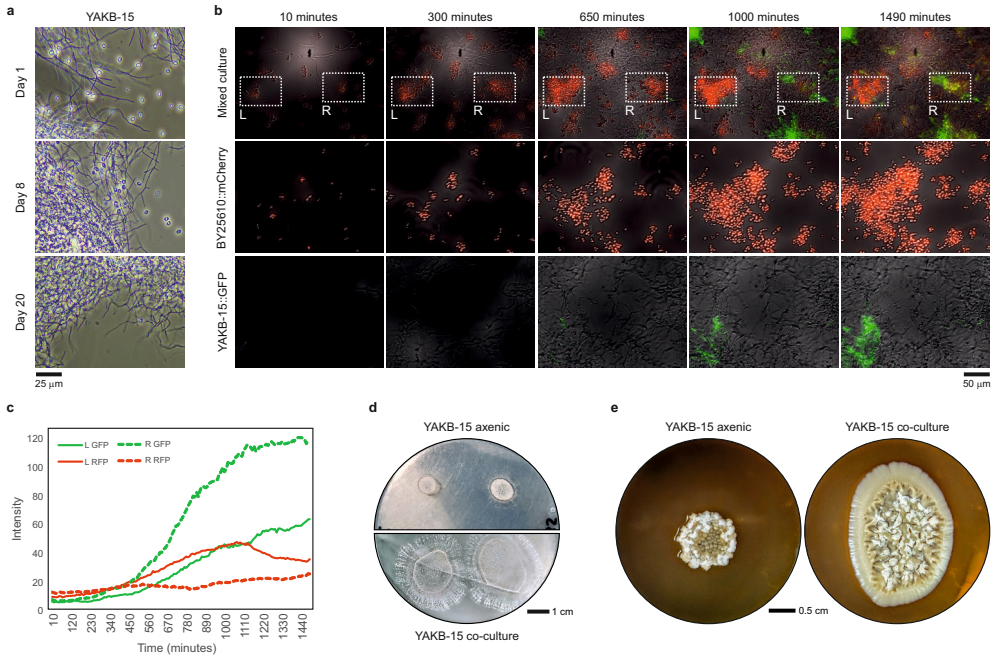


Figure 14. Microscopic observations of *Streptomyces*-yeast interactions. (a) Initial observations of dead yeast cells disappearing from a culture of *S. lavendulae* YAKB-15 after 20 days. (b) Time course fluorescence microscopy of *S. lavendulae* YAKB-15 with a choD-GFP promoter probe and *Saccharomyces cerevisiae* BY25610 constitutively expressing mCherry (RFP). Two populations of yeast cells are marked with white boxes. The left (L) population was untouched by *Streptomyces*, while the right (R) population interacted with *Streptomyces* and disappeared. (c) Fluorescence intensities over the time course showed an inverse relationship between GFP and RFP in the R population, while in the L population the intensities remained largely the same. (d, e) *S. lavendulae* YAKB-15 colonies grew larger when co-cultured with yeast on both nutrient-less and nutrient-rich agar. Source: Article II.

To further characterize the microbe-microbe interactions, I cultured *S. lavendulae* YAKB-15 on a nutrient rich agar plate axenically and spotted in the middle of a yeast colony. I discerned no zone of inhibition, which provided further evidence that this interaction was based on intimate physical contact (**Figure 14e**). Additionally, the size of the *Streptomyces* colony in the co-culture was noticeably larger, indicating that *Streptomyces* could be consuming the yeast (**Figure 14e**). To test for the yeast-consuming capacity of *S. lavendulae* YAKB-15, I repeated the previously mentioned plating experiment, but on an agar plate without any nutrients as a negative control and with a lawn of yeast cells. I observed, *Streptomyces* colonies grew much larger when co-cultured with yeast (**Figure 14d**), which gave further evidence that *Streptomyces* was actually able to utilize yeast for nutrients and suggested that this interaction could be further classified as predation. Furthermore,

no zone of inhibition was observed, as is typical for diffusible inhibition or killing, again suggesting that physical contact was essential (**Figure 14d**).

Cholesterol oxidases and polyenes were found to be associated with mycelia rather than the supernatant. These data pointed to the possibility that *Streptomyces* was performing physical-contact-mediated predation. From the perspective of public goods,[206,207] this would make sense since the nutrients released from lysed yeast cells would have a limited potential to diffuse. As discussed in **Section 1.2.2**, the well-established microbial predator *M. xanthus* has displayed this behavior as has *Streptomyces*, although not in the context of predation. When viewing the *Streptomyces*-yeast interaction through the lens of predation, we found a framework to explain the similar regulation of cholesterol oxidase and polyenes that otherwise had no apparent connection.

4.1.2.1 Omics Insights into the Predatome

With predation in mind, I explored the *Streptomyces* genome for genes that encode enzymes capable of digesting the polysaccharide-rich yeast cell wall. The cell wall of *Sacc. cerevisiae* is mainly composed of β-glucans, mannoproteins and chitin. More specifically, the β-glucans have a β-1,3-linked backbone with β-1,6-linked branches,[364] and mannans have an α-1,6-linked backbone with α-1,2- and α-1,3-linked branches.[365] Sixteen CAZyme genes were identified, four β-glucanases, five α-mannosidases, and seven chitinases. The CAZyme database was used as a guide to identify the putative activity of the CAZymes encoded within the genome of *S. lavendulae* YAKB-15, in addition to a literature search. The degradation of yeast cell wall polysaccharides was best studied in the human gut bacteria *Bacteroides*. In *B. uniformis,* enzymes of the glycoside hydrolase (GH) families 3, 16, 158 have been shown to be responsible for the cleavage of β-1,3-glucans and GH30 for the cleavage of β-1,6-glucans.[366] In *B. thetaiotaomicron*, it has been shown that first GH99 and GH92 debranches α-mannan backbone allowing GH38 and GH76 to cleave the backbone.[367,368] Next, one CAZyme putatively capable of degrading a component of the yeast cell wall was chosen for heterologous expression and confirmation of its catabolic capability, giving evidence that *Streptomyces* is capable of digesting the yeast cell wall.

Each of the three chosen CAZymes was heterologously expressed in *E. coli* and purified. The purified CAZymes were then assayed with their respective pure substrates as well as with whole yeast cells. GH16_3 glucanase (YeeB3) had the highest activity, followed by GH18 chitinase (YeeC3), but GH92 mannosidase (YeeA3) showed no activity with commercial substrates. The lack of mannosidase activity could have been explained by the incorrect prediction of enzymatic activity. It was reported that of five predicted *Cellulosimicrobium cellulans* GH92 enzymes

only one had the correct predicted activity,[369] suggesting that the enzymatic activity is difficult to predict. Therefore, it is possible that the mannosidase we chose is capable of digesting some other part of yeast mannan that is different from our given substrate. Here, we showed that *Streptomyces* does indeed have the ability to digest the yeast cell wall.

Subsequently, we further looked at yeast as an elicitor, not only of cholesterol oxidase and polyenes, but also of CAZymes. I performed a time course differential transcriptomic analysis to observe if/when the cholesterol oxidase, the polyene, and the CAZyme genes were upregulated, and confirmed the data with proteomics (**Figure 15a**). In addition, the transcriptome-level response of yeast could also be observed. I found that the polyene BGC was upregulated around 24 hours, as were most CAZyme genes, while the cholesterol oxidase gene was upregulated much later, around 48 hours (**Figure 15a**). The fact that the cholesterol oxidase gene was upregulated after the polyene BGC precluded participation in the regulation of the polyene BGC and was in line with the gene inactivation experiment.

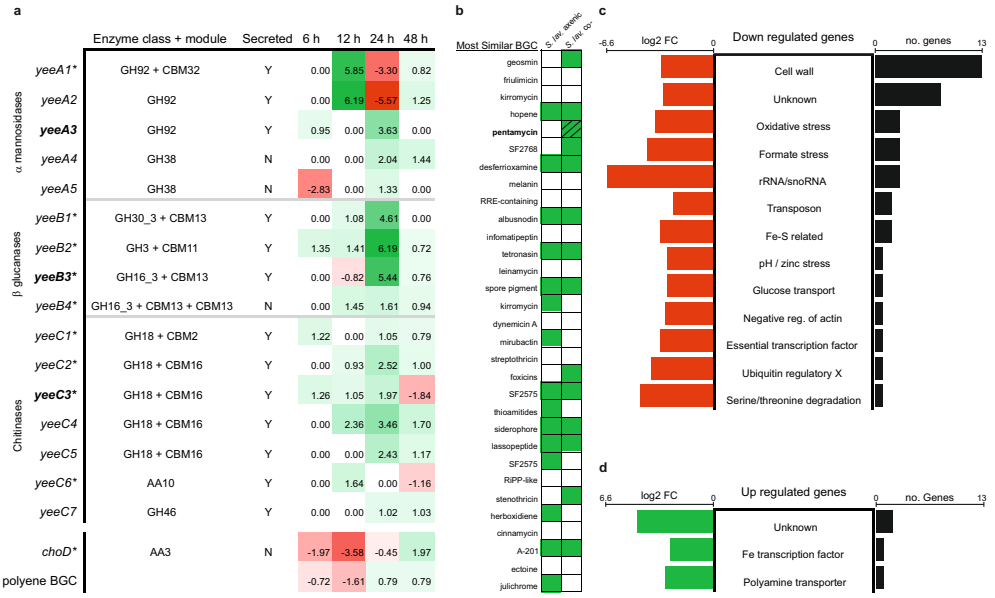


Figure 15. Genomic, transcriptomic, and proteomic insights into the predatome. (a) Differential expression (axenic versus co-culture) of *S. lavendulae* YAKB-15 genes producing putative yeast cell wall digesting enzymes; additionally, *choD* and the polyene clusters, which attack the yeast cell membrane, are shown. Numbers are log2 fold change, bolded genes were heterologously expressed, and genes marked with an * were confirmed as upregulated by proteomics. (b) Active *S. lavendulae* YAKB-15 clusters from axenic and yeast co-cultures, shown in green, and the identified compound is shaded. (c, d) *Saccharomyces cerevisiae* BY25610 mean differential expression between axenic and co-cultures and the number of genes grouped by cellular processes. Source: Article II.

Under closer examination, I found that two out of three GH92 mannosidase genes were strongly upregulated earlier, after 12 hours, followed by strong downregulation, after 24 hours (**Figure 15a**). The putative CAZyme products of these genes were predicted to debranch yeast mannans. I hypothesized, *Streptomyces* may be able to sense the resulting oligomannan branches and then repress the expression of the debranching mannosidases, while also increasing the expression of mannosidases further down the catabolic pathway. A recent study showed that cello-oligosaccharides could elicit a number of molecules from the virulome of *S. scabiei*.[370] These results also showed that CAZyme genes related to the digestion of the yeast cell wall digestion were upregulated before the polyene BGC and cholesterol oxidase gene, suggesting that they also participate in the attack against yeast.

RNA-Seq data also revealed significant changes in the transcription of several cryptic BGCs in *Streptomyces lavendulae* YAKB-15. In axenic cultures 15/31 (48%) of the BGCs were active, while 45% were active under co-culture conditions (**Figure 15b**). Additionally, 31% of the inactive BGCs became active, 40% of the active BGCs became inactive, 29% of the BGCs remained active, and 35% remained silent (**Figure 15b**). Of the five yeast-elicited BGCs, only the SF2768-like diisonitrile NP has been reported to have antifungal activity,[371] in addition to the observed pentamycin (**Figure 15b**). Geosmin was another yeast-elicited NP. The remining two yeast-elicited NPs were foxicins-like and stenothricin-like (**Figure 15b**). Foxicin A has been reported to be a siderophore that aids in iron acquisition, and additionally has antibiotic activity, both of which could provide fitness advantages against yeast.[372] Lastly, stenothricin has been shown to have wide antibiotic activities.[373] Together, these results suggested that *Streptomyces* predation on yeast is highly complex, as shown by its probable multipronged attack.

The response of yeast to the attack from *Streptomyces* was also observed on the transcriptomic level. I found that downregulation dominated the transcriptional response. The largest group of genes downregulated were related to cell wall biosynthesis (**Figure 15c**). The strongest downregulated genes were related to rRNA and snoRNA (**Figure 15c**). The general pattern of downregulation resembled the metabolic suppression of cryptobiosis as studied in tardigrades, where they become dormant under hostile conditions.[374] Indeed, yeast is commonly known to undergo cryptobiosis, as seen in most grocery stores in the form of dry yeast, where its metabolism shifts and is suppressed as it is desiccated.[375]

Only four genes were upregulated in yeast, two genes of unknown function, one Fe transcription factor, and one polyamine transporter (**Figure 15d**). Intracellular polyamines were likely depleted as a result of polyenes forming pores in the cell membrane, causing leakage. Polyamine depletion was shown to result in the alteration of the yeast cell wall, making it thicker, more heterogeneous, and irregular in shape.[376] Microscopic observations were in congruence, as the cell wall was

visibly thicker and irregular in shape (**Figure 14a**). Furthermore, polyamines have been shown in general to play important roles in the yeast response to stress.[377] Upregulation of a polyamine transporter gene appeared to have been the only active response yeast had to the *Streptomyces* attack.

4.1.2.2 *Streptomyces* Predation is Widespread

Furthermore, our objective was to show that the predatory phenomenon is widespread. We chose seven of our in-house strains and found that all except one strain, *S. lividans* TK24, displayed the characteristic sequestering and subsequent disappearance of yeast in liquid cultures, which provided evidence that the predatory phenomenon of *Streptomyces* is widespread. Furthermore, all tested strains showed the production of at least CAZymes or the potential to produce polyenes or cholesterol oxidase, except *S. lividans* TK24.

4.2 Genome Sequencing of Actinobacteria

Sequencing of *S. lavendulae* YAKB-15 resulted in 15,984,844 reads. The reads were normalized, error corrected, and trimmed to 7,196,183 reads, which were *de novo* assembled into 98 contigs. The contigs were then contiguated into 73 scaffolds with an N50 of 447,215 bp. The final genome assembly was 7.8 Mbp with a G + C content of 72.2% and a median coverage of 199x. The draft genome was assessed for completeness using BUSCO and resulted in 36 out of 40 (90%) single-copy orthologs being found. The genome was annotated and resulted in 7,001 CDSs.

Sequencing of *Streptomyces* sp. MBK6 resulted in 4,530,672 reads. The reads were error corrected and trimmed to 4,379,905 reads, which were *de novo* assembled into 65 contigs. The contigs were contiguated into 25 scaffolds with an N50 of 664,525 bp. The final assembly of the genome was 7.6 Mbp with a G + C content of 72.7% and a median coverage of 144x. The draft genome was assessed for completeness using BUSCO and resulted in 40 out of 40 (100%) single-copy orthologs being found. The genome was annotated and resulted in 6,742 CDSs.

The *Streptomyces* genomes fall well within the normal bounds of size, G + C content, and linearity.[233] The genomes also contain an average number of CDSs.[378] The coverage of the genomes was also comparable to public assemblies,[379] providing high confidence in each nucleotide. The organization of the genomes could not be assessed due to the fragmented nature of the assemblies. However, it is likely that these genomes possess the typical organization of the genus, where the central region is conserved while the arms are variable.

The rare Actinobacteria *Streptomonospora* sp. PA3 genome sequencing resulted in 7,003,312 reads. The reads were error corrected and trimmed to 6,887,114 reads,

which were *de novo* assembled into 24 contigs with an N50 of 575,293 bp. The final genome assembly was 5.7 Mbp with a G + C content of 72.9% and a median coverage of 292x. The draft genome was assessed using BUSCO and resulted in 40 out of 40 (100%) single-copy orthologs being found. The genome was annotated and resulted in 5,079 CDSs.

The *Streptomonospora* genome is relatively small compared to the other eight publicly available genomes in NCBI and only one was smaller (*Streptomonospora salina*, 5.6 Mbp). This small size may be related to the isolation source of the strain. PA3 was isolated from a marine sponge and is a putative defensive symbiont. In general, symbionts have smaller genomes due to genome erosion, i.e., they progressively lose functional genes as they streamline their metabolism with their host.[380] Genome erosion is additionally characterized by a bias towards A + T content. This A + T content bias is not observed in PA3, however, the same was observed in *S. philanthi*, where its G + C content bias was maintained despite its genome being eroded.[381] PA3 showed an increased osmoregularity potential compared to the terrestrial model Actinobacteria *S. coelicolor* A3(2). PA3 has evolved to be a halophile, as reflected in its marine environment. PA3 appears to have an increased capacity for the uptake and synthesis of betaine-type[382,383] osmoregulatory compounds. In particular, four gene products for the uptake of choline were identified, which could be converted to trimethylglycine by choline dehydrogenase and betaine aldehyde dehydrogenase. Furthermore, PA3 may encode a second pathway, which is not present in *S. coelicolor* A3(2), to produce betaine through sequential methylation of glycine, by glycine *N*-methyltransferase and dimethylglycine *N*-methyltransferase. However, gene products for the utilization of L-ectoine[382,384] were similar to those of *S. coelicolor* A3(2).

These genomes represent high-quality additions to our collective scientific knowledge. All three of these draft genomes have an N50 greater than 400 kb, making it likely that even the largest BGCs (quinolidomicin, >200 kb) can be found.[385] Moreover, the genome of the rare Actinobacteria *Streptomonospora* sp. PA3 is an essential addition to understanding this rare genus, since only a total of nine species have been sequenced to date.

4.3 Genome Mining to Discover Aromatic Polyketides

Genome mining was performed on four strains included in this work. The predicted putative BGCs are presented in **Table 2**. These predictions aided to link an observed NP to a BGC. Once this link was established, further characterization of the BGC became possible.

Table 2. Putative BGCs of strains from this study. The BGC names represent the most similar known cluster. Putative BGCs with no similar cluster are named after the predicted type. BGCs marked in red had their product detected.

