

# UC San Diego

## UC San Diego Electronic Theses and Dissertations

### Title

A Bacterial Cytological Profile (BCP) Guided Approach to Mining Bacterial Natural Products and Synthetic Compound Libraries for Antibiotic Discovery

### Permalink

<https://escholarship.org/uc/item/6jk0t24x>

### Author

Montano, Elizabeth T

### Publication Date

2020

### Supplemental Material

<https://escholarship.org/uc/item/6jk0t24x#supplemental>

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA SAN DIEGO

A Bacterial Cytological Profile (BCP) Guided Approach to Mining Bacterial Natural  
Products and Synthetic Compound Libraries for Antibiotic Discovery

A dissertation submitted in partial satisfaction of the  
requirements for the degree Doctor of Philosophy

in

Biology

by

Elizabeth Terese Montano

Committee in charge:

Professor Joseph Pogliano, Chair  
Professor Victor Nizet, Co-Chair  
Professor Eric Allen  
Professor James Golden  
Professor Paul Jensen  
Professor Kit Pogliano

2020



Copyright

Elizabeth Terese Montano, 2020

All rights reserved.

The Dissertation of Elizabeth Terese Montano is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

---

---

---

---

---

Co-Chair

---

Chair

University of California San Diego

2020

## **DEDICATION**

For my Lord and Savior, Jesus Christ.

For my wife, Monica.

For my grandparents, Cres and Elizabeth.

For my parents, Victor and Christine.

For my brother, Eric.

## TABLE OF CONTENTS

<b>SIGNATURE PAGE</b> .....	<b>iii</b>
<b>DEDICATION</b> .....	<b>iv</b>
<b>TABLE OF CONTENTS</b> .....	<b>v</b>
<b>LIST OF FIGURES</b> .....	<b>viii</b>
<b>LIST OF TABLES</b> .....	<b>x</b>
<b>LIST OF SUPPLEMENTAL FILES</b> .....	<b>xii</b>
<b>ACKNOWLEDGEMENTS</b> .....	<b>xiii</b>
<b>VITA</b> .....	<b>xv</b>
<b>ABSTRACT OF THE DISSERTATION</b> .....	<b>xvi</b>
<b>Chapter 1: Introduction to the dissertation</b> .....	<b>1</b>
1.1 Discovery of natural product antibiotics.....	2
1.2 Antibiotic resistance.....	5
1.3 Caves harbor great potential for drug discovery.....	5
1.3.1 Cave microbial ecology .....	5
1.3.2 Bacteria in caves .....	6
1.4 Role of antibiotics among cave bacteria .....	9
1.4.1 Caves are an ideal ecological niche to look for antibiotics.....	9
1.4.2 Selective pressures in caves that may affect antibiotic production.....	10
1.4.3 Communication through sublethal concentrations of antibiotics as an adaptation to abiotic and biotic factors in caves .....	11
1.4.4 The Atacama Desert is a hyper-arid environment that might contain similar microbes.....	13
1.5 References.....	14
<b>Chapter 2: Diversity and antibacterial activity of bacteria cultured from four semi-         arid carbonate limestone caves (Carlsbad, NM, USA)</b> .....	<b>23</b>
2.1 Abstract.....	24
2.2 Introduction.....	24
2.3 Materials and methods .....	27
2.3.1 Study sites: four subterranean limestone caves .....	27
2.3.2 Description of sample sites in each cave .....	28
2.3.3 Sample collection.....	28
2.3.4 Isolation and maintenance of bacterial cave strains.....	29
2.3.5 Bacterial genomic DNA extraction and quantification for 16S rRNA PCR amplification and sequencing .....	29
2.3.6 16S rRNA PCR amplification and sequencing.....	30
2.3.7 Phylogenetic analyses of bacterial isolates .....	30
2.3.8 Bacterial cytological profiling (BCP) on LB, ISP2, AIA .....	31
2.3.9 Genomic DNA extraction for PacBio whole-genome sequencing .....	32

2.3.10 Whole-genome sequencing, assembly, and annotation .....	32
2.3.11 Genome mining for secondary metabolic biosynthetic gene clusters.....	32
2.4 Results and discussion .....	33
2.4.1 Study sites: 4 carbonate limestone caves.....	33
2.4.2 Bacterial sample collection.....	33
2.4.3 Isolation, taxonomic identification, and genus level phylogeny.....	34
2.4.4 Bacterial cytological profiling of cave bacteria natural products on LB plates	34
2.4.5 Pathogen inhibition by cave isolates.....	37
2.4.6 Bacterial cytological profiling of cave bacteria natural products on different media.....	37
2.4.7 Genome sequencing.....	38
2.4.8 Biosynthetic capacity of 4 cave bacteria.....	38
2.5 Acknowledgements.....	38
2.6 Figures.....	40
2.7 Tables.....	53
2.8 References.....	65

<b>Chapter 3: Novel antibiotics inhibit DNA replication by targeting thymidylate kinase in <i>Escherichia coli</i> .....</b>	<b>72</b>
3.1 Abstract.....	73
3.2 Introduction.....	73
3.3 Materials and methods .....	75
3.3.1 Synthetic compound library screen.....	75
3.3.2 Minimum inhibitory concentration (MIC) determination.....	75
3.3.3 Bacterial cytological profiling (BCP) .....	76
3.3.4 Cell viability.....	76
3.3.5 Resistant mutant generation.....	77
3.3.6 Genomic DNA extraction and quantification .....	78
3.3.7 Genome sequencing, assembly, and variant analyses.....	78
3.3.8 Primer design and polymerase chain reaction (PCR) .....	80
3.3.9 <i>E. coli</i> genome manipulation by P1 virulent transduction.....	80
3.3.10 TMK expression in <i>E. coli</i> JP313 $\Delta tolC$ .....	82
3.3.11 Molecular docking .....	83
3.3.12 <i>E. coli</i> TMK IC <sub>50</sub> determination .....	83
3.4 Results and discussion .....	85
3.4.1 BCP identifies the cellular pathway(s) inhibited by two PAINS.....	85
3.4.2 Viability of <i>E. coli</i> $\Delta tolC$ cells treated with compounds 1 and 2.....	86
3.4.3 Isolation of resistant mutations identifies Tmk as a potential target .....	86
3.4.4 Docking of compound 1 to thymidylate kinase .....	87
3.4.5 Screening analogs of compounds 1 and 2.....	88
3.4.6 Mutant allele Tmk R2 confers resistance.....	88
3.4.7 Tmk overexpression confers resistance .....	89
3.3.8 Compounds inhibit thymidylate kinase <i>in Vitro</i> .....	89

3.5 Acknowledgements.....	91
3.6 Figures.....	92
3.7 Tables.....	96
3.8 References.....	103
<b>Chapter 4: Isolation and characterization of <i>Streptomyces sp.</i> bacteriophages and the biosynthetic arsenals of their associated hosts .....</b>	<b>105</b>
4.1 Abstract.....	106
4.2 Introduction.....	106
4.3 Materials and methods .....	108
4.3.1 Sample collection and site description.....	108
4.3.2 Isolation of <i>Streptomyces sp.</i> ....	109
4.3.3 Phage isolation and purification.....	109
4.3.4 Bacterial genomic DNA extraction and quantification for 16S rRNA PCR amplification and sequencing .....	110
4.3.5 Bacterial genomic DNA extraction for PacBio whole-genome sequencing...110	
4.3.6 Bacterial whole-genome sequencing, assembly, and annotation.....	111
4.3.7 Genomic analysis of secondary metabolite biosynthetic potential .....	111
4.3.8 Phage genomic DNA extraction .....	111
4.3.9 Phage genome sequencing, assembly, and annotation.....	111
4.3.10 16S rRNA PCR amplification and sequencing.....	112
4.3.11 CRISPR-Cas sequence analysis and predictions .....	112
4.3.12 Phylogenetic analyses of bacterial isolates .....	113
4.3.13 Cross-streak method for assessing antibacterial production potential .....	113
4.3.14 Bacterial cytological profiling (BCP) on plates.....	114
4.3.15 Host range experiment .....	114
4.3.16 Transmission electron microscopy .....	115
4.3.17 Strains used in this study .....	115
4.4 Results and discussion .....	116
4.4.1 Isolation and antibacterial activities of <i>Streptomyces sp.</i> ....	116
4.4.2 Mechanistic analysis of natural products produced by four <i>Streptomyces</i> isolates.....	117
4.4.3 Genomic analysis of four <i>Streptomyces</i> isolates.....	118
4.4.4 Antimicrobial activity of four <i>Streptomyces</i> isolates against clinically relevant pathogens .....	120
4.4.5 Phage isolation and host range.....	121
4.4.6 Characterization of CRISPR elements in the genomes of our phages.....	122
4.5 Acknowledgements.....	124
4.6 Figures.....	126
4.7 Tables.....	134
4.8 References.....	147
<b>Chapter 5: Concluding remarks.....</b>	<b>154</b>

## LIST OF FIGURES

### Chapter 2

<b>Figure 2.1</b>	Map displaying the location of four carbonate caves in the Guadalupe Mountains (Carlsbad Caverns National Park, New Mexico, USA).....	40
<b>Figure 2.2</b>	The maximum likelihood phylogeny of 334 isolates cultured from caves and 100 bacterial type strain reference sequences .....	41
<b>Figure 2.3</b>	Phylum level bacterial diversity of the isolates cultured from four carbonate limestone caves.....	42
<b>Figure 2.4</b>	Phylum level bacterial diversity of the isolates cultured from carbonate limestone caves .....	43
<b>Figure 2.5</b>	Screening cave bacteria for antibiotic production .....	44
<b>Figure 2.6</b>	Cave bacteria produce Gram-negative inhibitory bioactive compounds using a variety of mechanisms to target different cellular pathways.....	45
<b>Figure 2.7</b>	The maximum likelihood phylogeny of 84 isolates cultured from caves and 17 bacterial type strain reference sequences .....	46
<b>Figure 2.8</b>	<i>Bacillus toyonensis</i> strains exhibit different phenotypes when tested against <i>E. coli tolC</i> on LB.....	47
<b>Figure 2.9</b>	The maximum likelihood phylogeny of 9 isolates cultured from caves and 6 bacterial type strain reference sequences .....	48
<b>Figure 2.10</b>	BCP analysis of 5 <i>Bacillus</i> cave strains.....	49
<b>Figure 2.11</b>	BCP analysis of 4 Actinobacteria cave strains .....	50
<b>Figure 2.12</b>	Distribution of biosynthetic class among secondary metabolic gene clusters predicted in the genomes of 9 cave isolates.....	51
<b>Figure 2.13</b>	Sequence comparison of predicted BGCs to known secondary metabolic gene clusters.....	52

### Chapter 3

<b>Figure 3.1</b>	Cytological profiles of <i>E. coli</i> JP313 $\Delta tolC$ treated with compounds 1 and 2.....	92
<b>Figure 3.2</b>	Cell viability of <i>E. coli</i> JP313 $\Delta tolC$ .....	93
<b>Figure 3.3</b>	<i>E. coli</i> thymidylate kinase with compound 1 docked to the active site .....	94
<b>Figure 3.4</b>	Cytological profiles of <i>E. coli</i> JP313 $\Delta tolC$ treated with 9 5-Benzylidenerhodanine analogs .....	95

### Chapter 4

<b>Figure 4.1</b>	Phylogenetic tree of <i>Streptomyces</i> bacteria isolated from soil samples ..	126
<b>Figure 4.2</b>	Inhibition of bacterial growth by <i>Streptomyces</i> isolates .....	127
<b>Figure 4.3</b>	Differential BCP phenotypes of <i>E. coli</i> JP313 $\Delta tolC$ exposed to natural products produced by four <i>Streptomyces</i> soil isolates grown on different solid media .....	128
<b>Figure 4.4</b>	Genome characteristics of <i>Streptomyces sp.</i> DF, SFW, QF2, and JS isolated from soil samples.....	129
<b>Figure 4.5</b>	Comparison of BGCs encoded in the genomes of bacterial soil isolates and the predicted most similar, previously characterized BGC with an antibacterial product.....	130
<b>Figure 4.6</b>	Characterization of four <i>Streptomyces</i> phage isolated from soil samples .....	131
<b>Figure 4.7</b>	Genomic maps of phages showing regions containing sequence similarity to spacers found within the CRISPRs of strains QF2, DF, and SFW .....	132
<b>Figure 4.8</b>	Class I, Type I-E CRISPR-Cas system encoded in the WGS of strain QF2 .....	133



## LIST OF TABLES

### Chapter 2

<b>Table 2.1</b>	Specialized growth media used to culture and isolate bacteria from semi-arid limestone caves.....	53
<b>Table 2.2</b>	Genus classifications of cave isolates based on maximum likelihood analysis of 16S ribosomal RNA sequence data .....	54
<b>Table 2.3</b>	100 Reference type strain 16S ribosomal RNA sequences used in maximum likelihood analysis and classification of cave isolates.....	62
<b>Table 2.4</b>	Inhibition of pathogens by bacterial isolated from caves. ....	64

### Chapter 3

<b>Table 3.1</b>	Chemical structures and MICs of compounds 1 and 2 against <i>E. coli</i> JP313 $\Delta tolC$ .....	96
<b>Table 3.2</b>	Minimum inhibitory concentration (MIC) of <i>E. coli</i> JP313 $\Delta tolC$ and two mutants treated with compounds 1 and 2.....	97
<b>Table 3.3</b>	Minimum inhibitory concentration of <i>E. coli</i> JP313 $\Delta tolC$ [2 A69T] and 8 5-Benzylidenerhodanine derivatives.....	98
<b>Table 3.4</b>	Chemical structures and minimum inhibitory concentrations of 9 analogs against <i>E. coli</i> JP313 $\Delta tolC$ .....	99
<b>Table 3.5</b>	Results of pRSFDuet-1 plasmid overexpression of <i>E. coli</i> MG1655 TMK and variants in <i>E. coli</i> JP313 $\Delta tolC$ .....	100
<b>Table 3.6</b>	IC <sub>50</sub> values calculated from an <i>in vitro</i> assay of <i>E. coli</i> TMK and 12 structural analogs .....	101
<b>Table 3.7</b>	Plasmids and strains used in this study .....	102

### Chapter 4

<b>Table 4.1</b>	Top NCBI BlastN hits of the 16S rRNA gene sequences used to create the phylogenetic tree containing eight <i>Streptomyces</i> isolates .....	134
<b>Table 4.2</b>	BGCs encoded within the draft genome sequence of strain QF2 .....	135
<b>Table 4.3</b>	BGCs encoded within the draft genome sequence of strain JS .....	136
<b>Table 4.4</b>	BGCs encoded within the draft genome sequence of strain SFW .....	137
<b>Table 4.5</b>	BGCs encoded within the draft genome sequence of strain DF .....	138
<b>Table 4.6</b>	Inhibition of growth of clinically relevant pathogens by <i>Streptomyces</i> strains DF, SFW, QF2, and JS .....	139
<b>Table 4.7</b>	Summary of the NCBI WGS annotations of four phage isolates .....	140

<b>Table 4.8</b>	Functions of the putative proteins encoded within the genomes of four phage isolates .....	141
<b>Table 4.9</b>	General characteristics of predicted CRISPR-Cas systems within the genomes of strains DF, SFW, QF2, and JS .....	142
<b>Table 4.10</b>	Characteristics of the 38 CRISPRs predicted in the draft genome sequence of strain QF2 .....	143
<b>Table 4.11</b>	Characteristics of the spacers within the genomes of strain DF, SFW, and QF2 that have sequence similarity to at least one of the four phage isolates .....	144
<b>Table 4.12</b>	Characteristics of the 12 CRISPRs predicted in the complete genome sequence of strain DF.....	145
<b>Table 4.13</b>	Characteristics of the 11 CRISPRs predicted in the draft genome sequence of strain SFW .....	146

## **LIST OF SUPPLEMENTAL FILES**

### **Chapter 3**

Supplemental figures and tables (SI Figures 1-3, SI Table 1)

## ACKNOWLEDGEMENTS

This research was made possible by the community that supported me throughout graduate school. I want to thank my advisor Dr. Joe Pogliano for allowing me to join his lab family, supporting my research, improving my presentation and interview skills, encouraging the publication of my research by hosting weekend paper-athons, and for his endless enthusiasm. I would also like to thank the rest of my dissertation committee for their help and guidance throughout my graduate studies. I'd like to acknowledge all the past and present members of both Pogliano labs for sharing their expertise and creating a fun and positive work environment. I especially want to thank Jason Nideffer who joined my research journey, and with a stellar positive attitude, took the lead on experiments, did a few he'd rather not, made 10 versions of the same figure, served as my interpreter for Joe, and was essential to the publication process. I also want to acknowledge Eray Enustun and Joseph Sugie for working very hard on our publications. My warmest thanks go to Dr. Alan Derman, Hannah Tsunemoto, and Dr. Vorrapon Chaikeratisak for seeing me through the hard times, and for being there to celebrate the best! Finally, I want to express my deepest gratitude to my wife Monica and family for being a constant beacon of hope, inspiration, love, and encouragement.

Chapter 2, in part, is being formulated into a manuscript in preparation for publication of the material. Elizabeth T. Montañó, Joseph Sugie, Eray Enustun, Julia Busch, Diana E. Northup, Kit Pogliano, Joe Pogliano, 2020. The dissertation author was the primary investigator and author of this material.

Chapter 3, in full, has been formulated into a manuscript that will be submitted for publication in 2020. Elizabeth T. Montaña, Jason F. Nideffer, Joseph Sugie, Eray Enustun, Adam B. Shapiro, Alan I. Derman, Kit Pogliano, Joe Pogliano. The dissertation author was the primary investigator and author of this material.

Chapter 4, in full, has been submitted for publication of the material as it may appear in Nature Science Reports, 2020. Elizabeth T. Montaña, Jason F. Nideffer, Lauren Brumage, Marcella Erb, Alan I. Derman, John Paul Davis, Elena Estrada, Sharon Fu, Danielle Le, Aishwarya Vuppala, Cassidy Tran, Elaine Luterstein, Shivani Lakkaraju, Sriya Panchagnula, Caroline Ren, Jennifer Doan, Sharon Tran, Jamielyn Soriano, Yuya Fujita, Pranathi Gutala, Quinn Fujii, Minda Lee, Anthony Bui, Carleen Villarreal, Samuel R. Shing, Sean Kim, Danielle Freeman, Vipula Racha, Alicia Ho, Prianka Kumar, Kian Falah, Eray Enustun, Amy Prichard, Ana Gomez, Kanika Khanna, Shelly Trigg, Kit Pogliano, Joe Pogliano. The dissertation author was the primary investigator and author on these studies.

## **VITA**

- 2010 Bachelor of Science, University of New Mexico
- 2020 Doctor of Philosophy, University of California San Diego

## **ABSTRACT OF THE DISSERTATION**

A Bacterial Cytological Profile (BCP) Guided Approach to Mining Bacterial Natural Products and Synthetic Compound Libraries for Antibiotic Discovery

by

Elizabeth Terese Montano

Doctor of Philosophy in Biology

University of California San Diego, 2020

Professor Joseph Pogliano, Chair  
Professor Victor Nizet, Co-Chair

The exponential increase in multi-drug resistant pathogens is unmet by the number of effective antibiotics to treat their infections, posing a serious threat to human global health. To address this issue, we need new sources and better tools to discovery

new antibiotics. Bacterial cytological profiling (BCP) is a new tool that rapidly identifies the mechanism of action of bioactive compounds and is applicable to natural product discovery and serves as a guide to increase the hit to lead ratio among chemical library screens. This dissertation includes five chapters that explore bacterial natural products and synthetic small molecules, guided by BCP, for the discovery of new antibacterials. The first chapter introduces potential sources for novel discoveries and describes the utility of BCP, followed by three research chapters and a summary chapter.

Chapter 2 describes the diversity and antibacterial activity of bacteria cultured and isolated from four limestone caves in Carlsbad Caverns National Park (NM, USA). A genus level phylogenetic tree was generated to display the diversity of the bacteria. The antibacterial activities were characterized via bacterial cytological profiling (BCP) and the cross streak method. The biosynthetic potential of four Actinobacteria and five *Bacillus sp.* was investigated via genome mining for biosynthetic gene clusters.

Chapter 3 presents the discovery of novel antibiotics that inhibit DNA replication by targeting *E. coli* thymidylate kinase (TMK). BCP was used to identify these molecules from a synthetic small molecule library screen and to guide the structure activity relationship (SAR) studies aimed to identify compounds with improved bioactivity that also target DNA.

The final research chapter presents a study of the isolation and characterization of *Streptomyces sp.* bacteriophages and describes the biosynthetic arsenals of their associated hosts. Four *Streptomyces sp.* were isolated from soil samples and their bioactivity was determined against strains of *E. coli* and *B. subtilis*. Additionally,



bacteriophages were isolated from the soil samples, their genomes were sequenced, assembled, and annotated, and their host range was described. The bacterial genomes were assessed for known phage immunity factors to further characterize the host range of each bacteriophage and elucidate their potential use for *Streptomyces* genetic manipulation.

## **Chapter 1: Introduction to the dissertation**



## 1.1 Discovery of natural product antibiotics

Secondary metabolic natural products are non-essential molecules biosynthesized by living organisms<sup>1</sup>. In certain microenvironments, these metabolites serve as an important evolutionary strategy by increasing the overall fitness of producers and function as niche specific drivers of microbial evolution<sup>2</sup>. The specific mechanisms by which these molecules function in nature is largely unknown however some have been proposed including the formation and maintenance of symbiotic relationships among producers and non-producers in biofilms<sup>3</sup>, and competitive biological warfare<sup>4,5</sup>. Soil dwelling Actinobacteria, particularly *Streptomyces sp.*, have been the leading source of clinically useful compounds with novel chemistries and differential bioactivities including immunosuppressants, anti-cancer agents, antifungals, and antibiotics<sup>6-15</sup>.

The discovery of antibiotics has been one of the greatest contributions to increasing the longevity of the human lifespan<sup>16</sup>. Prior to the discovery of antibiotics in the late 19<sup>th</sup> and 20<sup>th</sup> centuries, the leading cause of death was untreated bacterial infections including pneumonia, diarrhea and diphtheria (cholera), and tuberculosis<sup>17</sup>. The advent of medical bacteriology sprung from the correlation between disease and the bacteria that cause them and marked the dawn of a new era in human global health. The correlation was a result of Robert Koch's discoveries of the anthrax disease cycle (*Bacillus anthracis*, 1876) and the bacteria responsible for tuberculosis (*Mycobacterium tuberculosis*, 1882) and cholera (*Vibrio cholerae*, 1883)<sup>18,19</sup>. These findings led to efforts to discover anti-infective drugs to cure these diseases.

The first clinically useful antibiotics were chemically synthesized and discovered by screening synthetic molecular libraries for bioactivity. Paul Ehrlich discovered

salvarsan (1909), the first antibiotic, using a mouse model to systematically screen a synthetic compound library for selective toxicity<sup>17,20-22</sup>. Similar to Ehrlich's approach, Gerhard Domagk discovered Prontosil (1928)<sup>23</sup>, a broad-spectrum sulfa drug that replaced salvarsan<sup>24,25</sup>, by screening azo related dyes against pathogenic bacteria. The first clinically useful natural product antibiotic was discovered serendipitously. Alexander Fleming discovered penicillin (1928), produced by the fungus *Penicillium nonatum*, after making an astute observation of a zone of lysis surrounding a fungal contaminant on a plate containing *Staphylococci*<sup>25</sup>. Penicillin was later developed as a potent antibiotic. Its first use in 1941 saved someone's life, and since then, countless others.

Following the discovery and successful clinical application of the first natural product antibiotic, and Rene Dubos' discovery of tyrothricin (1939, *Bacillus brevis*), Selman Waksman began a systematic study of soil-dwelling microbes for drug discovery, and identified members of the genus *Streptomyces* (belonging to the phylum Actinobacteria) as the most prolific producers of bioactive natural products, including antibiotics. This led to the discovery of neomycin (1949, *Streptomyces fradiae*), and the first compound active against *M. tuberculosis*, streptomycin (1943, *Streptomyces griseus*)<sup>26</sup>. These discoveries spurred the onset of the 'Golden Age' (1940-1970) of antibiotics<sup>27</sup>, the time period wherein natural products made by soil-dwelling microbes were mined for drug discovery yielding 270 new drugs and 11 structurally distinct classes, including several antibiotics produced by soil bacteria; gentamicin (*Micromonopora*)<sup>28,29</sup>, chlortetracycline (*Streptomyces aurefaciens*<sup>30</sup>), erythromycin A (*Streptomyces erythreus*<sup>31</sup>), oxytetracycline (*Streptomyces rimosus*<sup>32,33</sup>), tetracycline

(*Streptomyces aureofaciens*<sup>34</sup>), chloramphenicol (*Streptomyces venezuelae*<sup>35,36</sup>), kanamycin (*Streptomyces kanamyceticus*<sup>37</sup>), lincomycin (*Streptomyces lincolnensis*<sup>38</sup>), vancomycin (*Amycopolis orientalis*<sup>39</sup>), rifamycin (*Amycopolis mediterranei*<sup>40,41</sup>), novobiocin (*Streptomyces spheroides*<sup>42</sup>), specinomycin (*Streptomyces spectabilis*<sup>37</sup>), and daptomycin (*Streptomyces roseosorus*<sup>43,44</sup>). In addition to treating infectious diseases, antibiotics made many modern medical procedures possible, including high-risk surgical procedures, and cancer treatment<sup>27</sup>.

The ‘Golden Age’ drew to an end when the historically successful methods for natural product drug discovery were no longer useful. The traditional workflow included (1) the collection of a wide variety of soil samples from local and exotic locations, (2) isolation of the most ‘interesting’ microbes (identified on the basis of colony formation, and color), (3) fermentation, (4) crude extract screening (5) activity guided fractionation, (6) compound isolation, and (7) structure elucidation. A steep decrease in the number of novel drug discoveries<sup>45,46</sup> resulted from re-isolations of known bacterial strains and bioactive compounds became more frequent, making the task more laborious, and yielded fewer and fewer novel compounds<sup>47</sup>.

The years following the ‘Golden Age’ marked a new paradigm, ushering in completely new approaches to drug discovery. Microbes that are being mined for natural products come from novel taxa and unmined ecosystems in geographical locations that harbor unique bacterial communities. The advent and relatively low costs of whole genome sequencing has opened up novel discovery methods including investigating the natural product biosynthetic potential before applying the traditional methods for identifying the bioactive compound of interest (genome mining). Novel methods to speed

up the identification and purification of natural products with interesting bioactivities include molecular structure prediction on the basis of a natural product gene cluster, cloning a natural product cluster into an expression vector and transforming it into a strain genetically modified not to produce bioactive natural products. This process can also lead to cluster refactoring to produce a scaffold of interest bypassing synthetic chemistry.<sup>5,6,48,49,7-10,12-15</sup> Finally, fluorescence microscopy has been used as a powerful tool to rapid identify the mechanism of action of bioactive compounds via phenotypes of treated bacteria, which is still a major bottleneck in the discovery pipeline<sup>49,50</sup>.

## **1.2 Antibiotic resistance**

With the discovery of antibiotics came resistance. The pathogenic bacteria responsible for causing infections possess a variety of modes of resistance (inherent or acquired through mutations and genetic exchange) to a broad range of antibiotics<sup>51,52</sup>. Persistent overuse and misuse of these precious compounds over the past three decades further diversified the pool of acquired modes of resistance and greatly contributed to the rapid and widespread emergence of multi-drug resistant pathogens<sup>53,54</sup>. Antibiotic resistance is spread rapidly by horizontal transfer and clonal expansion, and as a result, populations of resistant pathogens are exponentially increasing, and the probability of newly- and re- emerging pathogens is increased significantly. As a result, diseases that were once easily treatable are deadly again because the available treatment options are limited or non-existent due to a lack of new antibiotics to treat them<sup>55</sup>.

## **1.3 Caves harbor great potential for drug discovery**

### **1.3.1 Cave microbial ecology.**

Caves select for diverse bacteria because of the location of caves in the subsurface, the lack of sunlight, limited primary nutrient sources that occur inorganically, rapidly changing host rock conditions over short distances (reduced compounds involved in redox reactions), and zonation<sup>56-61</sup>. Biosynthesis of natural products may be an important evolutionary strategy by which cave bacteria establish and maintain symbiotic relationships. For example, the natural product desferroxamine, which is produced by many *Streptomyces*, is required for some bacteria to be cultured *in vitro*<sup>3</sup>. This symbiotic relationship likely formed as an evolutionary mechanism to efficiently utilize the limited resources available in the cave environment. Cave bacteria are known to have evolved to use extremely small amounts of carbon<sup>62</sup>, fix CO<sub>2</sub> present in the cave, and mine the bedrock for inorganic electron sources<sup>63</sup> so working in symbiotic teams to gain fitness in an oligotrophic environment is not a far stretch for microbial evolution. The biosynthesis of natural products by bacteria in nature may be a mechanism of communication<sup>64,65</sup> ecological space competition, and perhaps both<sup>66,67</sup>. In tight knit microbial communities where symbioses occur<sup>68</sup>, the use of natural products for communication and weaponry may become more necessary than in other instances. Caves are well preserved ecosystems where phylogenetically novel bacteria and producers of novel natural products likely reside. Cave microbial ecosystems may represent a new source of antibiotics, and much remains to be explored. Less than fifty percent of caves in well studied areas have been subject to research and even fewer for natural products<sup>59</sup>.