***S. LAVENDULAE* YAKB-15**	***S. PEUCETIUS* ATCC 27952**	***STREPTOMYCES* SP. MBK6**	***STREPTOMONOSPORA* SP. PA3**
geosmin	unknown terpene	cyclomarin	arginomycin
friulimicins	geosmin	hopene	isorenieratene
kirromycin	sch-47554/47555	grincamycin	cinerubin B
hopene	salinichelins	unknown terpene	unknown lanthipeptide
pentamycin	daptomycin	unknown bacteriocin	fosfomycin
SF2768	ectoine	heronamide	unknown terpene
desferrioxamine b	streptoseomycin	unknown siderophore	fuscachelin
melanin	unknown terpene	albaflavenone	ebelactone
unknown RRE-containing	salinosporamide A	undecylprodigiosin	ectoine
albusnodin	unknown lanthipeptide	ectoine	
informatipeptin	unknown lasso peptide	paromomycin	
tetronasin	zorbamycin	melanin	
leinamycin	cytorhodin	daptomycin	
spore pigment	atratumycin	desferrioxamine B	
kirromycin	ficellomycin	informatipeptin	
dynemicin A	simocyclinone D8	asukamycin	
mirubactin	pentamycin	himastatin	
streptothricin	melanin	griselimycin	
foxicins	sapB	coelibactin	
SF2575	hopene	chalcomycin	
unknown thioamitides	stenothricin	brasilicardin A	
unknown siderophore		unknown terpene	
unknown lasso peptide		spore pigment	
SF2575		carotenoid	
stenothricin		zorbamycin	
unknown RiPP-like		pheganomycin	
herboxidiene			
cinnamycin			
A-201A			
ectoine			
julichrome Q3-3			

4.3.1 Genome Mining from Soil Isolates

As discussed previously, the production of prodigiosins in *S. coelicolor* is regulated by nutrient starvation with GlcNAc as an elicitor, and prodigiosins are used as apoptotic agents.[49,53] In Article II, we found, *S. coelicolor* did not sequester yeast cells, however, in MBK6, yeast cells were sequestered and prodigiosin production was elicited by yeast cells, which advised that prodigiosins were utilized as antifungals. In fact, prodigiosins have been shown to have antifungal activity.[386]

Streptomyces sp. MBK6 was isolated from the soil in Turku, Finland. When cultivated with dead yeast, the characteristic sequestering of yeast cells was observed, as well as production of a red pigment (**Figure 16**). Chemical analysis of culture extracts revealed the red pigment was two NPs, metacycloprodigiosin and undecylprodigiosin (**Figure 17**). I then sequenced, assembled, annotated, and genome mined MBK6 and found a prodigiosin BGC.

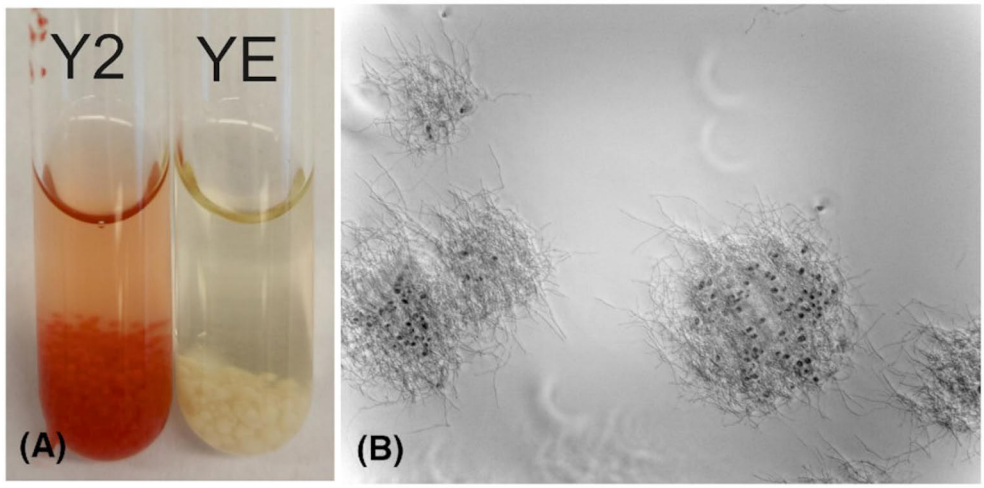


Figure 16. *Streptomyces* sp. MBK6-yeast interactions. (A) MBK6 cultures with yeast (Y2) and without yeast (YE) after four days. The red color indicated prodigiosins were produced. (B) MBK6 sequestered nearly all yeast cells after three days. Phase contrast microscopy, 100x magnification. The black dots are yeast cells. Source: Article III.

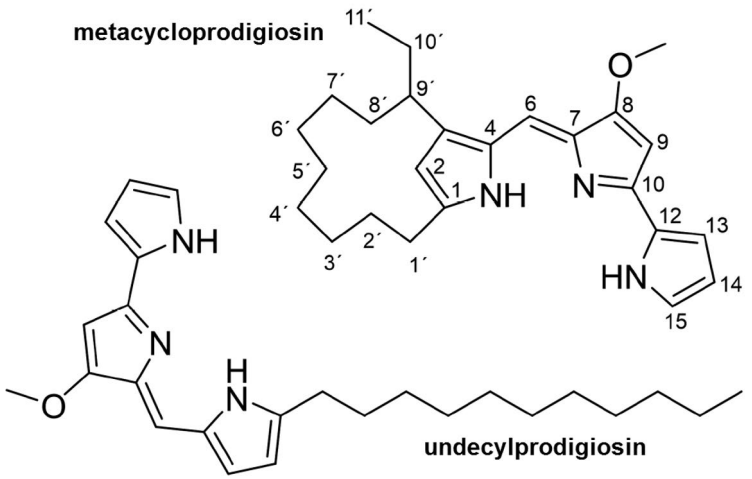


Figure 17. Compounds produced by MBK6 when cultured with yeast cells. Modified from: Article III.

We found that the *mbk* prodigiosin BGC was highly similar to the *red* BGC of *S. coelicolor* (**Figure 18A**).[387] However, dissimilarity was found between the *redZ* and *mbkZ* master regulator genes. Therefore, we characterized the effect of *mbkZ* on prodigiosin production in MBK and *S. coelicolor* in the presence and absence of yeast cells, and on the binding of MbkZ to the promoters of the BGC-specific regulatory genes. Interestingly, the N-terminal area of redZ was not conserved.[388] The difference in the nucleotide sequence with mbkZ suggested a different regulation of cognate prodigiosin BGC between the strains. On the amino acid level (**Figure 18**), the receiver domain showed weak similarity, while the highest similarity was within the DNA-binding domain.

To test these differences, we transformed *mbkZ* with its own promoter (*mbkZp*) and/or a constitutive promoter (*ermEp*) into *S. coelicolor*. The constructs were also transformed into MBK6, however, only the one containing *ermEp-mbkZp* was successful. Subsequently, each transformed strain, along with the wildtype strains, was cultured in parallel in media with whole yeast cells or yeast extract. The resulting levels of prodigiosin production were then assayed.

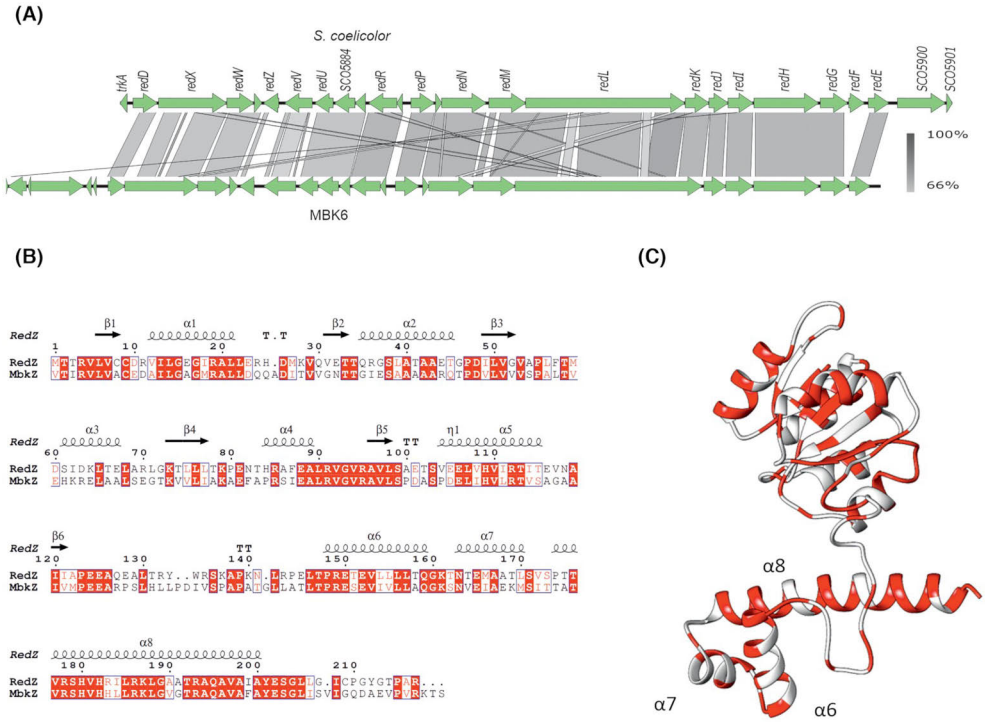


Figure 18. Comparison between MBK6 and *S. coelicolor* prodigiosin clusters and key regulator genes and proteins. (A) Comparison of the MBK6 prodigiosin cluster and the *S. coelicolor red* cluster. (B) Alignment of the key regulator protein sequences MbkZ and RedZ. Secondary structures are derived from the model in C. (C) Secondary homology structure of RedZ, using 4LDZ (Bacillus subtilis response regulator DesR) as a template. Amino acids conserved with MbkZ shown in red. Source: Article III.

The results showed that both wildtype strains produced considerably more prodigiosins in the medium containing yeast cell than in the medium containing yeast extract. In *S. coelicolor*, the construct with *mbkZp* stimulated production in both media, while production was repressed with the constructs containing *ermEp* and *ermEp-mbkZp* in both media. In MBK6, the construct with *ermEp-mbkZp* stimulated production in both media.

On the basis of the effects of *mbkZ* on prodigiosin production, we then investigated the binding of MbkZ to the promoter regions of *redD* and *mbkD*, i.e., *Streptomyces* antibiotic regulatory proteins (SARP). The results showed that MbkZ could bind to both promoter regions. This indicated that the cluster-situated regulatory network is likely similar in MBK6 and *S. coelicolor*. MbkZ binds to the promoter region of mbkD, the same as RedZ binds to the promoter region of *redD*, in turn activating the BGC. However, this also indicated that the regulatory network at the global level is more complex. The repression of prodigiosin production in *S. coelicolor* by expression of the *ermEp-mbkZp* and *ermEp* constructs may have been

due to binding of MbkZ to the promoter region of redD, but MbkZ not having the ability to function with *S. coelicolor* transcription machinery, thereby preventing the activation of the BGC. The stimulation of prodigiosin production in *S. coelicolor* by the expression of the *mbkZp* construct may have been due to a gene dose effect, particularly the presence of multiple *mbkZp* sequences in the multi-copy number plasmid.[389] These *trans* acting elements may bind to the natural repressors of the prodigiosin BGC, which would allow RedZ to be expressed and then activate *redD* transcription.

Together, these results showed that the prodigiosin BGCs of *S. coelicolor* and MBK6 were regulated differently at the global level and suggested that the concomitant prodigiosins have different ecological roles (**Figure 19**).

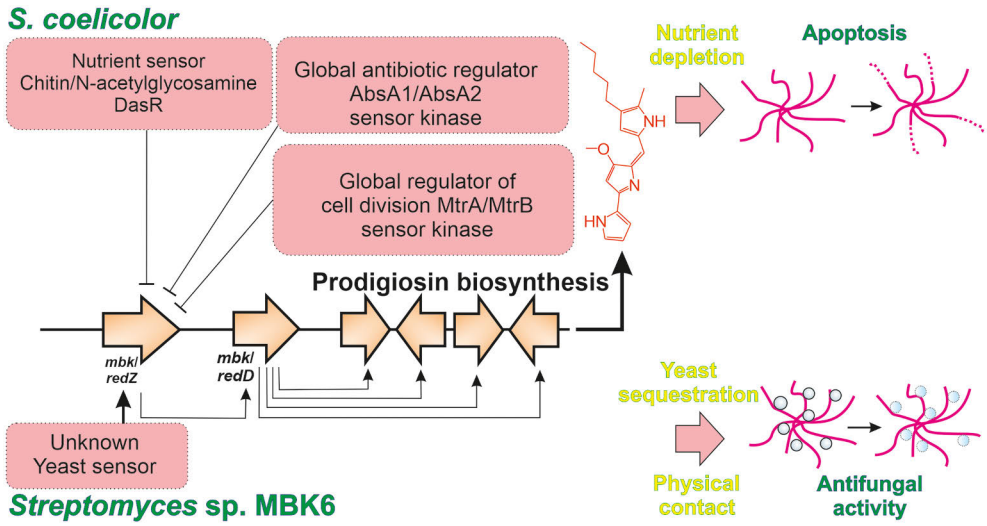


Figure 19. Putative differential regulation of prodigiosin biosynthesis in MBK6 and *S. coelicolor*. Source: Article III.

4.3.2 Genome Mining from Rare Actinobacteria

One approach to discover novel biologically active NPs is to focus on rare microbial genera.[390] Studies have shown that sponge-symbiotic Actinobacteria may represent an unexploited resource for novel biologically active NPs.[391] Therefore, in Article IV, we bioprospected the Persian Gulf for sponge-symbiotic Actinobacteria. Forty-five strains were isolated from 12 sponge species. These isolates were tested for their antimicrobial potency against seven bacterial pathogens (*Staphylococcus aureus*, *Bacillus cereus*, *Pseudomonas aeruginosa*, *Proteus vulgaris*, *Salmonella enterica*, *Klebsiella pneumonia*, and *E. coli*). Twenty-two isolates showed positive bioactivity. The isolates were then identified by their 16S rRNA genes. Most of the isolates were

identified as *Nocardiopsis* and *Streptomyces*, while the rest were identified as rare actinomycetes, i.e., *Saccharopolyspora*, *Amycolatopsis*, *Streptomonospora* and *Actinomadura*.

Next, we focused on aromatic polyketides since these have only been isolated from a narrow range of species. Most of these type II polyketides, for example, tetracycline and doxorubicin,[299] have been isolated from terrestrial *Streptomyces* species. Even within these species, the production of type II polyketides is rather rare, with an average of less than two BGCs per species.[392] In contrast, type I polyketides, e.g., erythromycin, and nonribosomal peptides, e.g., vancomycin and penicillin, can be widely found from diverse microbial phyla, e.g., Actinobacteria, Cyanobacteria, and Firmicutes.[393,394]

The genomic potential of these strains to produce aromatic polyketides was probed by searching for the presence of KSα genes. KSα was chosen because it is essential for the biosynthesis of aromatic polyketides, together with KSβ.[392] Twenty-three isolates were found to contain a KSα gene. The KSα gene sequences were compared to experimentally characterized sequences from the MIBiG database. Most sequences clustered with sequences responsible for spore pigment biosynthesis, which showed that KSα genes involved in antibiotic biosynthesis were scarce (**Figure 20**). Further analyses revealed that only five sequences were involved in the biosynthesis of aromatic polyketide. Three were from *Streptomyces* sp. (SM9, SM14, and PA13), one was from *Nocardiopsis* sp. SM36, and one from *Streptomonospora* sp. PA3 (**Figure 20**).

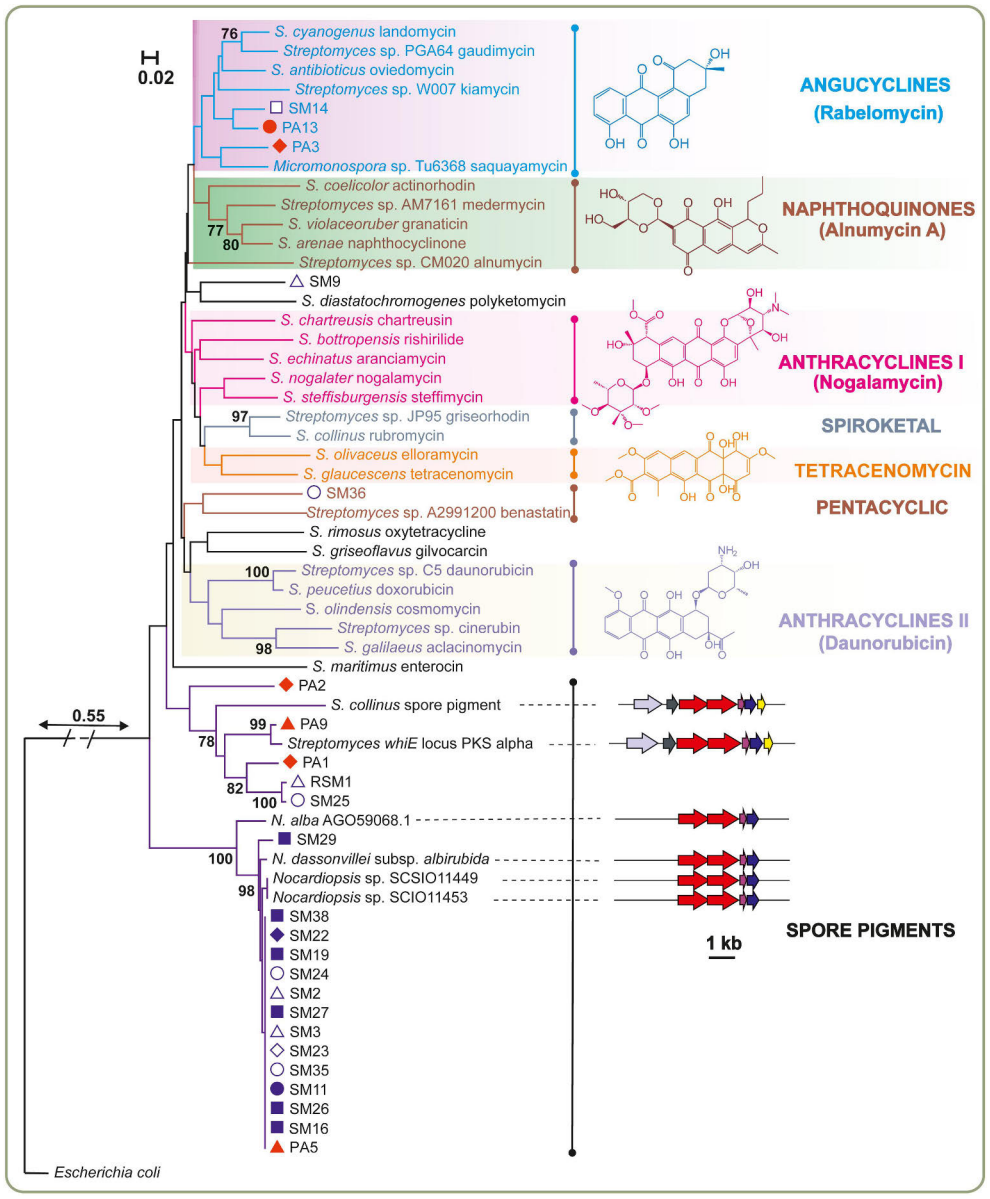


Figure 20. Potential of sponge-associated Actinobacteria to produce aromatic polyketides by genotyping. KSα domain amino acid sequences were used to construct the phylogenetic tree. Next to the strain names are the identified or predicted compound names and the representative chemical structures. The red, white, and blue symbols adjacent to the isolated strains correspond to its sponge host (Article IV, Figure 1). Source: Article IV.

PA3 was chosen for further study because, to our knowledge, no polyketides have been isolated from this genus. Furthermore, only two NPs have been isolated from *Streptomonospora*, streptomonomicin lasso peptides[390] and marinopyrone α-pyrones[395]. We cultured the strain on a large scale and two NPs were isolated. The first compound was identified as 1-hydroxy-4-methoxy-2-naphthoic acid, the second compound was identified as a novel aromatic polyketide and dubbed persiamycin A (**Figure 21A**). The bioactivity of persiamycin A was then tested, and we found it was able to inhibit both Gram-positive and Gram-negative bacteria. It was also cytotoxic to breast carcinoma tumor cell lines.

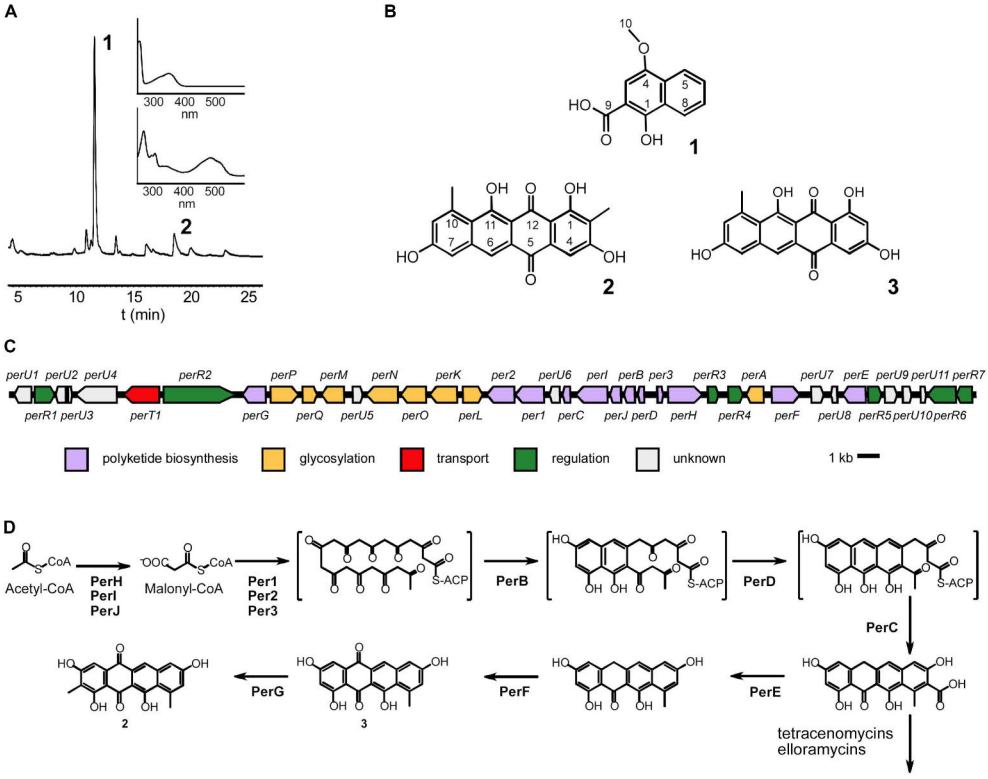


Figure 21. *Streptomonospora* sp. PA3 metabolites and persiamycin A biosynthesis. (A) HPLC trace, shown at 256 nm, and UV/VIS spectra of the purified compounds (**1**) 1-hydroxy-4-methoxy-2-naphthoic acid and (**2**) persiamycin A. (B) Chemical structures of (**1**), (**2**), and (**3**) tetracenomycin D, a related compound. (C) Schematic representation of the persiamycin biosynthetic gene cluster. (D) Putative biosynthetic steps for the production of (**2**), with apparent divergence of the tetracenomycin/elloramycins pathways after the fourth ring cyclization. Source: Article IV.

We then wanted to identify the cognate BGC. I sequenced the genome of PA3, followed by assembly, annotation, and BGC prediction. The genome was revealed

to contain only seven BGCs, which was consistent with its relatively small genome size. The relatively small genome size could also be related to the symbiotic lifestyle of PA3, as discussed previously. Only one BGC was predicted to produce an aromatic polyketide (**Figure 21C**), which showed the greatest similarity to the characterized cinerubin BGC. Further analysis demonstrated that a set of aromatases/cyclases, *perABCD*, could be homologous to those of the tetracenomycin biosynthetic pathway (**Figure 21D**). In addition, the structure of persiamycin A was similar to that of tetracenomycin D isolated from *S. glaucescens* (**Figure 21B**).[396] However, the biosynthesis of persiamycin appears to have diverged from the tetracenomycin pathway during the tailoring steps. PerE could catalyze the decarboxylation, PerF is likely a mono-oxygenase, and PerG could catalyze C-methylation (**Figure 21D**). Surprisingly, the PA3 KSα gene clustered with the angucycline clade instead of grouping with the tetracenomycin clade (**Figure 20**). This may reflect the divergence of the *Streptomonospora* genus from the reference *Streptomyces* strains.

Our initial aim was successful. We isolated and characterized a novel rare species of Actinobacteria, *Streptomonospora* sp. PA3, which provided us with a novel aromatic polyketide. Interestingly, the most surprising finding was that the tetracenomycin-like KSα gene of PA3 was more closely related to the angucycline KSα genes (**Figure 20**). As the PA3 KSα gene is the first described from the genus *Streptomonospora*, it is currently not possible to determine whether the unexpected sequence similarity was the result of the divergence of the *Streptomonospora* genus or whether our observation was an outlier. This could represent an exciting opportunity to bioprospect for additional *Streptomonospora* species with aromatic polyketides to unravel this mystery.

4.3.3 Genome Mining from Sequence Databases

Anthracyclines are clinically relevant anticancer agents that are synthesized by type II polyketide pathways. In general, the various genes in these pathways are surprisingly similar. It would appear that a limited set of genes is used for the generation of the great chemical diversity of anthracyclines.[397] The key to this chemical diversity lies in the tailoring steps. The varied modifications are the result of the tailoring biosynthetic enzymes having evolved different substrate specificities and catalytic properties. Unlike enzymes found in primary metabolism, these secondary metabolism enzymes are non-essential and therefore are not restricted to conserve their functions.[320] As a result, secondary metabolism proteins are able to rapidly acquire novel functionalities even without a gene duplication event. This has been demonstrated in anthracycline biosynthesis through homologous polyketide

cyclases acting as mono-oxygenases[398] and *vice versa*[399], and in the conversion of methyltransferases to hydroxylases[400].

In Article V, we mined sequenced *Streptomyces* genomes for the presence of cryptic anthracycline gene clusters. We particularly focused on genes responsible for the tailoring steps and traced the evolution of anthracycline methyltransferase-like proteins. The concomitant identified BGCs were narrowed down to 12 by only selecting BGCs that contained homologs for *rdmB*, 10-hydroxylase, and *rdmC*, aclacinomycin 15-methylesterase (**Figure 22A**). Until now, only two protein subfamilies have been reported, the only 10-hydroxylation subfamily[401,402] and the bifunctional 4-O-methylation/10-decarboxylation subfamily[403]. Genome mining allowed us to discover two new clades of methyltransferase proteins and we wished to investigate if these subfamilies could catalyze novel chemistry.

We then experimentally probed the activities of the newly identified members of the methyltransferase-like family. ZamB, EamK, TamK, and CalMB originating from *S. zinciresistens* K2,[404] *S. erythrochromogenes* NRRL B-2112,[405] *S. tsukubaensis* NRRL 18488,[406] and *Streptomyces* sp. CcalMP8W,[407] respectively, were selected. DnrK from *S. peucetius* and RdmB from *S. purpurascens* were also included. The results of the enzymatic assays showed that only DnrK catalyzed the canonical reaction of the protein family, 4-O-methylation. RdmB, ZamB, and CalMB had 10-hydroxylation activity. Surprisingly, TamK and EamK had 10-decarboxylation activity (**Figure 22B**).

Analysis of culture extracts led to the detection of komodoquinone B from *S. erythrochromogenes* NRRL B-2112 and *Streptomyces* sp. NRRL S-378 cultures, which confirmed the biological relevance of the 10-decarboxylation activity. This led to the identification of two BGCs responsible for the production of komodoquinones.

Another interesting aspect was discovered through comparative phylogenetic analysis of both the methyltransferases and the BGCs. The resulting phylogenetic trees mirrored each other exceptionally well, and four distinct clades were revealed (**Figure 22B, C**). As expected, one clade was represented by the bifunctional DnrK 4-O-methylation/10-decarboxylation subtype and another by the RdmB 10-hydroxylation subtype. The clade containing CalMB indicated a second branch of 10-hydroxylating enzymes, while the 10-decarboxylating enzymes consisting EamK/TamK provided another unique clade. Together, these results indicated that these so-called methyltransferases have evolved *in situ* in their respective BGCs. Furthermore, these results excluded the possibility for HGT, since the evolution of these methyltransferases appears to stringently follow the evolution of their BGCs. Interestingly, only a minority of the members of the methyltransferase-like family are, in effect, methyltransferases. It also appears that 10-hydroxylation is the major activity of this family. Additionally, 10-hydroxylation activity may provide some

evolutionary advantage, as the phylogenetic analysis pointed towards two independent routes for this activity (**Figure 22B**).

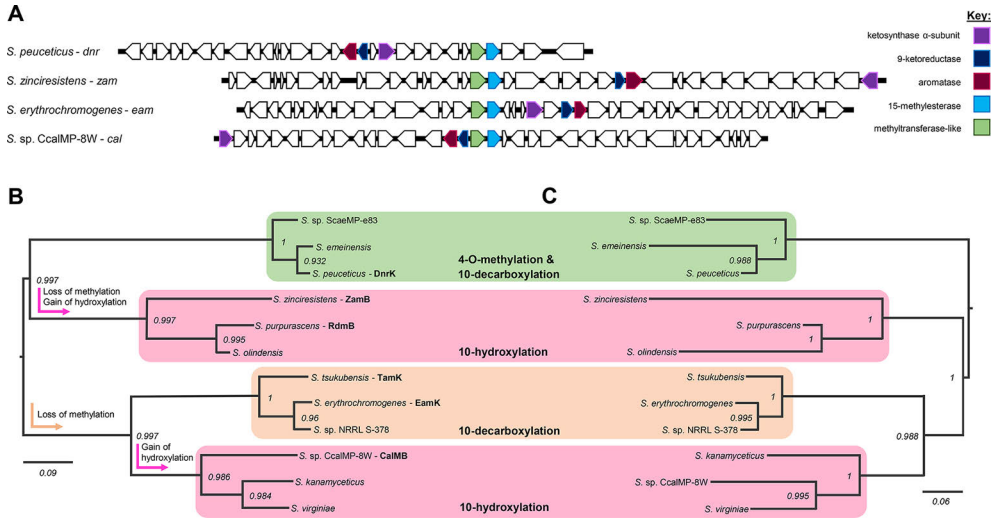


Figure 22. Evolution of anthracycline gene clusters and their methyltransferases. (A) The organization of anthracycline gene clusters. (B) Phylogenetic tree of anthracycline methyltransferases. (C) Phylogenetic tree of anthracycline gene clusters, represented by concatenated protein sequences involved in the formation of the core anthracycline carbon frame. Source: Article V.

5 Conclusion and Future Perspective

AMR is an imminent threat that may kill 10 million people per year and cost $100 trillion by 2050.[1,3] Actinobacteria have provided us with most of our clinically used antimicrobials for almost 100 years, with their ability to synthesize biologically active NPs.[18] In fact, Actinobacteria may have even provided ancient people with ethnomedicines as far back as 2 BC.[408] However, we have exhausted our ability to discover new NPs with our current methods, but the abundance of diverse BGCs in sequence databases indicates that the chemical space NPs occupy remains largely untouched.

Here, we used the framework of ecology to elicit the production of NPs. We showed that *Streptomyces* should be classified as an opportunistic predator and this predatory behavior could be exploited to elicit the production of antifungal NPs and enzymes. Additionally, our model of *Streptomyces* predation provides a possible explanation for how *Streptomyces* is able to prey on yeast at the molecular level (**Figure 23**). From this predatory classification we also provided evidence that prodigiosin NPs have a different ecological function from those previously reported and gained insight into their differential regulation.

Beyond finding a new way of eliciting the production of NPs and enzymes, we used genome mining as a tool to enhance our drug discovery efforts, leading to the discovery of a new NP and two new BGCs. In combination with genome mining, we bioprospected an underrepresented ecological niche and discovered a new species of rare Actinobacteria, *Streptomonospora*, which produces the novel antibacterial NP persiamycin. We also discovered an unusual phylogenetic grouping of a core gene from the persiamycin gene cluster, which suggested evolutionary divergence from the related *Streptomyces* genes. Finally, we showed that genome mining and phylogenetic analysis of BGCs tailoring enzymes could be used as a guide for NP discovery efforts, which resulted in the discovery of two new BGCs. Further phylogenetic analysis of the respective BGCs revealed the tailoring genes evolved into three subfamilies *in situ* in their respective BGCs.

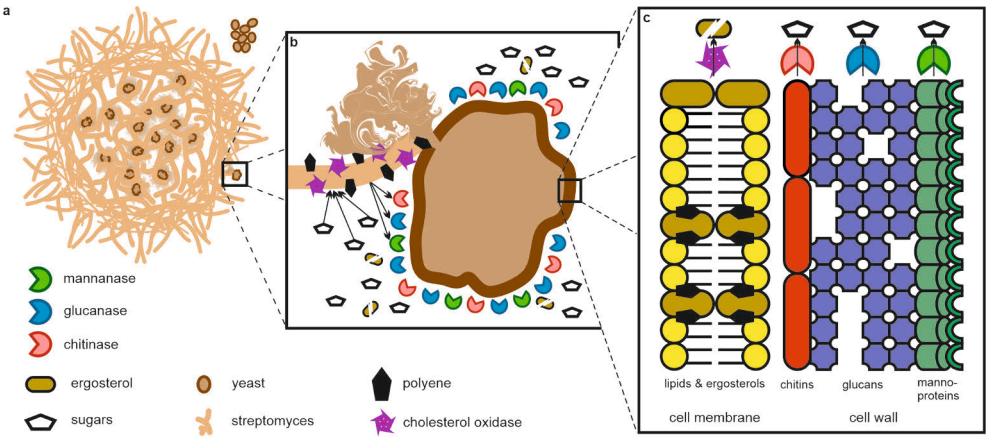


Figure 23. Multipronged attack on yeast by *Streptomyces* during predation. (a) *Streptomyces* sequesters yeast, by making physical contact and the yeast subsequently disappear; the untouched yeast remains intact. (b) Cell-associated cholesterol oxidase and polyenes are putative weapons that kill yeast, while secreted enzymes are putatively responsible for digestion. (c) More specifically, cholesterol oxidase and polyenes attack the cell membrane by disrupting its integrity and forming pores. The enzymes attack and digest the components of the yeast cell wall. Source: Article II.

The highlight of my thesis is uncovering *Streptomyces* predatory abilities. I believe this discovery has major implications for the future of drug discovery, which I would like to speculate on. As Einstein said: "The important thing is not to stop questioning. Curiosity has its own reason for existing. One cannot help but be in awe when one contemplates the mysteries of eternity… Never lose a holy curiosity."

Could the predatory behavior of *Streptomyces* be exploited to elicit the production of novel NPs and enzymes and reinvigorate the *Streptomyces* drug discovery pipeline? If so these novel NPs and enzymes could be developed into drugs to fight AMR. Currently, enzymes aren't commonly used as antimicrobial drugs. However, excitingly, enzymes have also been shown to have antimicrobial properties and are emerging as a novel strategy to fight microbes.[197,409] In fact, antimicrobial enzymes are the core of the topical veterinary drugs ZYMOX® and Oratene® and are touted to not harm the normal bacterial flora of the animal or contribute to AMR, as bacteria cannot easily acquire resistance to them.