### **1.3.2 Bacteria in caves**

Caves are multifarious environments populated by a variety of novel bacterial species nested in phyla that are known producers of natural products. Caves are inhabited



by a multitude of novel *Actinobacteria*<sup>69</sup>. There is strong potential for discovering novel adaptations and secondary metabolite pathways used to synthesize antibiotics within microbial genomes residing deep within Lechuguilla Cave, the deepest cave in the continental US<sup>70</sup>.

Data resulting from 16S rDNA sequencing of cultured bacterial cave isolates (from lava tubes in Hawaii and carbonate caves in New Mexico, USA) has revealed many novel species including those belonging to entirely new genera and families. Among these are many belonging to phyla that were previously determined as the most prolific natural product producers, including: Actinobacteria,  $\beta$ -Proteobacteria,  $\delta$ -Proteobacteria, Cyanobacteria, and the less-studied phyla Verrucomicrobia and Chloroflexi. The Actinobacteria comprised the greatest portion of the diversity found among all six phyla and are known as the most prolific producers of natural products with antibiotic activity<sup>71</sup>.

Distinct microbial communities exist in subsurface caves<sup>72-74</sup>. A diverse community of microorganisms also inhabit carbonate caves, deep within the extreme aphotic and oligotrophic subsurface<sup>69,75,76</sup>. Study results from 454 Life Sciences next generation pyro-sequencing of bacterial isolates in Lava Beds National Monument have revealed substantial differences in the phyla found in the cave versus surface soils. The ecological parameters that exist in caves and function as the drivers of evolution have previously been investigated for cave-adapted animals (troglobionts)<sup>77</sup> but little research exists on how microorganisms adapt to the abiotic and biotic selective pressures of caves in karstic and volcanic landscapes<sup>78</sup>. Microbial adaptations to the extreme subsurface environment may include slowed growth rates, unusual phenotypes, and metabolic

modification (e.g. antibiotics) etc. for survival and reproduction optimization. The following is an exploration of the possible observations of evolution among bacteria that are a result of cave specific selective pressures that may influence antibiotic production among these organisms.

Bacterial inhabitants of limestone caves have been investigated using culture-independent techniques and revealed the presence of novel species within the phyla Actinobacteria, Proteobacteria, Archaea, Planctomyces, Chloroflexi, Acidobacteria, Thiobacillus, Flexibacter, Cytophaga, and Bacteroides. This technique usually involves the extraction of stable marker compounds that are useful in representing the community and variable across interested microbial communities. Phylogenetic information is also contained in the markers allowing for less ambiguous profiling. For example, biomarkers exist for RNA and DNA molecules. This involves nucleic acid extraction, PCR amplification, and Genetic Fingerprinting Techniques for community profiling, sequencing, and phylogenetic analysis<sup>74</sup>.

Denaturing Gradient Gel Electrophoresis is used to identify bacterial species present in a mixed sample. Whereas this method is becoming replaced by more modern technology such as whole genome sequencing, it remains less expensive than community sequencing and is still used today<sup>79-81</sup>. Upon confirmation of a successful PCR, the collective PCR products from each sample are separated via DGGE to obtain an overall gene diversity fingerprint. DGGE separates the PCR products over an increasing urea-formamide gradient acrylamide gel based on their unique denaturation point. When a pool of double stranded DNA is exposed to an increasing concentration of urea in the acrylamide gel, the fragments will denature into single strands at a specific urea

concentration for each individual represented in the sample. The resulting banding patterns allow for separation of individual species within the community structure.

## **1.4 Role of antibiotics among cave bacteria**

### **1.4.1. Caves are an ideal ecological niche to look for antibiotics**

There is strong evidence for bacteria producing novel antimicrobial compounds in caves, and this ecosystem has been largely unexplored in a systematic manner for novel antibiotics in the past. Resistance genes recovered from bacterial isolates in Lechuguilla, a 4 million year old carbonate cave provide circumstantial evidence for the production of antimicrobial compounds for which the bacteria have synthetic pathway genes<sup>78,82-91</sup>. Bacterial species belonging to phyla that are known prolific natural product producers have been cultured and sequenced from the cave environment. Caves are a promising habitat wherein to look for novel antibiotic production because they are a terrestrial dark life ecosystem with distinct ecological pressures from the surface that may foster the evolution of novel mechanisms responsible for the production of novel natural products. For example, Naowarat Cheeptham and her team of researchers have discussed the potential for antibiotic production to occur among bacteria inhabiting cave biofilms. Their hypothesis is backed by the need to secrete antimicrobial compounds for the purpose of inhibiting nearby cells because competition for resources in a high cell density biofilm is fierce<sup>74</sup>. Previous studies indicate that antibiotic production is higher in a biofilm than among planktonic cells making biofilms in cave a promising microenvironment wherein to search for novel antibiotics<sup>82</sup>. A 1999 study by Burgess et al. reports that bacteria producing natural products are more frequently associated with in marine biofilms than in planktonic populations.<sup>74,78,91,92,83-90</sup>

While soil samples have been analyzed from remote and exotic terrain, caves are ideal study sites for antibiotic research because in addition to soil they house a variety of distinct micro-habitats and to date remain understudied for these purposes. These subterranean locations have the potential to enhance our understanding of diversity, biotechnological potential (fuel sources, medicine), and ecology (nutrient cycling, adaptation). In a recent paper by Bull and Asenjo, Marcel Jaspars of Aberdeen University references new bioactive compounds produced by microbes from volcanic springs, industrial waste sites, deep-seas, cryo-environments, and areas of high saline, pH, and metal contamination, providing support for this notion<sup>93</sup>. Researchers are waiting for a breakthrough to emerge from novel environmental sources because despite the last 15-20 years of sampling these kinds of natural habitats, no commercially viable compound has yet emerged.

#### **1.4.2. Selective pressures in caves that may affect antibiotic production**

Abiotic and biotic factors associated with natural habitats work together and act on their inhabitants as the driving forces of evolution. Evolutionary adaptations include metabolic functions (nutrient and energy sources), phenotypic expressions, community structure, distribution, and diversity. Microorganisms provide the greatest contribution to the evolutionary diversity that exists among living organisms. Specifically, Bacteria and Archaea are the domains that are comprised of microorganisms that have had the longest historical time to evolve as a result of the aforementioned abiotic and biotic selective pressures. One study investigating novel drug production among psychrophilic and psychrotrophic bacterial strains (*Serratia proteamaculans*) isolated from soil in the Isla de los Estados-Ushuaia Argentina reported the influence of two key abiotic ecological

factors on antibiotic production<sup>94,95</sup>. The temperature and nutrient availability during growth conditions are essential to creating the ideal conditions for antibiotic production. Isolates were grown in a variety of nutrients coupled with varying temperatures (4-45°C). The antimicrobial compound (Serraticin A) was produced at lower temperatures (4-8°C) whereas no production occurred at higher temperatures (30,37,45°C) despite the fact that cells were visibly present (perhaps not yet in stationary phase when bioactive secondary metabolites including antibiotics are most often produced) at these temperatures. This indicated adaptation to a cold environment and demonstrates the role of growth conditions on antimicrobial production. In addition to the low temperature that mimics the natural cold ecological conditions from where the bacterium was isolated, nutrients played a key role in antibiotic production. Antibiotic production was the highest when the isolate was grown in LB a nutrient rich media. This is counter intuitive because isolates producing similar antibiotics to Serraticin A (microcins) required minimal or nutrient poor media to obtain maximum production<sup>95</sup>. Ecological factors may play an active role in bacterial antibiotic production in nature.

#### **1.4.3 Communication through sublethal concentrations of antibiotics as an adaptation to abiotic and biotic factors in caves**

Support for the production of antibiotics for the purpose of bacterial communication exists in a paper by Deziel<sup>96</sup>. They propose that antibiotics are signaling molecules that are toxic at high doses. This theory stems from experiments where microbes were exposed to antibiotics at sublethal concentrations. The target bacteria did not ignore these compounds, rather the compounds elicited a response. These responses included the formation of biofilms, activation of the cells' SOS response, and becoming

more or less virulent. Further support for the communication hypothesis comes from Linares et al. (2006)<sup>64</sup> who state that microbes generate signals to coordinate mutually beneficial activities. Examples of these activities include antibiotics that kill prey and suppress or deter competitors. Sublethal concentrations of these antibiotics can elicit changes in microbial metabolism, behavioral change (trigger fleeing, hiding in a biofilm, and manipulation). Antibiotics tobramycin (produced by *Streptomyces tenebrarius*) and tetracycline (produced by *Streptomyces aureofaciens*) are effective against *Pseudomonas aeruginosa* by causing cell death and inhibiting their growth respectively. These antibiotics also elicit an increase in motility (triggers fleeing) and biofilm formation respectively. These are members of the Actinobacteria, the most prevalent group of bacteria inhabiting many caves.

There could be many antimicrobial compounds present at sublethal concentrations in an oligotrophic environment where communication may be more important. If we use the biofilm formation among cave bacteria as an example, we can speculate that this formation may have occurred because of the presence of many bioactive metabolites including antimicrobials present at low concentrations in the environment. Biofilm formation may reduce the potency of an antibiotic attack, in turn increasing the fitness of all members in the biofilm. If the above is true of biofilm formation and the widespread presence of antimicrobial compounds, the greatest variety in caves may be revealed in these biofilms. Additionally, many antibiotics produced in nature are potentially only produced in the presence of other community members. Maintaining the community could therefore be key to increasing the array of antibiotic biosynthesis in a laboratory setting.<sup>97-99</sup>

#### **1.4.4 The Atacama Desert is a hyper arid environment that might contain similar microbes.**

In addition to caves, other potential areas for exploring for antibiotics that are underexplored relative to the soil include deserts, deep seas, and coral reefs. A recent study of the hyper arid climactic region in the Atacama Desert of northern Chile revealed high microbial diversity with members that are known sources of natural products. Similar to caves located in the semi-arid region of the United States, the Actinobacteria contain a large proportion of the microbial diversity found in the Atacama Desert. Within this Desert ecosystem, this phylum revealed species that produce bioactive secondary metabolites (natural products) that display a wide range of beneficial activities including anti-inflammatory, anti-cancer, and antibiotic. This study states found increased levels of taxonomic diversity among the Actinobacteria is likely paired with increased potential to produce novel bioactive compounds. Of the twelve genera isolated and sequenced from the Atacama Desert, there was a high representation of species from the *Streptomyces* that contained a distinct clade (likely comprised of salar-adapted ecovars) that displayed antimicrobial activity with novel mechanisms of action. This clade was used as the determinant of the group of interest<sup>100</sup>. Further, Jaspar and his group, in order to compliment his organism selection (*Modestobacter*) used MALDI-TOF mass spectrometry and its potency demonstrated by the identification of a novel peptide produced by members within this clade. The genome of *Streptomyces sp.* C34 was mined and targeted the gene encoding 3-amino-5-hydroxybenzoic acid (AHBA) synthase revealing ansamycins. Among these, chaxamycin D displayed antibacterial activity against MRSA and inhibit Hsp90 (potential antitumor agent) while other chemical

entities were promising antibiotics. In order to identify as many of the chemicals with antibiotic activity produced by *Streptomyces sp.* C34, fermentation conditions were varied to express genes that are silent under other conditions and revealed the production of a novel lasso peptide, and chaxalactins, rare molecules that are active against Gram positive bacteria. The extreme nature of the Atacama Desert is similar to that of semi-arid caves in nutrient and liquid water availability. The region has had over 100 My of aridity and 10-15 My of aridity with an annual average rainfall of <2 mm, low concentrations of carbon, and microbial communities with high oxidation capacity making the Atacama Desert comparable to the subsurface cave environment in semi-arid climates.<sup>93</sup>

## 1.5 References

1. Vining, L. Functions Of Secondary Metabolites. *Annu. Rev. Microbiol.* (1990). doi:10.1146/annurev.micro.44.1.395
2. Weiss, G., Kovalerchick, D., Lieman-Hurwitz, J., Murik, O., De Philippis, R., Carmeli, S., Sukenik, A. & Kaplan, A. Increased algicidal activity of *Aeromonas veronii* in response to *Microcystis aeruginosa*: interspecies crosstalk and secondary metabolites synergism. *Environ. Microbiol.* (2019). doi:10.1111/1462-2920.14561
3. D'Onofrio, A., Crawford, J. M., Stewart, E. J., Witt, K., Gavrish, E., Epstein, S., Clardy, J. & Lewis, K. Siderophores from Neighboring Organisms Promote the Growth of Uncultured Bacteria. *Chem. Biol.* (2010). doi:10.1016/j.chembiol.2010.02.010
4. Bednarek, P. Chemical warfare or modulators of defence responses - the function of secondary metabolites in plant immunity. *Curr. Opin. Plant Biol.* (2012). doi:10.1016/j.pbi.2012.03.002
5. Patin, N. V., Floros, D. J., Hughes, C. C., Dorrestein, P. C. & Jensen, P. R. The role of inter-species interactions in *Salinispora* specialized metabolism. *Microbiol. (United Kingdom)* (2018). doi:10.1099/mic.0.000679
6. Busch, J., Agarwal, V., Schorn, M., Machado, H., Moore, B. S., Rouse, G. W., Gram, L. & Jensen, P. R. Diversity and distribution of the *bmp* gene cluster and its Polybrominated products in the genus *Pseudoalteromonas*. *Environ. Microbiol.* (2019). doi:10.1111/1462-2920.14532
7. Tuttle, R. N., Demko, A. M., Patin, N. V., Kapon, C. A., Donia, M. S., Dorrestein, P. & Jensen, P. R. Detection of natural products and their producers in ocean sediments. *Appl. Environ. Microbiol.* (2019). doi:10.1128/AEM.02830-18



8. Castro-Falcón, G., Millán-Aguíñaga, N., Roullier, C., Jensen, P. R. & Hughes, C. C. Nitrosopyridine Probe to Detect Polyketide Natural Products with Conjugated Alkenes: Discovery of Novodaryamide and Nocarditriene. *ACS Chem. Biol.* (2018). doi:10.1021/acscchembio.8b00598
9. Kim, M. C., Machado, H., Jang, K. H., Trzoss, L., Jensen, P. R. & Fenical, W. Integration of Genomic Data with NMR Analysis Enables Assignment of the Full Stereostructure of Neaumycin B, a Potent Inhibitor of Glioblastoma from a Marine-Derived Micromonospora. *J. Am. Chem. Soc.* (2018). doi:10.1021/jacs.8b04848
10. Letzel, A. C., Li, J., Amos, G. C. A., Millán-Aguíñaga, N., Ginigini, J., Abdelmohsen, U. R., Gaudêncio, S. P., Ziemert, N., Moore, B. S. & Jensen, P. R. Genomic insights into specialized metabolism in the marine actinomycete *Salinispora*. *Environ. Microbiol.* (2017). doi:10.1111/1462-2920.13867
11. Zhou, X., Liang, Z., Li, K., Fang, W., Tian, Y., Luo, X., Chen, Y., Zhan, Z., Zhang, T., Liao, S., Liu, S., Liu, Y., Fenical, W. & Tang, L. Exploring the Natural Piericidins as Anti-Renal Cell Carcinoma Agents Targeting Peroxiredoxin 1. *J. Med. Chem.* (2019). doi:10.1021/acs.jmedchem.9b00598
12. Kim, M. C., Cullum, R., Machado, H., Smith, A. J., Yang, I., Rodvold, J. J. & Fenical, W. Photopiperazines A-D, Photosensitive Interconverting Diketopiperazines with Significant and Selective Activity against U87 Glioblastoma Cells, from a Rare, Marine-Derived Actinomycete of the Family Streptomycetaceae. *J. Nat. Prod.* (2019). doi:10.1021/acs.jnatprod.9b00429
13. Ryu, M. J., Hwang, S., Kim, S., Yang, I., Oh, D. C., Nam, S. J. & Fenical, W. Meroindenon and Merochlorins e and f, Antibacterial Meroterpenoids from a Marine-Derived Sediment Bacterium of the Genus *Streptomyces*. *Org. Lett.* (2019). doi:10.1021/acs.orglett.9b01440
14. Lima, S., Chekan, J., McKinnie, S., Alvarenga, D., Luhavaya, H., Dörr, F., Etchegaray, A., Pinto, E., Moore, B. & Fiore, M. Cyanobacterial biosynthesis of a potent neurotoxic organophosphate. in *3rd Int. Conf. Nat. Prod. Discov. Dev. Genomic Era* (SIMB, 2020).
15. Hughes, C. C. & Fenical, W. Antibacterials from the sea. *Chem. - A Eur. J.* (2010). doi:10.1002/chem.201001279
16. Hutchings, M., Truman, A. & Wilkinson, B. Antibiotics: past, present and future. *Curr. Opin. Microbiol.* (2019). doi:10.1016/j.mib.2019.10.008
17. Zaffiri, L., Gardner, J. & Toledo-Pereyra, L. H. History of antibiotics. from salvarsan to cephalosporins. *J. Investig. Surg.* (2012). doi:10.3109/08941939.2012.664099
18. Knight, D. C. *Robert Koch: father of bacteriology*. (Pickle Partners Publishing, 2019).
19. Blevins, S. M. & Bronze, M. S. Robert Koch and the “golden age” of bacteriology.

- Int. J. Infect. Dis.* (2010). doi:10.1016/j.ijid.2009.12.003
20. Bosch, F. & Rosich, L. The contributions of Paul Ehrlich to pharmacology: A tribute on the occasion of the centenary of his Nobel Prize. *Pharmacology* (2008). doi:10.1159/000149583
  21. Drews, J. Paul Ehrlich: Magister Mundi. *Nat. Rev. Drug Discov.* (2004). doi:10.1038/nrd1498
  22. Gelpi, A., Gilbertson, A. & Tucker, J. D. Magic bullet: Paul Ehrlich, Salvarsan and the birth of venereology. *Sex. Transm. Infect.* (2015). doi:10.1136/sextrans-2014-051779
  23. FOUTS, J. R., KAMM, J. J. & BRODIE, B. B. Enzymatic reduction of prontosil and other azo dyes. *J. Pharmacol. Exp. Ther.* (1957).
  24. Otten, H. Domagk and the development of the sulphonamides. *J. Antimicrob. Chemother.* (1986). doi:10.1093/jac/17.6.689
  25. Bentley, R. Different roads to discovery; Prontosil (hence sulfa drugs) and penicillin (hence  $\beta$ -lactams). *J. Ind. Microbiol. Biotechnol.* (2009). doi:10.1007/s10295-009-0553-8
  26. Waksman, S. A., Schatz, A. & Reynolds, D. M. Production of antibiotic substances by actinomycetes. *Ann. N. Y. Acad. Sci.* (2010). doi:10.1111/j.1749-6632.2010.05861.x
  27. Brown, E. D. & Wright, G. D. Antibacterial drug discovery in the resistance era. *Nature* **529**, 336–343 (2016).
  28. Lee, B. K., Condon, R. G., Wagman, G. H. & Katz, E. Micromonospora produced gentamicin components. *Antimicrob. Agents Chemother.* (1976). doi:10.1128/AAC.9.1.151
  29. Bundy, K. J., Williams, C. J. & Luedemann, R. E. Stress-enhanced ion release - the effect of static loading. *Biomaterials* (1991). doi:10.1016/0142-9612(91)90108-M
  30. Duggar, B. M. Aureomycin: a product of the continuing search for new antibiotics. *Ann. N. Y. Acad. Sci.* (1948). doi:10.1111/j.1749-6632.1948.tb27262.x
  31. Sneader, W. *Drug Discovery: A History*. *Drug Discov. A Hist.* (2006). doi:10.1002/0470015535
  32. Finlay, A. C., Hobby, G. L., P'An, S. Y., Regna, P. P., Routien, J. B., Seeley, D. B., Shull, G. M., Sobin, B. A., Solomons, I. A., Vinson, J. W. & Kane, J. H. Terramycin, a new antibiotic. *Science* (80- ). (1950). doi:10.1126/science.111.2874.85
  33. Blackwood, R. K., Beereboom, J. J., Rennhard, H. H., von Wittenau, M. S. & Stephens, C. R. 6-Methylenetetraacyclines. III. 1 Preparation and Properties 2. *J. Am. Chem. Soc.* **85**, 3943–3953 (1963).

34. Conover, L. H., Moreland, W. T., English, A. R., Stephens, C. R. & Pilgrim, F. J. Terramycin. XI. Tetracycline. *J. Am. Chem. Soc.* (1953). doi:10.1021/ja01114a537
35. EHRLICH, J., BARTZ, Q. R., SMITH, R. M., JOSLYN, D. A. & BURKHOLDER, P. R. Chloromycetin, a New Antibiotic From a Soil Actinomycete. *Science (80- )*. (1947). doi:10.1126/science.106.2757.417
36. Rebstock, M. C., Crooks, H. M., Controulis, J. & Bartz, Q. R. Chloramphenicol (Chloromycetin). IV. Chemical Studies. *J. Am. Chem. Soc.* (1949). doi:10.1021/ja011175a065
37. UMEZAWA, H., UEDA, M., MAEDA, K., YAGISHITA, K., KONDO, S., OKAMI, Y., UTAHARA, R., OSATO, Y., NITTA, K. & TAKEUCHI, T. Production and isolation of a new antibiotic: kanamycin. *J. Antibiot. (Tokyo)*. (1957).
38. Mason, D. J., Dietz, A. & DeBoer, C. Lincomycin, a new antibiotic. I. Discovery and biological properties. *Antimicrob. Agents Chemother* **1962**, 554–559 (1962).
39. MCCORMICK, M. H., MCGUIRE, J. M., PITTENGER, G. E., PITTENGER, R. C. & STARK, W. M. Vancomycin, a new antibiotic. I. Chemical and biologic properties. *Antibiot. Annu.* (1955).
40. Maggi, N., Pasqualucci, C. R., Ballotta, R. & Sensi, P. Rifampicin: a new orally active rifampin. *Chemotherapy* (1966). doi:10.1159/000220462
41. Sensi, P. History of the development of rifampin. *Rev. Infect. Dis.* (1983). doi:10.1093/clinids/5.Supplement\_3.S402
42. SMITH, C. G., DIETZ, A., SOKOLSKI, W. T. & SAVAGE, G. M. Streptonivcin, a new antibiotic. I. Discovery and biologic studies. *Antibiot. Chemother.* (1956).
43. Eliopoulos, G. M., Willey, S., Reiszner, E., Spitzer, P. G., Caputo, G. & Moellering, R. C. In vitro and in vivo activity of LY 146032, a new cyclic lipopeptide antibiotic. *Antimicrob. Agents Chemother.* (1986). doi:10.1128/AAC.30.4.532
44. Debono, M., Barnhart, M., Carrell, C. B., Hoffmann, J. A., Occolowitz, J. L., Abbott, B. J., Fukuda, D. S., Hamill, R. L., Biemann, K. & Herlihy, W. C. A21978C, a complex of new acidic peptide antibiotics: Isolation, chemistry, and mass spectral structure elucidation. *J. Antibiot. (Tokyo)*. (1987). doi:10.7164/antibiotics.40.761
45. Dougherty, T. J. & Pucci, M. J. *Antibiotic discovery and development. Antibiot. Discov. Dev.* (2014). doi:10.1007/978-1-4614-1400-1
46. Foley, T. L. & Simeonov, A. Targeting iron assimilation to develop new antibacterials. *Expert Opin. Drug Discov.* (2012). doi:10.1517/17460441.2012.708335
47. Bernardini, S., Tiezzi, A., Laghezza Masci, V. & Ovidi, E. Natural products for human health: an historical overview of the drug discovery approaches. *Nat. Prod.*

Res. (2018). doi:10.1080/14786419.2017.1356838

48. Jensen, P. R. Natural Products and the Gene Cluster Revolution. *Trends Microbiol.* (2016). doi:10.1016/j.tim.2016.07.006
49. La Clair, J. J., Fenical, W. & Costa-Lotuf, L. V. Elucidating the mode of action of marine natural products through an Immunoaffinity Fluorescent (IAF) approach. *J. Braz. Chem. Soc.* (2016). doi:10.5935/0103-5053.20160148
50. Shen, B. A New Golden Age of Natural Products Drug Discovery. *Cell* (2015). doi:10.1016/j.cell.2015.11.031
51. Džidić, S., Šušković, J. & Kos, B. Antibiotic resistance mechanisms in bacteria: Biochemical and genetic aspects. *Food Technol. Biotechnol.* (2008).
52. Depardieu, F., Podglajen, I., Leclercq, R., Collatz, E. & Courvalin, P. Modes and modulations of antibiotic resistance gene expression. *Clin. Microbiol. Rev.* (2007). doi:10.1128/CMR.00015-06
53. Ventola, C. L. The antibiotic resistance crisis: part 1: causes and threats. *Pharm. Ther.* **40**, 277 (2015).
54. Rather, I. A., Kim, B. C., Bajpai, V. K. & Park, Y. H. Self-medication and antibiotic resistance: Crisis, current challenges, and prevention. *Saudi J. Biol. Sci.* (2017). doi:10.1016/j.sjbs.2017.01.004
55. Spagnolo, F. & Dykhuizen, D. E. Antibiotic Resistance Increases Evolvability and Maximizes Opportunities Across Fitness Landscapes. *bioRxiv* (2019). doi:10.1101/750729
56. Howarth, R., Unz, R. F., Seviour, E. M., Seviour, R. J., Blackall, L. L., Pickup, R. W., Jones, J. G., Yaguchi, J. & Head, I. M. Phylogenetic relationships of filamentous sulfur bacteria (*Thiothrix* spp. and Eikelboom type 021N bacteria) isolated from wastewater-treatment plants and description of *Thiothrix eikelboomii* sp. nov., *Thiothrix unzii* sp. nov., *Thiothrix fructosivorans* sp. *Int. J. Syst. Bacteriol.* **49**, 1817–1827 (1999).
57. Sket, B. The ecology of anchihaline caves. *Trends Ecol. Evol.* **11**, 221–225 (1996).
58. Bilandžija, H., Četković, H. & Jeffery, W. R. Evolution of albinism in cave planthoppers by a convergent defect in the first step of melanin biosynthesis. *Evol. Dev.* **14**, 196–203 (2012).
59. Lee, N. M.; Meisinger, D. B.; Aubrecht, R. ., Kovacik, L.; Saiz-Jimenez, C.; Baskar, S. . & Baskar, R.; Liebl, W.; Porter, M. L.; Engel, A. S. *Life at extremes.* (Cambridge: CAB International, 2012).
60. Gabriel, C. R. & Northup, D. E. in *Cave Microbiomes A Nov. Resour. Drug Discov.* (ed. Cheeptham, N.) 85–108 (Springer New York, 2013). doi:10.1007/978-1-4614-5206-5\_5
61. Saiz-Jimenez, C. in *Cave Microbiomes A Nov. Resour. Drug Discov.* (ed.