Even more important is genome mining, which has proven itself as an invaluable method that aids in drug discovery. Fundamentally, all NPs start from the genome and from the genome we have gained insight into the biosynthetic logic. When combined with other methods such as phylogenetic analysis it becomes possible to not only predict what clusters are present in a given genome but to predict how different a NP a BGC can produce. Although, how to elicit a given NP remains elusive. Currently, we are still revitalizing the Golden Age of antibiotic discovery since we have yet to find a broadly applicable and bountiful method as the Waksman

Platform. Could bioprospecting be the missing link to regain a method as prolific as the Waksman Platform? They are both based on isolating new microbial species but differ in that bioprospecting is organized and systematic of the isolation source. Combining bioprospecting and genome mining could identify areas such as under explored ecosystems or specific ecological niches as places that are enriched with new BGCs and ultimately new NPs. Another possibility lies in synthetic biology, where, in the future, we could rationally engineer bespoke (un)natural products based on the knowledge we gained from their biosynthetic logic. Will one of these fields of research be the missing link or a field still unknown? Regardless, I believe genome mining is the penultimate step in reviving the Golden Age of antibiotic discovery.

Summa summarum, "I would rather have questions that can't be answered than answers that can't be questioned" (Richard Feynman).

Acta est fabula plaudite

Acknowledgements

The experimental work in my thesis was conducted in the Antibiotic Biosynthesis Engineering (ABE) Laboratory at the Department of Life Technologies, University of Turku, Finland, from 2017 to 2023. I thank the Finnish Functional Genomics Centre, Cell Imaging and Cytometry Core, and Turku Proteomics Facility of the Turku Bioscience Centre (with the support of Biocenter Finland) for additional experimental support. Computational resources were graciously provided by CSC – IT Center for Science, Finland. The personal financial support of the Turku University Foundation and the Finnish Cultural Foundation is gratefully acknowledged. Additional financial support was provided by the Academy of Finland, the Novo Nordisk Foundation, and the Sigrid Jusélius Foundation. I greatly appreciate the opportunity to attend so many international conferences and public outreach events, which were made possible by the generous travel grants of the University of Turku Graduate School and the Computational and Molecular Methodologies for Life Sciences research program.

I wish to show my appreciation to Dr. Miia Mäkelä and Dr. Gerald Lackner for their patience, understanding, and constructive comments during the pre-examination of my thesis. I am honored and humbled to have Prof. Istvan Molnar as my opponent.

The kind words and support of my advisors, Dr. Martti Tolvanen and Dr. Pekka Rapu, are greatly appreciated. I am grateful to Prof. Jyrki Heino for leading our department and creating a productive research environment.

I would like to express my deepest gratitude to Prof. Mikko Metsä-Ketelä for his superb supervision, advise, and patience. He provided me with the opportunity to learn many new fields of research and the freedom to become an independent scientist, for which he has my everlasting gratitude, and I am honored to be a part of his upcoming projects.

I am grateful to Dr. Tatiana Buldakova for initiating the collaboration, which brought me into the ABE Lab, and her continued support and kindness. I appreciate the assistance of Dr. Petja Rosenqvist with analytical chemistry in the Department of Chemistry. I am thankful to Dr. Soheila Matroodi for reinvigorating one of my favorite projects. I would also like to thank all of my coauthors and collaborators.

I would like to thank my fellow ABE members, both former and present, for sharing their knowledge, fun coffee breaks, and lunch company. I would like to give special thanks to Kaisa Palmu for her patience and guidance, both in and out of the lab. I sincerely thank Dr. Jarmo Niemi for his vast knowledge and mind like a steel trap. I appreciate the hard work of Mitchell Laughlin and Arina Koroleva in our many shared projects and Heli Tirkkonen for helping me with analytical chemistry.

I would like to acknowledge Anu Hirvensalo, Heli Kuusela, Hannele Heinonen, Petra Sirviö, and Dr. Mika Keränen for their technical support. I am profoundly appreciative of Janne Mäkinen, Dr. Vilja Siitonen, and Dr. Erika Artukka for their vital efforts to keep the labs in order. A special thanks is given to Dr. Sema Sirin for encouraging me while I was writing my thesis.

I give my warmest thanks to my friends, especially the Einola family, Dr. Janne Atosuo, Ville Kyllönen, and Daniel Lövdahl, for helping me realize a life outside of the university and making me feel like a part of Finnish society. I am deeply indebted to my undergraduate mentor, Prof. Deborah Kimbrell (University of California, Davis), who always encouraged me to pursue a doctoral degree and provided me with so many opportunities. This endeavor would not have been possible without Dr. Amir Akhgari, who has been a remarkable colleague, a memorable friend, and a non-biological brother.

Last but not least, I am incredibly grateful to my family for their unconditional support and love. Words fail to express how much I love and appreciate my parents, Joyce and Rick, my sister and her family, Stephanie, Chio, Lily and Kai, and my grandparents, Pat and Tom, Dorothy and Yas, for everything they have provided me. I truly would not have been able to achieve this much without them.

Turku, Friday the sixteenth of June, 2023

Keith Yamada

List of References

1. Murray, C. J. *et al.* Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. *Lancet* **399**, 629–655 (2022).
2. O'neill, J. Tackling drug-resistant infections globally: final report and recommendations. (2016).
3. Jonas, O. B., Irwin, A., Berthe, F. C. J., Le Gall, F. G. & Marquez, P. V. Drug-resistant infections : a threat to our economic future (Vol. 2) : final report (English). (2017).
4. Reardon, S. WHO warns against 'post-antibiotic' era. *Nature* (2014). doi:10.1038/nature.2014.15135
5. Kwon, J. H. & Powderly, W. G. The post-antibiotic era is here. *Science (80-.).* **373**, 471–471 (2021).
6. Mendelson, M., Sharland, M. & Mpundu, M. Antibiotic resistance: calling time on the 'silent pandemic'. *JAC-Antimicrobial Resist.* **4**, (2022).
7. Newman, D. J. & Cragg, G. M. Natural products as sources of new drugs over the nearly four decades from 01/1981 to 09/2019. *J. Nat. Prod.* **83**, 770–803 (2020).
8. Davies, J. & Davies, D. Origins and evolution of antibiotic resistance. *Microbiol. Mol. Biol. Rev.* **74**, 417–433 (2010).
9. Ventola, C. L. The antibiotic resistance crisis: part 1: causes and threats. *P T* **40**, 277–83 (2015).
10. Aminov, R. I. A brief history of the antibiotic era: lessons learned and challenges for the future. *Front. Microbiol.* **1**, (2010).
11. Ehrlich, P. Experimental researches on specific therapy. in *The Collected Papers of Paul Ehrlich* 106–117 (Elsevier, 1960). doi:10.1016/B978-0-08-009056-6.50015-4
12. Jones, D. S., Podolsky, S. H. & Greene, J. A. The burden of disease and the changing task of medicine. *N. Engl. J. Med.* **366**, 2333–2338 (2012).
13. Fleming, A. On the antibacterial action of cultures of a *Penicillium*, with special reference to their use in the isolation of *B. influenzae*. *Br. J. Exp. Pathol.* **10**, 226–236 (1929).
14. Dosani, S. Penicillin Man: Alexander Fleming and the antibiotic revolution. *BMJ* **330**, 50.2 (2005).
15. Fleming, A. Streptococcal meningitis treated with penicillin. *Lancet* **242**, 434–438 (1943).
16. Domagk, G. A contribution to the chemotherapy of bacterial infections. *Clin. Infect. Dis.* **8**, 163–166 (1986).
17. Wainwright, M. & Kristiansen, J. E. On the 75th anniversary of Prontosil. *Dye. Pigment.* **88**, 231–234 (2011).
18. Waksman, S. A. Decomposition of the various chemical constituents etc. of complex plant materials by pure cultures of fungi and bacteria. *Arch. Mikrobiol.* **2**, 136–154 (1931).
19. Schatz, A., Bugle, E. & Waksman, S. A. Streptomycin, a substance exhibiting antibiotic activity against Gram-positive and Gram-negative bacteria.*. *Exp. Biol. Med.* **55**, 66–69 (1944).
20. Lyddiard, D., Jones, G. L. & Greatrex, B. W. Keeping it simple: lessons from the golden era of antibiotic discovery. *FEMS Microbiol. Lett.* **363**, fnw084 (2016).
21. WHO. The top 10 causes of death. (2020). Available at: https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death. (Accessed: 25th October 2022)

22. Barber, M. & Rozwadowska-Dowzenko, M. Infection by penicillin-resistant staphylococci. *Lancet* **252**, 641–644 (1948).
23. Iskandar, K. *et al.* Antibiotic discovery and resistance: The chase and the race. *Antibiotics* **11**, 182 (2022).
24. Watanabe, Y., Martini, J. E. J. & Ohmoto, H. Geochemical evidence for terrestrial ecosystems 2.6 billion years ago. *Nature* **408**, 574–578 (2000).
25. Atanasov, A. G., Zotchev, S. B., Dirsch, V. M. & Supuran, C. T. Natural products in drug discovery: advances and opportunities. *Nat. Rev. Drug Discov.* **20**, 200–216 (2021).
26. Chen, Y., Garcia de Lomana, M., Friedrich, N.-O. & Kirchmair, J. Characterization of the chemical space of known and readily obtainable natural products. *J. Chem. Inf. Model.* **58**, 1518–1532 (2018).
27. Bentley, S. D. *et al.* Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). *Nature* **417**, 141–147 (2002).
28. Behie, S. W., Bonet, B., Zacharia, V. M., McClung, D. J. & Traxler, M. F. Molecules to ecosystems: Actinomycete natural products *in situ*. *Front. Microbiol.* **7**, 1–11 (2017).
29. van Bergeijk, D. A., Terlouw, B. R., Medema, M. H. & van Wezel, G. P. Ecology and genomics of Actinobacteria: new concepts for natural product discovery. *Nat. Rev. Microbiol.* **18**, 546–558 (2020).
30. Scott, J. J. *et al.* Bacterial protection of beetle-fungus mutualism. *Science (80-.).* **322**, 63–63 (2008).
31. Book, A. J. *et al.* Cellulolytic *Streptomyces* strains associated with herbivorous insects share a phylogenetically linked capacity to degrade lignocellulose. *Appl. Environ. Microbiol.* **80**, 4692–4701 (2014).
32. Conn, V. M., Walker, A. R. & Franco, C. M. M. Endophytic Actinobacteria induce defense pathways in *Arabidopsis thaliana*. *Mol. Plant-Microbe Interact.* **21**, 208–18 (2008).
33. Currie, C. R., Poulsen, M., Mendenhall, J., Boomsma, J. J. & Billen, J. Coevolved crypts and exocrine glands support mutualistic bacteria in fungus-growing ants. *Science (80-.).* **311**, 81–83 (2006).
34. Johansen, K. S. Discovery and industrial applications of lytic polysaccharide mono-oxygenases. *Biochem. Soc. Trans.* **44**, 143–149 (2016).
35. Seipke, R. F., Kaltenpoth, M. & Hutchings, M. I. Streptomyces as symbionts: an emerging and widespread theme? *FEMS Microbiol. Rev.* **36**, 862–876 (2012).
36. Pérez, J., Moraleda-Muñoz, A., Marcos-Torres, F. J. & Muñoz-Dorado, J. Bacterial predation: 75 years and counting! *Environ. Microbiol.* **18**, 766–779 (2016).
37. Pérez, J., Contreras-Moreno, F. J., Marcos-Torres, F. J., Moraleda-Muñoz, A. & Muñoz-Dorado, J. The antibiotic crisis: How bacterial predators can help. *Comput. Struct. Biotechnol. J.* **18**, 2547–2555 (2020).
38. Valdezate, S. *Nocardia*, *Rhodococcus*, *Streptomyces* and other aerobic actinomycetes. in *Encyclopedia of Infection and Immunity* 589–613 (Elsevier, 2022). doi:10.1016/B978-0-12-818731-9.00099-9
39. Herbrík, A. *et al.* A human lung-associated *Streptomyces* sp. TR1341 produces various secondary metabolites responsible for virulence, cytotoxicity and modulation of immune response. *Front. Microbiol.* **10**, (2020).
40. Sharma, A. *Streptomyces*. in *Encyclopedia of Food Microbiology* 2134–2138 (Elsevier, 1999). doi:10.1006/rwfm.1999.1545
41. Chater, K. F. & Hopwood, D. A. *Streptomyces*. in *Bacillus subtilis and Other Gram-Positive Bacteria* 83–99 (ASM Press, 2014). doi:10.1128/9781555818388.ch6
42. Barka, E. A. *et al.* Taxonomy, physiology, and natural products of Actinobacteria. *Microbiol. Mol. Biol. Rev.* **80**, 1–43 (2016).
43. Lewis, K. Platforms for antibiotic discovery. *Nat. Rev. Drug Discov.* **12**, 371–387 (2013).

44. Salama, S., Habib, M. H., Hatti-Kaul, R. & Gaber, Y. Reviewing a plethora of oxidative-type reactions catalyzed by whole cells of *Streptomyces* species. *RSC Adv.* **12**, 6974–7001 (2022).
45. Yagüe, P. *et al.* Transcriptomic analysis of *Streptomyces coelicolor* differentiation in solid sporulating cultures: First compartmentalized and second multinucleated mycelia have different and distinctive transcriptomes. *PLoS One* **8**, e60665 (2013).
46. van der Heul, H. U., Bilyk, B. L., McDowall, K. J., Seipke, R. F. & van Wezel, G. P. Regulation of antibiotic production in Actinobacteria: new perspectives from the post-genomic era. *Nat. Prod. Rep.* **35**, 575–604 (2018).
47. Flärdh, K. & Buttner, M. J. *Streptomyces* morphogenetics: dissecting differentiation in a filamentous bacterium. *Nat. Rev. Microbiol.* **7**, 36–49 (2009).
48. Świątek, M. A. *et al.* Functional analysis of the N-acetylglucosamine metabolic genes of *Streptomyces coelicolor* and role in control of development and antibiotic production. *J. Bacteriol.* **194**, 1136–44 (2012).
49. Rigali, S. *et al.* Feast or famine: The global regulator DasR links nutrient stress to antibiotic production by *Streptomyces*. *EMBO Rep.* **9**, 670–675 (2008).
50. Gomez-Escribano, J. P. *et al.* Structure and biosynthesis of the unusual polyketide alkaloid coelimycin P1, a metabolic product of the *cpk* gene cluster of *Streptomyces coelicolor* M145. *Chem. Sci.* **3**, 2716 (2012).
51. Claessen, D., Rozen, D. E., Kuipers, O. P., Søgaard-Andersen, L. & van Wezel, G. P. Bacterial solutions to multicellularity: a tale of biofilms, filaments and fruiting bodies. *Nat. Rev. Microbiol.* **12**, 115–124 (2014).
52. Manteca, A., Mäder, U., Connolly, B. A. & Sanchez, J. A proteomic analysis of *Streptomyces coelicolor* programmed cell death. *Proteomics* **6**, 6008–6022 (2006).
53. Tenconi, E., Traxler, M. F., Hoebreck, C., van Wezel, G. P. & Rigali, S. Production of prodiginines is part of a programmed cell death process in *Streptomyces coelicolor*. *Front. Microbiol.* **9**, (2018).
54. Zhang, Z. *et al.* Antibiotic production in *Streptomyces* is organized by a division of labor through terminal genomic differentiation. *Sci. Adv.* **6**, 1–9 (2020).
55. Zacharia, V. M. *et al.* Genetic network architecture and environmental cues drive spatial organization of phenotypic division of labor in *Streptomyces coelicolor*. *MBio* **12**, 1–18 (2021).
56. Traxler, M. F. & Rozen, D. E. Ecological drivers of division of labour in *Streptomyces*. *Curr. Opin. Microbiol.* **67**, 102148 (2022).
57. West, S. A. & Cooper, G. A. Division of labour in microorganisms: an evolutionary perspective. *Nat. Rev. Microbiol.* **14**, 716–723 (2016).
58. Giri, S., Waschina, S., Kaleta, C. & Kost, C. Defining division of labor in microbial communities. *J. Mol. Biol.* **431**, 4712–4731 (2019).
59. Cimermancic, P. *et al.* Insights into secondary metabolism from a global analysis of prokaryotic biosynthetic gene clusters. *Cell* **158**, 412–421 (2014).
60. Doroghazi, J. R. & Metcalf, W. W. Comparative genomics of actinomycetes with a focus on natural product biosynthetic genes. *BMC Genomics* **14**, 611 (2013).
61. Sriswasdi, S., Yang, C. & Iwasaki, W. Generalist species drive microbial dispersion and evolution. *Nat. Commun.* **8**, 1162 (2017).
62. Moran, N. A. Microbial Minimalism. *Cell* **108**, 583–586 (2002).
63. Khalil, M., Lerat, S., Beaudoin, N. & Beaulieu, C. The plant pathogenic bacterium *Streptomyces scabies* degrades the aromatic components of potato periderm via the β-ketoadipate pathway. *Front. Microbiol.* **10**, (2019).
64. Heine, D. *et al.* Chemical warfare between leafcutter ant symbionts and a co-evolved pathogen. *Nat. Commun.* **9**, 2208 (2018).
65. Becher, P. G. *et al.* Developmentally regulated volatiles geosmin and 2-methylisoborneol attract a soil arthropod to *Streptomyces* bacteria promoting spore dispersal. *Nat. Microbiol.* **5**, 821–829 (2020).

66. Zaroubi, L. *et al.* The ubiquitous soil terpene geosmin acts as a warning chemical. *Appl. Environ. Microbiol.* **88**, (2022).
67. Jones, S. E. *et al. Streptomyces* exploration is triggered by fungal interactions and volatile signals. *Elife* **6**, 1–21 (2017).
68. Jones, S. E. *et al. Streptomyces* volatile compounds influence exploration and microbial community dynamics by altering iron availability. *MBio* **10**, (2019).
69. Locey, K. J. & Lennon, J. T. Scaling laws predict global microbial diversity. *Proc. Natl. Acad. Sci.* **113**, 5970–5975 (2016).
70. Lennon, J. T. & Locey, K. J. More support for Earth's massive microbiome. *Biol. Direct* **15**, 5 (2020).
71. Microbiology by numbers. *Nat. Rev. Microbiol.* **9**, 628–628 (2011).
72. O'Malley, M. A. 'Everything is everywhere: but the environment selects': ubiquitous distribution and ecological determinism in microbial biogeography. *Stud. Hist. Philos. Sci. Part C Stud. Hist. Philos. Biol. Biomed. Sci.* **39**, 314–325 (2008).
73. Cockell, C. S. Are microorganisms everywhere they can be? *Environ. Microbiol.* **23**, 6355–6363 (2021).
74. Junkins, E. N., McWhirter, J. B., McCall, L.-I. & Stevenson, B. S. Environmental structure impacts microbial composition and secondary metabolism. *ISME Commun.* **2**, 15 (2022).
75. Fondi, M. *et al.* "Every gene is everywhere but the environment selects": Global geolocalization of gene sharing in environmental samples through network analysis. *Genome Biol. Evol.* **8**, 1388–1400 (2016).
76. Shaffer, J. P. *et al.* Standardized multi-omics of Earth's microbiomes reveals microbial and metabolite diversity. *Nat. Microbiol.* **7**, 2128–2150 (2022).
77. Lewin, G. R. *et al.* Evolution and ecology of Actinobacteria and their bioenergy applications. *Annu. Rev. Microbiol.* **70**, 235–254 (2016).
78. Madrova, P. *et al.* A short-term response of soil microbial communities to cadmium and organic substrate amendment in long-term contaminated soil by toxic elements. *Front. Microbiol.* **9**, 1–12 (2018).
79. Olajuyigbe, F. M. & Ehiosun, K. I. Assessment of crude oil degradation efficiency of newly isolated Actinobacteria reveals untapped bioremediation potentials. *Bioremediat. J.* **20**, 133–143 (2016).
80. Rathore, D. S., Sheikh, M. & Singh, S. P. Marine Actinobacteria: New horizons in bioremediation. in *Recent Developments in Microbial Technologies* 425–449 (Springer, Singapore, 2021). doi:10.1007/978-981-15-4439-2_20
81. Polti, M. A., Aparicio, J. D., Benimeli, C. S. & Amoroso, M. J. Simultaneous bioremediation of Cr(VI) and lindane in soil by Actinobacteria. *Int. Biodeterior. Biodegradation* **88**, 48–55 (2014).
82. Tan, L. T.-H., Chan, K.-G., Lee, L.-H. & Goh, B.-H. *Streptomyces* bacteria as potential probiotics in aquaculture. *Front. Microbiol.* **7**, (2016).
83. Cockell, C. S., Kelly, L. C. & Marteinsson, V. Actinobacteria —An ancient phylum active in volcanic rock weathering. *Geomicrobiol. J.* **30**, 706–720 (2013).
84. Anwar, S., Ali, B. & Sajid, I. Screening of rhizospheric actinomycetes for various *in-vitro* and *In-vivo* plant growth promoting (PGP) traits and for agroactive compounds. *Front. Microbiol.* **7**, 1–11 (2016).
85. Anilkumar, R. R., Edison, L. K. & Pradeep, N. S. Exploitation of fungi and Actinobacteria for sustainable agriculture. in *Microbial Biotechnology* (eds. Patra, J. K., Vishnuprasad, C. N. & Das, G.) **1**, 135–162 (Springer Singapore, 2017).
86. Xiong, Y.-W. *et al.* Enhancement of growth and salt tolerance of tomato seedlings by a natural halotolerant actinobacterium *Glutamicibacter halophytocola* KLBMP 5180 isolated from a coastal halophyte. *Plant Soil* **445**, 307–322 (2019).

87. Shimizu, M. Endophytic actinomycetes: Biocontrol agents and growth promoters. in *Bacteria in Agrobiology: Plant Growth Responses* 201–220 (Springer Berlin Heidelberg, 2011). doi:10.1007/978-3-642-20332-9_10
88. Law, J. W.-F. *et al.* The potential of *Streptomyces* as biocontrol agents against the rice blast fungus, *Magnaporthe oryzae* (*Pyricularia oryzae*). *Front. Microbiol.* **8**, (2017).
89. Zamoum, M. *et al.* Biocontrol capacities and plant growth-promoting traits of endophytic Actinobacteria isolated from native plants of Algerian Sahara. *J. Plant Dis. Prot.* **122**, 215–223 (2015).
90. Shivlata, L. & Satyanarayana, T. Thermophilic and alkaliphilic Actinobacteria: biology and potential applications. *Front. Microbiol.* **6**, (2015).
91. Jang, H.-D. & Chen, K.-S. Production and characterization of thermostable cellulases from metagenomic gene libraries. *World J. Microbiol. Biotechnol.* 263–268 (2003).
92. Yang, C.-H. & Liu, W.-H. Cloning and characterization of a maltotriose-producing α-amylase gene from *Thermobifida fusca*. *J. Ind. Microbiol. Biotechnol.* **34**, 325–330 (2007).
93. Tsuchiya, K., Sakashita, H., Nakamura, Y. & Kimura, T. Production of thermostable alkaline protease by alkalophilic *Thermoactinomyces* sp. HS682. *Agric. Biol. Chem.* **55**, 3125–3127 (1991).
94. Zhao, L. *et al.* Expression of the *Thermobifida fusca* xylanase Xyn11A in *Pichia pastoris* and its characterization. *BMC Biotechnol.* **15**, 18 (2015).
95. Vaijayanthi, G., Vijayakumar, R. & Dhanasekaran, D. Actinobacteria — A biofactory of novel enzymes. in *Actinobacteria - Basics and Biotechnological Applications* (InTech, 2016). doi:10.5772/61436
96. Suzuki, K., Yahara, S., Maehata, K. & Uyeda, M. Isoaurostatin, a novel topoisomerase inhibitor produced by *Thermomonospora alba*. *J. Nat. Prod.* **64**, 204–207 (2001).
97. Suzuki, K., Nagao, K., Monnai, Y., Yagi, A. & Uyeda, M. Topostatin, a novel inhibitor of topoisomerases I and II produced by *Thermomonospora alba* strain No. 1520. I. Taxonomy, fermentation, isolation and biological activities. *J. Antibiot. (Tokyo).* **51**, 991–998 (1998).
98. Wagg, C., Dudenhöffer, J., Widmer, F. & Heijden, M. G. A. Linking diversity, synchrony and stability in soil microbial communities. *Funct. Ecol.* **32**, 1280–1292 (2018).
99. Kinkel, L. L., Schlatter, D. C., Bakker, M. G. & Arenz, B. E. *Streptomyces* competition and co-evolution in relation to plant disease suppression. *Res. Microbiol.* **163**, 490–499 (2012).
100. Turroni, F., van Sinderen, D. & Ventura, M. Genomics and ecological overview of the genus *Bifidobacterium*. *Int. J. Food Microbiol.* **149**, 37–44 (2011).
101. Loria, R., Kers, J. & Joshi, M. Evolution of plant pathogenicity in *Streptomyces*. *Annu. Rev. Phytopathol.* **44**, 469–87 (2006).
102. Schrey, S. D. *et al.* Production of fungal and bacterial growth modulating secondary metabolites is widespread among mycorrhiza-associated streptomycetes. *BMC Microbiol.* **12**, 164 (2012).
103. Kaltenpoth, M. Actinobacteria as mutualists: general healthcare for insects? *Trends Microbiol.* **17**, 529–35 (2009).
104. Chanson, A., Moreau, C. S. & Duplais, C. Assessing biosynthetic gene cluster diversity of specialized metabolites in the conserved gut symbionts of herbivorous turtle ants. *Front. Microbiol.* **12**, (2021).
105. Hardoim, P. R. *et al.* The hidden world within plants: Ecological and evolutionary considerations for defining functioning of microbial endophytes. *Microbiol. Mol. Biol. Rev.* **79**, 293–320 (2015).
106. Castillo, U. F. *et al.* Munumbicins, wide-spectrum antibiotics produced by *Streptomyces* NRRL 30562, endophytic on *Kennedia nigriscans*. *Microbiology* **148**, 2675–2685 (2002).
107. Castillo, U. F. *et al.* Kakadumycins, novel antibiotics from *Streptomyces* sp. NRRL 30566, an endophyte of *Grevillea pteridifolia*. *FEMS Microbiol. Lett.* **224**, 183–190 (2003).
108. Ezra, D. *et al.* Coronamycins, peptide antibiotics produced by a verticillate *Streptomyces* sp. (MSU-2110) endophytic on *Monstera* sp. *Microbiology* **150**, 785–793 (2004).

109. Chen, P. *et al.* Community composition and metabolic potential of endophytic Actinobacteria from coastal salt marsh plants in Jiangsu, China. *Front. Microbiol.* **10**, (2019).
110. Donabedian, H. & Boyle, M. D. P. Clot formation by group A *Streptococci*. *Infect. Immun.* **66**, 2362–2364 (1998).
111. Huish, S., Thelwell, C. & Longstaff, C. Activity regulation by fibrinogen and fibrin of streptokinase from *Streptococcus pyogenes*. *PLoS One* **12**, e0170936 (2017).
112. Boukhatem, Z. F., Merabet, C. & Tsaki, H. Plant growth promoting Actinobacteria, the most promising candidates as bioinoculants? *Front. Agron.* **4**, (2022).
113. Zhang, Y. *et al. Streptomyces* application triggers reassembly and optimization of the rhizosphere microbiome of cucumber. *Diversity* **13**, 413 (2021).
114. Kumar, A. & Singh, J. Biofilms forming microbes: Diversity and potential application in plant–microbe interaction and plant growth. in 173–197 (2020). doi:10.1007/978-3-030-38453-1_6
115. Gong, Y. *et al.* Antifungal volatile organic compounds from *Streptomyces setonii* WY228 control black spot disease of sweet potato. *Appl. Environ. Microbiol.* **88**, (2022).
116. Lyu, A., Yang, L., Wu, M., Zhang, J. & Li, G. High efficacy of the volatile organic compounds of *Streptomyces yanglinensis* 3-10 in suppression of *Aspergillus* contamination on peanut kernels. *Front. Microbiol.* **11**, (2020).
117. Cordovez, V. *et al.* Diversity and functions of volatile organic compounds produced by *Streptomyces* from a disease-suppressive soil. *Front. Microbiol.* **6**, (2015).
118. Yamada, Y. *et al.* Terpene synthases are widely distributed in bacteria. *Proc. Natl. Acad. Sci.* **112**, 857–862 (2015).
119. Cuervo, L., Méndez, C., Salas, J. A., Olano, C. & Malmierca, M. G. Volatile compounds in actinomycete communities: A new tool for biosynthetic gene cluster activation, cooperative growth promotion, and drug discovery. *Cells* **11**, 3510 (2022).
120. Book, A. J. *et al.* Evolution of high cellulolytic activity in symbiotic *Streptomyces* through selection of expanded gene content and coordinated gene expression. *PLoS Biol.* **14**, e1002475 (2016).
121. Takasuka, T. E., Book, A. J., Lewin, G. R., Currie, C. R. & Fox, B. G. Aerobic deconstruction of cellulosic biomass by an insect-associated *Streptomyces*. *Sci. Rep.* **3**, 1030 (2013).
122. Ohashi, K. *et al.* Mannose- and mannobiose-specific responses of the insect-associated cellulolytic bacterium *Streptomyces* sp. strain SirexAA-E. *Appl. Environ. Microbiol.* **87**, (2021).
123. De Fine Licht, H. H., Boomsma, J. J. & Tunlid, A. Symbiotic adaptations in the fungal cultivar of leaf-cutting ants. *Nat. Commun.* **5**, 5675 (2014).
124. Currie, C. R. & Stuart, A. E. Weeding and grooming of pathogens in agriculture by ants. *Proc. R. Soc. London. Ser. B Biol. Sci.* **268**, 1033–1039 (2001).
125. Sen, R. *et al.* Generalized antifungal activity and 454-screening of *Pseudonocardia* and *Amycolatopsis* bacteria in nests of fungus-growing ants. *Proc. Natl. Acad. Sci.* **106**, 17805–17810 (2009).
126. Carr, G., Derbyshire, E. R., Caldera, E., Currie, C. R. & Clardy, J. Antibiotic and antimalarial quinones from fungus-growing ant-associated *Pseudonocardia* sp. *J. Nat. Prod.* **75**, 1806–1809 (2012).
127. Oh, D.-C., Poulsen, M., Currie, C. R. & Clardy, J. Dentigerumycin: a bacterial mediator of an ant-fungus symbiosis. *Nat. Chem. Biol.* **5**, 391–393 (2009).
128. Van Arnam, E. B. *et al.* Selvamicin, an atypical antifungal polyene from two alternative genomic contexts. *Proc. Natl. Acad. Sci.* **113**, 12940–12945 (2016).
129. Goldstein, S. L. & Klassen, J. L. *Pseudonocardia* symbionts of fungus-growing ants and the evolution of defensive secondary metabolism. *Front. Microbiol.* **11**, (2020).
130. Chapman, J. Uncovering the molecular mechanisms behind mycetoma. (Newcastle University, 2021).
131. Cano-Jiménez, E. *et al.* Farmer's lung disease. A review. *Arch. Bronconeumol. (English Ed.* **52**, 321–328 (2016).

132. Gallo, R. L. & Hooper, L. V. Epithelial antimicrobial defence of the skin and intestine. *Nat. Rev. Immunol.* **12**, 503–516 (2012).
133. Bolourian, A. & Mojtahedi, Z. *Streptomyces*, shared microbiome member of soil and gut, as 'old friends' against colon cancer. *FEMS Microbiol. Ecol.* **94**, (2018).
134. Huang, Y. J. *et al.* The airway microbiome in patients with severe asthma: Associations with disease features and severity. *J. Allergy Clin. Immunol.* **136**, 874–884 (2015).
135. Collado, M. C., Rautava, S., Aakko, J., Isolauri, E. & Salminen, S. Human gut colonisation may be initiated *in utero* by distinct microbial communities in the placenta and amniotic fluid. *Sci. Rep.* **6**, 23129 (2016).
136. Rea, D. *et al.* Microbiota effects on cancer: from risks to therapies. *Oncotarget* **9**, 17915–17927 (2018).
137. Simmons, T. L. *et al.* Biosynthetic origin of natural products isolated from marine microorganism–invertebrate assemblages. *Proc. Natl. Acad. Sci.* **105**, 4587–4594 (2008).
138. Webster, N. S. *et al.* Deep sequencing reveals exceptional diversity and modes of transmission for bacterial sponge symbionts. *Environ. Microbiol.* (2009). doi:10.1111/j.1462-2920.2009.02065.x
139. Taylor, M. W., Radax, R., Steger, D. & Wagner, M. Sponge-associated microorganisms: Evolution, ecology, and biotechnological potential. *Microbiol. Mol. Biol. Rev.* **71**, 295–347 (2007).
140. Siegl, A. & Hentschel, U. PKS and NRPS gene clusters from microbial symbiont cells of marine sponges by whole genome amplification. *Environ. Microbiol. Rep.* **2**, 507–513 (2009).
141. Fieseler, L. *et al.* Widespread occurrence and genomic context of unusually small polyketide synthase genes in microbial consortia associated with marine sponges. *Appl. Environ. Microbiol.* **73**, 2144–2155 (2007).
142. Hentschel, U. *et al.* Molecular evidence for a uniform microbial community in sponges from different oceans. *Appl. Environ. Microbiol.* **68**, 4431–4440 (2002).
143. Khan, S. T., Takagi, M. & Shin-ya, K. Actinobacteria associated with the marine sponges *Cinachyra* sp., *Petrosia* sp., and *Ulosa* sp. and their culturability. *Microbes Environ.* **27**, 99–104 (2012).
144. Thomas, T. *et al.* Diversity, structure and convergent evolution of the global sponge microbiome. *Nat. Commun.* **7**, 11870 (2016).
145. Morales-Castilla, I., Matias, M. G., Gravel, D. & Araújo, M. B. Inferring biotic interactions from proxies. *Trends Ecol. Evol.* **30**, 347–356 (2015).
146. Martin, B. D. & Schwab, E. Current usage of symbiosis and associated terminology. *Int. J. Biol.* **5**, (2012).
147. Pacheco, A. R. & Segrè, D. A multidimensional perspective on microbial interactions. *FEMS Microbiol. Lett.* **366**, (2019).
148. Kumbhar, C., Mudliar, P., Bhatia, L., Kshirsagar, A. & Watve, M. Widespread predatory abilities in the genus *Streptomyces*. *Arch. Microbiol.* **196**, 235–248 (2014).
149. Zeng, Y. *et al.* A *Streptomyces globisporus* strain kills *Microcystis aeruginosa* via cell-to-cell contact. *Sci. Total Environ.* **769**, 144489 (2021).
150. Patin, N. V. *et al.* Effects of actinomycete secondary metabolites on sediment microbial communities. *Appl. Environ. Microbiol.* **83**, (2017).
151. Zeph, L. R. & Casida, L. E. Gram-negative versus Gram-positive (actinomycete) nonobligate bacterial predators of bacteria in soil. *Appl. Environ. Microbiol.* **52**, 819–823 (1986).
152. Sydney, N. *et al.* The genetics of prey susceptibility to myxobacterial predation: A review, including an investigation into *Pseudomonas aeruginosa* mutations affecting predation by *Myxococcus xanthus*. *Microb. Physiol.* **31**, 57–66 (2021).
153. Pérez-Brocal, V., Latorre, A. & Moya, A. Symbionts and pathogens: What is the difference? in *Current topics in microbiology and immunology* **358**, 215–243 (2011).
154. Chanway, C. P. Endophytes: they're not just fungi! *Can. J. Bot.* **74**, 321–322 (1996).