- Cheeptham, N.) 69–84 (Springer New York, 2013). doi:10.1007/978-1-4614-5206-5\_4
62. Laiz, L., Gonzalez-Delvalle, M., Hermosin, B., Ortiz-Martinez, A. & Saiz-Jimenez, C. Isolation of cave bacteria and substrate utilization at different temperatures. *Geomicrobiol. J.* (2003). doi:10.1080/713851125
  63. Engel, A. S., Porter, M. L., Stern, L. A., Quinlan, S. & Bennett, P. C. Bacterial diversity and ecosystem function of filamentous microbial mats from aphotic (cave) sulfidic springs dominated by chemolithoautotrophic “Epsilonproteobacteria.” *FEMS Microbiol. Ecol.* (2004). doi:10.1016/j.femsec.2004.07.004
  64. Linares, J. F., Gustafsson, I., Baquero, F. & Martinez, J. L. Antibiotics as intermicrobiol signaling agents instead of weapons. *Proc. Natl. Acad. Sci. U. S. A.* (2006). doi:10.1073/pnas.0608949103
  65. Mlot, C. Antibiotics in nature: beyond biological warfare. *Science (80-. ).* (2009). doi:10.1126/science.324\_1637
  66. Fajardo, A., Linares, J. F. & Martínez, J. L. Towards an ecological approach to antibiotics and antibiotic resistance genes. *Clin. Microbiol. Infect.* (2009). doi:10.1111/j.1469-0691.2008.02688.x
  67. Fajardo, A. & Martínez, J. L. Antibiotics as signals that trigger specific bacterial responses. *Curr. Opin. Microbiol.* (2008). doi:10.1016/j.mib.2008.02.006
  68. Bewley, C. A., Holland, N. D. & Faulkner, D. J. Two classes of metabolites from *Theonella swinhoei* are localized in distinct populations of bacterial symbionts. *Experientia* (1996). doi:10.1007/BF01925581
  69. Northup, D. E., Barns, S. M., Yu, L. E., Spilde, M. N., Schelble, R. T., Dano, K. E., Crossey, L. J., Connolly, C. A., Boston, P. J., Natvig, D. O. & Dahm, C. N. Diverse microbial communities inhabiting ferromanganese deposits in Lechuguilla and Spider Caves. *Environ. Microbiol.* **5**, 1071–1086 (2003).
  70. Gillieson, D. *Caves: processes, development and management.* (John Wiley & Sons, 2009).
  71. Yucel, S. & Yamac, M. Selection of *Streptomyces* isolates from Turkish karstic caves against antibiotic resistant microorganisms. *Pak. J. Pharm. Sci.* (2010).
  72. Northup, D. E. & Lavoie, K. H. Geomicrobiology of caves: A review. *Geomicrobiol. J.* (2001). doi:10.1080/01490450152467750
  73. Barton, H. A. & Northup, D. E. Geomicrobiology in cave environments: Past, current and future perspectives. *J. Cave Karst Stud.* (2007).
  74. Cheeptham, N. *Cave microbiomes: A novel resource for drug discovery.* **1**, (Springer Science & Business Media, 2012).
  75. Rosado, A. S., Duarte, G. F., Seldin, L. & Van Elsas, J. D. Genetic diversity of

- nifH gene sequences in *Paenibacillus azotofixans* strains and soil samples analyzed by denaturing gradient gel electrophoresis of PCR-amplified gene fragments. *Appl. Environ. Microbiol.* (1998). doi:10.1128/aem.64.8.2770-2779.1998
76. Barton, H. A. Introduction to cave microbiology: A review for the non-specialist. *J. Cave Karst Stud.* (2006).
  77. Culver, D.C., and Pipan, T. *The Biology of Caves and their Subterranean Habitats: Biology of Habitats.* (Oxford University Press, 2009).
  78. Bhullar, K., Waglechner, N., Pawlowski, A., Koteva, K., Banks, E. D., Johnston, M. D., Barton, H. A. & Wright, G. D. Antibiotic resistance is prevalent in an isolated cave microbiome. *PLoS One* **7**, 1–11 (2012).
  79. Dong, X. & Reddy, G. B. Soil bacterial communities in constructed wetlands treated with swine wastewater using PCR-DGGE technique. *Bioresour. Technol.* (2010). doi:10.1016/j.biortech.2009.09.071
  80. Green, S. J., Leigh, M. B. & Neufeld, J. D. in *Handb. Hydrocarb. Lipid Microbiol.* (2010). doi:10.1007/978-3-540-77587-4\_323
  81. van der Kraan, G. M., Buijzen, F., Kuijvenhoven, C. A. T. & Muyzer, G. in *Appl. Microbiol. Mol. Biol. Oilf. Syst.* (2010). doi:10.1007/978-90-481-9252-6\_5
  82. Mearns-Spragg, A., Bregu, M., Boyd, K. G. & Burgess, J. G. Cross-species, induction and enhancement of antimicrobial activity produced by epibiotic bacteria from marine algae and invertebrates, after exposure to terrestrial bacteria. *Let. Appl. Microbiol.* (1998). doi:10.1046/j.1472-765X.1998.00416.x
  83. Milward, A. E. Characterization of the Antimicrobial Secondary Metabolites Produced by the Cave-Dwelling *Streptomyces* ICC1 Strain. (2019).
  84. Avguštin, J. A., Etrič, A. & Ašić, L. Screening the cultivable cave microbial mats for the production of antimicrobial compounds and antibiotic resistance. *Int. J. Speleol.* (2019). doi:10.5038/1827-806X.48.3.2272
  85. Gosse, J. T., Ghosh, S., Sproule, A., Overy, D., Cheeptham, N. & Boddy, C. N. Whole genome sequencing and metabolomic study of cave *Streptomyces* isolates ICC1 and ICC4. *Front. Microbiol.* (2019). doi:10.3389/fmicb.2019.01020
  86. Maciejewska, M., Adam, D., Martinet, L., Naômé, A., Całusińska, M., Delfosse, P., Carnol, M., Barton, H. A., Hayette, M. P., Smargiasso, N., De Pauw, E., Hanikenne, M., Baurain, D. & Rigali, S. A phenotypic and genotypic analysis of the antimicrobial potential of cultivable *Streptomyces* isolated from cave moonmilk deposits. *Front. Microbiol.* (2016). doi:10.3389/fmicb.2016.01455
  87. Ghosh, S., Paine, E., Wall, R., Kam, G., Lauriente, T., Sa-ngarmangkang, P.-C., Horne, D. & Cheeptham, N. In Situ Cultured Bacterial Diversity from Iron Curtain Cave, Chilliwack, British Columbia, Canada. *Diversity* **9**, 36 (2017).
  88. Banerjee, S., Jha, D. K. & Joshi, S. R. in *Microb. Divers. Ecosyst. Sustain. Biotechnol. Appl.* (2019). doi:10.1007/978-981-13-8487-5\_1

89. Wiseschart, A., Mhuantong, W., Tangphatsornruang, S., Chantasingh, D. & Pootanakit, K. Shotgun metagenomic sequencing from Manao-Pee cave, Thailand, reveals insight into the microbial community structure and its metabolic potential. *BMC Microbiol.* (2019). doi:10.1186/s12866-019-1521-8
90. Wiseschart, A., Mhuanthong, W., Thongkam, P., Tangphatsornruang, S., Chantasingh, D. & Pootanakit, K. Bacterial Diversity and Phylogenetic Analysis of Type II Polyketide Synthase Gene from Manao-Pee Cave, Thailand. *Geomicrobiol. J.* (2018). doi:10.1080/01490451.2017.1411993
91. Milward, A. Optimizing the Production of Antimicrobial Secondary Metabolites Produced by the Cave-Dwelling Streptomyces S1 Strain. (2018).
92. Bachmann, B. O. Ask first and shoot later-Anticancer natural product discovery from cave actinomycetes using Multiplexed Activity Metabolomics. in *SIMB Annu. Meet. Exhib. 2019* (SIMB, 2019).
93. Bull, A. T. & Asenjo, J. A. Microbiology of hyper-arid environments: Recent insights from the Atacama Desert, Chile. *Antonie van Leeuwenhoek, Int. J. Gen. Mol. Microbiol.* **103**, 1173–1179 (2013).
94. Sánchez, L. A., Gómez, F. F. & Delgado, O. D. Cold-adapted microorganisms as a source of new antimicrobials. *Extremophiles* (2009). doi:10.1007/s00792-008-0203-5
95. Sánchez, L. A., Hedström, M., Delgado, M. A. & Delgado, O. D. Production, purification and characterization of serraticin A, a novel cold-active antimicrobial produced by *Serratia proteamaculans* 136. *J. Appl. Microbiol.* (2010). doi:10.1111/j.1365-2672.2010.04720.x
96. Déziel, E., Lépine, F., Milot, S., He, J., Mindrinos, M. N., Tompkins, R. G. & Rahme, L. G. Analysis of *Pseudomonas aeruginosa* 4-hydroxy-2-alkylquinolines (HAQs) reveals a role for 4-hydroxy-2-heptylquinoline in cell-to-cell communication. *Proc. Natl. Acad. Sci. U. S. A.* (2004). doi:10.1073/pnas.0307694100
97. Harris, K. D. & Kolodkin-Gal, I. Applying the handicap principle to biofilms: condition-dependent signalling in *Bacillus subtilis* microbial communities. *Environ. Microbiol.* (2019). doi:10.1111/1462-2920.14497
98. Santillan, E., Seshan, H., Constancias, F., Drautz-Moses, D. I. & Wuertz, S. Frequency of disturbance alters diversity, function, and underlying assembly mechanisms of complex bacterial communities. *npj Biofilms Microbiomes* (2019). doi:10.1038/s41522-019-0079-4
99. Gu, L., Wu, J., Hua, Z. & Chu, K. The response of nitrogen cycling and bacterial communities to *E. coli* invasion in aquatic environments with submerged vegetation. *J. Environ. Manage.* **261**, 110204 (2020).
100. Santhanam, R., Rong, X., Huang, Y., Andrews, B. A., Asenjo, J. A. & Goodfellow, M. *Streptomyces bullii* sp. nov., isolated from a hyper-arid Atacama

Desert soil. *Antonie van Leeuwenhoek*, *Int. J. Gen. Mol. Microbiol.* (2013).  
doi:10.1007/s10482-012-9816-x



**Chapter 2: Diversity and antibacterial activity of bacteria cultured from four semi-arid carbonate limestone caves (Carlsbad, NM, USA)**

## 2.1 Abstract

The culturable diversity of bacteria isolated from four carbonate limestone caves, Lechuguilla, Spider, Carlsbad Cavern, and Backcountry, was examined (Carlsbad Caverns National Park, New Mexico, USA). I purified and taxonomically characterized 334 isolates belonging to four phyla, **150** Actinobacteria, **95** Firmicutes, **82** Proteobacteria and **7** Bacteroidetes. I screened them for antibiotic production, determined the mechanism of action (MOA) of the antibiotics they produce and on the basis of phylogenetic relatedness and bioactivity prioritized a group of 9 isolates consisting of five *Bacillus sp.* and four Actinobacteria for downstream analyses. I focused on these strains by screening them for pathogen inhibition, identifying the MOA of NPs expressed when grown on two additional media, sequenced their genomes and mined them for biosynthetic gene clusters associated with secondary metabolic natural products (NPs). A summary of the results is provided below.

## 2.2 Introduction

Multidrug resistant bacterial pathogens have created an urgent need for the discovery of new antibiotics<sup>1-4</sup>. Microbial derived natural products (NPs) are a rich source of biologically active compounds<sup>5-7</sup>. The global distribution of NP biosynthetic potential is not fully understood, however it is widely accepted that in some ecological niches and among select bacterial lineages it is enriched<sup>8</sup>, while in others it is significantly reduced<sup>9</sup>. Soil dwelling *Streptomyces sp.*, have long been recognized as prolific producers of bioactive compounds including two-thirds of the antibiotics currently on the market<sup>10-13</sup>. Recent evidence suggests that more extreme habitats also harbor the potential for drug discovery<sup>14</sup>, including hydrothermal vents<sup>15</sup>, cryo-

environments<sup>16-18</sup>, deep-seas<sup>19-22</sup>, and subterranean caves<sup>23-26</sup>. The characterization of the natural products biosynthesized by bacteria isolated from underexplored locations increases the likelihood of discovering antibiotics with novel pharmacores and mechanisms of action<sup>27</sup>.

Caves are naturally occurring and geographically widespread reservoirs of bacterial biodiversity<sup>28-30</sup> and represent a unique, underexploited habitat in comparison to the surface soil from which Actinobacteria have previously been isolated and characterized for the production of antibiotics<sup>11,31</sup>. Specifically, the dark zone of caves is the most unique in comparison to surface soils because they are characteristically oligotrophic as a result of the lack of sunlight and therefore photosynthesis. Because of this, primary nutrient sources are limited and occur inorganically, resulting in a microbial community dominated by chemolithoautotrophs<sup>32-37</sup>. The dark zone is also uniquely well preserved as it is secluded from weathering, maintains a relatively stable temperature and is moisture saturated<sup>28,38</sup>. The known diversity of bacteria in limestone caves in semi-arid climactic regions include phyla with members that are known prolific producers of clinically useful natural products including Actinobacteria,  $\beta$ -Proteobacteria,  $\delta$ -Proteobacteria, Cyanobacteria, and the less-studied phyla Verrucomicrobia and Chloroflexi<sup>39-41</sup>. The ecological parameters that exist in caves and function as the drivers of evolutionary speciation and metabolic diversification have previously been investigated for cave-adapted animals (troglobionts)<sup>42</sup>, but less is known about how microorganisms adapt to the abiotic and biotic selective pressures of caves in karstic landscapes<sup>31</sup>. Similar to marine Actinobacteria, functional bacterial adaptations to the

extreme subsurface environment may include secondary metabolic modifications resulting in the biosynthesis of new bioactive compounds<sup>43,44</sup>.

While the NP biosynthetic potential existing in the dark zones of caves is largely unknown, a few collections of bacteria from caves are currently being mined for bioactive natural products. For example, the Bachmann group identified novel hypogeamycins from the Actinomycete *Nonomuraea specus* isolated from Hardin's cave system near Ashland Tennessee<sup>45</sup>. Additionally, four different compounds with various bioactivities were isolated from *Streptomyces* species cultured from the biggest conglomeratic Karstic Cave in Siberia<sup>46</sup>. Also, a pigment with a phenooxazinone chromophore was extracted from a *Streptomyces sp.* from Borra Caves in India and is active against MRSA and TB<sup>47</sup>. Finally, novel polyketide glycosides, Cervimycins A-D, that are active against MRSA and VRE were discovered from *Streptomyces tendae* isolated from the Grotta dei Cervi in Italy<sup>25,48</sup>. Despite these recent efforts, to date, less than fifty percent of caves have been subjected to extensive scientific research, and even fewer for antibiotic discovery, leaving much to be discovered in this extreme subterranean habitat<sup>49</sup>.

Bacterial Cytological Profiling (BCP), is a whole cell screening technique that can be used to rapidly identify bacteria that produce inhibitory compounds. BCP uses fluorescent microscopy to image *E. coli* cells that have been exposed to a potential antibiotic. The images of the treated cells are then quantitated to generate a cytological profile, which can then be compared to profiles of untreated cells grown under the same condition and a library of profiles for known antibiotics. BCP is accurate enough to identify most of the major pathways by which antibiotics are known to act and has the

ability to distinguish between molecules acting on different steps within the same pathway. BCP also allows the identification of the MOA for mixtures of compounds, including unpurified antibiotics found within crude natural product extracts. BCP is especially useful for NP discovery because it can be performed directly on petri plates, serving both as a dereplication filter *and* a discovery method to rapidly prioritize strains for further studies<sup>50-60</sup>.

Here I report on the cultivation-based bacterial diversity of Lechuguilla, Spider, Carlsbad Cavern, and Backcountry Caves. I evaluate the potential for bacteria cultured from four semi-arid carbonate caves to produce novel antibiotics and explore the possibility for cave specific ecological parameters to help characterize the areas within the caves where these bacteria reside. This research uses a combination of geomicrobiological, phylogenetic, and bioinformatic approaches to characterize the biosynthetic potential of cave bacteria for natural product antibiotic discovery. A total of 334 bacterial strains were screened for antimicrobial activity resulting in a select group of 95 displaying antibiotic activity. We assessed their ability to produce compounds with bioactivity against pathogenic bacterial strains and used BCP to characterize their activity. This work highlights a collection of bacteria that can be explored further for the discovery of potentially novel natural products.

## **2.3 Materials and methods**

### **2.3.1 Study sites: four subterranean limestone caves**

Four caves located in the karstic landscape of Carlsbad Caverns National Park (CCNP) were chosen for this study; Lechuguilla [Lech], Carlsbad Cavern [CC], Spider [S], and Backcountry [BC]. CCNP is situated in the semi-arid climactic region of the

Guaralupe Mountains in the Chihuahuan Desert (NM, USA). (32°10'31"N, 104°26'38"W). The caves were formed by sulfuric acid speleogenesis, range in temperature from 15.4°C to 18.3°C and maintain ~99% relative humidity. Each cave was chosen on the basis of varied age, depth, length, microhabitats, and anthropogenic impact.<sup>38</sup>

### **2.3.2 Description of sample sites in each cave**

Within each cave, sample sites were identified along the length (23 m to 209 km), and at varying depths (-1.5 m to -300 m). The sampling sites were chosen on the basis of differing amounts of two biologically important chemical elements, carbon (4.98% to 9.25%) and nitrogen (0.04% to 2.99%), and on the basis of diverse microhabitats including ferromanganese deposits (FMD), speleothems, moonmilk, water, rock, and soils containing surface soil and decomposed wall rock.<sup>38</sup>

### **2.3.3 Sample collection**

Bacterial cultures were obtained aseptically by dipping a sterile rayon-tipped swab in sterile deionized water, rubbing the sample site surface, and streaking it onto a variety of solid agar growth media. All cultures were incubated in the cave for the first 24 hours then maintained in the lab, in the dark, at cave temperature (~20°C). Upon arrival in the lab, the culture plates were scraped and glycerol stocked for long-term storage at -80°C. Media included those known to support the growth of the oligotrophic, slow growing, cave bacteria including, low nutrient ½ R2A, Actinomycete Isolation Agar (AIA), Humic Acid Vitamin Agar with antibiotics, RASS, and Bennett's (for *Actinoplanes* sp.) (Table 1). On site sampling in Backcountry Cave (BCC) varied within 5-6 m of depth below the surface every 3 meters, Spider Cave (SP) sample sites varied

from 38-39 m of depth every 10 m, whereas sampling in Lechuguilla was done at every 50 m of increasing depth below the surface beginning at -150 m and ending at -300 m. Sampling in Carlsbad Cavern was done in LHT passage, at -228 m every 50 m along the length of the passage. All samples were taken in the aphotic zone, except for entrance samples in BCC and Spider Cave.<sup>40</sup>

### **2.3.4 Isolation and maintenance of bacterial cave strains**

To obtain individual bacteria in pure culture, mixed cave cultures were subcultured onto specialized growth media and grown in the dark at room temperature for two to three months. Upon the growth of individual colonies, different colony types were streak isolated onto the same media that they were initially grown on. Streak isolations were repeated three times or more until considered pure then put into a freezer stock maintained at -80°C.

### **2.3.5 Bacterial genomic DNA extraction and quantification for 16S rRNA PCR amplification and sequencing**

Crude genomic DNA was extracted from >500 bacterial isolates using a protocol adapted from a QIAGEN DNeasy Blood and Tissue kit. Ten microliters of the resulting genomic DNA was diluted then used as the template in the polymerase chain reaction (PCR) to amplify the bacterial 16S rRNA gene. Strains were cultured overnight at 30°C in 5 mL of LB broth while rolling. Cells were pelleted (16,000 x g, 3 min) from 1 ml of culture, re-suspended in 180 µl of lysis buffer (prepared in house), and incubated at 37°C for 45 min after which 25 µl of proteinase K (20 mg/mL) and 200 µl Buffer AL (Qiagen) was added. The samples were vortexed at maximum speed for 20 sec, incubated at 56°C for 30 min, and 200 µl of ethanol (96-100%) was added. The samples were vortexed at

maximum speed for 30 sec, added to a DNeasy Mini spin column, centrifuged (16,000 x g, 1 min), and the supernatant was discarded. Buffer AW2 (Qigen) was added (500 µl), followed by centrifugation (20,000 x g, 3 min). The DNeasy Mini spin column was placed into a sterile 2 mL microcentrifuge tube, and the gDNA was eluted in 100 µl of AE Buffer by centrifugation (20,000 x g, 1 min) following a 1 min incubation at room temperature. The gDNA concentration was quantified (1 µl sample volume) with a Thermo Scientific™ NanoDrop™ One Microvolume UV-Vis Spectrophotometer (840274100) and stored at -20°C.

### **2.3.6 16S rRNA PCR amplification and sequencing**

16S ribosomal DNA templates (~1,465 bp) were amplified using Q5 high fidelity PCR (New England Biolabs) with the universal primer set 27F (5'-AGAGTTTGATCCTGGCTCAG-3') and 1492R (5'-GGTTACCTTGTTACGACTT-3')<sup>61</sup>. Each PCR mixture (50 µl) contained 100 ng of template gDNA, 500 pmol of each primer, and 200 µM dNTPs. PCR thermocycling conditions were as follows: 30 seconds of initial denaturation at 98°C, 30 cycles of denaturation at 98°C for 10 seconds, annealing for 15 seconds at 60°C, extension at 72°C for 1.5 minutes, and a final extension at 72°C for 5 minutes then held at 4°C. PCR products were purified with the oligonucleotide cleanup protocol as described in the Monarch PCR & DNA Cleanup Kit 5 µg user manual (NEB #T1030). Resulting 1.5kB amplicons were electrophoresed in 1X TAE buffer on a 1% agarose gel containing 2µl of Ethidium bromide and visualized under UV transillumination. Clean PCR products were sequenced using Sanger methods by Eton Biosciences (<https://www.etonbio.com/>) and trimmed for quality before analysis.

### **2.3.7 Phylogenetic analyses of bacterial isolates**



16S rRNA sequences were trimmed on both ends, (5' and 3') in Geneious Prime using the Trim Ends function with an error probability limit set at 0.05, which trims regions with more than a 5% chance of an error per base. Reference sequences were obtained using a nucleotide BLAST (blastn) against the reference RNA sequences (refseq\_rna) database. The nucleotide sequences used to determine the genus level phylogeny were aligned using MUSCLE. Maximum likelihood analysis of the data was run using RAxML with 100 rapid bootstrap replicates and the GTR+G model<sup>62,63</sup>. All trees were visualized using FigTree v1.4.2.

### **2.3.8 Bacterial cytological profiling (BCP) on LB, ISP2, AIA**

The cave bacteria were streaked in the middle of LB, AIA, and ISP2 plates and incubated for one week at 30°C. To test for antibiotic activity, 10 µl of an exponentially growing culture of *E.coli tolC* was then spotted in a line perpendicular to the cave strain, incubated for two hours, and the imaged using BCP<sup>50</sup>. The plates were allowed to incubate for 3 hours and then one side of the agar containing the *E. coli* cells was cut out, stained with fluorescent dyes [1 µg/mL FM4-64 (red), 2 µg/mL DAPI (blue), and 0.5 µM SYTOX (Green)] and imaged by fluorescence microscopy. A 10 µl sample of *E. coli* grown alone on a petri plate served as a control. The remaining plate was left to grow overnight and the next day the plates were assessed for killing activity by looking for a zone of inhibition of the *E. coli* strain (Figure 5). The 95 cave bacteria that were positive for next day killing were all assigned a cell profile that matched a known MOA (Figure 6). These cave bacteria were then grouped based on species to investigate whether bacterial members belonging to the same species can produce molecules that act on a different cellular target in the same organism (Figure 7).

### **2.3.9 Genomic DNA extraction for PacBio whole-genome sequencing**

High molecular weight genomic DNA (20-160 kb) was extracted from 4 Actinobacteria (EM5, EM7, EM9, and EM10) and 5 *Bacillus* strains (EM1, EM2, EM3, EM6) with the QIAGEN-Genomic-tip 500/G kit (10262) according to the manufacturer's protocol for bacteria.

### **2.3.10 Whole-genome sequencing, assembly, and annotation**

The genome sequences of 4 Actinobacteria and 5 *Bacillus sp.* were generated using the Pacific Biosciences RS II (PacBio RS II) single molecule real-time (SMRT) sequencing platform at the IGM Genomics Center, University of California, San Diego, La Jolla, CA. Genome sequences were assembled using the HGAP protocol integrated in the PacBio RS II sequencer (smrt analysis v2.3.0/Patch5) resulting in a variable number (n = 1-10) of contigs per genome, and ranged in size from 5.58 to 8.54 Mb. The mauve contig mover was used to order the contigs of 8 draft genome sequences (genomes of strains EM1, EM2, EM3, EM6, EM5, EM7, and EM9) relative to a closely related reference sequence (*B. thuringiensis* HD-771, *S. griseus subsp. griseus* NBRC 13350, *S. fulvissimus* DSM 40593, and *S. coelicoflavus* ZG0656 respectively). DNA sequencing of strain EM10 resulted in a single contig and did not require reordering to restore gene synteny. Gene prediction and annotation were made with the Rapid Annotations using Subsystems Technology (RASTtk)<sup>64</sup> platform.

### **2.3.11 Genome mining for secondary metabolic gene clusters**

antiSMASH v5.0<sup>65</sup> was used to perform the prediction of biosynthetic gene clusters. Annotations were from AntiSMASH of secondary metabolite biosynthesis gene clusters (smBGCs).

## **2.4 Results and discussion**

### **2.4.1 Study sites: 4 carbonate limestone caves**

Four cave systems were identified as study sites and from their maps a systematic sampling method was set up. The caves are located in the semi-arid climactic region of the southwestern United States in the state of New Mexico in Carlsbad Caverns National Park (CCNP). These caves include Lechuguilla Cave, Left Hand Tunnel in Carlsbad Cavern, Spider Cave, and Backcountry Cave (Figure 1). The caves were systematically assessed along the length and at varying depths (-1.5 to -300 m) for areas of interest including; the Dark Zone, areas with low anthropogenic effect, and the presence of Moonmilk<sup>38</sup>.

### **2.4.2 Bacterial sample collection**

A total of 108 cultures were isolated from a variety of microhabitats including flowstone, stalactites, and moonmilk within all four caves using sterile swabs dipped into lab grade DI water streaked onto a variety of specialized growth media<sup>38</sup>. It is commonly accepted that less than two percent of the microorganisms occurring in nature have been successfully cultured. Recent culture independent phylogenetic studies focusing on microorganisms have provided support for this<sup>31,41,74,75,66-73</sup>. Growth media and conditions were therefore customized for improving culturability of the existing bacterial biodiversity in caves. Media included those known to support the growth of the oligotrophic, slow growing, cave bacteria including; low nutrient ½ R2A, Actinomycete Isolation Agar (AIA), Humic Acid Vitamin Agar with antibiotics, Oatmeal agar (for *Actinomadura sp.*), and Bennett's (for *Actinoplanes sp.*). All samples were incubated in the cave for the first 24 hours then maintained in the lab, in the dark, at cave temperature

(~20°C). Upon arrival in the lab, the growth plates were scraped and glycerol stocked for long-term storage at -80°C. The cave samples were streak isolated onto a variety of growth media including AIA with antibiotics to further select for *Streptomyces sp.*, the original growth media to isolate the highest level of diversity, and nutrient rich Luria Bertani (LB).

#### **2.4.3 Isolation, taxonomic identification, and genus level phylogeny**

A total of 334 cave strains were isolated and the 16S rRNA gene was sequenced revealing the taxonomic identification of each, and the diversity among the collection. The phylogenetic relatedness was determined among all 334 taxa and is displayed in a phylogenetic tree (Figure 2) Table 2 contains a list of the cave isolate strain names, and their corresponding genus level classification, and Table 3 contains a list of the 100 type strain reference sequences used to build the tree. The cave isolate strains classified in 25 genera, both Gram-positive and Gram-negative, spanning 4 phyla (Figures 2-4). Isolate A13A1 is the only strain in the *Agromyces* genus. The *Staphylococcus* strains fall inside the larger *Bacillus* clade; this could be an artifact of the limited sequence data used in the study and/or because those two genera are very highly similar based on 16S rRNA sequences.

#### **2.4.4 Bacterial cytological profiling of cave bacteria natural products on LB plates**

All 334 isolates were screened with tandem cross-streak inhibition assays on solid LB media to identify isolates that produce inhibitory natural products against *E. coli* JP313 *ΔtolC* (pump knock out mutant strain rendering it more sensitive to most antibiotics), and *E. coli* JP1033 *lptD* (lipopolysaccharide assembly protein mutant rendering it more sensitive to most antibiotics). A total of 95 strains (29%) were active

against 2 test stains of *E. coli*. 53 strains (16%) displayed a zone of inhibition when tested against *E. coli* JP1033 *lptD*, and 42 strains (13%) displayed a zone of inhibition when tested against *E. coli* JP313  $\Delta$ *tolC*, and 239 (72%) were inactive under these test conditions (Figure 6). An example of the *E. coli tolC* control and test plates are shown in Figure 5. Next, I assessed the taxonomic classification data gathered for all 95 (29%) isolates. Representatives from all four phyla were present among the 95 inhibitor strains and were heavily represented among the Firmicutes, the majority belonged to the genus *Bacillus*. Fifteen of the *Bacillus sp.* had very similar 16S rRNA sequences and grouped with *Bacillus toyonensis*, a known producer of an antibacterial lantibiotic (Figure 7). Among the Actinobacteria were four inhibitor strains that belonged to the genera *Streptomyces* and *Nocardiopsis*. (Figure 9, Table 4).

The likely MOA of compounds produced by the 95 strains that displayed a zone of inhibition against *E. coli* JP313  $\Delta$ *tolC* and *E. coli* JP1033 *lptD* on LB were studied using BCP. We found 7 major BCP phenotypes, including inhibitors of the cell envelope, protein translation, DNA replication, membrane activity, and one DNA intercalation phenotype. The most common MOA (28% of strains) was those producing multiple phenotypes and strains that produce a DNA intercalating agent followed closely by DNA replication inhibition were the least common under these growth conditions (Figure 6).

Because *Streptomyces sp.* are known prolific antibiotic producers and they displayed phenotypes indicative of producing bioactive molecules that target different major cellular pathways, four of the Actinobacteria were chosen for further characterization (Figure 9, Table 4). In addition, because so many Firmicutes displayed bioactivity against *E. coli*, I selected five *Bacillus strains* for further study. (Figure 9,

Table 4). *Bacillus sp.* are a prevalent group among the Firmicutes with antibacterial activity, and among these, we found 15 closely related *Bacillus toyonensis* isolates. *B. toyonensis* is a probiotic in livestock feed and was recently found to be a producer of a lantibiotic molecule with antibacterial effects. Fifteen closely related members of the Firmicutes tightly cluster with *Bacillus toyonensis*, and according to their 16S rRNA gene these cave isolates are the same species, however, I found that they have different MOA against *E. coli* JP313 *AtolC* and *E. coli* JP1033 *lptD* on LB (Figure 8) suggesting variation among the natural products produced. For example, one of the strains inhibited DNA replication while two others inhibited protein translation and cell wall biogenesis. Despite the close relationship based on 16S rRNA, our results reflect previous studies suggesting that bacteria with high 16S rRNA gene sequence similarity can exhibit different capacities to produce antibacterial molecules and it is worth further investigating many other closely related members of the collection. Strains of bacteria belonging to the same species that are capable of inhibiting different cellular pathways alludes to the evolution of potentially unique natural product genes present within the cave microbiome. The same species of bacteria cultured from an oligotrophic environment like caves may have evolved unique biosynthetic gene clusters that act on different cellular pathways within the same Gram-negative organism (such as *E. coli*) to make one variant better equipped to compete for space and resources than another.

I also examined the MOA using BCP of four Actinomycete strains that inhibited *E. coli tolC* growth. These strains displayed four different MOA, including cell wall, DNA intercalation, a mixture of cell and DNA, and a mixture of cell wall and membrane

activity (Table 4). Based on the BCP phenotypes of these strains, I identified four Actinomyces strains to prioritize for genome sequencing.