155. van Overbeek, L. S., Bergervoet, J. H. W., Jacobs, F. H. H. & van Elsas, J. D. The low-temperature-induced viable-but-nonculturable state affects the virulence of *Ralstonia solanacearum* biovar 2. *Phytopathology®* **94**, 463–469 (2004).
156. Kloepper, J. W., McInroy, J. A., Liu, K. & Hu, C.-H. Symptoms of fern distortion syndrome resulting from inoculation with opportunistic endophytic fluorescent *Pseudomonas* spp. *PLoS One* **8**, e58531 (2013).
157. Mougi, A. The roles of amensalistic and commensalistic interactions in large ecological network stability. *Sci. Rep.* **6**, 29929 (2016).
158. Corno, G., Caravati, E., Callieri, C. & Bertoni, R. Effects of predation pressure on bacterial abundance, diversity, and size-structure distribution in an oligotrophic system. *J. Limnol.* **67**, 107 (2008).
159. Chase, J. M. *et al.* The interaction between predation and competition: a review and synthesis. *Ecol. Lett.* **5**, 302–315 (2002).
160. Chesson, P. Mechanisms of maintenance of species diversity. *Annu. Rev. Ecol. Syst.* **31**, 343–366 (2000).
161. Jurkevitch, E. & Mitchell, R. J. *The Ecology of Predation at the Microscale*. *The Ecology of Predation at the Microscale* (Springer International Publishing, 2020). doi:10.1007/978-3-030-45599-6
162. Chesson, P. & Kuang, J. J. The interaction between predation and competition. *Nature* **456**, 235–238 (2008).
163. Hungate, B. A. *et al.* The functional significance of bacterial predators. *MBio* **12**, (2021).
164. Guerrero, R., Esteve, I., Pedros-Alio, C. & Gaju, N. Predatory bacteria in prokaryotic communities. *Ann. N. Y. Acad. Sci.* **503**, 238–250 (1987).
165. Davidov, Y., Huchon, D., Koval, S. F. & Jurkevitch, E. A new alpha-proteobacterial clade of *Bdellovibrio*-like predators: implications for the mitochondrial endosymbiotic theory. *Environ. Microbiol.* **8**, 2179–2188 (2006).
166. Cavalier-Smith, T. Predation and eukaryote cell origins: A coevolutionary perspective. *Int. J. Biochem. Cell Biol.* **41**, 307–322 (2009).
167. Mu, D.-S. *et al. Bradymonabacteria*, a novel bacterial predator group with versatile survival strategies in saline environments. *Microbiome* **8**, 126 (2020).
168. Wang, S., Mu, D.-S., Zheng, W.-S. & Du, Z.-J. Complete genome sequence of *Bradymonas sediminis* FA350T, the first representative of the order Bradymonadales. *Mar. Genomics* **46**, 62–65 (2019).
169. Rendulic, S. *et al.* A predator unmasked: Life cycle of *Bdellovibrio bacteriovorus* from a genomic perspective. *Science (80-.).* **303**, 689–692 (2004).
170. Moya, A., Gil, R. & Latorre, A. The evolutionary history of symbiotic associations among bacteria and their animal hosts: a model. *Clin. Microbiol. Infect.* **15**, 11–13 (2009).
171. McCutcheon, J. P. & Moran, N. A. Parallel genomic evolution and metabolic interdependence in an ancient symbiosis. *Proc. Natl. Acad. Sci.* **104**, 19392–19397 (2007).
172. Hobley, L. *et al.* Genome analysis of a simultaneously predatory and prey-independent, novel *Bdellovibrio bacteriovorus* from the River Tiber, supports *in silico* predictions of both ancient and recent lateral gene transfer from diverse bacteria. *BMC Genomics* **13**, 670 (2012).
173. Bobyk, M. A., Afinogenova, A. V., Dudinskaya, M. V., Lambina, V. A. & Kulaev, I. S. Detection of polyphosphates and enzymes of polyphosphate metabolism in *Bdellovibrio bacteriovorus*. *Zentralblatt für Bakteriol. Parasitenkunde, Infekt. und Hyg. Zweite Naturwissenschaftliche Abteilung Mikrobiol. der Landwirtschaft, der Technol. und des Umweltschutzes* **135**, 461–466 (1980).
174. Dugan, K. Pathogen Predators (Archived). (2015). Available at: https://www.darpa.mil/program/pathogen-predators. (Accessed: 10th October 2022)
175. Lambert, C., Chang, C.-Y., Capeness, M. J. & Sockett, R. E. The first bite — profiling the predatosome in the bacterial pathogen *Bdellovibrio*. *PLoS One* **5**, e8599 (2010).

176. Duncan, M. C. *et al.* High-throughput analysis of gene function in the bacterial predator *Bdellovibrio bacteriovorus*. *MBio* **10**, (2019).
177. Willis, A. R. *et al.* Injections of predatory bacteria work alongside host immune cells to treat *Shigella* infection in zebrafish larvae. *Curr. Biol.* **26**, 3343–3351 (2016).
178. Shatzkes, K. *et al.* Predatory bacteria attenuate *Klebsiella pneumoniae* burden in rat lungs. *MBio* **7**, (2016).
179. Iebba, V. *et al.* Higher prevalence and abundance of *Bdellovibrio bacteriovorus* in the human gut of healthy subjects. *PLoS One* **8**, e61608 (2013).
180. Atterbury, R. J. & Tyson, J. Predatory bacteria as living antibiotics – where are we now? *Microbiology* **167**, 1–8 (2021).
181. Landwehr, W., Wolf, C. & Wink, J. Actinobacteria and Myxobacteria—Two of the most important bacterial resources for novel antibiotics. in 273–302 (2016). doi:10.1007/82_2016_503
182. Reichenbach, H. Myxobacteria: A source of new antibiotics. *Trends Biotechnol.* **6**, 115–121 (1988).
183. Reichenbach, H., Lang, E., Schumann, P. & Spröer, C. *Byssovorax cruenta* gen. nov., sp. nov., nom. rev., a cellulose-degrading myxobacterium: Rediscovery of '*Myxococcus cruentus*' Thaxter 1897. *Int. J. Syst. Evol. Microbiol.* **56**, 2357–2363 (2006).
184. Berleman, J. E., Chumley, T., Cheung, P. & Kirby, J. R. Rippling is a predatory behavior in *Myxococcus xanthus*. *J. Bacteriol.* **188**, 5888–5895 (2006).
185. Anscombe, F. J. & Singh, B. N. Limitation of bacteria by micro-predators in soil. *Nature* **161**, 140–141 (1948).
186. Fiegna, F. & Velicer, G. J. Exploitative and hierarchical antagonism in a cooperative bacterium. *PLoS Biol.* **3**, e370 (2005).
187. Sudo, S. & Dworkin, M. Bacteriolytic enzymes produced by *Myxococcus xanthus*. *J. Bacteriol.* **110**, 236–245 (1972).
188. Thiery, S. & Kaimer, C. The predation strategy of *Myxococcus xanthus*. *Front. Microbiol.* **11**, (2020).
189. Xiao, Y., Gerth, K., Müller, R. & Wall, D. Myxobacterium-produced antibiotic TA (myxovirescin) inhibits type II signal peptidase. *Antimicrob. Agents Chemother.* **56**, 2014–21 (2012).
190. Cortina, N. S., Krug, D., Plaza, A., Revermann, O. & Müller, R. Myxoprincomide: a natural product from *Myxococcus xanthus* discovered by comprehensive analysis of the secondary metabolome. *Angew. Chem. Int. Ed. Engl.* **51**, 811–6 (2012).
191. Volz, C., Kegler, C. & Müller, R. Enhancer binding proteins act as hetero-oligomers and link secondary metabolite production to myxococcal development, motility, and predation. *Chem. Biol.* **19**, 1447–59 (2012).
192. Goldman, B. S. *et al.* Evolution of sensory complexity recorded in a myxobacterial genome. *Proc. Natl. Acad. Sci.* **103**, 15200–15205 (2006).
193. Nomoto, M. & Narahashi, Y. A proteolytic enzyme of *Streptomyces griseus*. *J. Biochem.* **46**, 653–667 (1959).
194. Busche, T. *et al.* Multi-omics and targeted approaches to determine the role of cellular proteases in *Streptomyces* protein secretion. *Front. Microbiol.* **9**, (2018).
195. Vela Gurovic, M. S., Díaz, M. L., Gallo, C. A. & Dietrich, J. Phylogenomics, CAZyome and core secondary metabolome of *Streptomyces albus* species. *Mol. Genet. Genomics* **296**, 1299–1311 (2021).
196. Otani, H., Udwary, D. W. & Mouncey, N. J. Comparative and pangenomic analysis of the genus *Streptomyces*. *Sci. Rep.* **12**, 18909 (2022).
197. Li, Z. *et al.* A novel outer membrane β-1,6-glucanase is deployed in the predation of fungi by myxobacteria. *ISME J.* **13**, 2223–2235 (2019).

198. Gómez-Santos, N., Glatter, T., Koebnik, R., Świątek-Połatyńska, M. A. & Søgaard-Andersen, L. A TonB-dependent transporter is required for secretion of protease PopC across the bacterial outer membrane. *Nat. Commun.* **10**, 1360 (2019).
199. Mendes, M. V. *et al.* Cholesterol oxidases act as signaling proteins for the biosynthesis of the polyene macrolide pimaricin. *Chem. Biol.* **14**, 279–290 (2007).
200. Canon, F., Nidelet, T., Guédon, E., Thierry, A. & Gagnaire, V. Understanding the mechanisms of positive microbial interactions that benefit lactic acid bacteria co-cultures. *Front. Microbiol.* **11**, (2020).
201. McBride, M. J. & Zusman, D. R. Behavioral analysis of single cells of *Myxococcus xanthus* in response to prey cells of *Escherichia coli*. *FEMS Microbiol. Lett.* **137**, 227–231 (1996).
202. Zhang, H., Rao, N. N., Shiba, T. & Kornberg, A. Inorganic polyphosphate in the social life of *Myxococcus xanthus* : Motility, development, and predation. *Proc. Natl. Acad. Sci.* **102**, 13416–13420 (2005).
203. Wu, C. *et al.* Expanding the chemical space for natural products by *Aspergillus-Streptomyces* co-cultivation and biotransformation. *Sci. Rep.* **5**, 10868 (2015).
204. Luti, K. J. K. & Mavituna, F. Elicitation of *Streptomyces coelicolor* with dead cells of *Bacillus subtilis* and *Staphylococcus aureus* in a bioreactor increases production of undecylprodigiosin. *Appl. Microbiol. Biotechnol.* **90**, 461–466 (2011).
205. Onaka, H., Mori, Y., Igarashi, Y. & Furumai, T. Mycolic acid-containing bacteria induce natural-product biosynthesis in *Streptomyces* species. *Appl. Environ. Microbiol.* **77**, 400–406 (2011).
206. Smith, P. & Schuster, M. Public goods and cheating in microbes. *Curr. Biol.* **29**, R442–R447 (2019).
207. Reintjes, G., Arnosti, C., Fuchs, B. & Amann, R. Selfish, sharing and scavenging bacteria in the Atlantic Ocean: a biogeographical study of bacterial substrate utilisation. *ISME J.* **13**, 1119–1132 (2019).
208. Marshall, R. C. & Whitworth, D. E. Is “wolf-pack” predation by antimicrobial bacteria cooperative? Cell behaviour and predatory mechanisms indicate profound selfishness, even when working alongside kin. *BioEssays* **41**, 1800247 (2019).
209. Faddetta, T. *et al. Streptomyces coelicolor* vesicles: Many molecules to be delivered. *Appl. Environ. Microbiol.* **88**, (2022).
210. Hoefler, B. C. *et al.* A link between linearmycin biosynthesis and extracellular vesicle genesis connects specialized metabolism and bacterial membrane physiology. *Cell Chem. Biol.* **24**, 1238-1249.e7 (2017).
211. Blackburn, S. A., Shepherd, M. & Robinson, G. K. The polyene antifungal candicidin is selectively packaged into membrane vesicles in *Streptomyces* S4. *Arch. Microbiol.* **204**, 289 (2022).
212. Rashid, M. H. *et al.* Polyphosphate kinase is essential for biofilm development, quorum sensing, and virulence of *Pseudomonas aeruginosa*. *Proc. Natl. Acad. Sci.* **97**, 9636–9641 (2000).
213. Rashid, M. H., Rao, N. N. & Kornberg, A. Inorganic polyphosphate is required for motility of bacterial pathogens. *J. Bacteriol.* **182**, 225–227 (2000).
214. Rashid, M. H. & Kornberg, A. Inorganic polyphosphate is needed for swimming, swarming, and twitching motilities of *Pseudomonas aeruginosa*. *Proc. Natl. Acad. Sci.* **97**, 4885–4890 (2000).
215. Zhang, H., Gómez-García, M. R., Brown, M. R. W. & Kornberg, A. Inorganic polyphosphate in *Dictyostelium discoideum* : Influence on development, sporulation, and predation. *Proc. Natl. Acad. Sci.* **102**, 2731–2735 (2005).
216. Kawakoshi, A. *et al.* Deciphering the genome of polyphosphate accumulating Actinobacterium *Microlunatus phosphovorus*. *DNA Res.* **19**, 383–394 (2012).
217. Chouayekh, H. & Virolle, M.-J. The polyphosphate kinase plays a negative role in the control of antibiotic production in *Streptomyces lividans*. *Mol. Microbiol.* **43**, 919–930 (2002).

218. Le Maréchal, P. *et al.* Comparative proteomic analysis of *Streptomyces* lividans wild-type and *ppk* mutant strains reveals the importance of storage lipids for antibiotic biosynthesis. *Appl. Environ. Microbiol.* **79**, 5907–5917 (2013).
219. Barron, G. L. Predatory fungi, wood decay, and the carbon cycle. *Biodiversity* **4**, 3–9 (2003).
220. Yang, Y., Yang, E., An, Z. & Liu, X. Evolution of nematode-trapping cells of predatory fungi of the *Orbiliaceae* based on evidence from rRNA-encoding DNA and multiprotein sequences. *Proc. Natl. Acad. Sci.* **104**, 8379–8384 (2007).
221. Degenkolb, T. & Vilcinskas, A. Metabolites from nematophagous fungi and nematicidal natural products from fungi as an alternative for biological control. Part I: metabolites from nematophagous ascomycetes. *Appl. Microbiol. Biotechnol.* **100**, 3799–3812 (2016).
222. Kuo, T.-H., Yang, C.-T., Chang, H.-Y., Hsueh, Y.-P. & Hsu, C.-C. Nematode-trapping fungi produce diverse metabolites during predator–prey interaction. *Metabolites* **10**, 117 (2020).
223. Thorn, R. G. & Barron, G. L. Carnivorous mushrooms. *Science (80-.).* **224**, 76–78 (1984).
224. Soares, F. E. de F., Sufiate, B. L. & de Queiroz, J. H. Nematophagous fungi: Far beyond the endoparasite, predator and ovicidal groups. *Agric. Nat. Resour.* **52**, 1–8 (2018).
225. Yang, C.-T. *et al.* Natural diversity in the predatory behavior facilitates the establishment of a robust model strain for nematode-trapping fungi. *Proc. Natl. Acad. Sci.* **117**, 6762–6770 (2020).
226. Chen, S.-A., Lin, H.-C. & Hsueh, Y.-P. The cAMP-PKA pathway regulates prey sensing and trap morphogenesis in the nematode-trapping fungus *Arthrobotrys oligospora*. *G3 Genes|Genomes|Genetics* **12**, (2022).
227. Westhoff, S., Kloosterman, A. M., van Hoesel, S. F. A., van Wezel, G. P. & Rozen, D. E. Competition Sensing Changes Antibiotic Production in *Streptomyces*. *MBio* **12**, 1–13 (2021).
228. Ueda, K. *et al.* Wide distribution of interspecific stimulatory events on antibiotic production and sporulation among *Streptomyces* species. *J. Antibiot. (Tokyo).* **53**, 979–982 (2000).
229. Slattery, M., Rajbhandari, I. & Wesson, K. Competition-mediated antibiotic induction in the marine bacterium *Streptomyces tenjimariensis*. *Microb. Ecol.* **41**, 90–96 (2001).
230. Vaz Jauri, P. & Kinkel, L. L. Nutrient overlap, genetic relatedness and spatial origin influence interaction-mediated shifts in inhibitory phenotype among *Streptomyces* spp. *FEMS Microbiol. Ecol.* **90**, 264–275 (2014).
231. Gonzalez, D., Sabnis, A., Foster, K. R. & Mavridou, D. A. I. Costs and benefits of provocation in bacterial warfare. *Proc. Natl. Acad. Sci.* **115**, 7593–7598 (2018).
232. Ventura, M. *et al.* Genomics of Actinobacteria : Tracing the evolutionary history of an ancient phylum. *Microbiol. Mol. Biol. Rev.* **71**, 495–548 (2007).
233. Verma, M. *et al.* Phylogenetic analyses of phylum Actinobacteria based on whole genome sequences. *Res. Microbiol.* **164**, 718–728 (2013).
234. Kirby, R. Chromosome diversity and similarity within the actinomycetales. *FEMS Microbiol. Lett.* **319**, 1–10 (2011).
235. Lorenzi, J.-N., Lespinet, O., Leblond, P. & Thibessard, A. Subtelomeres are fast-evolving regions of the *Streptomyces* linear chromosome. *Microb. genomics* **7**, (2019).
236. Barry, J. D., Ginger, M. L., Burton, P. & McCulloch, R. Why are parasite contingency genes often associated with telomeres? *Int. J. Parasitol.* **33**, 29–45 (2003).
237. Moxon, R., Bayliss, C. & Hood, D. Bacterial contingency loci: The role of simple sequence DNA repeats in bacterial adaptation. *Annu. Rev. Genet.* **40**, 307–333 (2006).
238. Tidjani, A.-R., Bontemps, C. & Leblond, P. Telomeric and sub-telomeric regions undergo rapid turnover within a *Streptomyces* population. *Sci. Rep.* **10**, 7720 (2020).
239. Volff, J.-N. & Altenbuchner, J. A new beginning with new ends: Linearisation of circular chromosomes during bacterial evolution. *FEMS Microbiol. Lett.* **186**, 143–150 (2000).
240. Pandza, S. *et al.* Recombination between the linear plasmid pPZG101 and the linear chromosome of *Streptomyces rimosus* can lead to exchange of ends. *Mol. Microbiol.* **28**, 1165–1176 (1998).
241. Cohan, F. M. Bacterial Species and Speciation. *Syst. Biol.* **50**, 513–524 (2001).

242. Lioy, V. S. *et al.* Dynamics of the compartmentalized *Streptomyces* chromosome during metabolic differentiation. *Nat. Commun.* **12**, 5221 (2021).
243. Chater, K. F. & Chandra, G. The evolution of development in *Streptomyces* analysed by genome comparisons. *FEMS Microbiol. Rev.* **30**, 651–672 (2006).
244. Letzel, A.-C. *et al.* Genomic insights into specialized metabolism in the marine actinomycete *Salinispora*. *Environ. Microbiol.* **19**, 3660–3673 (2017).
245. Adamek, M. *et al.* Comparative genomics reveals phylogenetic distribution patterns of secondary metabolites in *Amycolatopsis* species. *BMC Genomics* **19**, 426 (2018).
246. Bilyk, B., Horbal, L. & Luzhetskyy, A. Chromosomal position effect influences the heterologous expression of genes and biosynthetic gene clusters in *Streptomyces albus* J1074. *Microb. Cell Fact.* **16**, 5 (2017).
247. McDonald, B. R. & Currie, C. R. Lateral gene transfer dynamics in the ancient bacterial genus *Streptomyces*. *MBio* **8**, (2017).
248. Bruns, H. *et al.* Function-related replacement of bacterial siderophore pathways. *ISME J.* **12**, 320–329 (2018).
249. Kurosawa, K. *et al.* Rhodostreptomycins, antibiotics biosynthesized following horizontal gene transfer from *Streptomyces padanus* to *Rhodococcus fascians*. *J. Am. Chem. Soc.* **130**, 1126–1127 (2008).
250. Caicedo-Montoya, C., Manzo-Ruiz, M. & Ríos-Estepa, R. Pan-Genome of the genus *Streptomyces* and prioritization of biosynthetic gene clusters with potential to produce antibiotic compounds. *Front. Microbiol.* **12**, (2021).
251. Fookes, M. C. *et al. Mycoplasma genitalium*: whole genome sequence analysis, recombination and population structure. *BMC Genomics* **18**, 993 (2017).
252. Vernikos, G., Medini, D., Riley, D. R. & Tettelin, H. Ten years of pan-genome analyses. *Curr. Opin. Microbiol.* **23**, 148–154 (2015).
253. Ventura, M., Canchaya, C., Fitzgerald, G. F., Gupta, R. S. & van Sinderen, D. Genomics as a means to understand bacterial phylogeny and ecological adaptation: the case of bifidobacteria. *Antonie Van Leeuwenhoek* **91**, 351–372 (2007).
254. Beld, J., Sonnenschein, E. C., Vickery, C. R., Noel, J. P. & Burkart, M. D. The phosphopantetheinyl transferases: catalysis of a post-translational modification crucial for life. *Nat. Prod. Rep.* **31**, 61–108 (2014).
255. Romero-Rodríguez, A., Robledo-Casados, I. & Sánchez, S. An overview on transcriptional regulators in *Streptomyces*. *Biochim. Biophys. Acta - Gene Regul. Mech.* **1849**, 1017–1039 (2015).
256. Hoskisson, P. A. & Fernández-Martínez, L. T. Regulation of specialised metabolites in Actinobacteria - expanding the paradigms. *Environ. Microbiol. Rep.* **10**, 231–238 (2018).
257. Kong, D., Wang, X., Nie, J. & Niu, G. Regulation of antibiotic production by signaling molecules in *Streptomyces*. *Front. Microbiol.* **10**, (2019).
258. Niu, G., Chater, K. F., Tian, Y., Zhang, J. & Tan, H. Specialised metabolites regulating antibiotic biosynthesis in *Streptomyces* spp. *FEMS Microbiol. Rev.* **40**, 554–573 (2016).
259. Świątek-Połatyńska, M. A. *et al.* Genome-wide analysis of *in vivo* binding of the master regulator DasR in *Streptomyces coelicolor* identifies novel non-canonical targets. *PLoS One* **10**, e0122479 (2015).
260. Wang, L. *et al.* Autoregulation of antibiotic biosynthesis by binding of the end product to an atypical response regulator. *Proc. Natl. Acad. Sci.* **106**, 8617–8622 (2009).
261. Leskiw, B. K., Lawlor, E. J., Fernandez-Abalos, J. M. & Chater, K. F. TTA codons in some genes prevent their expression in a class of developmental, antibiotic-negative, *Streptomyces* mutants. *Proc. Natl. Acad. Sci.* **88**, 2461–2465 (1991).
262. Fernández-Moreno, M. A., Caballero, J., Hopwood, D. A. & Malpartida, F. The *act* cluster contains regulatory and antibiotic export genes, direct targets for translational control by the *bldA tRNA* gene of *Streptomyces*. *Cell* **66**, 769–780 (1991).

263. Chater, K. F. & Chandra, G. The use of the rare UUA codon to define “Expression Space” for genes involved in secondary metabolism, development and environmental adaptation in *Streptomyces*. *J. Microbiol.* **46**, 1–11 (2008).
264. Higo, A., Horinouchi, S. & Ohnishi, Y. Strict regulation of morphological differentiation and secondary metabolism by a positive feedback loop between two global regulators AdpA and BldA in *Streptomyces griseus*. *Mol. Microbiol.* **81**, 1607–1622 (2011).
265. O’Rourke, S. *et al.* Extracellular signalling, translational control, two repressors and an activator all contribute to the regulation of methylenomycin production in *Streptomyces coelicolor*. *Mol. Microbiol.* **71**, 763–778 (2009).
266. Wang, M. *et al.* Improvement of natamycin production by cholesterol oxidase overexpression in *Streptomyces gilvosporeus*. *J. Microbiol. Biotechnol.* **26**, 241–247 (2016).
267. Marushima, K., Ohnishi, Y. & Horinouchi, S. CebR as a master regulator for cellulose/cellooligosaccharide catabolism affects morphological development in *Streptomyces griseus*. *J. Bacteriol.* **191**, 5930–40 (2009).
268. Francis, I. M., Jourdan, S., Fanara, S., Loria, R. & Rigali, S. The cellobiose sensor CebR is the gatekeeper of *Streptomyces scabies* pathogenicity. *MBio* **6**, 2–7 (2015).
269. Rudd, B. A. M. & Hopwood, D. A. Genetics of actinorhodin biosynthesis by *Streptomyces coelicolor* A3(2). *J. Gen. Microbiol.* **114**, 35–43 (1979).
270. Feitelson, J. S., Malpartida, F. & Hopwood, D. A. Genetic and biochemical characterization of the *red* gene cluster of *Streptomyces coelicolor* A3(2). *Microbiology* **131**, 2431–2441 (1985).
271. Hopwood, D. A. & Wright, H. M. CDA is a new chromosomally-determined antibiotic from *Streptomyces coelicolor* A3(2). *Microbiology* **129**, 3575–3579 (1983).
272. Wright, L. F. & Hopwood, D. A. Identification of the antibiotic determined by the SCP1 plasmid of *Streptomyces coelicolor* A3 (2). *J. Gen. Microbiol.* **95**, 96–106 (1976).
273. Malpartida, F. & Hopwood, D. A. Molecular cloning of the whole biosynthetic pathway of a *Streptomyces* antibiotic and its expression in a heterologous host. *Nature* **309**, 462–464 (1984).
274. Hoskisson, P. A. & Seipke, R. F. Cryptic or silent? The known unknowns, unknown knowns, and unknown unknowns of secondary metabolism. *MBio* **11**, 1–5 (2020).
275. NCBI. Taxonomy browser: Actinobacteria. *NCBI* (2022). Available at: https://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?lvl=0&id=201174. (Accessed: 31st October 2022)
276. Chung, Y.-H. *et al.* Comparative genomics reveals a remarkable biosynthetic potential of the *Streptomyces* phylogenetic lineage associated with rugose-ornamented spores. *mSystems* **6**, (2021).
277. Blin, K., Shaw, S., Kautsar, S. A., Medema, M. H. & Weber, T. The antiSMASH database version 3: increased taxonomic coverage and new query features for modular enzymes. *Nucleic Acids Res.* **49**, D639–D643 (2021).
278. Kautsar, S. A. *et al.* MIBiG 2.0: a repository for biosynthetic gene clusters of known function. *Nucleic Acids Res.* (2019). doi:10.1093/nar/gkz882
279. Skinnider, M. A. *et al.* Genomes to natural products PRediction Informatics for Secondary Metabolomes (PRISM). *Nucleic Acids Res.* gkv1012 (2015). doi:10.1093/nar/gkv1012
280. Khaldi, N. *et al.* SMURF: Genomic mapping of fungal secondary metabolite clusters. *Fungal Genet. Biol.* **47**, 736–741 (2010).
281. Blin, K. *et al.* antiSMASH 6.0: improving cluster detection and comparison capabilities. *Nucleic Acids Res.* **49**, W29–W35 (2021).
282. Medema, M. H. *et al.* antiSMASH: rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. *Nucleic Acids Res.* **39**, W339–W346 (2011).
283. Johnston, C. W. *et al.* Assembly and clustering of natural antibiotics guides target identification. *Nat. Chem. Biol.* **12**, 233–239 (2016).

284. Dejong, C. A. *et al.* Polyketide and nonribosomal peptide retro-biosynthesis and global gene cluster matching. *Nat. Chem. Biol.* **12**, 1007–1014 (2016).
285. Johnston, C. W. *et al.* An automated Genomes-to-Natural Products platform (GNP) for the discovery of modular natural products. *Nat. Commun.* **6**, 8421 (2015).
286. Kautsar, S. A., van der Hooft, J. J. J., de Ridder, D. & Medema, M. H. BiG-SLiCE: A highly scalable tool maps the diversity of 1.2 million biosynthetic gene clusters. *Gigascience* **10**, (2021).
287. Gavriilidou, A. *et al.* Compendium of specialized metabolite biosynthetic diversity encoded in bacterial genomes. *Nat. Microbiol.* **7**, 726–735 (2022).
288. Kautsar, S. A., Blin, K., Shaw, S., Weber, T. & Medema, M. H. BiG-FAM: the biosynthetic gene cluster families database. *Nucleic Acids Res.* **49**, D490–D497 (2021).
289. Ziemert, N., Alanjary, M. & Weber, T. The evolution of genome mining in microbes – a review. *Nat. Prod. Rep.* **33**, 988–1005 (2016).
290. Sugimoto, Y. *et al.* A metagenomic strategy for harnessing the chemical repertoire of the human microbiome. *Science (80-.).* **366**, (2019).
291. Sélem-Mojica, N., Aguilar, C., Gutiérrez-García, K., Martínez-Guerrero, C. E. & Barona-Gómez, F. EvoMining reveals the origin and fate of natural product biosynthetic enzymes. *Microb. genomics* **5**, (2019).
292. Cruz-Morales, P. *et al.* Phylogenomic analysis of natural products biosynthetic gene clusters allows discovery of arseno-organic metabolites in model streptomycetes. *Genome Biol. Evol.* **8**, 1906–1916 (2016).
293. Staunton, J. & Weissman, K. J. Polyketide biosynthesis: a millennium review. *Nat. Prod. Rep.* **18**, 380–416 (2001).
294. Donadio, S., Staver, M., McAlpine, J., Swanson, S. & Katz, L. Modular organization of genes required for complex polyketide biosynthesis. *Science (80-.).* **252**, 675–679 (1991).
295. Kennedy, J. *et al.* Modulation of polyketide synthase activity by accessory proteins during lovastatin biosynthesis. *Science (80-.).* **284**, 1368–1372 (1999).
296. McDaniel, R., Ebert-Khosla, S., Fu, H., Hopwood, D. A. & Khosla, C. Engineered biosynthesis of novel polyketides: influence of a downstream enzyme on the catalytic specificity of a minimal aromatic polyketide synthase. *Proc. Natl. Acad. Sci.* **91**, 11542–11546 (1994).
297. Bao, W., Wendt-Pienkowski, E. & Hutchinson, C. R. Reconstitution of the iterative type II polyketide synthase for tetracenomycin F2 biosynthesis. *Biochemistry* **37**, 8132–8138 (1998).
298. Arthur, C. J. *et al.* The malonyl transferase activity of type II polyketide synthase acyl carrier proteins. *Chem. Biol.* **13**, 587–596 (2006).
299. Hertweck, C., Luzhetskyy, A., Rebets, Y. & Bechthold, A. type II polyketide synthases: gaining a deeper insight into enzymatic teamwork. *Nat. Prod. Rep.* **24**, 162–190 (2007).
300. Tsai, S.-C. (Sheryl). The structural enzymology of iterative aromatic polyketide synthases: A critical comparison with fatty acid synthases. *Annu. Rev. Biochem.* **87**, 503–531 (2018).
301. Sundaram, S. & Hertweck, C. On-line enzymatic tailoring of polyketides and peptides in thiotemplate systems. *Curr. Opin. Chem. Biol.* **31**, 82–94 (2016).
302. Thibodeaux, C. J., Melançon, C. E. & Liu, H. Natural-product sugar biosynthesis and enzymatic glycodiversification. *Angew. Chemie Int. Ed.* **47**, 9814–9859 (2008).
303. Challis, G. L. & Naismith, J. H. Structural aspects of non-ribosomal peptide biosynthesis. *Curr. Opin. Struct. Biol.* **14**, 748–756 (2004).
304. Winn, M., Fyans, J. K., Zhuo, Y. & Micklefield, J. Recent advances in engineering nonribosomal peptide assembly lines. *Nat. Prod. Rep.* **33**, 317–347 (2016).
305. Noda-Garcia, L. & Tawfik, D. S. Enzyme evolution in natural products biosynthesis: target- or diversity-oriented? *Curr. Opin. Chem. Biol.* **59**, 147–154 (2020).
306. Vining, L. C. Secondary metabolism, inventive evolution and biochemical diversity — a review. *Gene* **115**, 135–140 (1992).
307. Jenke-Kodama, H., Sandmann, A., Müller, R. & Dittmann, E. Evolutionary implications of bacterial polyketide synthases. *Mol. Biol. Evol.* **22**, 2027–2039 (2005).

308. Slot, J. C. & Rokas, A. Multiple GAL pathway gene clusters evolved independently and by different mechanisms in fungi. *Proc. Natl. Acad. Sci.* **107**, 10136–10141 (2010).
309. Kawaguchi, J., Mori, H., Iwai, N. & Wachi, M. A secondary metabolic enzyme functioned as an evolutionary seed of a primary metabolic enzyme. *Mol. Biol. Evol.* **39**, 1–12 (2022).
310. Caetano-Anollés, G. *et al.* The origin and evolution of modern metabolism. *Int. J. Biochem. Cell Biol.* **41**, 285–297 (2009).
311. Mazurenko, S., Prokop, Z. & Damborsky, J. Machine learning in enzyme engineering. *ACS Catal.* **10**, 1210–1223 (2020).
312. Newman, D. J. & Cragg, G. M. Natural products as sources of new drugs from 1981 to 2014. *J. Nat. Prod.* **79**, 629–661 (2016).
313. Fewer, D. P. & Metsä-Ketelä, M. A pharmaceutical model for the molecular evolution of microbial natural products. *FEBS J.* **287**, 1429–1449 (2020).
314. Kharel, M. K. *et al.* Angucyclines: Biosynthesis, mode-of-action, new natural products, and synthesis. *Nat. Prod. Rep.* **29**, 264–325 (2012).
315. Kallio, P. *et al.* Flavoprotein hydroxylase PgaE catalyzes two consecutive oxygen-dependent tailoring reactions in angucycline biosynthesis. *Biochemistry* **50**, 5535–5543 (2011).
316. Koskiniemi, H. *et al.* Crystal structures of two aromatic hydroxylases involved in the early tailoring steps of angucycline biosynthesis. *J. Mol. Biol.* **372**, 633–648 (2007).
317. Chen, Y. *et al.* Characterization of JadH as an FAD- and NAD(P)H-dependent bifunctional hydroxylase/dehydrase in jadomycin biosynthesis. *ChemBioChem* **11**, 1055–1060 (2010).
318. Patrikainen, P. *et al.* Structure-based engineering of angucyclinone 6-ketoreductases. *Chem. Biol.* **21**, 1381–1391 (2014).
319. Paananen, P. *et al.* Structural and functional analysis of angucycline C-6 ketoreductase LanV involved in landomycin biosynthesis. *Biochemistry* **52**, 5304–5314 (2013).
320. Metsä-Ketelä, M. Evolution inspired engineering of antibiotic biosynthesis enzymes. *Org. Biomol. Chem.* **15**, 4036–4041 (2017).
321. Tello, M., Kuzuyama, T., Heide, L., Noel, J. P. & Richard, S. B. The ABBA family of aromatic prenyltransferases: broadening natural product diversity. *Cell. Mol. Life Sci.* **65**, 1459–1463 (2008).
322. Kuzuyama, T., Noel, J. P. & Richard, S. B. Structural basis for the promiscuous biosynthetic prenylation of aromatic natural products. *Nature* **435**, 983–987 (2005).
323. Bonitz, T., Alva, V., Saleh, O., Lupas, A. N. & Heide, L. Evolutionary relationships of microbial aromatic prenyltransferases. *PLoS One* **6**, e27336 (2011).
324. Tanner, M. E. Mechanistic studies on the indole prenyltransferases. *Nat. Prod. Rep.* **32**, 88–101 (2015).
325. Metzger, U. *et al.* The structure of dimethylallyl tryptophan synthase reveals a common architecture of aromatic prenyltransferases in fungi and bacteria. *Proc. Natl. Acad. Sci.* **106**, 14309–14314 (2009).
326. Chen, R. *et al.* Molecular insights into the enzyme promiscuity of an aromatic prenyltransferase. *Nat. Chem. Biol.* **13**, 226–234 (2017).
327. Heide, L. Prenyl transfer to aromatic substrates: genetics and enzymology. *Curr. Opin. Chem. Biol.* **13**, 171–179 (2009).
328. Luk, L. Y. P. & Tanner, M. E. Mechanism of dimethylallyltryptophan synthase: Evidence for a dimethylallyl cation intermediate in an aromatic prenyltransferase reaction. *J. Am. Chem. Soc.* **131**, 13932–13933 (2009).
329. Grundmann, A., Kuznetsova, T., Afiyatullov, S. S. & Li, S.-M. FtmPT2, an N -Prenyltransferase from *Aspergillus fumigatus*, catalyses the last step in the biosynthesis of fumitremorgin B. *ChemBioChem* **9**, 2059–2063 (2008).
330. Takahashi, S. *et al.* Biochemical characterization of a novel indole prenyltransferase from *Streptomyces* sp. SN-593. *J. Bacteriol.* **192**, 2839–2851 (2010).