#### **2.4.5 Pathogen inhibition by cave isolates**

I began further characterizing the NP arsenal of the 9 most interesting isolates, five *Bacillus sp.* and four Actinobacteria, by assessing their biosynthetic potential to inhibit the growth of five known pathogens using the rapid cross-streaking method. All but three cave isolates inhibited at least one of the known pathogens including *E. coli* ATCC25922 (WT), *S. aureus* USA 300 1516 (MRSA), *P. aeruginosa* strain K2733 ( $\Delta$ pump), PA01 (WT), and P4 (MDR) (Table 4).

#### **2.4.6 Bacterial cytological profiling of cave bacteria natural products on different media**

Next, I assessed the potential for expressing different bioactive molecules under different growth conditions using BCP to identify the MOA of the NPs produced on two additional solid media. Initially I screened all cave isolates, including the 9 strains of interest, on LB media against *E. coli*  $\Delta$ tolC AD3644. I expanded the characterization of their biosynthetic capacities by evaluating them on Actino Isolation Agar (AIA) and the low nutrient media ISP2. This experiment was attempted to further diversify the phenotypes, and potentially induce the expression of diverse bioactive compounds among the 9 strains. This also allowed for the identification of the growth conditions under which each strain expresses the most interesting NPs resulting in unique BCP profiles as shown in Figure 10 (*Bacillus* strains) and Figure 11 (Actinobacteria strains), the BCP phenotypes of the strains depended upon the growth conditions. For example EM5

inhibited translation of *E.coli tolC* when grown on LB but inhibited cell division when grown on ISP2 (Figure 11).

#### **2.4.7 Genome sequencing**

The genomes of nine cave isolates (4 *Streptomyces* and 5 *Bacillus*) were sequenced using PacBio at the UCSD IGM sequencing Core and assembled into contigs. When the sequence data came back, I reorganized the contigs resulting from assembly to the whole genome sequence (WGS) of the most similar known bacterial isolate (based on 16S rRNA sequence) to restore gene synteny. Of these nine, I focused on the five Actinobacteria: strains EM5, EM7, EM9, and EM10. I annotated the genomes using RAST.

#### **2.4.8 Biosynthetic capacity of 4 cave bacteria.**

I used AntiSMASH 5.0 to identify predicted gene clusters associated with secondary metabolism (Figure 12), and using sequence comparison by pairwise amino acid alignment, began to characterize the clusters by determining the level of similarity or dissimilarity to known BGCs (Figure 13). Each strain encoded more than 10 BGCs, with many of the clusters not matching to previously described BGCs (Figures 12 and 13).

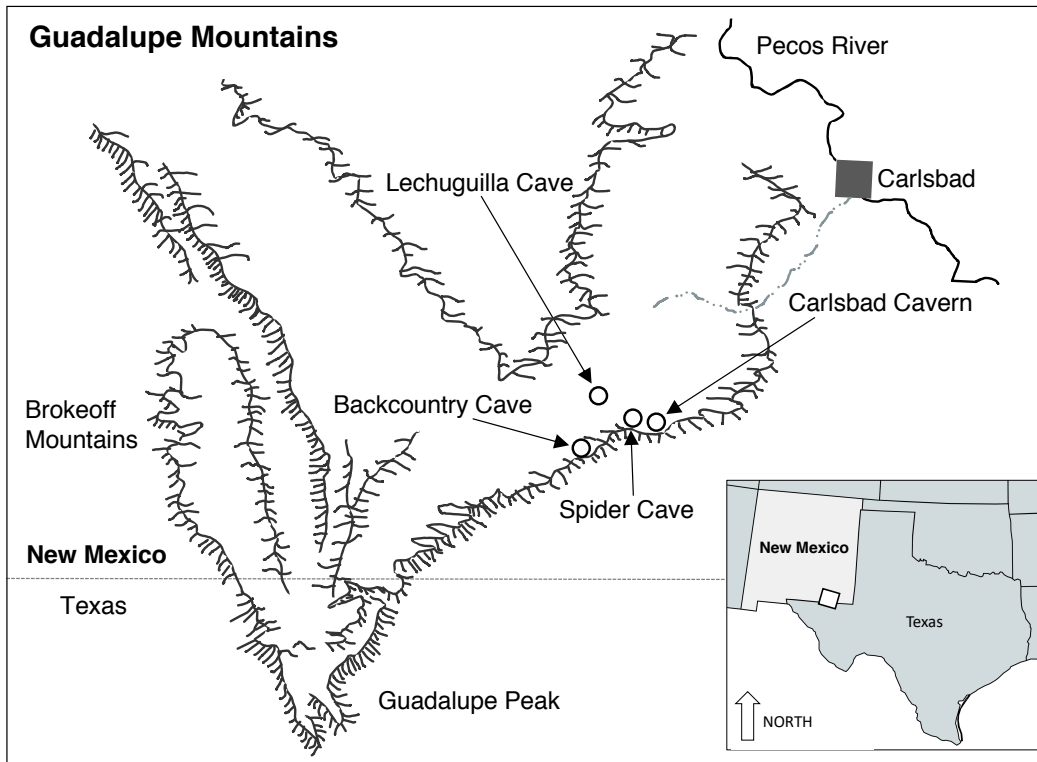
#### **2.5 Acknowledgements**

These studies were supported by the National Institute of Health (R01-AI113295). KP and JP have an equity interest in Linnaeus Bioscience Incorporated, and receive consulting income from the company. The terms of this arrangement have been reviewed and approved by the University of California, San Diego in accordance with its conflict of interest policies.

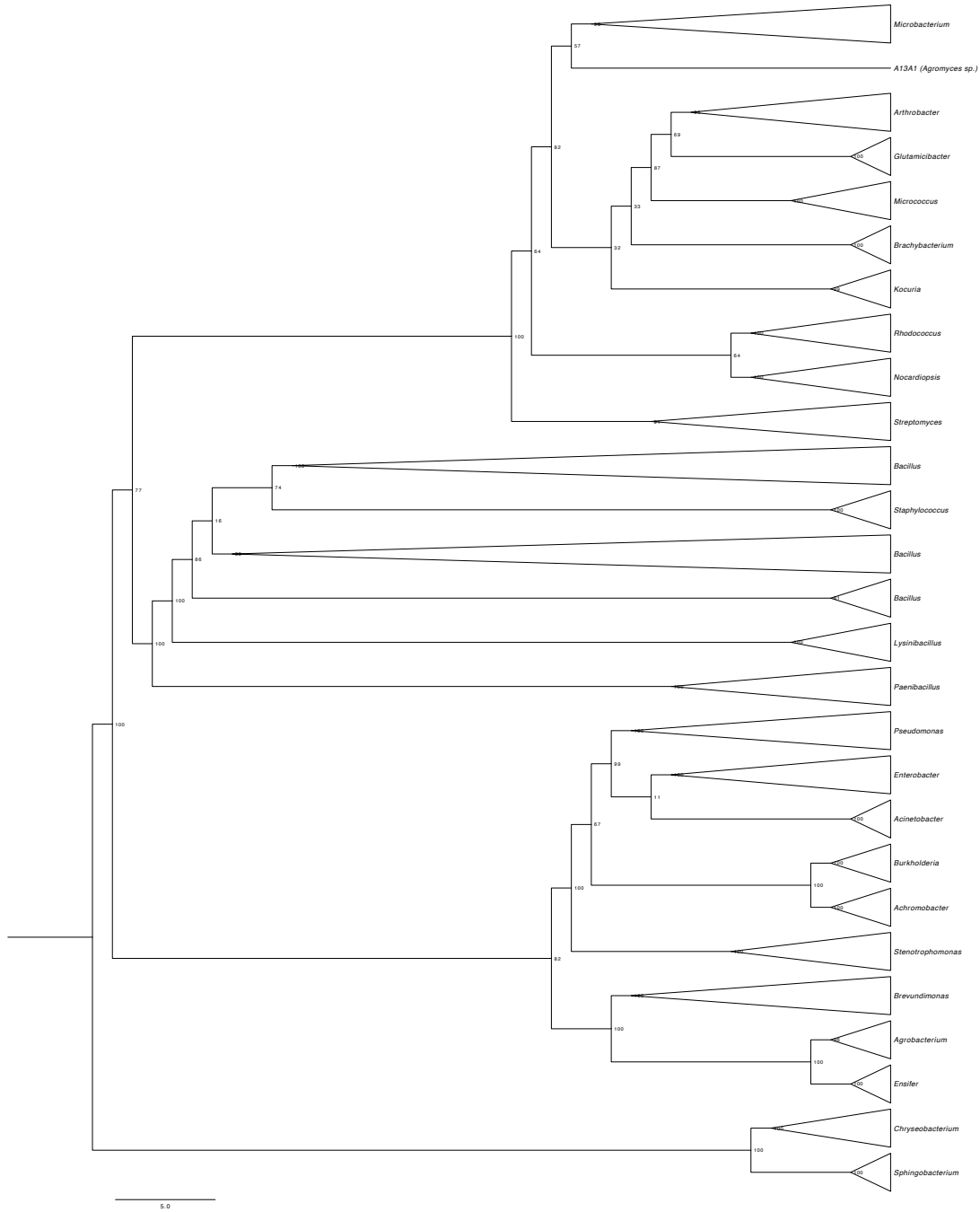


Chapter 2, in part, is being formulated into a manuscript in preparation for publication of the material. Elizabeth T. Montañó, Joseph Sugie, Eray Enustun, Julia Busch, Diana E. Northup, Kit Pogliano, Joseph Pogliano, 2020. The dissertation author was the primary investigator and author of this material.

## 2.6 Figures

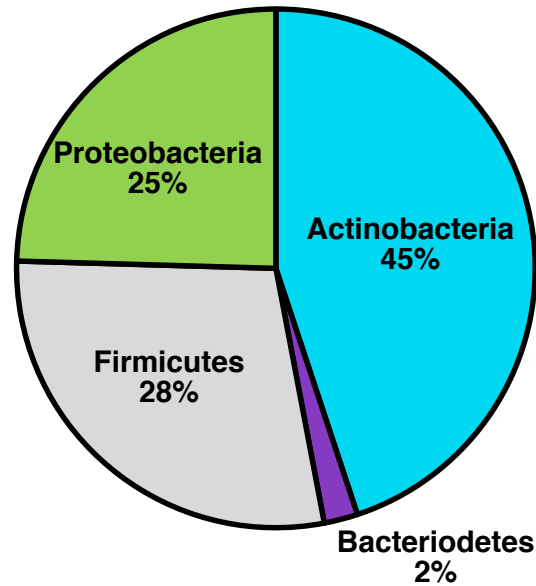


**Figure 2.1 Map displaying the location of four carbonate caves in the Guadalupe Mountains (Carlsbad Caverns National Park, New Mexico, USA).**

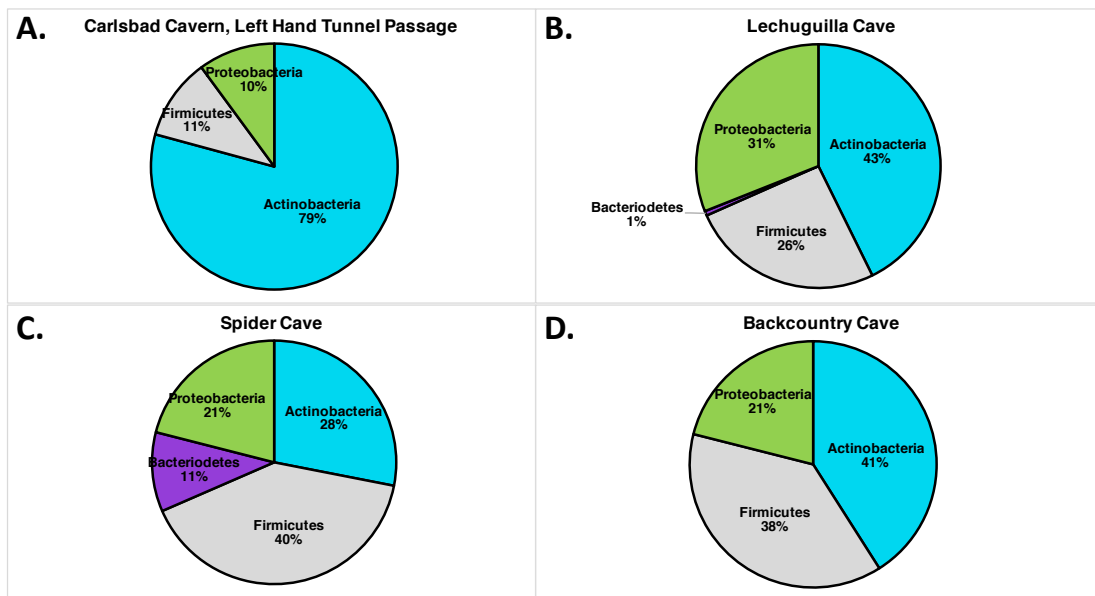


**Figure 2.2** The maximum likelihood phylogeny of 334 isolates cultured from caves and 100 bacterial type strain reference sequences. 16S rRNA sequences were aligned with MUSCLE and analyzed with RAxML. Genus level clades have been collapsed for viewing purposes.

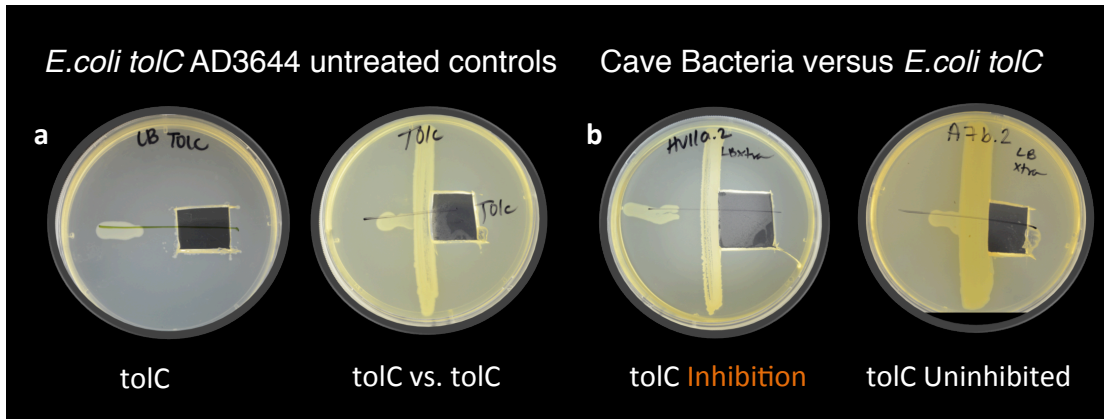
## Phyla of Bacteria Cultured from 4 Carbonate Caves



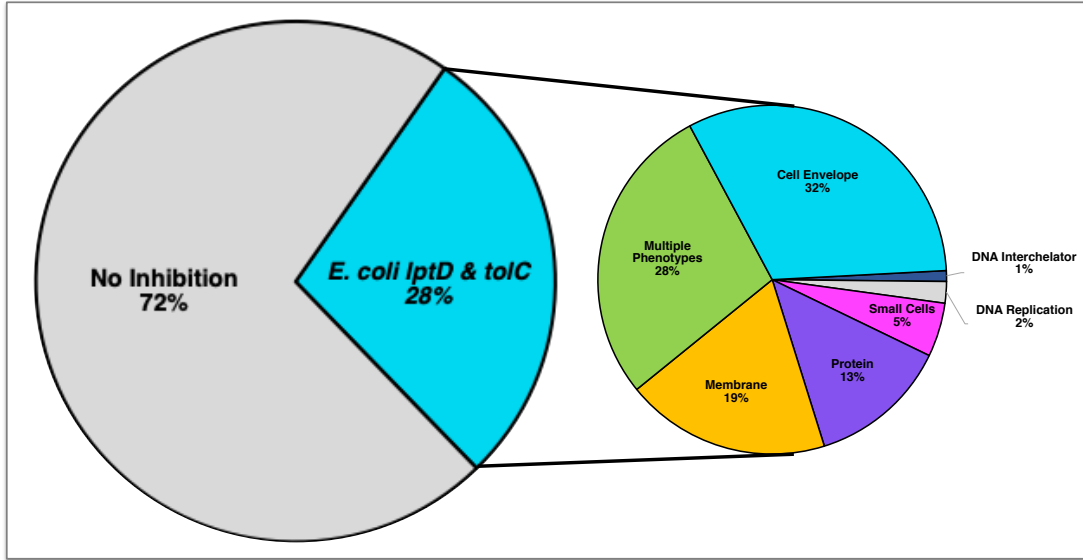
**Figure 2.3.** Phylum level bacterial diversity of the isolates cultured from four carbonate limestone caves; Lechuguilla, Backcountry, Spider, and Carlsbad Cavern. Members of the Actinobacteria were selectively cultured for this study.



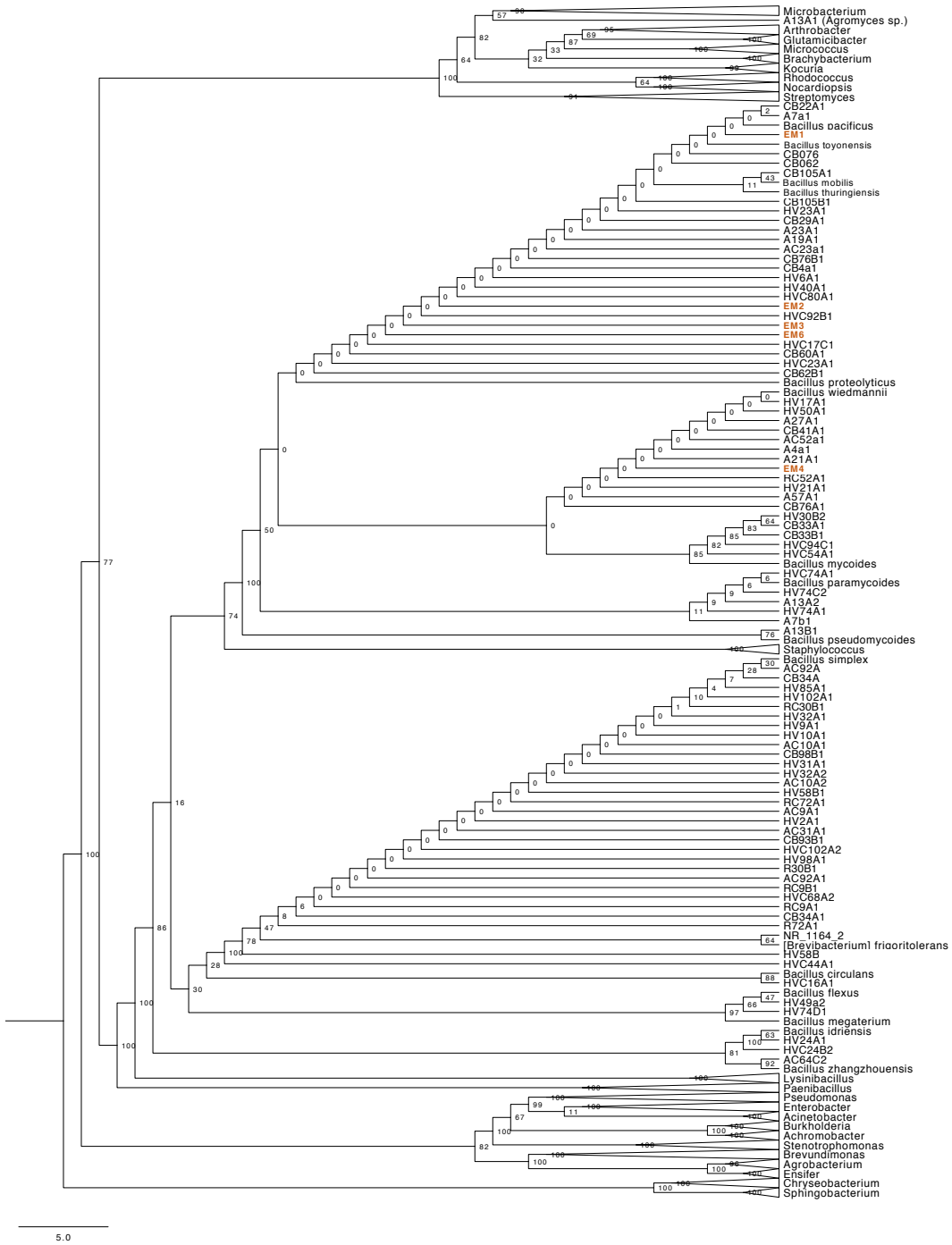
**Figure 2.4. Phylum level bacterial diversity of the isolates cultured from carbonate limestone caves. (A) Carlsbad Cavern, the Left Hand Tunnel Passage, (B) Lechuguilla Cave, (C) Spicer Cave, and (D) Backcountry Cave.**



**Figure 2.5. Screening cave bacteria for antibiotic production.** The cave bacteria is streaked in the middle of the plate and allowed to grow for 3 to 5 days. 10  $\mu$ l of an exponentially growing culture of *E. coli tolC* is then spotted in a line perpendicular to the cave strain, incubated for two hours, and the imaged using BCP<sup>23</sup>. Plates are incubated overnight at 30°C and the zone of inhibition on the plate is determined. **(A)** *E. coli tolC* control alone or *E. coli tolC* tested against *E. coli tolC* as a control. **(B)** *E. coli tolC* tested against two cave strains. *E. coli tolC* is inhibited by the cave strain HV11A2 (left) but not by A7B2 (right).

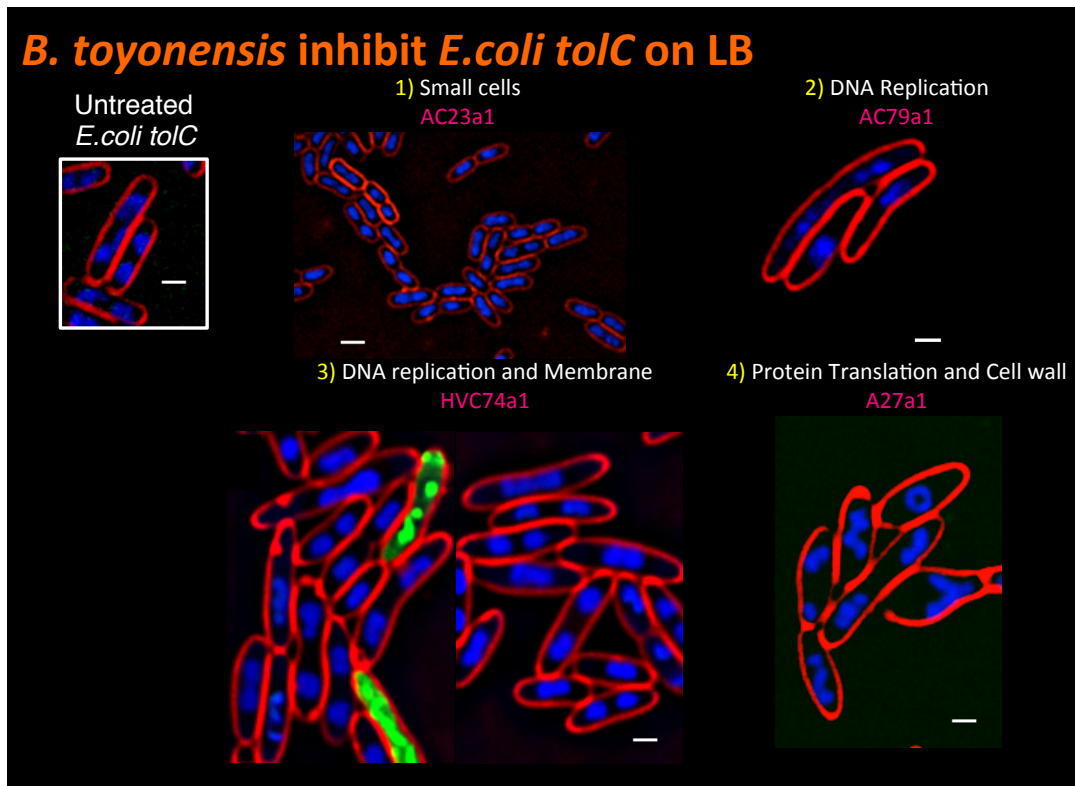


**Figure 2.6. Cave bacteria produce Gram-negative inhibitory bioactive compounds using a variety of mechanisms to target different cellular pathways.**

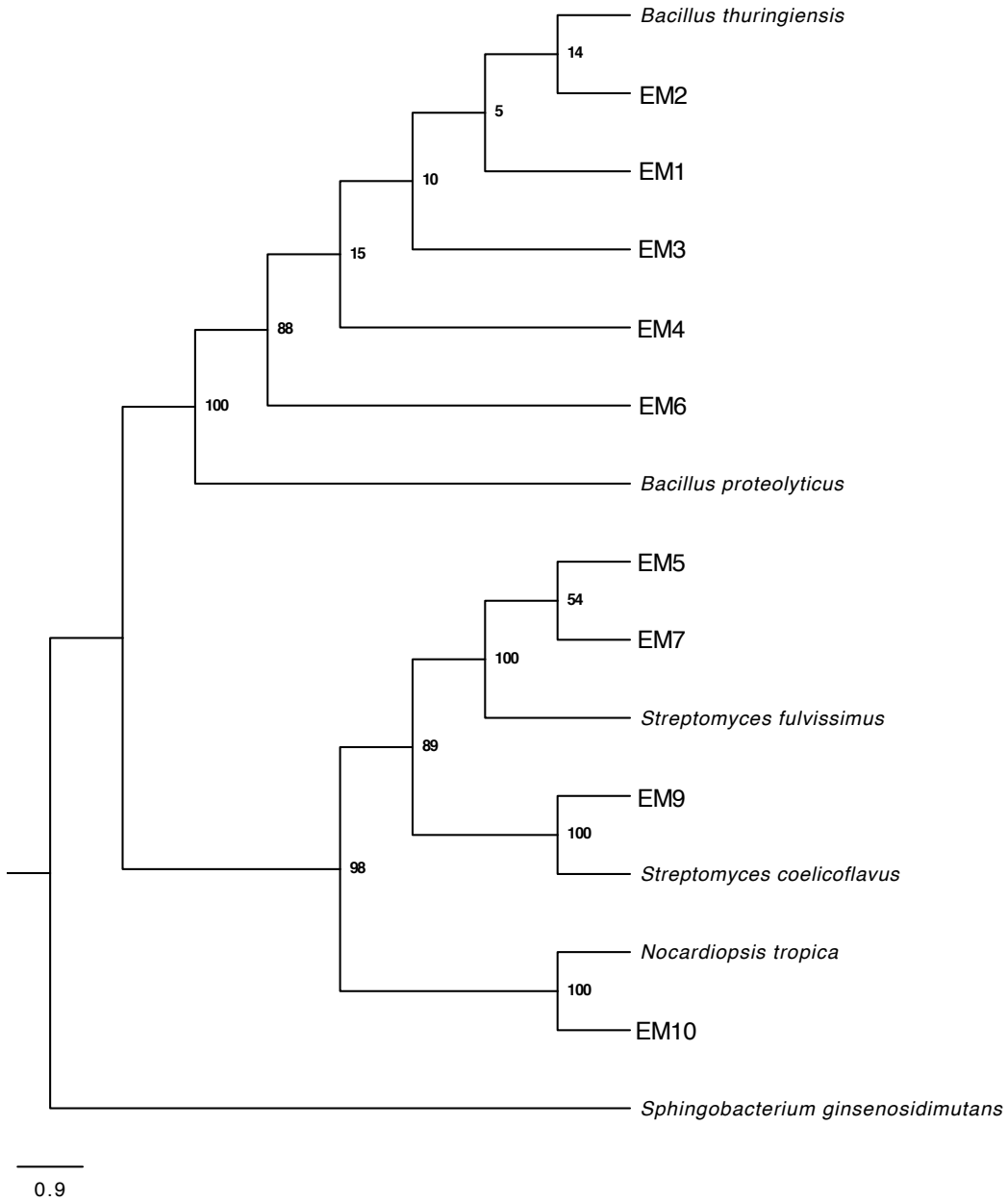


**Figure 2.7. The maximum likelihood phylogeny of 84 isolates cultured from caves and 17 bacterial type strain reference sequences.** 16S rRNA sequences were aligned with MUSCLE and analyzed with RAXML. Two closely related clades of *Bacillus* sp. isolates had differential inhibition abilities and major cellular targets against *E.coli tolC* when tested on LB with different cellular targets. Five isolates, highlighted in orange, from these clades were chosen for further study based on the difference in killing and MOA observed in BCP.

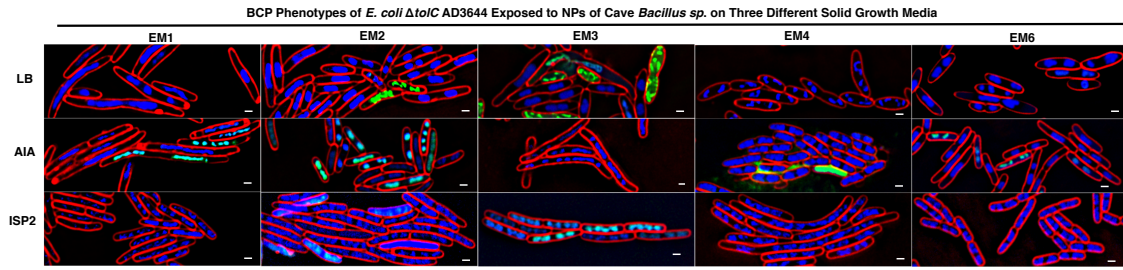




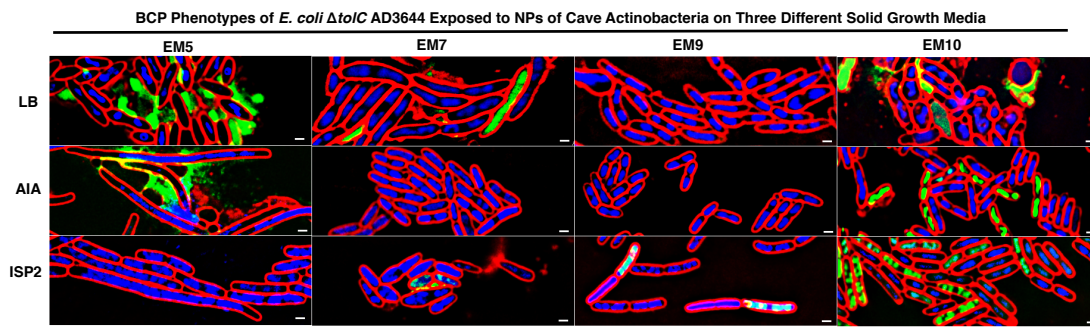
**Figure 2.8.** *Bacillus toyonensis* strains exhibit different phenotypes when tested against *E. coli tolC* on LB. (1) AC23a1 produces very small cells. (2) AC79a1 blocks DNA replication. (3) HVC74a1 targets DNA replication and membrane. (4) A27a1 targets protein translation and cell wall.



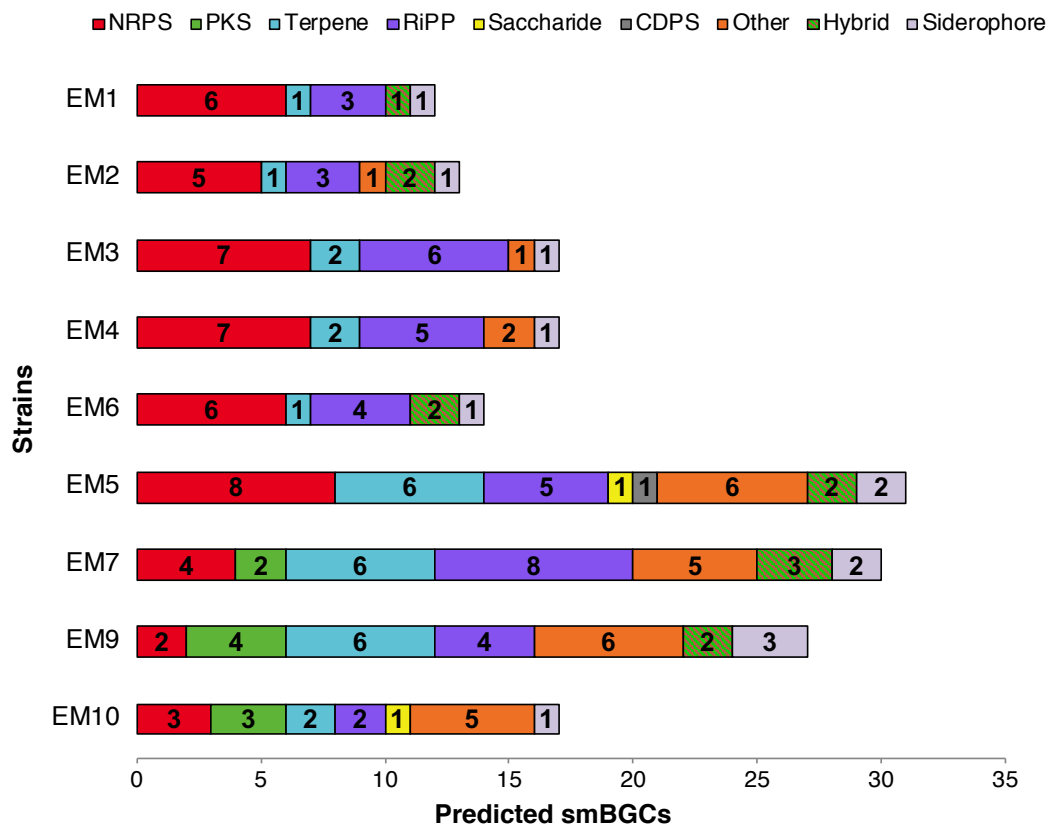
**Figure 2.9. The maximum likelihood phylogeny of 9 isolates cultured from caves and 6 bacterial type strain reference sequences. 16S rRNA sequences were aligned with MUSCLE and analyzed with RAxML.**



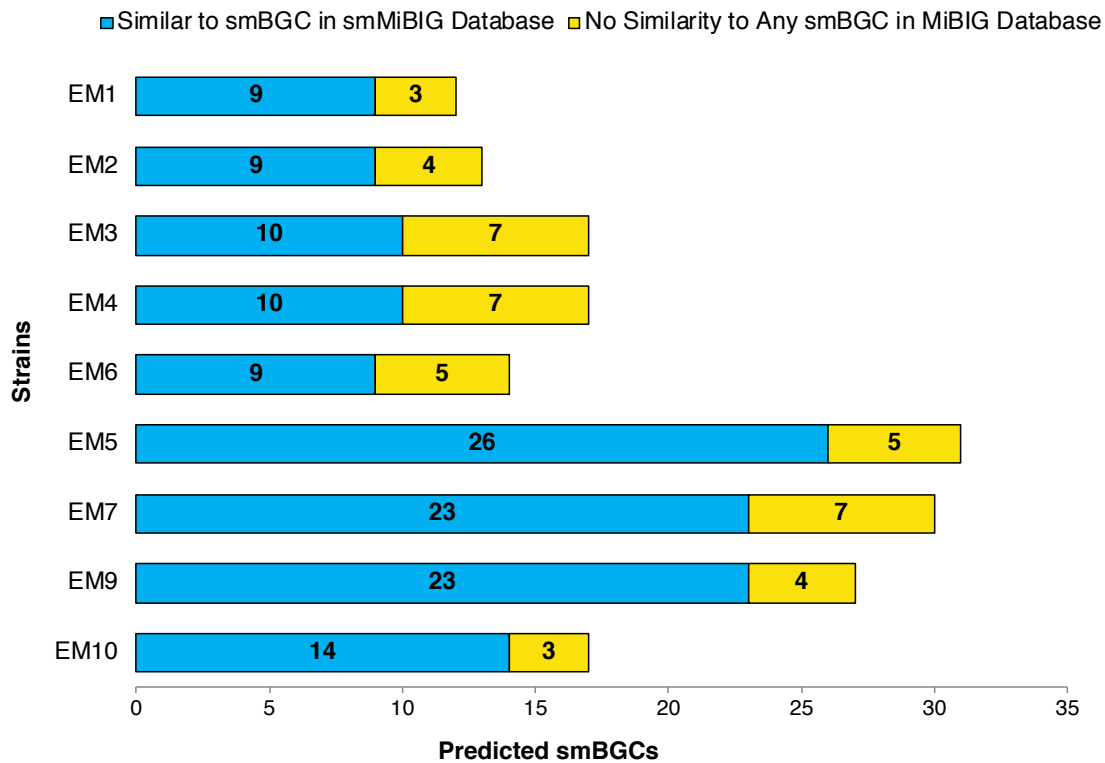
**Figure 2.10.** BCP analysis of 5 *Bacillus* cave strains.



**Figure 2.11. BCP analysis of 4 Actinobacteria cave strains.**



**Figure 2.12. Distribution of biosynthetic class among secondary metabolic gene clusters predicted in the genomes of 9 cave isolates.** The total number of BGCs associated with secondary metabolism according to antiSMASH5.0 and the distribution of their corresponding biosynthetic classes are shown.



**Figure 2.13. Sequence comparison of predicted BGCs to known secondary metabolic gene clusters.** Displaying the number of BGCs predicted in 9 bacterial genomes that share some amount of sequence similarity with clusters in the MiBIG database.

## 2.7 Tables

**Table 2.1. Specialized growth media used to culture and isolate bacteria from semi-arid limestone caves.**

Medium Name	Medium Composition (1 L)
<b>1/2 R2A</b>	yeast extract 0.5g, proteose peptone No.3 0.5g, casamino acids 0.5g, dextrose 0.5g, soluble starch 0.5g, sodium pyruvate 0.3g, dipotassium phosphate 0.3g, magnesium sulfate 0.05g, agar 30g. (pH 7.2± 0.2). ± pulverized limesone
<b>Bennet's</b>	yeast extract 1g, glucose 10g, casein 2g, nutrient agar with beef extract 23g. (pH 7.2± 0.2)
<b>HVA</b>	Na <sub>2</sub> HPO <sub>4</sub> 0.5g, KCl 1.71g, MgSO <sub>4</sub> *7H <sub>2</sub> O 0.05g, FeSO <sub>4</sub> *7H <sub>2</sub> O 0.01g, CaCO <sub>3</sub> 0.02g, humic acid 1g, agar 18g.
<b>RASS</b>	starch orglycerol 12.5g, L-arginine 0.1g, K <sub>2</sub> HPO <sub>4</sub> 1g, NaCl 1g, MgSO <sub>4</sub> 7H <sub>2</sub> O 0.5g, trace elements 1 mL, agar 15g. (trace elements/100ml CuSO <sub>4</sub> 0.1g, MnSO <sub>4</sub> 0.1g, Fe <sub>2</sub> (SO <sub>4</sub> ) <sub>3</sub> -6H <sub>2</sub> O 1.0g, ZnSO <sub>4</sub> - 7H <sub>2</sub> O 0.1g)
<b>AIA</b>	sodium caseinate 2g, L-Asparagine 0.1g, sodium propionate 4g, dipotassium phosphate 0.5g, magnesium sulphate 0.1g, glycerol 5 mL, rifampicin 50 µg/mL-1, cycloheximide 100 µg/mL-1, agar 15g. (pH 8.1 ± 0.2)
<b>LB</b>	yeast extract 5g, tryptone 10g, NaCl 10g, agar 15g. (pH 7.0)
<b>ISP2</b>	yeast extract 4g, malt extract 10g, dextrose 4g, agar 2g. (pH 7.0)

**Table 2.2. Genus classifications of cave isolates based on maximum likelihood analysis of 16S ribosomal RNA sequence data.**

<b>Strain ID</b>	<b>Genus Classification</b>
A10A1	<i>Micrococcus</i>
A12A1	<i>Chryseobacterium</i>
A12B1	<i>Chryseobacterium</i>
A13A1	<i>Agromyces</i>
A13A2	<i>Bacillus</i>
A13B1	<i>Bacillus</i>
A17A1	<i>Bacillus</i>
A18A1	<i>Microbacterium</i>
A18A2	<i>Microbacterium</i>
A19A1	<i>Bacillus</i>
A21A1	<i>Bacillus</i>
A22A1	<i>Microbacterium</i>
A22A2	<i>Microbacterium</i>
A22C1	<i>Microbacterium</i>
A22D1	<i>Microbacterium</i>
A23A1	<i>Bacillus</i>
A25B1	<i>Microbacterium</i>
A25C1	<i>Microbacterium</i>
A26A1	<i>Arthrobacter</i>
A27A1	<i>Bacillus</i>
A30A1	<i>Enterobacter</i>
A34A1	<i>Arthrobacter</i>
A34B1	<i>Stenotrophomonas</i>
A42B1	<i>Enterobacter</i>
A43A1	<i>Pseudomonas</i>
A43B1	<i>Pseudomonas</i>
A4A1	<i>Bacillus</i>
A56C1	<i>Microbacterium</i>
A57A1	<i>Bacillus</i>
A61A1	<i>Rhodococcus</i>
A65A1	<i>Microbacterium</i>
A73A1	<i>Glutamicibacter</i>
A77A1	<i>Arthrobacter</i>
A7A1	<i>Bacillus</i>
A7A2	<i>Stenotrophomonas</i>
A7B1	<i>Bacillus</i>
A7B2	<i>Stenotrophomonas</i>
A82B1	<i>Kocuria</i>
AC106A1	<i>Paenibacillus</i>
AC106A2	<i>Paenibacillus</i>
AC10A1	<i>Bacillus</i>
AC10A2	<i>Bacillus</i>



**Table 2.2. Genus classifications of cave isolates (continued).**

<b>Strain ID</b>	<b>Genus Classification</b>
AC12B1	<i>Chryseobacterium</i>
AC23a1	<i>Bacillus</i>
AC24B1	<i>Paenibacillus</i>
AC24B2	<i>Paenibacillus</i>
AC30B1	<i>Enterobacter</i>
AC31A1	<i>Bacillus</i>
AC38A1	<i>Stenotrophomonas</i>
AC42A2	<i>Enterobacter</i>
AC43A1	<i>Bacillus</i>
AC52a1	<i>Bacillus</i>
AC56B	<i>Arthrobacter</i>
AC57A1	<i>Microbacterium</i>
AC58B1	<i>Microbacterium</i>
AC61A1	<i>Lysinibacillus</i>
AC61A2	<i>Paenibacillus</i>
AC61A3	<i>Paenibacillus</i>
AC64B1	<i>Microbacterium</i>
AC64C1	<i>Arthrobacter</i>
AC64C2	<i>Bacillus</i>
AC64C3	<i>Arthrobacter</i>
AC64D1	<i>Microbacterium</i>
AC67A1	<i>Stenotrophomonas</i>
AC67B1	<i>Stenotrophomonas</i>
AC68A1	<i>Brevundimonas</i>
AC69B1	<i>Enterobacter</i>
AC73A1	<i>Glutamicibacter</i>
AC79A1	<i>Bacillus</i>
AC81B1	<i>Microbacterium</i>
AC83B1	<i>Arthrobacter</i>
AC83B2	<i>Microbacterium</i>
AC85A1	<i>Arthrobacter</i>
AC85A2	<i>Arthrobacter</i>
AC86C1	<i>Microbacterium</i>
AC86D1	<i>Microbacterium</i>
AC86D2	<i>Brevundimonas</i>
AC88A1	<i>Arthrobacter</i>
AC8A1	<i>Microbacterium</i>
AC8A2	<i>Rhodococcus</i>
AC8B1	<i>Enterobacter</i>
AC8B2	<i>Microbacterium</i>
AC91A1	<i>Arthrobacter</i>
AC92A	<i>Bacillus</i>

**Table 2.2. Genus classifications of cave isolates (continued).**

<b>Strain ID</b>	<b>Genus Classification</b>
AC92A1	<i>Bacillus</i>
AC92C1	<i>Lysinibacillus</i>
AC94B1	<i>Brevundimonas</i>
AC94B2	<i>Microbacterium</i>
AC95B1	<i>Paenibacillus</i>
AC95B2	<i>Paenibacillus</i>
AC9A1	<i>Bacillus</i>
ARC1	<i>Streptomyces</i>
ARC10	<i>Streptomyces</i>
ARC11	<i>Streptomyces</i>
ARC12	<i>Streptomyces</i>
ARC13	<i>Streptomyces</i>
ARC14	<i>Streptomyces</i>
ARC15	<i>Streptomyces</i>
ARC16	<i>Achromobacter</i>
ARC18	<i>Burkholderia</i>
ARC2	<i>Streptomyces</i>
ARC21	<i>Streptomyces</i>
ARC22	<i>Streptomyces</i>
ARC25	<i>Nocardiopsis</i>
ARC26	<i>Nocardiopsis</i>
ARC27	<i>Nocardiopsis</i>
ARC28	<i>Streptomyces</i>
ARC29	<i>Streptomyces</i>
ARC3	<i>Streptomyces</i>
ARC31	<i>Streptomyces</i>
ARC32	<i>Streptomyces</i>
ARC33	<i>Streptomyces</i>
ARC34	<i>Streptomyces</i>
ARC35	<i>Nocardiopsis</i>
ARC36	<i>Nocardiopsis</i>
ARC37	<i>Nocardiopsis</i>
ARC38	<i>Streptomyces</i>
ARC4	<i>Streptomyces</i>
ARC40	<i>Streptomyces</i>
ARC41	<i>Streptomyces</i>
ARC42	<i>Streptomyces</i>
ARC5	<i>Streptomyces</i>
ARC6	<i>Streptomyces</i>
ARC7	<i>Streptomyces</i>
R30B1	<i>Bacillus</i>
R30B2	<i>Stenotrophomonas</i>

**Table 2.2. Genus classifications of cave isolates (continued).**

<b>Strain ID</b>	<b>Genus Classification</b>
R55A1	<i>Microbacterium</i>
R72A1	<i>Bacillus</i>
R87A1	<i>Microbacterium</i>
RC26A1	<i>Pseudomonas</i>
RC30B1	<i>Bacillus</i>
RC34B1	<i>Brevundimonas</i>
RC43A1	<i>Brevundimonas</i>
RC43A2	<i>Brevundimonas</i>
RC43B1	<i>Microbacterium</i>
RC43B2	<i>Lysinibacillus</i>
RC45B1	<i>Microbacterium</i>
RC52A1	<i>Bacillus</i>
RC53A1	<i>Microbacterium</i>
RC53A2	<i>Microbacterium</i>
RC55A1	<i>Microbacterium</i>
RC59A1	<i>Microbacterium</i>
RC72A1	<i>Bacillus</i>
RC83B1	<i>Brevundimonas</i>
RC8A1	<i>Pseudomonas</i>
RC9A1	<i>Bacillus</i>
RC9B1	<i>Bacillus</i>
CB062A	<i>Bacillus</i>
CB076B	<i>Bacillus</i>
CB102A1	<i>Bacillus</i>
CB103A1	<i>Paenibacillus</i>
CB105A1	<i>Bacillus</i>
CB105B1	<i>Bacillus</i>
CB106A1	<i>Micrococcus</i>
CB106A2	<i>Brevundimonas</i>
CB106B1	<i>Brevundimonas</i>
CB10A1	<i>Lysinibacillus</i>
CB112A1	<i>Pseudomonas</i>
CB112B	<i>Paenibacillus</i>
CB113A1	<i>Paenibacillus</i>
CB113B1	<i>Paenibacillus</i>
CB11A1	<i>Paenibacillus</i>
CB12A1	<i>Burkholderia</i>
CB18A1	<i>Paenibacillus</i>
CB18A2	<i>Paenibacillus</i>
CB21A1	<i>Pseudomonas</i>
CB22A1	<i>Bacillus</i>
CB23B	<i>Paenibacillus</i>

**Table 2.2. Genus classifications of cave isolates (continued).**

<b>Strain ID</b>	<b>Genus Classification</b>
CB23B1	<i>Paenibacillus</i>
CB24C1	<i>Acinetobacter</i>
CB28A1	<i>Bacillus</i>
CB29A1	<i>Bacillus</i>
CB32A1	<i>Enterobacter</i>
CB33A1	<i>Bacillus</i>
CB33B1	<i>Bacillus</i>
CB34A	<i>Bacillus</i>
CB34A1	<i>Bacillus</i>
CB39A1	<i>Enterobacter</i>
CB40A1	<i>Enterobacter</i>
CB41A1	<i>Bacillus</i>
CB42A1	<i>Acinetobacter</i>
CB43A1	<i>Acinetobacter</i>
CB43B1	<i>Microbacterium</i>
CB43B2	<i>Pseudomonas</i>
CB4a1	<i>Bacillus</i>
CB52A1	<i>Lysinibacillus</i>
CB53A1	<i>Enterobacter</i>
CB56A1	<i>Stenotrophomonas</i>
CB58A1	<i>Rhodococcus</i>
CB60A1	<i>Bacillus</i>
CB62A1	<i>Bacillus</i>
CB62B1	<i>Bacillus</i>
CB63A1	<i>Pseudomonas</i>
CB64A1	<i>Enterobacter</i>
CB67A1	<i>Pseudomonas</i>
CB67A2	<i>Pseudomonas</i>
CB69A1	<i>Stenotrophomonas</i>
CB75B1	<i>Pseudomonas</i>
CB76A1	<i>Bacillus</i>
CB76B1	<i>Bacillus</i>
CB8A1	<i>Paenibacillus</i>
CB8B1	<i>Achromobacter</i>
CB8C1	<i>Achromobacter</i>
CB91A1	<i>Paenibacillus</i>
CB93A1	<i>Stenotrophomonas</i>
CB93B1	<i>Bacillus</i>
CB94A1	<i>Enterobacter</i>
CB94B1	<i>Enterobacter</i>
CB97A1	<i>Stenotrophomonas</i>
CB97B1	<i>Enterobacter</i>

**Table 2.2. Genus classifications of cave isolates (continued).**

<b>Strain ID</b>	<b>Genus Classification</b>
CB98B1	<i>Bacillus</i>
CB99A1	<i>Burkholderia</i>
HV102A1	<i>Bacillus</i>
HV102A2	<i>Microbacterium</i>
HV108A1	<i>Paenibacillus</i>
HV10A1	<i>Bacillus</i>
HV11A1	<i>Paenibacillus</i>
HV11A2	<i>Paenibacillus</i>
HV12A1	<i>Agrobacterium</i>
HV17A1	<i>Bacillus</i>
HV20A1	<i>Lysinibacillus</i>
HV20A2	<i>Lysinibacillus</i>
HV21A1	<i>Bacillus</i>
HV22A1	<i>Brachybacterium</i>
HV22A2	<i>Brachybacterium</i>
HV23A1	<i>Bacillus</i>
HV24A1	<i>Bacillus</i>
HV25B1	<i>Microbacterium</i>
HV25C1	<i>Pseudomonas</i>
HV27B1	<i>Microbacterium</i>
HV2A1	<i>Bacillus</i>
HV30B2	<i>Bacillus</i>
HV31A1	<i>Bacillus</i>
HV32A1	<i>Bacillus</i>
HV32A2	<i>Bacillus</i>
HV40A1	<i>Bacillus</i>
HV41A1	<i>Lysinibacillus</i>
HV43A1	<i>Brevundimonas</i>
HV43A2	<i>Brevundimonas</i>
HV43B1	<i>Pseudomonas</i>
HV44B1	<i>Microbacterium</i>
HV47A1	<i>Microbacterium</i>
HV49A1	<i>Ensifer</i>
HV49A2	<i>Bacillus</i>
HV50A1	<i>Bacillus</i>
HV55A1	<i>Microbacterium</i>
HV56A1	<i>Microbacterium</i>
HV58B	<i>Bacillus</i>
HV58B1	<i>Bacillus</i>
HV61A1	<i>Rhodococcus</i>
HV66A1	<i>Paenibacillus</i>
HV66B1	<i>Micrococcus</i>

**Table 2.2. Genus classifications of cave isolates (continued).**

<b>Strain ID</b>	<b>Genus Classification</b>
HV68A1	<i>Staphylococcus</i>
HV69A1	<i>Agrobacterium</i>
HV69B1	<i>Sphingobacterium</i>
HV69B2	<i>Brevundimonas</i>
HV6A1	<i>Bacillus</i>
HV71A1	<i>Microbacterium</i>
HV74A1	<i>Bacillus</i>
HV74B	<i>Brevundimonas</i>
HV74C1	<i>Paenibacillus</i>
HV74C2	<i>Bacillus</i>
HV74D1	<i>Bacillus</i>
HV76A1	<i>Microbacterium</i>
HV77B1	<i>Microbacterium</i>
HV83A1	<i>Rhodococcus</i>
HV83B1	<i>Rhodococcus</i>
HV83C1	<i>Microbacterium</i>
HV85A1	<i>Bacillus</i>
HV87B1	<i>Microbacterium</i>
HV88A1	<i>Pseudomonas</i>
HV88A2	<i>Microbacterium</i>
HV94C1	<i>Brevundimonas</i>
HV98A1	<i>Bacillus</i>
HV9A1	<i>Bacillus</i>
HVC102A2	<i>Bacillus</i>
HVC106B1	<i>Paenibacillus</i>
HVC106B2	<i>Paenibacillus</i>
HVC108A1	<i>Paenibacillus</i>
HVC12A	<i>Chryseobacterium</i>
HVC12A1	<i>Chryseobacterium</i>
HVC16A1	<i>Bacillus</i>
HVC16A2	<i>Rhodococcus</i>
HVC17C1	<i>Bacillus</i>
HVC23A1	<i>Bacillus</i>
HVC24B1	<i>Brevundimonas</i>
HVC24B2	<i>Bacillus</i>
HVC33B1	<i>Microbacterium</i>
HVC33B2	<i>Microbacterium</i>
HVC34A1	<i>Stenotrophomonas</i>
HVC34A2	<i>Pseudomonas</i>
HVC43B1	<i>Pseudomonas</i>
HVC44A1	<i>Bacillus</i>
HVC44A2	<i>Microbacterium</i>

**Table 2.2. Genus classifications of cave isolates (continued).**

<b>Strain ID</b>	<b>Genus Classification</b>
HVC46A	<i>Microbacterium</i>
HVC46A1	<i>Microbacterium</i>
HVC47A1	<i>Microbacterium</i>
HVC48A1	<i>Ensifer</i>
HVC51A1	<i>Brevundimonas</i>
HVC53A1	<i>Microbacterium</i>
HVC54A1	<i>Bacillus</i>
HVC55A1	<i>Microbacterium</i>
HVC56A1	<i>Pseudomonas</i>
HVC59A1	<i>Microbacterium</i>
HVC61A1	<i>Paenibacillus</i>
HVC61A2	<i>Rhodococcus</i>
HVC65A1	<i>Pseudomonas</i>
HVC65A2	<i>Pseudomonas</i>
HVC66A1	<i>Paenibacillus</i>
HVC68A1	<i>Staphylococcus</i>
HVC68A2	<i>Bacillus</i>
HVC68B1	<i>Microbacterium</i>
HVC69B1	<i>Rhodococcus</i>
HVC72A1	<i>Kocuria</i>
HVC74A1	<i>Bacillus</i>
HVC7B1	<i>Stenotrophomonas</i>
HVC80A1	<i>Bacillus</i>
HVC81A1	<i>Paenibacillus</i>
HVC81D1	<i>Microbacterium</i>
HVC81D2	<i>Pseudomonas</i>
HVC83A1	<i>Pseudomonas</i>
HVC83A2	<i>Brevundimonas</i>
HVC86A1	<i>Brevundimonas</i>
HVC8A1	<i>Microbacterium</i>
HVC90A1	<i>Paenibacillus</i>
HVC90B1	<i>Microbacterium</i>
HVC91A1	<i>Arthrobacter</i>
HVC92A1	<i>Sphingobacterium</i>
HVC92B1	<i>Bacillus</i>
HVC94C1	<i>Bacillus</i>
HVC94D1	<i>Microbacterium</i>
HVC94D2	<i>Microbacterium</i>
HVC97A1	<i>Micrococcus</i>
HVC97B1	<i>Microbacterium</i>

**Table 2.3. 100 Reference type strain 16S ribosomal RNA sequences used in maximum likelihood analysis and classification of cave isolates.**

<b>Accession</b>
NR_024640
NR_024957
NR_025083
NR_025228
NR_025517
NR_025548
NR_025589
NR_025922
NR_026162
NR_026194
NR_026234
NR_026342
NR_026452
NR_027199
NR_028993
NR_029109
NR_037024
NR_041175
NR_041210
NR_041396
NR_041546
NR_042092
NR_042136
NR_042263
NR_042568
NR_043268
NR_043535
NR_043846
NR_043854
NR_044365
NR_044524
NR_044525
NR_044906
NR_075062
NR_102890
NR_103947
NR_104689
NR_104776
NR_104795
NR_104978
NR_104982
NR_108441
NR_108531
NR_108689
NR_112067
NR_112192
NR_112261
NR_112306
NR_112403



**Table 2.3. 100 Reference type strain 16S ribosomal RNA sequences (continued).**

<b>Accession</b>
NR 112429
NR 112527
NR 112569
NR 112632
NR 112746
NR 112780
NR 113341
NR 113611
NR 113627
NR 113648
NR 113732
NR 113800
NR 113828
NR 113855
NR 113893
NR 113990
NR 113991
NR 114435
NR 114710
NR 114963
NR 114986
NR 115064
NR 116448
NR 116722
NR 116873
NR 117268
NR 117621
NR 117644
NR 118008
NR 118146
NR 118617
NR 121761
NR 125618
NR 133968
NR 134084
NR 134114
NR 134118
NR 145647
NR 145660
NR 145886
NR 146667
NR 148610
NR 148786
NR 152692
NR 156872
NR 157010
NR 157731
NR 157733
NR 157734
NR 157735
NR 163642

**Table 2.4. Inhibition of pathogens by bacterial isolated from caves.** Isolates EM4, EM9, and EM10 had no inhibitory activity against test pathogens, whereas 4 isolates inhibited the clinically relevant *E. coli* ATCC 25922 strain (EM1-3 and EM6), and five inhibited a MDR clinical isolate of MRSA (USA300 TCH1516) (EM1-2, EM5-7).