331. Rudolf, J. D., Wang, H. & Poulter, C. D. Multisite prenylation of 4-substituted tryptophans by dimethylallyltryptophan synthase. *J. Am. Chem. Soc.* **135**, 1895–1902 (2013).
332. Elshahawi, S. I. *et al.* Structure and specificity of a permissive bacterial C-prenyltransferase. *Nat. Chem. Biol.* **13**, 366–368 (2017).
333. Gallagher, K. A. & Jensen, P. R. Genomic insights into the evolution of hybrid isoprenoid biosynthetic gene clusters in the MAR4 marine streptomycete clade. *BMC Genomics* **16**, 960 (2015).
334. Kieser, T., Bibb, M., Buttner, M., Chater, K. & Hopwood, D. *Practical Streptomyces Genetics*. (The John Innes Foundation, 2000).
335. Nikodinovic, J., Barrow, K. D. & Chuck, J. A. High yield preparation of genomic DNA from *Streptomyces*. *Biotechniques* **35**, 932–936 (2003).
336. Andrews, S. FASTQC A Quality Control tool for High Throughput Sequence Data. *Babraham Institute* (2015). Available at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/Help/3 Analysis Modules/. (Accessed: 9th February 2016)
337. Coil, D., Jospin, G. & Darling, A. E. A5-miseq: an updated pipeline to assemble microbial genomes from Illumina MiSeq data. *Bioinformatics* **31**, 587–589 (2015).
338. Assefa, S., Keane, T. M., Otto, T. D., Newbold, C. & Berriman, M. ABACAS: algorithm-based automatic contiguation of assembled sequences. *Bioinformatics* **25**, 1968–1969 (2009).
339. Tsai, I. J., Otto, T. D. & Berriman, M. Improving draft assemblies by iterative mapping and assembly of short reads to eliminate gaps. *Genome Biol.* **11**, R41 (2010).
340. Brettin, T. *et al.* RASTtk: A modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. *Sci. Rep.* **5**, 8365 (2015).
341. Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V & Zdobnov, E. M. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* **31**, 3210–3212 (2015).
342. Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M. & Barton, G. J. Jalview version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**, 1189–1191 (2009).
343. Sievers, F. *et al.* Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**, 539 (2011).
344. Price, M. N., Dehal, P. S. & Arkin, A. P. FastTree 2 – approximately maximum-likelihood trees for large alignments. *PLoS One* **5**, e9490 (2010).
345. Huson, D. H. & Scornavacca, C. Dendroscope 3: An interactive tool for rooted phylogenetic trees and networks. *Syst. Biol.* **61**, 1061–1067 (2012).
346. Robert, X. & Gouet, P. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* **42**, W320–W324 (2014).
347. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
348. Langmead, B., Wilks, C., Antonescu, V. & Charles, R. Scaling read aligners to hundreds of threads on general-purpose processors. *Bioinformatics* **35**, 421–432 (2019).
349. Anders, S., Pyl, P. T. & Huber, W. HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–169 (2015).
350. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR : a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).
351. Kallio, M. A. *et al.* Chipster: user-friendly analysis software for microarray and other high-throughput data. *BMC Genomics* **12**, 507 (2011).
352. Amos, G. C. A. *et al.* Comparative transcriptomics as a guide to natural product discovery and biosynthetic gene cluster functionality. *Proc. Natl. Acad. Sci.* **114**, E11121–E11130 (2017).
353. Zhou, S. *et al.* Pentamycin biosynthesis in Philippine *Streptomyces* sp. S816: Cytochrome P450-catalyzed installation of the C-14 hydroxyl group. *ACS Chem. Biol.* **14**, 1305–1309 (2019).

354. Dhakal, D. *et al.* Complete genome sequence of *Streptomyces peucetius* ATCC 27952, the producer of anticancer anthracyclines and diverse secondary metabolites. *J. Biotechnol.* **267**, 50–54 (2018).
355. Song, E. *et al.* Proteomic approach to enhance doxorubicin production in panK-integrated *Streptomyces peucetius* ATCC 27952. *J. Ind. Microbiol. Biotechnol.* **38**, 1245–1253 (2011).
356. Worsley, S. F. *et al. Streptomyces* endophytes promote host health and enhance growth across plant species. *Appl. Environ. Microbiol.* **86**, (2020).
357. Li, Z., Rawlings, B. J., Harrison, P. H. & Vederas, J. C. Production of new polyene antibiotics by *Streptomyces cellulosae* after addition of ethyl(Z)-16-phenylhexadec-9-enoate. *J. Antibiot. (Tokyo).* **42**, 577–584 (1989).
358. Crawford, J. M., Vagstad, A. L., Ehrlich, K. C. & Townsend, C. A. Starter unit specificity directs genome mining of polyketide synthase pathways in fungi. *Bioorg. Chem.* **36**, 16–22 (2008).
359. Xu, L.-H., Fushinobu, S., Ikeda, H., Wakagi, T. & Shoun, H. Crystal structures of cytochrome P450 105P1 from *Streptomyces avermitilis* : Conformational flexibility and histidine ligation State. *J. Bacteriol.* **191**, 1211–1219 (2009).
360. Xu, L.-H. *et al.* Regio- and stereospecificity of filipin hydroxylation sites revealed by crystal structures of cytochrome P450 105P1 and 105D6 from *Streptomyces avermitilis*. *J. Biol. Chem.* **285**, 16844–16853 (2010).
361. Corbin, D. R., Grebenok, R. J., Ohnmeiss, T. E., Greenplate, J. T. & Purcell, J. P. Expression and chloroplast targeting of cholesterol oxidase in transgenic tobacco plants. *Plant Physiol.* **126**, 1116–1128 (2001).
362. Zygmunt, W. A. & Tavormina, P. A. Steroid interference with antifungal activity of polyene antibiotics. *Appl. Microbiol.* **14**, 865–869 (1966).
363. O'Connor, T. J., Boyd, D., Dorer, M. S. & Isberg, R. R. Aggravating genetic interactions allow a solution to redundancy in a bacterial pathogen. *Science (80-.).* **338**, 1440–1444 (2012).
364. Lipke, P. N. & Ovalle, R. Cell wall architecture in yeast: New structure and new challenges. *J. Bacteriol.* **180**, 3735–3740 (1998).
365. Kocourek, J. & Ballou, C. E. Method for fingerprinting yeast cell wall mannans. *J. Bacteriol.* **100**, 1175–1181 (1969).
366. Singh, R. P., Rajarammohan, S., Thakur, R. & Hassan, M. Linear and branched β- glucans degrading enzymes from versatile *Bacteroides uniformis* JCM 13288 T and their roles in cooperation with gut bacteria. *Gut Microbes* **12**, 1826761 (2020).
367. Cuskin, F. *et al.* Human gut *Bacteroidetes* can utilize yeast mannan through a selfish mechanism. *Nature* **517**, 165–169 (2015).
368. Chen, J. *et al.* Alpha- and beta-mannan utilization by marine *Bacteroidetes*. *Environ. Microbiol.* **20**, 4127–4140 (2018).
369. Tiels, P. *et al.* A bacterial glycosidase enables mannose-6-phosphate modification and improved cellular uptake of yeast-produced recombinant human lysosomal enzymes. *Nat. Biotechnol.* **30**, 1225–1231 (2012).
370. Deflandre, B. *et al.* The virulome of *Streptomyces scabiei* in response to cello-oligosaccharide elicitors. *Microb. Genomics* **8**, (2022).
371. Xu, Y. & Tan, D. S. Total synthesis of the bacterial diisonitrile chalkophore SF2768. *Org. Lett.* **21**, 8731–8735 (2019).
372. Greule, A. *et al.* Wide distribution of foxicin biosynthetic gene clusters in *Streptomyces* strains – An unusual secondary metabolite with various properties. *Front. Microbiol.* **8**, (2017).
373. Liu, W.-T. *et al.* MS/MS-based networking and peptidogenomics guided genome mining revealed the stenothricin gene cluster in *Streptomyces roseosporus*. *J. Antibiot. (Tokyo).* **67**, 99–104 (2014).
374. Møbjerg, N. & Neves, R. C. New insights into survival strategies of tardigrades. *Comp. Biochem. Physiol. Part A Mol. Integr. Physiol.* **254**, 110890 (2021).
375. Wright, J. C. Cryptobiosis 300 years on from van Leuwenhoek: What have we learned about tardigrades? *Zool. Anzeiger - A J. Comp. Zool.* **240**, 563–582 (2001).

376. Miret, J. J., Solari, A. J., Barderi, P. A. & Goldemberg, S. H. Polyamines and cell wall organization in *Saccharomyces cerevisiae*. *Yeast* **8**, 1033–1041 (1992).
377. Valdés-Santiago, L. & Ruiz-Herrera, J. Stress and polyamine metabolism in fungi. *Front. Chem.* **1**, (2014).
378. Lee, N. *et al.* Thirty complete *Streptomyces* genome sequences for mining novel secondary metabolite biosynthetic gene clusters. *Sci. Data* **7**, 55 (2020).
379. Page, A. J. *et al.* Robust high-throughput prokaryote *de novo* assembly and improvement pipeline for Illumina data. *Microb. Genomics* **2**, (2016).
380. Noh, S. Linear paths to genome reduction in a defensive symbiont. *Proc. Natl. Acad. Sci.* **118**, (2021).
381. Nechitaylo, T. Y. *et al.* Incipient genome erosion and metabolic streamlining for antibiotic production in a defensive symbiont. *Proc. Natl. Acad. Sci.* **118**, (2021).
382. da Costa, M. S., Santos, H. & Galinski, E. A. An overview of the role and diversity of compatible solutes in Bacteria and Archaea. in 117–153 (1998). doi:10.1007/BFb0102291
383. Zou, H. *et al.* The metabolism and biotechnological application of betaine in microorganism. *Appl. Microbiol. Biotechnol.* **100**, 3865–3876 (2016).
384. Czech, L., Poehl, S., Hub, P., Stöveken, N. & Bremer, E. Tinkering with osmotically controlled transcription allows enhanced production and excretion of ectoine and hydroxyectoine from a microbial cell factory. *Appl. Environ. Microbiol.* **84**, (2018).
385. Hashimoto, T. *et al.* Biosynthesis of quinolidomicin, the largest known macrolide of terrestrial origin: Identification and heterologous expression of a biosynthetic gene cluster over 200 kb. *Org. Lett.* **20**, 7996–7999 (2018).
386. Stankovic, N., Senerovic, L., Ilic-Tomic, T., Vasiljevic, B. & Nikodinovic-Runic, J. Properties and applications of undecylprodigiosin and other bacterial prodigiosins. *Appl. Microbiol. Biotechnol.* **98**, 3841–3858 (2014).
387. Cerdeño, A. M., Bibb, M. J. & Challis, G. L. Analysis of the prodiginine biosynthesis gene cluster of *Streptomyces coelicolor* A3(2): new mechanisms for chain initiation and termination in modular multienzymes. *Chem. Biol.* **8**, 817–829 (2001).
388. Guthrie, E. P. *et al.* A response-regulator-like activator of antibiotic synthesis from *Streptomyces coelicolor* A3(2) with an amino-terminal domain that lacks a phosphorylation pocket. *Microbiology* **144**, 727–738 (1998).
389. Kieser, T., Hopwood, D. A., Wright, H. M. & Thompson, C. J. pIJ101, a multi-copy broad host-range *Streptomyces* plasmid: Functional analysis and development of DNA cloning vectors. *Mol. Gen. Genet. MGG* **185**, 223–238 (1982).
390. Metelev, M. *et al.* Structure, bioactivity, and resistance mechanism of streptomonomicin, an unusual lasso peptide from an understudied halophilic actinomycete. *Chem. Biol.* **22**, 241–250 (2015).
391. Montalvo, N. F., Mohamed, N. M., Enticknap, J. J. & Hill, R. T. Novel actinobacteria from marine sponges. *Antonie Van Leeuwenhoek* **87**, 29–36 (2005).
392. Metsä-Ketelä, M. *et al.* Molecular evolution of aromatic polyketides and comparative sequence analysis of polyketide ketosynthase and 16S ribosomal DNA genes from various *Streptomyces* species. *Appl. Environ. Microbiol.* **68**, 4472–4479 (2002).
393. Siezen, R. J. & Khayatt, B. I. Natural products genomics. *Microb. Biotechnol.* **1**, 275–282 (2008).
394. Mahajan, Girish, B. Antibacterial agents from actinomycetes - A review. *Front. Biosci.* **E4**, 240 (2012).
395. Lee, J. *et al.* Marinopyrones A–D, α-pyrones from marine-derived actinomycetes of the family Nocardiopsaceae. *Tetrahedron Lett.* **57**, 1997–2000 (2016).
396. Rohr, J. *et al.* Metabolic products of microorganisms. 249. Tetracenomycins B3 and D3, key intermediates of the elloramycin and tetracenomycin C biosynthesis. *J. Antibiot. (Tokyo).* **41**, 1066–1073 (1988).

397. Metsä-Ketelä, M., Niemi, J., Mäntsälä, P. & Schneider, G. Anthracycline biosynthesis: Genes, enzymes and mechanisms. in 101–140 (2007). doi:10.1007/128_2007_14
398. Siitonen, V. *et al.* Divergent non-heme iron enzymes in the nogalamycin biosynthetic pathway. *Proc. Natl. Acad. Sci.* **113**, 5251–5256 (2016).
399. Wang, Y. *et al.* Modifications of aclacinomycin T by aclacinomycin methyl esterase (RdmC) and aclacinomycin-10-hydroxylase (RdmB) from *Streptomyces purpurascens*. *Biochim. Biophys. Acta - Protein Struct. Mol. Enzymol.* **1480**, 191–200 (2000).
400. Jansson, A., Koskiniemi, H., Mäntsälä, P., Niemi, J. & Schneider, G. Crystal structure of a ternary complex of DnrK, a methyltransferase in daunorubicin biosynthesis, with bound products. *J. Biol. Chem.* **279**, 41149–41156 (2004).
401. Jansson, A., Niemi, J., Lindqvist, Y., Mäntsälä, P. & Schneider, G. Crystal structure of aclacinomycin-10-hydroxylase, a S -adenosyl- l -methionine-dependent methyltransferase homolog involved in anthracycline biosynthesis in *Streptomyces purpurascens*. *J. Mol. Biol.* **334**, 269–280 (2003).
402. Jansson, A. *et al.* Aclacinomycin 10-hydroxylase is a novel substrate-assisted hydroxylase requiring S-adenosyl-l-methionine as cofactor. *J. Biol. Chem.* **280**, 3636–3644 (2005).
403. Grocholski, T., Dinis, P., Niiranen, L., Niemi, J. & Metsä-Ketelä, M. Divergent evolution of an atypical S -adenosyl- l -methionine–dependent monooxygenase involved in anthracycline biosynthesis. *Proc. Natl. Acad. Sci.* **112**, 9866–9871 (2015).
404. Lin, Y. *et al.* Draft genome of *Streptomyces zinciresistens* K42, a novel metal-resistant species isolated from copper-zinc mine tailings. *J. Bacteriol.* **193**, 6408–6409 (2011).
405. Doroghazi, J. R. *et al.* A roadmap for natural product discovery based on large-scale genomics and metabolomics. *Nat. Chem. Biol.* **10**, 963–968 (2014).
406. Barreiro, C. *et al.* Draft genome of *Streptomyces tsukubaensis* NRRL 18488, the producer of the clinically important immunosuppressant tacrolimus (FK506). *J. Bacteriol.* **194**, 3756–3757 (2012).
407. Cruz-Morales, P. *et al.* Actinobacteria phylogenomics, selective isolation from an iron oligotrophic environment and siderophore functional characterization, unveil new desferrioxamine traits. *FEMS Microbiol. Ecol.* **93**, (2017).
408. Ramachanderan, R. & Schaefer, B. Tetracycline antibiotics. *ChemTexts* **7**, 18 (2021).
409. Thallinger, B., Prasetyo, E. N., Nyanhongo, G. S. & Guebitz, G. M. Antimicrobial enzymes: An emerging strategy to fight microbes and microbial biofilms. *Biotechnol. J.* **8**, 97–109 (2013).
410. Keane, R. & Berleman, J. The predatory life cycle of *Myxococcus xanthus*. *Microbiology* **162**, 1–11 (2016).
411. Deane, C. D. & Mitchell, D. A. Lessons learned from the transformation of natural product discovery to a genome-driven endeavor. *J. Ind. Microbiol. Biotechnol.* **41**, 315–331 (2014).
412. Hackl, S. & Bechthold, A. The gene *bldA* , a regulator of morphological differentiation and antibiotic production in *Streptomyces*. *Arch. Pharm. (Weinheim).* **348**, 455–462 (2015).
413. Kwan, D. H. & Schulz, F. The stereochemistry of complex polyketide biosynthesis by modular polyketide synthases. *Molecules* **16**, 6092–6115 (2011).
414. Shrestha, B. *et al.* Engineering *Streptomyces peucetius* for doxorubicin and daunorubicin biosynthesis. in 191–209 (2019). doi:10.1007/978-3-030-01881-8_7
415. Karas, J. A. *et al.* Structure–activity relationships of daptomycin lipopeptides. *J. Med. Chem.* **63**, 13266–13290 (2020)

TURUN
YLIOPISTO
UNIVERSITY
OF TURKU

ISBN 978-951-29-9287-4 (PRINT)
ISBN 978-951-29-9288-1 (PDF)
ISSN 0082-7002 (Print)
ISSN 2343-3175 (Online)

Painosalama Oy, Turku, Finland 2023