Isolate	Cave and Microhabitat	<i>E.coli</i> $\Delta$ tolC AD3644 MOA (LB)	Pathogen Inhibition
<b>EM1</b> <i>Bacillus sp.</i>	Lechuguilla Cave column	DNA Replication	<i>E. coli</i> ATCC 25922 (WT) <i>S. aureus</i> USA 300 TCH1516 (MRSA)
<b>EM2</b> <i>Bacillus sp.</i>	Lechuguilla Red Fe-Mn deposits and crystals	DNA Replication and Membrane	<i>E. coli</i> ATCC 25922 (WT) <i>S. aureus</i> USA 300 TCH1516 (MRSA)
<b>EM3</b> <i>Bacillus sp.</i>	Back Country Cave Rock	Cell wall or Membrane	<i>E. coli</i> ATCC 25922 (WT)
<b>EM4</b> <i>Bacillus sp.</i>	Back Country Cave Rock	Cell wall and DNA	None
<b>EM6</b> <i>Bacillus sp.</i>	Lechuguilla Red FeMn deposits and crystals	Cell Wall and DNA	<i>E. coli</i> ATCC 25922 (WT) <i>S. aureus</i> USA 300 TCH1516 (MRSA)
<b>EM5</b> <i>S. griseus</i>	Carlsbad Cavern stalagmite	Cell wall, Membrane and Protein translation	<i>S. aureus</i> USA 300 TCH1516 (MRSA) <i>P. aeruginosa</i> K2733 ( $\Delta$ pump)
<b>EM7</b> <i>S. fulvissimus</i>	Carlsbad Cavern water droplet	Cell wall / Cell envelope	<i>S. aureus</i> USA 300 TCH1516 (MRSA)
<b>EM9</b> <i>S. lividans</i>	Carlsbad Cavern cave formation	Cell Wall	None
<b>EM10</b> <i>N. umidischolae</i>	Carlsbad Cavern cave column	Cell wall and DNA	None

## 2.8 References

1. Woc-Colburn, L. & Francisco, D. M. A. in *Highly Infect. Dis. Crit. Care* 139–146 (Springer, 2020).
2. Davies, J. & Davies, D. Microbiology and Molecular Biology Reviews Microbiology and Molecular Biology Reviews Origins and Evolution of Antibiotic Resistance. *Mol. Biol. Rev. Microbiol. Mol. Biol. Rev* (2010). doi:10.1128/MMBR.00016-10
3. Felden, B. & Cattoira, V. Bacterial Adaptation to Antibiotics through Regulatory RNAs. *Antimicrob. Agents Chemother.* (2018). doi:10.1128/AAC.02503-17
4. Strachan, C. R. & Davies, J. The whys and wherefores of antibiotic resistance. *Cold Spring Harb. Perspect. Med.* (2017). doi:10.1101/cshperspect.a025171
5. Clardy, J., Fischbach, M. A. & Walsh, C. T. New antibiotics from bacterial natural products. *Nat. Biotechnol.* (2006). doi:10.1038/nbt1266
6. Newman, D. J. & Cragg, G. M. Natural Products as Sources of New Drugs from 1981 to 2014. *J. Nat. Prod.* **79**, 629–661 (2016).
7. Monciardini, P., Iorio, M., Maffioli, S., Sosio, M. & Donadio, S. Discovering new bioactive molecules from microbial sources. *Microb. Biotechnol.* **7**, 209–220 (2014).
8. Genilloud, O. Actinomycetes: Still a source of novel antibiotics. *Nat. Prod. Rep.* **34**, 1203–1232 (2017).
9. Smanski, M. J., Schlatter, D. C. & Kinkel, L. L. Leveraging ecological theory to guide natural product discovery. *J. Ind. Microbiol. Biotechnol.* **43**, 115–128 (2016).
10. Fischbach, M. A. & Walsh, C. T. Antibiotics for emerging pathogens. *Science (80- .)* **325**, 1089–1093 (2009).
11. Hover, B. M., Kim, S. H., Katz, M., Charlop-Powers, Z., Owen, J. G., Ternei, M. A., Maniko, J., Estrela, A. B., Molina, H., Park, S., Perlin, D. S. & Brady, S. F. Culture-independent discovery of the malacidins as calcium-dependent antibiotics with activity against multidrug-resistant Gram-positive pathogens. *Nat. Microbiol.* **3**, 415–422 (2018).
12. Ling, L. L., Schneider, T., Peoples, A. J., Spoering, A. L., Engels, I., Conlon, B. P., Mueller, A., Schäberle, T. F., Hughes, D. E., Epstein, S., Jones, M., Lazarides, L., Steadman, V. A., Cohen, D. R., Felix, C. R., Fetterman, K. A., Millett, W. P., Nitti, A. G., Zullo, A. M., Chen, C. & Lewis, K. A new antibiotic kills pathogens

- without detectable resistance. *Nature* **517**, 455–459 (2015).
13. Yamanaka, K., Reynolds, K. A., Kersten, R. D., Ryan, K. S., Gonzalez, D. J., Nizet, V., Dorrestein, P. C. & Moore, B. S. Direct cloning and refactoring of a silent lipopeptide biosynthetic gene cluster yields the antibiotic taromycin A. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 1957–1962 (2014).
  14. Bull, A. T. & Asenjo, J. A. Microbiology of hyper-arid environments: Recent insights from the Atacama Desert, Chile. *Antonie van Leeuwenhoek, Int. J. Gen. Mol. Microbiol.* **103**, 1173–1179 (2013).
  15. Shi, Y., Pan, C., Auckloo, B. N., Chen, X., Chen, C. T. A., Wang, K., Wu, X., Ye, Y. & Wu, B. Stress-driven discovery of a cryptic antibiotic produced by *Streptomyces* sp. WU20 from Kueishantao hydrothermal vent with an integrated metabolomics strategy. *Appl. Microbiol. Biotechnol.* **101**, 1395–1408 (2017).
  16. Li, W. ting, Luo, D., Huang, J. ning, Wang, L. lin, Zhang, F. guo, Xi, T., Liao, J. min & Lu, Y. yuan. Antibacterial constituents from Antarctic fungus, *Aspergillus sydowii* SP-1. *Nat. Prod. Res.* **32**, 662–667 (2018).
  17. Shah, M., Sun, C., Sun, Z., Zhang, G., Che, Q., Gu, Q., Zhu, T. & Li, D. Antibacterial Polyketides from Antarctica Sponge-Derived Fungus *Penicillium* sp. HDN151272. *Mar. Drugs* **18**, (2020).
  18. Atalah, J., Blamey, L., Muñoz-Ibacache, S., Gutierrez, F., Urzua, M., Encinas, M. V., Páez, M., Sun, J. & Blamey, J. M. Isolation and characterization of violacein from an Antarctic *Iodobacter*: a non-pathogenic psychrotolerant microorganism. *Extremophiles* **24**, 43–52 (2020).
  19. Han, Y., Wang, Y., Yang, Y. & Chen, H. Shellmycin A – D , Novel Bioactive Tetrahydroanthra-. (2020).
  20. Zhang, J., Li, B., Qin, Y., Karthik, L., Zhu, G., Hou, C., Jiang, L., Liu, M., Ye, X., Liu, M., Hsiang, T., Dai, H., Zhang, L. & Liu, X. A new abyssomicin polyketide with anti-influenza A virus activity from a marine-derived *Verrucospora* sp. MS100137. *Appl. Microbiol. Biotechnol.* 1533–1543 (2020). doi:10.1007/s00253-019-10217-2
  21. Balansa, W., Wodi, S. I. M., Rieuwpassa, F. J. & Ijong, F. G. Agelasines B , D and antimicrobial extract of a marine sponge *Agelas* sp . from Tahuna Bay , Sangihe Islands , Indonesia. **21**, 699–706 (2020).
  22. Le, T. C., Lee, E. J., Lee, J., Hong, A., Yim, C. Y., Yang, I., Choi, H., Chin, J., Cho, S. J., Ko, J., Hwang, H., Nam, S. J. & Fenical, W. Saccharoquinoline, a cytotoxic alkaloidal meroterpenoid from marine-derived bacterium *Saccharomonospora* sp. *Mar. Drugs* **17**, 4–11 (2019).

23. Bukelskis, D., Dabkevičienė, D., Lukosevičiūtė, L., Bucelis, A., Kriaučiūnas, I., Lebedeva, J. & Kuisiėnė, N. Screening and Transcriptional Analysis of Polyketide Synthases and Non-ribosomal Peptide Synthetases in Bacterial Strains From Krubera–Voronja Cave. *Front. Microbiol.* **10**, 1–11 (2019).
24. Long, Y., Jiang, J., Hu, X., Zhou, J., Hu, J. & Zhou, S. Actinobacterial community in Shuanghe Cave using culture-dependent and -independent approaches. *World J. Microbiol. Biotechnol.* **35**, (2019).
25. Rangseekaew, P. & Pathom-Aree, W. Cave actinobacteria as producers of bioactive metabolites. *Front. Microbiol.* **10**, (2019).
26. Syiemiong, D. & Jha, D. K. Antibacterial potential of Actinobacteria from a Limestone Mining Site in Meghalaya, India. *J. Pure Appl. Microbiol.* **13**, 789–802 (2019).
27. Pye, C. R., Bertin, M. J., Lokey, R. S., Gerwick, W. H. & Linington, R. G. Retrospective analysis of natural products provides insights for future discovery trends. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 5601–5606 (2017).
28. Schabereiter-gurtner, C., Piñar, G., Rölleke, S. & Bustillo, T. Culture-independent analyses of bacterial communities on paleolithic paintings and surrounding rock walls in karstic caves. Altamira, Tito Bustillo, La Garma and Llonín. 1–3 (2002).
29. Schabereiter-Gurtner, C., Saiz-Jimenez, C., Piñar, G., Lubitz, W. & Rölleke, S. Phylogenetic diversity of bacteria associated with Paleolithic paintings and surrounding rock walls in two Spanish caves (Llonín and La Garma). *FEMS Microbiol. Ecol.* **47**, 235–247 (2004).
30. Zimmermann, J., Gonzalez, J. M., Saiz-Jimenez, C. & Ludwig, W. Detection and phylogenetic relationships of highly diverse uncultured acidobacterial communities in Altamira Cave using 23S rRNA sequence analyses. *Geomicrobiol. J.* **22**, 379–388 (2005).
31. Bhullar, K., Waglechner, N., Pawlowski, A., Koteva, K., Banks, E. D., Johnston, M. D., Barton, H. A. & Wright, G. D. Antibiotic resistance is prevalent in an isolated cave microbiome. *PLoS One* **7**, 1–11 (2012).
32. Howarth, R., Unz, R. F., Seviour, E. M., Seviour, R. J., Blackall, L. L., Pickup, R. W., Jones, J. G., Yaguchi, J. & Head, I. M. Phylogenetic relationships of filamentous sulfur bacteria (*Thiothrix* spp. and Eikelboom type 021N bacteria) isolated from wastewater-treatment plants and description of *Thiothrix eikelboomii* sp. nov., *Thiothrix unzii* sp. nov., *Thiothrix fructosivorans* sp. *Int. J. Syst. Bacteriol.* **49**, 1817–1827 (1999).

33. Sket, B. The ecology of anchihaline caves. *Trends Ecol. Evol.* **11**, 221–225 (1996).
34. Bilandžija, H., Četković, H. & Jeffery, W. R. Evolution of albinism in cave planthoppers by a convergent defect in the first step of melanin biosynthesis. *Evol. Dev.* **14**, 196–203 (2012).
35. Lee, N. M.; Meisinger, D. B.; Aubrecht, R. ., Kovacik, L.; Saiz-Jimenez, C.; Baskar, S. . & Baskar, R.; Liebl, W.; Porter, M. L.; Engel, A. S. *Life at extremes*. (Cambridge: CAB International, 2012).
36. Gabriel, C. R. & Northup, D. E. in *Cave Microbiomes A Nov. Resour. Drug Discov.* (ed. Cheeptham, N.) 85–108 (Springer New York, 2013). doi:10.1007/978-1-4614-5206-5\_5
37. Saiz-Jimenez, C. in *Cave Microbiomes A Nov. Resour. Drug Discov.* (ed. Cheeptham, N.) 69–84 (Springer New York, 2013). doi:10.1007/978-1-4614-5206-5\_4
38. Montano, E. T. & Henderson, L. O. in (2013). doi:10.1007/978-1-4614-5206-5\_6
39. Engel, A. S. in *Geomicrobiol. Mol. Environ. Perspect.* (2010). doi:10.1007/978-90-481-9204-5\_10
40. Chen, Y., Wu, L., Boden, R., Hillebrand, A., Kumaresan, D., Moussard, H., Baciú, M., Lu, Y. & Murrell, J. C. Life without light: Microbial diversity and evidence of sulfur- and ammonium-based chemolithotrophy in Movile Cave. *ISME J.* (2009). doi:10.1038/ismej.2009.57
41. Northup, D. E., Melim, L. A., Spilde, M. N., Hathaway, J. J. M., Garcia, M. G., Moya, M., Stone, F. D., Boston, P. J., Dapkevicius, M. L. N. E. & Riquelme, C. Lava cave microbial communities within mats and secondary mineral deposits: Implications for life detection on other planets. *Astrobiology* (2011). doi:10.1089/ast.2010.0562
42. Culver, D.C., and Pipan, T. *The Biology of Caves and their Subterranean Habitats: Biology of Habitats*. (Oxford University Press, 2009).
43. Penn, K., Jenkins, C., Nett, M., Udvary, D. W., Gontang, E. A., McGlinchey, R. P., Foster, B., Lapidus, A., Podell, S., Allen, E. E., Moore, B. S. & Jensen, P. R. Genomic islands link secondary metabolism to functional adaptation in marine Actinobacteria. *ISME J.* **3**, 1193–1203 (2009).
44. Ziemert, N., Lechner, A., Wietz, M., Millañ-Aguiñaga, N., Chavarria, K. L. & Jensen, P. R. Diversity and evolution of secondary metabolism in the marine actinomycete genus *Salinispora*. *Proc. Natl. Acad. Sci. U. S. A.* **111**, (2014).

45. Derewacz, D. K., McNeese, C. R., Scalmani, G., Covington, C. L., Shanmugam, G., Marnett, L. J., Polavarapu, P. L. & Bachmann, B. O. Structure and stereochemical determination of hypogeamins from a cave-derived actinomycete. *J. Nat. Prod.* **77**, 1759–1763 (2014).
46. Axenov-Gribanov, D. V., Voytsekhovskaya, I. V., Tokovenko, B. T., Protasov, E. S., Gamaiunov, S. V., Rebets, Y. V., Luzhetskyy, A. N. & Timofeyev, M. A. Actinobacteria isolated from an underground lake and moonmilk speleothem from the biggest conglomeratic karstic cave in Siberia as sources of novel biologically active compounds. *PLoS One* **11**, 1–12 (2016).
47. Radhakrishnan, M., Raman, V. A., Bharathi, S., Balagurunathan, R. & Kumar, V. Anti MRSA and antitubercular activity of phenoxazinone containing molecule from Borra Caves *Streptomyces* sp. BCA1. **5**, 13040 (2014).
48. Herold, K., Gollmick, F. A., Groth, I., Roth, M., Menzel, K. D., Möllmann, U., Gräfe, U. & Hertweck, C. Cervimycin A-D: A polyketide glycoside complex from a cave bacterium can defeat vancomycin resistance. *Chem. - A Eur. J.* **11**, 5523–5530 (2005).
49. Bell, E. *Life at extremes: environments, organisms, and strategies for survival (Vol. 1)*. (Cabi, 2012).
50. Nonejuie, P., Burkart, M., Pogliano, K. & Pogliano, J. Bacterial cytological profiling rapidly identifies the cellular pathways targeted by antibacterial molecules. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 16169–16174 (2013).
51. Nonejuie, P., Trial, R. M., Newton, G. L., Lamsa, A., Ranmali Perera, V., Aguilar, J., Liu, W. T., Dorrestein, P. C., Pogliano, J. & Pogliano, K. Application of bacterial cytological profiling to crude natural product extracts reveals the antibacterial arsenal of *Bacillus subtilis*. *J. Antibiot. (Tokyo)*. **69**, 353–361 (2016).
52. Peters, C. E., Lamsa, A., Liu, R. B., Quach, D., Sugie, J., Brumage, L., Pogliano, J., Lopez-Garrido, J. & Pogliano, K. Rapid Inhibition Profiling Identifies a Keystone Target in the Nucleotide Biosynthesis Pathway. *ACS Chem. Biol.* **13**, 3251–3258 (2018).
53. Quach, D. T., Sakoulas, G., Nizet, V., Pogliano, J. & Pogliano, K. Bacterial Cytological Profiling (BCP) as a Rapid and Accurate Antimicrobial Susceptibility Testing Method for *Staphylococcus aureus*. *EBioMedicine* **4**, 95–103 (2016).
54. Lamsa, A., Lopez-Garrido, J., Quach, D., Riley, E. P., Pogliano, J. & Pogliano, K. Rapid Inhibition Profiling in *Bacillus subtilis* to Identify the Mechanism of Action of New Antimicrobials. *ACS Chem. Biol.* **11**, 2222–2231 (2016).
55. Htoo, H. H., Brumage, L., Chaikeratisak, V., Tsunemoto, H., Sugie, J.,

- Tribuddharat, C., Pogliano, J. & Nonejuie, P. Bacterial Cytological Profiling as a Tool To Study Mechanisms of Action of Antibiotics That Are Active against *Acinetobacter baumannii*. *Antimicrob. Agents Chemother.* **63**, 1–11 (2019).
56. Mohammad, H., Younis, W., Ezzat, H. G., Peters, C. E., Abdelkhalek, A., Cooper, B., Pogliano, K., Pogliano, J., Mayhoub, A. S. & Seleem, M. N. Bacteriological profiling of diphenylureas as a novel class of antibiotics against methicillin-resistant *Staphylococcus aureus*. *PLoS One* (2017). doi:10.1371/journal.pone.0182821
  57. Liu, W.-T., Lamsa, A., Wong, W. R., Boudreau, P. D., Kersten, R., Peng, Y., Moree, W. J., Duggan, B. M., Moore, B. S. & Gerwick, W. H. MS/MS-based networking and peptidogenomics guided genome mining revealed the stenothricin gene cluster in *Streptomyces roseosporus*. *J. Antibiot. (Tokyo)*. **67**, 99–104 (2014).
  58. Rajput, A., Poudel, S., Tsunemoto, H., Meehan, M., Szubin, R., Olson, C. A., Lamsa, A., Seif, Y., Dillon, N., Vrbanac, A., Sugie, J., Dahesh, S., Monk, J. M., Dorrestein, P. C., Knight, R., Nizet, V., Palsson, B. O., Feist, A. M. & Pogliano, J. Profiling the effect of nafcillin on HA-MRSA D712 using bacteriological and physiological media. *Sci. data* (2019). doi:10.1038/s41597-019-0331-z
  59. Poudel, S., Tsunemoto, H., Meehan, M., Szubin, R., Olson, C. A., Lamsa, A., Seif, Y., Dillon, N., Vrbanac, A., Sugie, J., Dahesh, S., Monk, J. M., Dorrestein, P. C., Pogliano, J., Knight, R., Nizet, V., Palsson, B. O. & Feist, A. M. Characterization of CA-MRSA TCH1516 exposed to nafcillin in bacteriological and physiological media. *Sci. data* (2019). doi:10.1038/s41597-019-0051-4
  60. Lamsa, A., Liu, W. T., Dorrestein, P. C. & Pogliano, K. The *Bacillus subtilis* cannibalism toxin SDP collapses the proton motive force and induces autolysis. *Mol. Microbiol.* **84**, 486–500 (2012).
  61. Lane, D. J. *16S/23S rRNA Sequencing*. In: *Stackebrandt, E. and Goodfellow, M., Eds., Nucleic Acid Techniques in Bacterial Systematics*. John Wiley Sons, Print. (1991).
  62. Edgar, R. C. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* (2004). doi:10.1093/nar/gkh340
  63. Stamatakis, A. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* (2014). doi:10.1093/bioinformatics/btu033
  64. Brettin, T., Davis, J. J., Disz, T., Edwards, R. A., Gerdes, S., Olsen, G. J., Olson, R., Overbeek, R., Parrello, B., Pusch, G. D., Shukla, M., Thomason, J. A., Stevens, R., Vonstein, V., Wattam, A. R. & Xia, F. RASTtk: A modular and extensible implementation of the RAST algorithm for building custom annotation pipelines



and annotating batches of genomes. *Sci. Rep.* **5**, (2015).

65. Blin, K., Shaw, S., Steinke, K., Villebro, R., Ziemert, N., Lee, S. Y., Medema, M. H. & Weber, T. antiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. *Nucleic Acids Res.* **47**, W81–W87 (2019).
66. Hill, G. T., Mitkowski, N. A., Aldrich-Wolfe, L., Emele, L. R., Jurkonie, D. D., Ficke, A., Maldonado-Ramirez, S., Lynch, S. T. & Nelson, E. B. Methods for assessing the composition and diversity of soil microbial communities. *Appl. Soil Ecol.* (2000). doi:10.1016/S0929-1393(00)00069-X
67. Muyzer, G., De Waal, E. C. & Uitterlinden, A. G. Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. *Appl. Environ. Microbiol.* (1993). doi:10.1128/aem.59.3.695-700.1993
68. Pace, N. R. A molecular view of microbial diversity and the biosphere. *Science* (80-. ). (1997). doi:10.1126/science.276.5313.734
69. Malik, S., Beer, M., Megharaj, M. & Naidu, R. The use of molecular techniques to characterize the microbial communities in contaminated soil and water. *Environ. Int.* (2008). doi:10.1016/j.envint.2007.09.001
70. Cheeptham, N., Dapkevicius, L., Whitfield, Northup, D., Saiz-Jimenez & Gabriel. *Cave Microbiomes: A Novel Resource for Drug Discovery.* (2013).
71. Giraffa, G. & Neviani, E. DNA-based, culture-independent strategies for evaluating microbial communities in food-associated ecosystems. *Int. J. Food Microbiol.* (2001). doi:10.1016/S0168-1605(01)00445-7
72. Hirsch, P. R., Mauchline, T. H. & Clark, I. M. Culture-independent molecular techniques for soil microbial ecology. *Soil Biol. Biochem.* (2010). doi:10.1016/j.soilbio.2010.02.019
73. Kirk, J. L., Beaudette, L. A., Hart, M., Moutoglis, P., Klironomos, J. N., Lee, H. & Trevors, J. T. Methods of studying soil microbial diversity. *J. Microbiol. Methods* (2004). doi:10.1016/j.mimet.2004.04.006
74. Leckie, S. E. Methods of microbial community profiling and their application to forest soils. *For. Ecol. Manage.* (2005). doi:10.1016/j.foreco.2005.08.007
75. Maukonen, J. & Saarela, M. Microbial communities in industrial environment. *Curr. Opin. Microbiol.* (2009). doi:10.1016/j.mib.2009.04.002

**Chapter 3: Novel antibiotics inhibit DNA replication by targeting thymidylate  
kinase in *Escherichia coli***

### 3.1 Abstract

The discovery of new antibiotics is a pressing need perpetuated by an increasing prevalence of multi-drug resistant (MDR) pathogens. Here we report the identification of a chemical series of pan assay interference compounds (PAINS), which we have classified according to their activity and mechanism of action in *E. coli*  $\Delta$ tolC. PAINS have been largely avoided in the field of antibiotic discovery due to their promiscuity, which makes them very difficult to characterize. Despite this, however, these understudied molecules exhibit considerable antibacterial activity and show promise as potential starting points for new efforts in antibiotic development. We initially used bacterial cytological profiling (BCP) to determine that two PAINS inhibit the DNA synthesis and cell wall biogenesis pathways. We conducted genetic studies that identified thymidylate kinase (an enzyme essential for DNA synthesis) as one of the primary intracellular targets of these two compounds. An activity screen of 29 structural analogs yielded nine active molecules, on which we performed BCP to identify their relative *in vivo* specificities for the DNA replication pathway. One molecule in particular appeared to potently and selectively inhibit DNA replication. Thymidylate kinase was confirmed as the molecular target of this compound and the other analogs via resistance studies, plasmid overexpression, and an *in vitro* biochemical assay. These inhibitors of bacterial thymidylate kinase may be useful scaffolds that could be explored further to develop novel antibacterial drugs active against multi-drug resistant pathogens.

### 3.2 Introduction

Due to an increasing emergence of MDR bacteria, there is a desperate need for new antibiotics to effectively treat persistent infections.<sup>1-3</sup> While the rate at which

pathogens are developing resistance to clinically administered antibiotics appears continues to increase, many companies once actively involved in antibiotic discovery have left the industry. As a result, the rate at which antibiotics are achieving clinical status is not fast enough to keep pace with evolving resistance. For example, in 2015, ceftazidime-avibactam-resistant *Klebsiella pneumoniae* was identified within just months of the antibiotic's approval for use in the clinic.<sup>4</sup> While there is no way to predict whether or not a particular antibiotic will retain long-term efficacy against MDR pathogens in the clinic, the research and development of potential drug candidates is critical for combatting one of the greatest threats to global health today.

Vast libraries of natural products and/or synthetic molecules are frequently screened in order to identify those with antibacterial activity.<sup>5</sup> These screens have often identified a number of active molecules;<sup>6</sup> however, molecules are rarely pursued if their mechanisms of action (MOA) are difficult to classify. Consequently, many of these molecules do not make it to the development stage of antibiotic research.<sup>6</sup> For example, pan assay interference compounds (PAINS) are often discarded from antibiotic studies because they are notoriously difficult to study due to their tendency to bind promiscuously to many different targets.<sup>7</sup> In fact, filters have been developed for removing PAINS from compound libraries prior to screening.<sup>8</sup> For these reasons, many PAINS remain unstudied.

Previously, we developed bacterial cytological profiling (BCP), a method for rapidly screening antibiotics in order to identify their broader MOA and the cellular pathway(s) that they inhibit<sup>9-12</sup>. This fluorescence microscopy technique allows us to assess the specificity of antibiotics *in vivo*, and thus, it allows us to effectively study

molecules with multiple cellular targets that would otherwise present an insurmountable challenge. When coupled with SAR analyses, BCP allows the identification of PAINS analogs that have a greater specificity for a single target pathway.

Here, we conducted an antibiotic screen and identified a chemical series of PAINS with antibacterial activity against *E. coli*  $\Delta tolC$ . Using BCP, we discovered that some of these molecules appear to inhibit both cell wall biogenesis and DNA replication *in vivo*, while other structural analogs exhibit greater specificity for the DNA synthesis pathway. *In vivo* and *in vitro* experiments revealed that these molecules specifically inhibit thymidylate kinase (TMPK), an essential enzyme that catalyzes the synthesis of thymidine 5'-diphosphate from thymidine 5'-monophosphate.<sup>13</sup> Conserved by many bacterial species, TMPK is a promising target for new antibiotics. Thus, our study, which identifies a chemical backbone with specific activity against TMPK, highlights a novel path toward studying a family of antibiotics that has long been considered too much of a “pain” to study.

### **3.3 Materials and methods**

#### **3.3.1 Synthetic compound library screen**

A synthetic small molecule library consisting of ~1,800 compounds was obtained from the ChemBridge EXPRESS-pick library stock collection and screened in 96-well plates at 100  $\mu\text{g}/\text{mL}$  in LB to identify compounds active against *E. coli* JP313  $\Delta tolC$  (i.e. “hits”).

#### **3.3.2 Minimum inhibitory concentration (MIC) determination**

The minimum inhibitory concentration (MIC) was determined using the broth microdilution method performed in triplicate. All compounds used in this study were

purchased from ChemBridge in 5 mg increments and resuspended in dimethylsulfoxide (DMSO). Concentrated stocks of each compound were prepared at 20 mM and stored at -80°C. Prior to the start of each experiment, new experimental stocks were prepared in DMSO to the desired concentration. A 1:100 dilution of an overnight culture of *E. coli* JP313  $\Delta tolC$  was prepared in LB liquid media and grown at 30°C to an OD<sub>600</sub> between 0.2 and 0.4. With the exception of a media control column, 1  $\mu$ L of cells diluted to an OD<sub>600</sub> of 0.05 was added to a 96-well plate containing 2-fold serial dilutions of eight different starting concentrations of each compound in 100  $\mu$ L LB. An initial cell density count was determined using a Tecan reader, then the plates were incubated at 30°C for 24 hours while shaking. After 24 hours the optical density was determined and bioactivity was calculated by subtracting the initial OD<sub>600</sub> reading at T<sup>0</sup> from the final OD<sub>600</sub> reading at T<sup>24</sup>. Bioactive compounds were defined as resulting in an OD<sub>600</sub> 0.5 and below, indicating inhibition.

### **3.3.3 Bacterial cytological profiling (BCP)**

High-resolution fluorescence microscopy and BCP was conducted with *E. coli* JP313  $\Delta tolC$  as previously described by Nonejuie et. al<sup>22</sup>. Briefly, overnight cultures were diluted 1:100 in LB and incubated while rolling at 30°C until cells reached an OD<sub>600</sub> between 0.15-0.17. 300  $\mu$ L of cells was added to compounds previously prepared to the desired test concentration then incubated while rolling at 30°C. Microscopy images were taken after 30 minutes, and 2 hours, and in the case of compounds 1 and 2, also at 1 and 4 hours, post-treatment. Samples were prepared for microscopy by adding 20  $\mu$ L of cells to 1  $\mu$ L of a fluorescent dye mixture containing FM4-64, DAPI, and SYTOX-green.

### **3.3.4 Cell viability**

Viability counts were performed in triplicate for *E. coli* JP313  $\Delta tolC$  treated with compounds 1 and 2 at each imaging time point in a 96-well plate. 10  $\mu\text{L}$  of cells was added to 90  $\mu\text{L}$  of 1X Tbase (10X Tbase = 20g  $(\text{NH}_4)_2\text{SO}_4$ , 140 g  $\text{K}_2\text{HPO}_4$ , 60 g  $\text{KH}_2\text{PO}_4$ , 10 g  $\text{Na}_3\text{Citrate}\cdot\text{H}_2\text{O}$  per L) and serially diluted 1:10 to  $10^{-8}$ . A multichannel pipettor was used to add 5  $\mu\text{L}$  of each dilution onto an LB solid agar plate. The plates were incubated at room temperature overnight, then individual colonies were counted. To calculate the colony forming units (CFU) per mL, colony counts were multiplied by the volume plated, the resulting value was divided by the DMSO control then  $\log_{10}$  transformed. This was performed in triplicate and the resulting values were averaged then plotted with standard error values.

### **3.3.5 Resistant mutant generation**

Mutants resistant to compounds 1 and 2 were obtained via serial passaging. A single colony of *E. coli* JP313  $\Delta tolC$  was added to 5 mL of LB and grown overnight at 30°C while rolling. A 1:500 dilution of the overnight culture was prepared in 6 mL LB and incubated at 30°C while rolling until the cells reached an  $\text{OD}_{600}$  of 0.2-0.23. In a small glass culture tube, 1  $\mu\text{L}$  of cells was added to 1 mL of LB and compound at a concentration of 1/2x MIC. The samples were incubated at 30°C for 1 day while rolling. If growth occurred at the starting concentration, a new culture was started by adding 1  $\mu\text{L}$  of cells to 1 mL of LB containing a slightly higher concentration of compound. Serial passaging was performed until the highest concentration of each compound was reached. When the mutant strains no longer grew overnight in a higher concentration of compound, the cultures were purified on plates containing the compound grown to saturation for the preparation of glycerol stocks for long term storage at -80°C. The

extent to which each mutant was resistant to their corresponding compound was determined by identifying the MIC and comparing it to the MIC of the parent *E. coli* JP313  $\Delta tolC$  strain.

### **3.3.6 Genomic DNA extraction and quantification**

Genomic DNA was extracted using a protocol adapted for use with Qiagen's DNeasy® Blood & Tissue Kit. Bacterial strains were grown from individual colonies in 5 mL of LB overnight at 30°C while rolling. 500  $\mu$ L of the overnight culture was pelleted (16,000 x g, 3 min), re-suspended in 180  $\mu$ L of a lysozyme-RNase mixture [100 mg lysozyme, 5  $\mu$ L RNase A (100 mg/mL) 5 mL lysis buffer (20mM Tris·Cl pH 8.0, 2mM sodium EDTA, 1.2% Triton®X-100, H<sub>2</sub>O)], and incubated at 37°C for 45 minutes. Next, 25  $\mu$ L of proteinase K (20 mg/mL) and 200  $\mu$ L of AL Buffer (Qiagen) was added to each sample followed by a 20 second vortex at maximum speed, and a 30 minute incubation at 56°C. Post-incubation, 200  $\mu$ L of lab grade ethanol (96-100%) was added to each sample followed by a 30 second vortex at maximum speed. The samples were then added to a DNeasy Mini spin column, centrifuged (16,000 x g, 1 min), and the supernatant was discarded. 500  $\mu$ L of AW2 Buffer (Qigen) was added to the column, followed by centrifugation (20,000 x g, 3 min). The DNeasy Mini spin column was placed into a sterile 2 mL microcentrifuge tube, incubated for 1 minute at room temperature, and the genomic DNA (gDNA) was eluted in 100  $\mu$ L of AE Buffer by centrifugation (20,000 x g, 1 min). The gDNA concentration was quantified (1  $\mu$ l sample volume) with a Thermo Scientific™ NanoDrop™ One Microvolume UV-Vis Spectrophotometer (840274100) and stored at -20°C.

### **3.3.7 Genome sequencing, assembly, and variant analyses**



The genome sequences of two *E. coli* JP313  $\Delta tolC$  mutants were generated using the Illumina MiSeq sequencing platform with a V2 500 paired-end read kit at the La Jolla Institute for Allergy and Immunology Research (LIAI), San Diego, CA. All sequence processing and analyses were performed in Geneious Prime. Two FastQ files containing forward and reverse sequence read lists were generated for each genome. The files were imported into Geneious and simultaneously paired (input of expected insert size of 500 bp), resulting in a single paired read list. The sequences were trimmed using the BBDuk v38.37 plugin. BBDuk identifies and trims adapters based on presets for Illumina adapters and paired read overhangs, trims the end of sequences based on quality (Q score), and discards short reads and their associated pair mate. Adapter sequences trimmed using presets for Illumina adapters were trimmed from the right end of the sequence with the Kmer length set at 27 with a maximum substitution allotment of 1. A minimum overlap value of 25 was set for adapter trimming based on paired read overhangs. Low quality sequences were trimmed on both ends with the minimum quality score set at 30, and reads shorter than 30 bp were discarded. After trimming, the paired-end reads were merged into a single read with BBMerge. The resulting merged and unmerged sequences were mapped to the previously sequenced and annotated genome of the *E. coli* JP313  $\Delta tolC$  parent strain using the Geneious assembler set to medium sensitivity, and to find structural variants, short insertions, and deletions of any size. Further variation analysis was performed with the resulting alignment using the Geneious SNP finder set to use the bacterial genetic code and the following parameters; a minimum variant frequency of 0.25, maximum variant P-value of 6, minimum strand-bias P-value of 5 when exceeding 65% bias, and to calculate an approximate P-value for identified

variants. Manual inspection of these resulting data led to the identification of mutations in present in each genome.

### **3.3.8 Primer design and polymerase chain reaction (PCR)**

Thymidylate kinase DNA templates (~1,099 bp) were amplified from *E. coli* JP313  $\Delta tolC$  strains with a Q5 high-fidelity DNA PCR (New England Biolabs) with primer set EM029-TMKF1 (5'-ATGGCAAACCTTTCTCGTG-3') and EM030-TMKR1 (5'-GGTGTAGTAATCGGGATG-3'). Each PCR mixture (50  $\mu$ L) contained 100 ng of template genomic DNA, 500 pmol of each primer, and 200  $\mu$ M dNTPs. PCR thermocycling conditions were as follows: 30 seconds of initial denaturation at 98°C, 30 cycles of denaturation at 98°C for 10 seconds, annealing for 15 seconds at 61°C, extension at 72°C for 2 minutes, and a final extension at 72°C for 5 minutes then held at 4°C. PCR products were purified with the oligonucleotide cleanup protocol as described in the Monarch PCR & DNA Cleanup Kit 5  $\mu$ g user manual (NEB #T1030). Clean PCR products were sequenced using Sanger methods by Eton Biosciences (<https://www.etonbio.com/>) and trimmed for quality prior to analysis.

### **3.3.9 *E. coli* genome manipulation by P1 virulent transduction**

P1 $vir$  bacteriophage was used to move ~100 kb portions of the *E. coli* genome from one genetic variant to another. The lysates that were used to infect recipient strains were prepared from phage grown on the following four *E. coli* strains: MG1655 CGSC 9032 ( $\Delta fhuE764::kan$ ), *E. coli* JP313  $\Delta tolC$ , *E. coli* JP313  $\Delta tolC$  R.mut 1 [TMK Q45P], *E. coli* JP313  $\Delta tolC$  R.mut. 2 [TMK A69T].

Lysates of a virulent P1 mutant phage (P1 $vir$ ) were prepared in liquid. First, the *E. coli* nonlysogenic donor strains containing the genetic marker to be transduced, were

prepared for infection by growing a single colony at 37°C while rolling in 7 mL LB to an OD<sub>600</sub> of 0.1. Glucose and CaCl<sub>2</sub> were added to the culture to final concentrations of 0.2% and 5 mM, respectively, followed by a 10 minute incubation while rolling at 37°C. Second, the donor strains were infected with 60 μL of a high titer (10<sup>9</sup>-10<sup>10</sup> pfu/mL) P1*vir* stock lysate and incubated at 37°C while rolling for ~3 hours, or until the culture became clear indicating bacterial lysis. Next, to ensure complete lysis of the bacterial cells, 60 μL of CHCl<sub>3</sub> was added to the cultures followed by a 10 minute incubation at 37°C while rolling, and a 2 hour incubation at 4°C. Finally, the lysates were prepared by separating the phage lysate from the bacterial cell debris via a 30 minute centrifugation (1,300 rpm) at 4°C. The supernatant containing the lysate was added to a sterile 5 ml screw-cap vial and serially diluted 1:100 in 3mL of LB to a final dilution of 10<sup>-3</sup>. After the addition of 60 μL of CHCl<sub>3</sub> to all dilutions (10<sup>0</sup>-10<sup>-3</sup>), the lysates were stored at 4°C.

Transductions were performed with the lysates prepared from donor strains and recipient strains of interest. First, an overnight culture of the recipient strain was prepared by inoculating 5 mL LB with a single colony followed by incubation at 37°C while rolling. 1 mL of the overnight culture was pelleted (16,000 x g, 3 min) and re-suspended in 1 mL of filter sterilized MC buffer (xx). Second, the recipient strains were infected by the previously prepared P1*vir* lysates. All transductions were performed by infecting 100 μL of the recipient strain in MC buffer with 100 μL of all four dilutions (10<sup>0</sup>-10<sup>-3</sup>) of phage lysate separately, resulting in 4 infections/recipient strain. Experimental controls included 100 μL of the 4 phage lysate dilutions and the recipient strain in MC buffer separately. The cultures were incubated for 30 minutes in a 37°C water bath followed by the addition of 1 M NaCitrate (200 μL to transductions and 100 μL to controls), and LB

(1 mL). Next, the cultures were incubated for 1 hour at 37°C while rolling, then the cells were pelleted (16,000 x g, 3 min), resuspended in 300  $\mu$ L of 0.1 M NaCitrate, and spread onto antibiotic plates. After an overnight incubation at 37°C, resultant transductants were purified via patching onto plates previously spread with 300  $\mu$ L of 0.1 M NaCitrate (LB and LB with the antibiotic used for selection).

### 3.3.10 TMK expression in *E. coli* JP313 $\Delta tolC$

Three full length *E. coli* K-12 MG1655 thymidylate kinase (WT, Q45P, A69T) nucleotide sequences were synthesized and cloned into the plasmid expression vector pRSFDuet-1 (GenScript). pRSFDuet-1 is under the control of the T7 lac promoter and induced by lactose, provides a high level of protein expression. Plasmids were transformed into *E. coli* JP313  $\Delta tolC$  resulting in the following recombinant *E. coli* strains: JP313  $\Delta tolC$ -pRSFDuet, JP313  $\Delta tolC$ -pRSFDuet-*tmk*<sup>+</sup>, JP313  $\Delta tolC$ -pRSFDuet-*tmk*-134, and JP313  $\Delta tolC$ -pRSFDuet-*tmk*-203.

Proteins were expressed in *E. coli* JP313  $\Delta tolC$  grown in LB media in the presence of kanamycin (30  $\mu$ g/mL). Cells were initially grown in 5 mL of LB at 30°C while rolling until an OD<sub>600</sub> 0.1-0.2 was reached. The cells were diluted to OD<sub>600</sub> 0.05 and 1  $\mu$ L was added to each well except column 12, the media control. The assay was performed with eight different concentrations of each compound, serially diluted 2-fold, in 96-well plates in 100  $\mu$ L of LB with kanamycin 30  $\mu$ g/mL. Cell densities were determined using a Tecan reader before and after a 24 hour incubation at 30°C while shaking. The effect of protein expression on the activity of each compound was calculated by subtracting the starting cell densities from the final, then calculating the resulting MIC. MICs were determined as the lowest concentration of compound resulting

in an OD<sub>600</sub> 0.5 and below, indicating inhibition. Elevated MICs against these strains were observed for compounds targeting TMK.

### 3.3.11 Molecular docking

The ligand-protein docking was performed using Autodock Vina (v1.1.2). The crystal structure of Escherichia coli thymidylate kinase (TmkK) was obtained in complex with P1-(5'-adenosyl)-P5-(5'-thymidyl)pentaphosphate and P1-(5'-adenosyl)P5-[5'-(3'-azido-3'-deoxythymidine)] pentaphosphate from Protein Data Bank (PDB code: 4TMK, <http://www.rcsb.org/>). Water molecules, cofactors, and the originally docked ligand were removed while polar hydrogens and partial charges were added all with Autodock Tools (v1.3.6) which is available as a part of MGLtools v1.3.6 made available by The Scripps Research Institute (<http://mgltools.scripps.edu/downloads>). Docking was carried out following the standard procedures for Autodock Vina and the resulting docking configuration was viewed with Chimera X (<https://www.rbvi.ucsf.edu/chimerax/>)<sup>23</sup>.

### 3.3.12 *E. coli* TMK IC<sub>50</sub> determination

*E. coli* TMK activity was measured using the ATPlite luminescence assay (PerkinElmer cat# 6016739) in 384-well (Corning 3825) format. The following solutions were prepared prior to experimentation; buffer (50 mM HEPES (pH 8.0), 25 mM NaAcetate, 10 mM MgCl<sub>2</sub>, 5 mM dithioereitol, 0.01% TX-100, and 0.5 mM EDTA in water), 3X TDP/ADP (195 μM TDP and 30 μM ADP in buffer), and buffer + 1% (v/v) DMSO (70 μL of DMSO in 7 mL of buffer). The 13 compounds (1, 2, 8, 9, 11-16, 27, 28, and 31) included in this assay were prepared as follows: 10 mM stocks in DMSO, and 100 μM working stocks in buffer. In preparation for the assay, a 384-well dilution plate was set up encompassing columns 1-13 and rows A-L. With the exception of row B, all

wells were filled with 30  $\mu\text{L}$  of buffer + 1% (v/v) DMSO. Each well in row B (1-13) was filled with 60  $\mu\text{L}$  of one of the 13 compounds (100  $\mu\text{M}$ ), and all of the remaining wells (columns 1-13, rows C-L) were filled with 30  $\mu\text{L}$  of the buffer + 1% (v/v) DMSO solution. A series of 2-fold dilutions was made for each compound by taking 30  $\mu\text{L}$  from row B and mixing it with row C, then taking 30  $\mu\text{L}$  from row C and mixing it with row D, this dilution method was continued to row L and the remaining extra 30  $\mu\text{L}$  was discarded. The concentrations of each compound in the dilution plate ranged from 100  $\mu\text{M}$  to 0.098  $\mu\text{M}$ . Enzyme inhibition studies were performed in two 384-well plates, plate one (+tmk) contained a 10.05  $\mu\text{L}$  reaction mixture consisting of 3.35  $\mu\text{L}$  transferred from each well of the dilution plate, 3.35  $\mu\text{L}$  of 3X TDP/ADP, 3.35  $\mu\text{L}$  of *E. coli* TMK (1200 pM), and in row M 10  $\mu\text{L}$  of buffer was added for background subtraction. Plate two (-tmk) contained a 10.05  $\mu\text{L}$  reaction mixture consisting of 3.35  $\mu\text{L}$  transferred from each well of the dilution plate, 3.35  $\mu\text{L}$  of 3X TDP/ADP, and 3.35  $\mu\text{L}$  of buffer. The final test concentrations of each compound ranged from 33.3  $\mu\text{M}$  to 0.033  $\mu\text{M}$ . The plates were incubated for one hour at room temperature, then the experiment was quenched with the addition of 5  $\mu\text{L}$  of ATPlite followed by a 5 minute incubation in the dark at room temperature. The results of the experiment were identified via luminescence detected by a PherastarFS plate reader using the "ATP 384" method with the following settings: 1 second measurement with the LUM Plus module, a 0.1 second settling time, gain set to 3203 and a focal height of 11.1 mm. The  $\text{IC}_{50}$  values were determined by subtracting the average value of row M in plate 1 (the background) from all wells in plates 1 and 2, then correcting for signal suppression in the data from plate 1 (+tmk) using the data from plate 2 (-tmk) at each concentration for each compound. Next, the percent inhibition was

calculated for each inhibitor concentration as the amount of enzyme activity inhibited in the test wells as compared to the complete reaction wells. A minimum was calculated as the average of the -tmk data with no compound (plate 2, row A), and a maximum was the +tmk data with no compound (plate 1, row A), the minimum and maximum percentages are from 0 and 100 respectively. Finally, the data analysis was done using Excel-fit software (ID Business Solutions Ltd., U.K.) with equation 205 by nonlinear regression (curve fitting). The percent of TMK inhibition for each test concentration (n=11) was plotted using the minimum and maximum percentages from 0 to 100 against the logarithm of inhibitor concentrations [ $\mu\text{M}$ ]. The  $\text{IC}_{50}$  values of each compound was calculated from the regression fit as the concentration at which 50% inhibition of the TMK enzyme activity was observed.

### **3.4 Results and discussion**

#### **3.4.1 BCP identifies the cellular pathway(s) inhibited by two PAINS**

A library of 1,798 synthetic small molecules was screened in order to identify compounds that inhibit the growth of *E. coli*  $\Delta\text{tolC}$ . This screen allowed us to identify a chemical series that shared a similar backbone structure, which included a rhodanine moiety and are known as pan-assay interference compounds, or PAINS<sup>9</sup>. Two compounds in this series are shown in Table 1. Their MICs against *E. coli*  $\Delta\text{tolC}$  were determined to be 11.65  $\mu\text{M}$  and 6.05  $\mu\text{M}$ , respectively. To rapidly probe the MOA of these compounds, we performed BCP. *E. coli*  $\Delta\text{tolC}$  cells were treated with each compound at concentrations equivalent to 1X, 3X, and 5X MIC for durations of 30 minutes, one hour, two hours, and four hours. Following treatment, the cells were stained with fluorescent dyes and imaged using high resolution fluorescence microscopy (Figure

1). Consistent with the knowledge that molecules belonging to the PAINS family of antibiotics potentially target multiple pathways, we found that compounds 1 and 2 inhibited both DNA replication and cell wall biogenesis. As evident in the resulting cytological profiles, compound 1 induced cell filamentation with chromosomal segregation defects at 3X and 5X MIC after two hours, and at these concentrations, some cells also appeared to be swollen, permeabilized, or even lysed (Figure 1A). Compound 2 induced similar DNA replication defects after one hour of treatment at all tested concentrations, and a severe degree of cell lysis was observed after four hours of treatment (Figure 1B).

#### **3.4.2 Viability of *E. coli* $\Delta tolC$ cells treated with compounds 1 and 2**

Our viability experiments further characterized the antibacterial effects of compounds 1 and 2 against *E. coli*  $\Delta tolC$ . Compound 1 appeared largely bacteriostatic, for even when cells were treated at 5X MIC, viability never decreased considerably, and cells resumed growth on an agar plate shortly after treatment ceased (Figure 2A). Compound 2, on the other hand, demonstrated nearly a 100-fold reduction in *E. coli*  $\Delta tolC$  viable cell counts when treated at 5X MIC (Figure 2B). These data are consistent with our BCP observations because while both compounds inhibited DNA replication, compound 1 appeared to inhibit cell wall biogenesis to a lesser degree, evidenced by a low frequency of cell lysis.

#### **3.4.3 Isolation of resistant mutations identifies Tmk as a potential target**

Though we suspected that compounds 1 and 2 inhibited the DNA replication pathway, we did not know which protein within the pathway was inhibited. Thus, in order to identify a potential enzymatic target, we isolated resistant mutants against both



compounds. Using serial passaging, one resistant mutant was successfully isolated for each compound. Mutant 1 (generated using compound 1) was six times more resistant to compound 1 than the *ΔtolC* parent strain (Table 2). This strain was also resistant (~ 3-fold) to compound 2. Mutant 2 (generated using compound 2), when compared to the *ΔtolC* parent, was approximately 10-fold resistant to both compounds. Whole genome sequencing revealed that both of these resistant mutants contained missense mutations in the *tmk* gene encoding thymidylate kinase, an enzyme essential to the DNA replication pathway (SI Figure 1). In mutant 1, proline replaced glutamine at residue 45, and in mutant 2, threonine replaced alanine at residue 69. Mutant 1 contained an additional mutation in the *dksA* gene, and mutant 2 contained an additional mutation in the *yjeA* gene. However, because *tmk* was the only gene that was mutated in both strains and because cross-resistance was conferred, we suspected that these mutated *tmk* alleles likely constituted the mutations conferring resistance to compound 1 and compound 2. Therefore, we moved the *tmkR2(A69T)* allele into a clean strain background via phage P1 transduction. The *tmkR2(A69T)* allele conferred 6-fold resistance to compound 1 and 2.4-fold resistance to compound 2 (Table 3). These results suggest that TMPK is one of the targets of these compounds *in vivo*.

#### **3.4.4 Docking of compound 1 to thymidylate kinase**

In order to better understand the molecular mechanism of our compounds, we used computational modeling to dock compound 1 to thymidylate kinase (Figure 3). This analysis predicted that the binding site of compound 1 overlaps considerably with the TMPK active site, forming hydrogen bonds with residues THR 17, LYS 16, ASP 99, and ARG 100. While binding to this location would presumably compete with substrate

binding and block its catalytic activity, neither of the resistant mutations were located at the enzymatic active site or the compound binding site. Therefore, it remains unclear how these mutations confer resistance to compounds 1 and 2.

#### **3.4.5 Screening analogs of compounds 1 and 2**

We screened 29 structural analogs of compounds 1 and 2 for their antibacterial activity against *E. coli AtolC*. The structures of these analogs, designated compounds 3 – 31, are described in SI Figure 2. Of these analogs, seven displayed antibacterial activity with an MIC less than 50  $\mu\text{M}$  (Table 4, SI Table 1). Compound 27 was the most potent, with an MIC as low as 1.87  $\mu\text{M}$ . As an additional screen, we performed BCP in order to determine if any of these analogs targeted the same major pathway (DNA replication) as compounds 1 and 2 (Figure 4). To do this, we treated *E. coli AtolC* with each of the seven active analogs at a concentration equivalent to 5X its MIC. Images were taken after 30 minutes and two hours of treatment. Many of the compounds appeared to inhibit DNA replication to some degree (compound 9, compound 11, compound 12, compound 27, and compound 28), with two compounds also causing lysis (compound 12 and compound 28). Compound 13 also induced cell lysis; however, it did not produce any phenotypes associated with the inhibition of DNA replication. All the analogs that exhibited an MIC lower than compound 1 induced DNA replication defects, while some of the compounds that exhibited higher MICs lost this phenotype. These data suggest that the inhibition of DNA replication might be the primary mechanism by which active analogs of compound 1 and compound 2 inhibit the growth of *E. coli*.

#### **3.4.6 Mutant allele Tmk R2 confers resistance**

To determine if TMPK might be the molecular target of these analogs, we tested their activity against the *tmkR2(A69T)* mutant strain. This mutant exhibited some degree of resistance to all of the analogs (Table 3). The MIC of compound 27, which appeared based on our BCP data to be the most specific inhibitor of DNA replication, was most drastically increased in the *E. coli* strain containing *tmkR2(A69T)*. The mutant strain was 10-fold resistant to compound 27, suggesting that this analog specifically inhibits thymidylate kinase *in vivo*. Conversely, those analogs that appeared not to inhibit DNA replication as explicitly as compound 27 experienced only a modest increase in MIC against the *tmkR2(A69T)* strain. For example, the MICs of compound 13 and compound 31 were only approximately 2-fold higher against this strain compared to the *E. coli*  $\Delta tolC$  strain containing the wild-type *tmk*.

#### **3.4.7 Tmk overexpression confers resistance**

To provide additional evidence that thymidylate kinase might be the target of this family of compounds, we tested whether or not TMPK overexpression conferred resistance. To accomplish this, we cloned the wild-type *tmk* gene and the *tmkR2(A69T)* and *tmkR1(Q45P)* resistant alleles on multi-copy number plasmids and measured the MICs of all nine of our active compounds against these engineered strains. Our experiments revealed that, typically a 4-fold increase in MIC was observed when wild-type *tmk* was overexpressed, and, in some cases, an even greater increase (up to 8-fold) was observed when *tmkR2(A69T)* or *tmkR1(Q45P)* was overexpressed (Table 5).

#### **3.4.8 Compounds inhibit thymidylate kinase *in Vitro***

Taken together, our genetic and cell biology studies suggested that thymidylate kinase is likely one of the *E. coli* enzymes inhibited by this family of molecules. We,

therefore, performed an assay to determine whether or not these compounds can inhibit purified *E. coli* thymidylate kinase *in vitro*. Using this assay, we obtained IC<sub>50</sub>s for all nine active compounds and, as well, for three hand-picked compounds that did not have MICs against *E. coli*  $\Delta tolC$  (compound 8, compound 14, and compound 16). The inactive compounds studied were selected because of their relative structural diversity. The IC<sub>50</sub> values for all 12 compounds ranged from a low of 250 nm (compound 27) to a high of 3.6  $\mu$ M (compound 16). The three compounds for which MICs could not be measured had some of the highest IC<sub>50</sub> values. Additionally, the data obtained for these compounds did not fit as well to a sigmoidal curve as the data obtained for other more active compounds, further suggesting the relative inactivity of these compounds (SI Figure 3). Despite this, our data did not display an apparent correlation between IC<sub>50</sub> values and MICs. For example, compound 2, which had one of the lowest MICs against *E. coli*  $\Delta tolC$ , had an IC<sub>50</sub> as high as 2.5  $\mu$ M. This result might be explained in part by the fact that the cell envelope is a significant barrier to antibiotic infiltration, and different structural features of our antibiotic analogs could affect their ability to enter the cell.<sup>14</sup> Taken together, these results are consistent with TMPK being one of the cellular targets of molecules within this family.

A collection of genetic, cell biological, and biochemical data suggests that many of the molecules within this family inhibit TMPK *in vivo*. However, some of these molecules likely have additional targets *in vivo*. For example, compound 2 appeared to inhibit both cell wall biogenesis and DNA replication *in vivo* and had a relatively high IC<sub>50</sub> (2.5 $\mu$ M) for the inhibition of TMPK *in vitro*. Compound 27, stands out among this group because it appeared to be relatively specific for TMPK, inducing a strong DNA

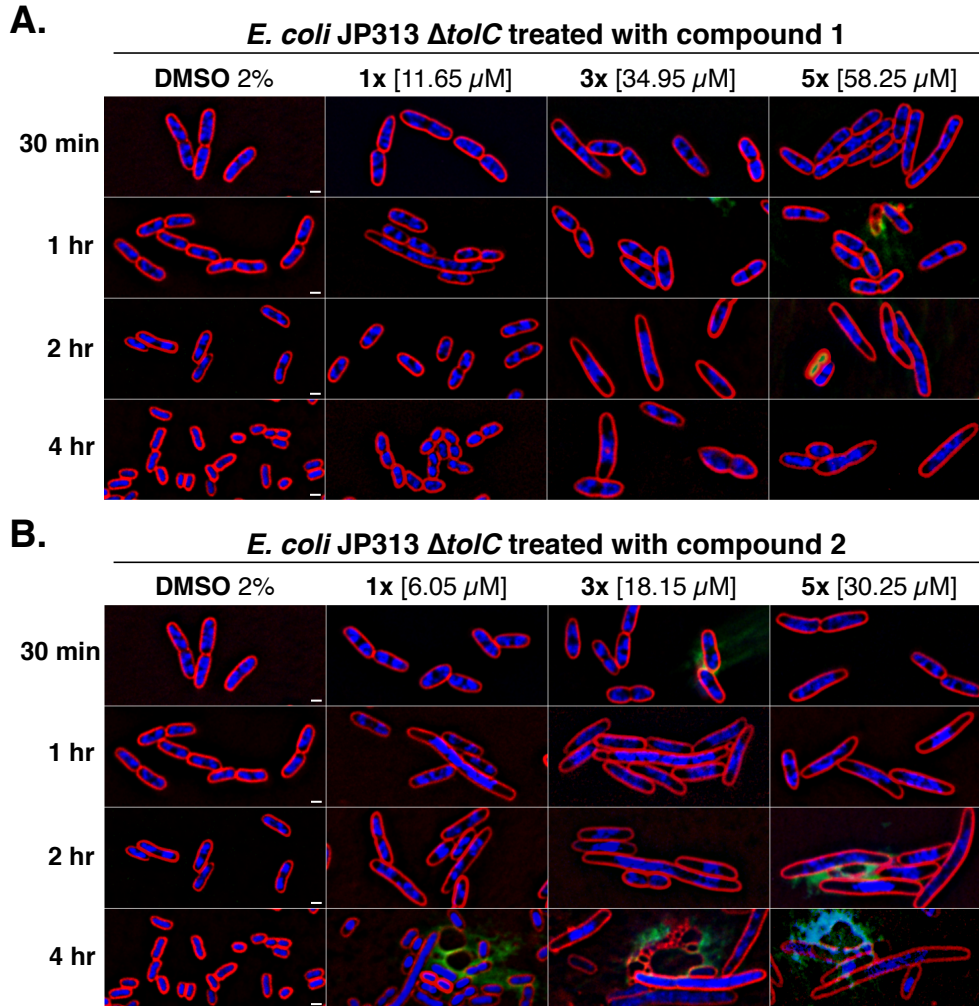
replication block *in vivo*, displaying the most potent IC<sub>50</sub> *in vitro*, and inhibiting growth of *E. coli*  $\Delta$ *tolC* with the lowest MIC. This work highlights the utility of combining an *in vivo* MOA assays with medicinal chemistry to study the antibacterial mechanisms of promiscuous molecules.

### **3.5 Acknowledgements**

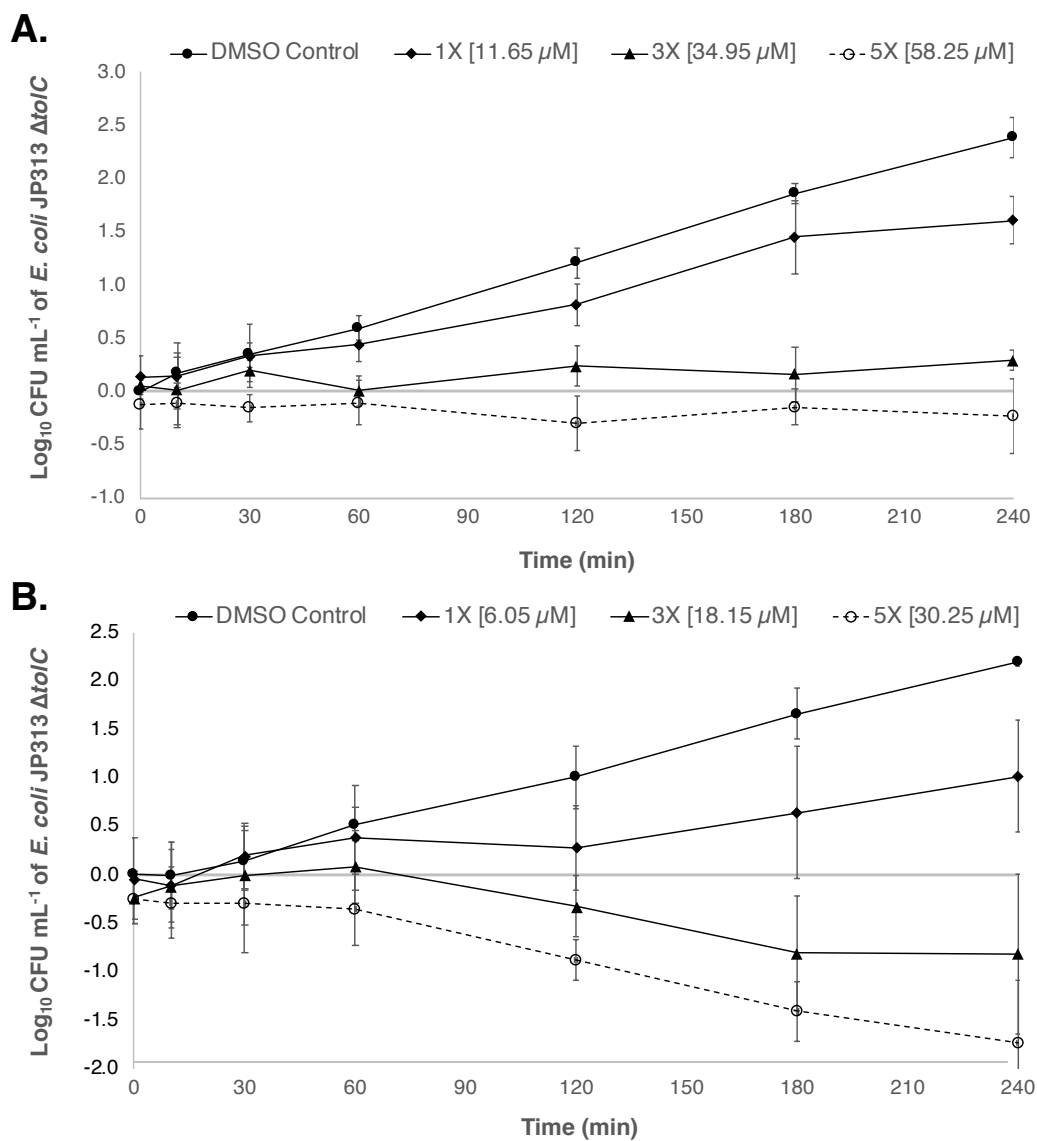
These studies were supported by the National Institute of Health (R01-AI113295). KP and JP have an equity interest in Linnaeus Bioscience Incorporated, and receive consulting income from the company. The terms of this arrangement have been reviewed and approved by the University of California, San Diego in accordance with its conflict of interest policies.

Chapter 3, in full, has been formulated into a manuscript that will be submitted for publication in 2020. Elizabeth T. Montañó, Jason F. Nideffer, Joseph Sugie, Eray Enustun, Adam B. Shapiro, Alan I. Derman, Kit Pogliano, Joe Pogliano. The dissertation author was the primary investigator and author of this material.

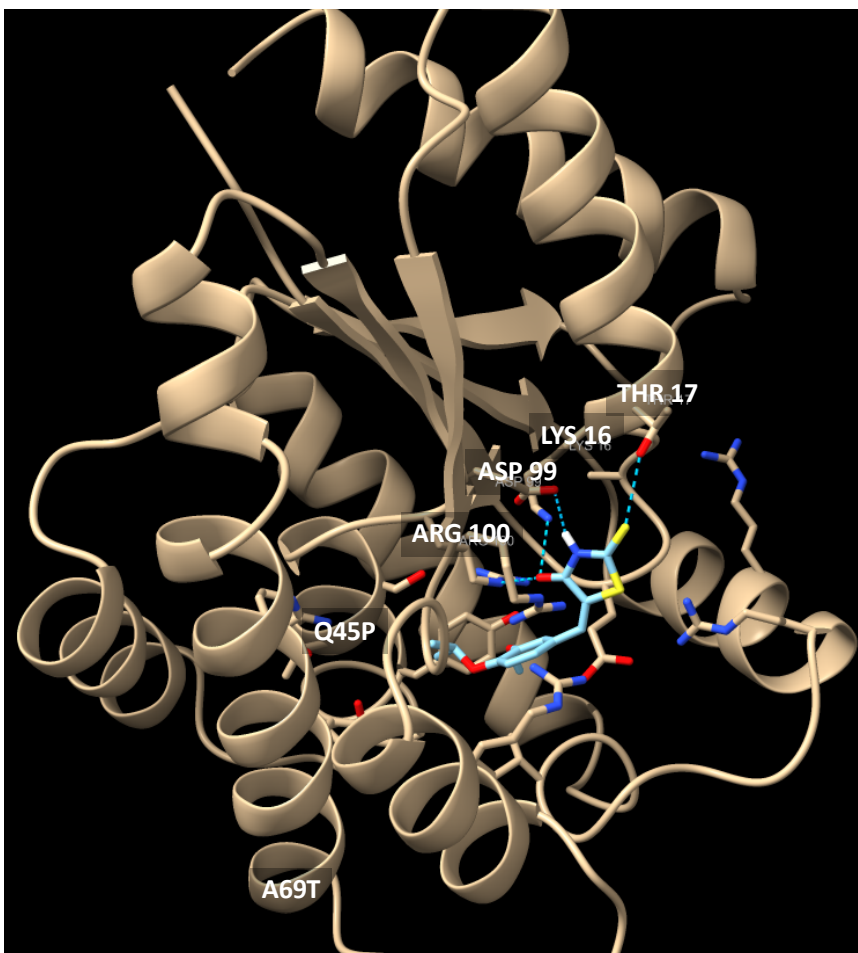
### 3.6 Figures



**Figure 3.1. Cytological profiles of *E. coli* JP313  $\Delta tolC$  treated with compounds 1 and 2. (A) Compound 1 at 1x, 3x, and 5x MIC for 10 minutes, 30 minutes, 1 hour, 2 hours, and 4 hours. (B) Compound 2 at 1x, 3x, and 5x MIC for 10 minutes, 30 minutes, 1 hour, 2 hours, and 4 hours. BCP images were collected after staining the cells with FM4-64 (red), DAPI (blue), and SYTOX-green (green). (Scale bar, 1  $\mu$ M).**

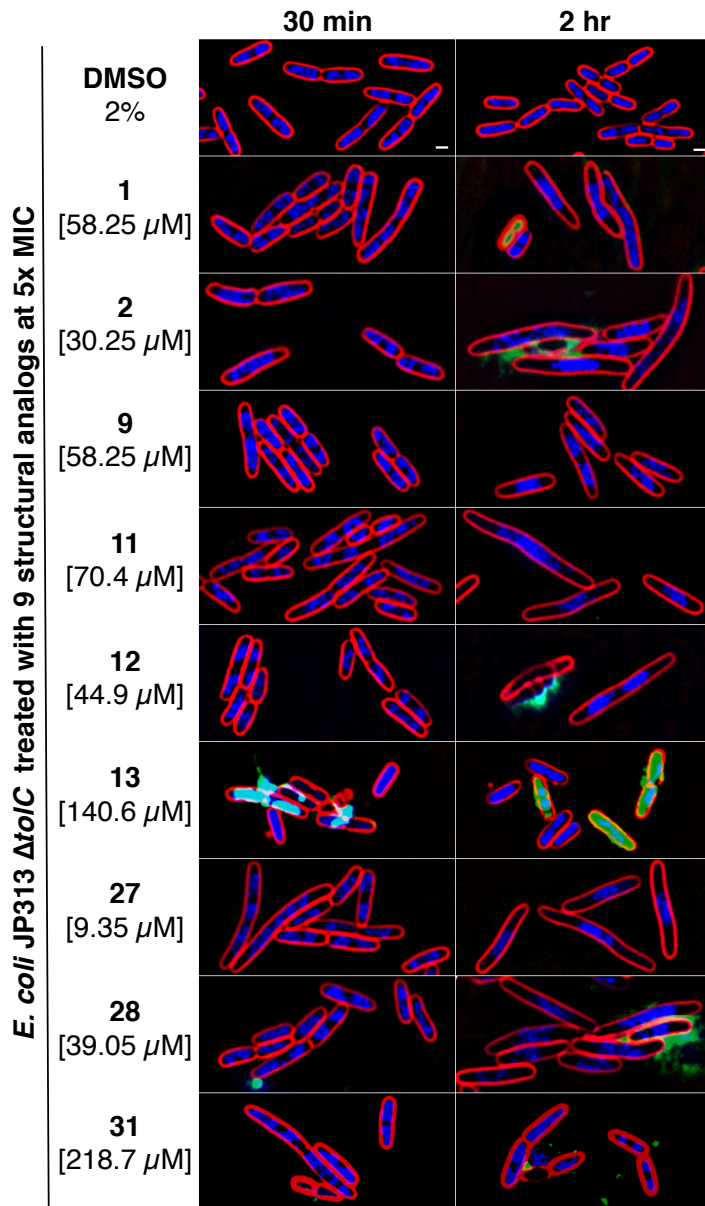


**Figure 3.2. Cell viability of *E. coli* JP313  $\Delta$ tolC.** Cells were treated with (A) Compound 1 and (B) Compound 2, at 1x, 3x, and 5x MIC for 4 hours and viability was measured via cells counts after 10 minutes, 30 minutes, 1 hour, 2 hours, and 4 hours.



**Figure 3.3.** *E. coli* thymidylate kinase with compound 1 docked to the active site. Amino acid residues with hydrogen bonds are displayed as well as the two mutations that arose in two different, non-clonal, *E. coli* JP313  $\Delta tolC$  strains conferring resistance to two different compounds that target *E. coli* TMK (compound 1 [Q45P] and compound 2 [A69T]). (PDB 4TMK)

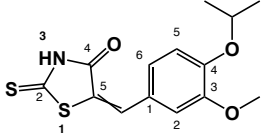
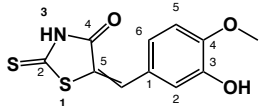




**Figure 3.4. Cytological profiles of *E. coli* JP313  $\Delta tolC$  treated with 9 5-Benzylidenerhodanine analogs.** *E. coli* JP313  $\Delta tolC$  DMSO control (2%), and BCP profiles of treatment with 9 structural analogs at 5X MIC, after 30 minutes and 2 hours. (Scale bar, 1  $\mu$ m).

### 3.7 Tables

**Table 3.1. Chemical structures and MICs of compounds 1 and 2 against *E. coli* JP313  $\Delta tolC$ .**

Compounds	Structure	<i>E. coli</i> JP313 $\Delta tolC$ MIC ( $\mu$ M)
1		10.8-12.5
2		5.86-6.25

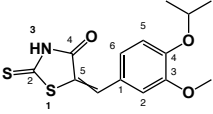
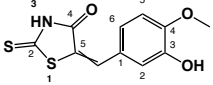
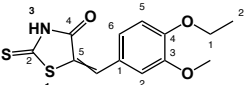
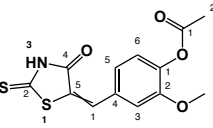
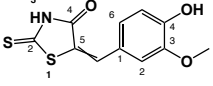
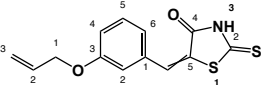
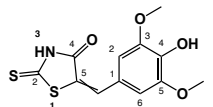
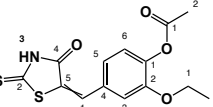
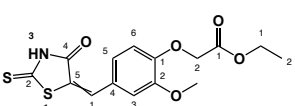
**Table 3.2. Minimum inhibitory concentration (MIC) of *E. coli* JP313  $\Delta tolC$  and two mutants treated with compounds 1 and 2.** MIC values were determined using the standard microdilution method performed in triplicate.

Strains	MIC ( $\mu$ M)	
	Compound 1	Compound 2
<i>E. coli</i> JP313 $\Delta tolC$	11.65	6.05
<i>E. coli</i> JP313 $\Delta tolC$ R. Mutant 1	75	17.18
<i>E. coli</i> JP313 $\Delta tolC$ R. Mutant 2	>100	81

**Table 3.3. Minimum inhibitory concentration of *E. coli* JP313  $\Delta tolC$  [2 A69T] and 8 5-Benzylidenerhodanine derivatives.**

<b>Compounds</b>	<b><i>E. coli</i> JP313 <math>\Delta tolC</math> [2 A69T] MIC (<math>\mu\text{M}</math>)</b>	<b>Fold Increase in MIC</b>
<b>1</b>	68.75	6
<b>2</b>	14.84	2.4
<b>9</b>	43.75	3.7
<b>11</b>	37.5	2.6
<b>12</b>	20.31	2.3
<b>13</b>	56.25	2
<b>27</b>	18.75	10
<b>28</b>	19.53	2.5
<b>31</b>	93.75	2.1

**Table 3.4. Chemical structures and minimum inhibitory concentrations of 9 structural analogs against *E. coli* JP313  $\Delta tolC$ .**

Compounds	Structure	<i>E. coli</i> JP313 $\Delta tolC$ MIC ( $\mu$ M)
1		10.8-12.5
2		5.86-6.25
9		10.8-12.5
11		14.06-14.1
12		8.59-9.37
13		28.125
27		1.87
28		7.81
31		43.75

**Table 3.5. Results of pRSFDuet-1 plasmid overexpression of *E. coli* MG1655 TMK and variants in *E. coli* JP313  $\Delta tolC$ .**

<i>E. coli</i> JP313 $\Delta tolC$ Strains	Compounds MIC ( $\mu$ M)								
	1	2	9	11	12	13	27	28	31
plasmid	15.5	7.75	15.5	15.5	7.75	62.6	3.88	7.75	31.25
plasmid + WT TMK	62	31	62	62	31	125	15.5	31	125
plasmid + R1 [Q45P] TMK	62	62	62	>62	>62	125	15.5	64	125
plasmid + R2 [A69T] TMK	62	62	62	>62	>62	125	15.5	64	125

**Table 3.6. IC<sub>50</sub> values calculated from an *in vitro* assay of *E. coli* TMK and 12 structural analogs.**

<b>Compounds</b>	<b><i>E. coli</i> TMK IC<sub>50</sub> (μM)</b>
<b>1</b>	0.39
<b>2</b>	2.5
<b>8</b>	1.2
<b>9</b>	0.41
<b>11</b>	0.48
<b>12</b>	0.5
<b>13</b>	1
<b>14</b>	4.2
<b>16</b>	5.6
<b>27</b>	0.25
<b>28</b>	0.47
<b>31</b>	0.68

**Table 3.7. Plasmids and strains used in this study.**

Plasmids and Strains	Relevant Genotype and Characteristics	Source
Plasmids		
pRSFDuet	T7 promoter, kan <sup>R</sup>	GenScript
pRSFDuet- <i>tmk</i> <sup>+</sup>	pRSFDuet with <i>E. coli</i> K-12 MG1655 <i>tmk</i> (WT) cloned	GenScript
pRSFDuet- <i>tmk</i> -134	pRSFDuet with <i>E. coli</i> K-12 MG1655 <i>tmk</i> (SNP at allele 134, an A to C transversion, resulting in a missense mutation of amino acid residue 45, Q to P), compound <b>1<sup>f</sup></b>	GenScript
pRSFDuet- <i>tmk</i> -205	pRSFDuet with <i>E. coli</i> K-12 MG1655 <i>tmk</i> (SNP at allele 205, an G to A transition, resulting in a missense mutation of amino acid residue 69, A to T), compound <b>2<sup>f</sup></b>	GenScript
Strains		
<i>E. coli</i> JP313 $\Delta tolC$	<i>E. coli</i> K-12 MC4100 with a $\Delta tolC5$ from strain CGSC5634	Previous study
JP313-pRSFDuet	<i>E. coli</i> JP313 $\Delta tolC$ with plasmid pRSFDuet	This study
JP313-pRSFDuet- <i>tmk</i> <sup>+</sup>	<i>E. coli</i> JP313 $\Delta tolC$ with plasmid -pRSFDuet- <i>tmk</i> <sup>+</sup>	This study
JP313-pRSFDuet- <i>tmk</i> -134	<i>E. coli</i> JP313 $\Delta tolC$ with plasmid -pRSFDuet- <i>tmk</i> -134	This study
JP313-pRSFDuet- <i>tmk</i> -205	<i>E. coli</i> JP313 $\Delta tolC$ with plasmid -pRSFDuet- <i>tmk</i> -205	This study
<i>E. coli</i> JW1088-5 (CGSC 9032)	<i>E. coli</i> MG1655 with genotype F-, $\Delta(araD-araB)567$ , $\Delta lacZ4787(::rrnB-3)$ , $\lambda'$ , $\Delta fhuE764::kan$ , <i>rph-1</i> , $\Delta(rhaD-rhaB)568$ , <i>hsdR514</i> .	The Coli Genetic Stock Center
JP313- <i>tmk</i> -134	<i>E. coli</i> JP313 $\Delta tolC$ , compound <b>1<sup>f</sup></b> (SNP at allele 134, an A to C transversion, resulting in a missense mutation of amino acid residue 45, Q to P)	This study
JP313- <i>tmk</i> -205	<i>E. coli</i> JP313 $\Delta tolC$ , compound <b>2<sup>f</sup></b> (SNP at allele 205, an G to A transition, resulting in a missense mutation of amino acid residue 69, A to T)	This study



### 3.8 References

1. Wencewicz, T. A. Crossroads of Antibiotic Resistance and Biosynthesis. *J. Mol. Biol.* **431**, 3370–3399 (2019).
2. Brown, E. D. & Wright, G. D. Antibacterial drug discovery in the resistance era. *Nature* **529**, 336–343 (2016).
3. Pogue, J. M., Kaye, K. S., Cohen, D. A. & Marchaim, D. Appropriate antimicrobial therapy in the era of multidrug-resistant human pathogens. *Clin. Microbiol. Infect.* **21**, 302–312 (2015).
4. Humphries, R. M., Yang, S., Hemarajata, P., Ward, K. W., Hindler, J. A., Miller, S. A. & Gregson, A. First report of ceftazidime-avibactam resistance in a KPC-3-expressing *Klebsiella pneumoniae* isolate. *Antimicrob. Agents Chemother.* (2015). doi:10.1128/AAC.01165-15
5. Newman, D. Screening and identification of novel biologically active natural compounds. *F1000Research* **6**, 1–13 (2017).
6. Hughes, J. P., Rees, S. S., Kalindjian, S. B. & Philpott, K. L. Principles of early drug discovery. *Br. J. Pharmacol.* **162**, 1239–1249 (2011).
7. Tomašić, T. & Peterlin Mašič, L. Rhodanine as a scaffold in drug discovery: A critical review of its biological activities and mechanisms of target modulation. *Expert Opin. Drug Discov.* **7**, 549–560 (2012).
8. Baell, J. B. & Holloway, G. A. New substructure filters for removal of pan assay interference compounds (PAINS) from screening libraries and for their exclusion in bioassays. *J. Med. Chem.* **53**, 2719–2740 (2010).
9. Nonejuie, P., Trial, R. M., Newton, G. L., Lamsa, A., Ranmali Perera, V., Aguilar, J., Liu, W. T., Dorrestein, P. C., Pogliano, J. & Pogliano, K. Application of bacterial cytological profiling to crude natural product extracts reveals the antibacterial arsenal of *Bacillus subtilis*. *J. Antibiot. (Tokyo)*. **69**, 353–361 (2016).
10. Lamsa, A., Lopez-Garrido, J., Quach, D., Riley, E. P., Pogliano, J. & Pogliano, K. Rapid Inhibition Profiling in *Bacillus subtilis* to Identify the Mechanism of Action of New Antimicrobials. *ACS Chem. Biol.* **11**, 2222–2231 (2016).
11. Htoo, H. H., Brumage, L., Chaikeratisak, V., Tsunemoto, H., Sugie, J., Tribuddharat, C., Pogliano, J. & Nonejuie, P. Bacterial Cytological Profiling as a Tool To Study Mechanisms of Action of Antibiotics That Are Active against *Acinetobacter baumannii*. *Antimicrob. Agents Chemother.* **63**, 1–11 (2019).
12. Nonejuie, P., Burkart, M., Pogliano, K. & Pogliano, J. Bacterial cytological

profiling rapidly identifies the cellular pathways targeted by antibacterial molecules. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 16169–16174 (2013).

13. Martínez-Botella, G., Breen, J. N., Duffy, J. E. S., Dumas, J., Geng, B., Gowers, I. K., Green, O. M., Guler, S., Hentemann, M. F., Hernandez-Juan, F. A., Joseph-McCarthy, D., Kawatkar, S., Larsen, N. A., Lazari, O., Loch, J. T., Macritchie, J. A., McKenzie, A. R., Newman, J. V., Olivier, N. B., Otterson, L. G., Owens, A. P., Read, J., Sheppard, D. W. & Keating, T. A. Discovery of selective and potent inhibitors of gram-positive bacterial thymidylate kinase (TMK). *J. Med. Chem.* (2012). doi:10.1021/jm3011806
14. Zgurskaya, H. I., López, C. A. & Gnanakaran, S. Permeability Barrier of Gram-Negative Cell Envelopes and Approaches to Bypass It. *ACS Infect. Dis.* **1**, 512–522 (2016).

**Chapter 4: Isolation and characterization of *Streptomyces* bacteriophages and the biosynthetic arsenals of their associated hosts**

## 4.1 Abstract

The threat to public health posed by drug-resistant bacteria is rapidly increasing, as some of healthcare's most potent antibiotics are becoming obsolete. Approximately two-thirds of the world's antibiotics are derived from natural products produced by *Streptomyces* encoded biosynthetic gene clusters (BGCs). Thus, in order to identify novel BGCs, we sequenced the genomes of four bioactive *Streptomyces* strains isolated from the soil in San Diego County and used bacterial cytological profiling (BCP) adapted for agar plate culturing in order to examine the mechanisms of bacterial inhibition exhibited by these strains. In the four strains, we identified known antibiotic-producing BGCs that partially explain the results of our mechanistic analyses. However, these known clusters could not fully account for the antibacterial activity exhibited by the strains, suggesting that novel clusters might encode antibiotics. Additionally, due to the utility of bacteriophage for genetically manipulating bacterial strains, we isolated four phages (BartholomewSD, IceWarrior, Shawty, and TrvxScott) against *S. platensis* and tested the host range of each phage against eight *Streptomyces* isolates. This study identified *Streptomyces* strains with the potential to produce novel chemical matter as well as four phages that may prove useful for genetically manipulating the strains encoding these clusters.

## 4.2 Introduction

Antibiotic discovery is an international priority requiring immediate action<sup>1</sup>. The increasing prevalence of multi-drug resistant (MDR) bacterial pathogens has resulted in an increased use of last-resort antibiotics<sup>1-3</sup>. Microbes that produce natural products are the most prolific source of clinically approved antibiotics<sup>4</sup>. In particular, soil dwelling

Actinobacteria, notably the *Streptomyces* sp., account for two-thirds of the antibiotics currently on the market<sup>5-7</sup>. Despite intensive studies, however, the full potential of microbes to produce natural products has not been fully realized<sup>8</sup>. Genome mining studies have shown that microbes encode many biosynthetic gene clusters (BGCs) that have not yet been characterized<sup>8</sup>. It is widely assumed that many of these clusters produce novel natural products and that further characterization of *Streptomyces* bacteria increases the probability of identifying molecules with unique chemical structures and new mechanisms of action<sup>9</sup>.

In addition to identifying *Streptomyces* strains containing potentially novel BGCs, it is necessary to improve on the conventional approaches used in natural product antibiotic discovery. One of the major stumbling blocks in natural product discovery is dereplication, since the isolation of bioactive molecules often yields antibiotics that have previously been discovered. We recently developed Bacterial Cytological Profiling (BCP) as a new whole-cell screening technique that can be used to rapidly identify the mechanism of action (MOA) of antibiotics<sup>10-15</sup>. BCP can accurately identify the pathway inhibited by antibacterial compounds present in unfractionated crude organic extracts and can be used to guide the purification of molecules with specific bioactivities<sup>10,14</sup>. BCP can also be used to screen bacterial strains directly on petri plates in order to identify and prioritize those strains that produce molecules with desired MOA<sup>14</sup>. In effect, BCP helps with the problem of dereplication by allowing for the determination of the MOA of antibiotics synthesized by a particular *Streptomyces* strain before labor-intensive antibiotic purification efforts are performed.

Since many BGCs are not expressed under laboratory conditions, genetic methods are often used to augment their expression and facilitate the identification and purification of their products<sup>16</sup>. Sometimes, increased expression can be achieved using techniques such as CRISPR/Cas, Red/ET recombineering, or plasmid cloning and overexpression<sup>16</sup>. However, there is still an occasional need to move large chromosomal regions from one strain to another via transduction in order to engineer strains optimally suited for antibiotic production. Transduction requires a phage capable of infecting the strain(s) of interest. Moreover, because phages generally display narrow host ranges<sup>17</sup> and relatively few *Streptomyces* phages have been isolated<sup>18</sup> compared to the large number of studied *Streptomyces* bacteria<sup>19</sup>, phages aptly suited for genetic manipulations are not available for the majority of antibiotic producing *Streptomyces* strains isolated.

Here we describe the isolation and characterization of *Streptomyces* strains and phages. We sequenced and analyzed the genomes of four phage isolates and also assessed their ability to infect different *Streptomyces* strains. Additionally, we used a combination of bioinformatics and BCP to characterize the antibiotic biosynthetic potential of four *Streptomyces* strains that displayed an ability to inhibit Gram-negative and Gram-positive growth. This work highlights a novel set of gene clusters and *Streptomyces* sp. phages that serve as a starting point for the isolation of potentially novel natural products.

## **4.3 Materials and methods**

### **4.3.1 Soil sample collection and site description**

Undergraduate students enrolled in the Phage Hunters Advancing Genomics and Evolutionary Science (PHAGE) class at UCSD collected soil samples for isolating

bacteria and their associated phages. Soil samples (approx. 30 mL) were collected around San Diego County (32.7157° N, 117.1611° W), California, USA.

#### **4.3.2 Isolation of *Streptomyces sp.***

Actinomycete isolation agar (AIA) plates (for one liter: sodium caseinate 2 g, L-Asparagine 0.1 g, sodium propionate 4 g, dipotassium phosphate 0.5 g, magnesium sulphate 0.1 g, glycerol 5 mL, rifampicin 50 µg mL<sup>-1</sup>, cycloheximide 100 µg mL<sup>-1</sup>, agar 15 g, pH 8.1) supplemented with 100 µg/ml of cycloheximide (CHX), 100 µg/ml of nystatin (NYS) and 10 µg/ml of rifampicin were used to select for Actinobacteria from soil samples. One gram of soil was added to the agar surface and streaked across the AIA-NYS-CHX plate and incubated for two days at 30°C. The plates were investigated for individual colonies with morphologies indicative of *Streptomyces sp.* (vegetative hyphae, aerial mycelium), those colonies were picked and purified at least four times on AIA-CHX plates.

#### **4.3.3 Phage isolation and purification**

Bacteriophages were isolated from soil samples with host, *Streptomyces platensis* JCM 4664 substr. MJ1A1<sup>41</sup>. An enrichment culture was prepared from 1 g soil and 2.5 ml of *S. platensis* added to 15 ml of Luria-Bertani (LB) medium (for one liter: tryptone 10 g, yeast extract 5 g, NaCl 10 g, agar 15 g, pH 7.0), followed by a 2-day incubation at 30°C with shaking. Phage were isolated from a 1.2 ml volume of enrichment culture that was centrifuged at maximum speed for 3 min, 1 ml of the resulting supernatant was filtered (0.22 µm filter), and 5 µl of the filtrate was spotted and then streaked onto an LB plate containing 100 µg/ml of cycloheximide. *S. platensis* (0.1 ml) was mixed with 4.5 ml of LB top agar 0.7%, poured over the streak plate and incubated for two days at 30°C.

Resultant plaques were re-streaked onto new LB plates containing 100 µg/ml of cycloheximide about 3-4 times for phage purification.

#### **4.3.4 Bacterial genomic DNA extraction and quantification for 16S rRNA PCR amplification and sequencing**

An adaptation of the DNeasy® Blood & Tissue Kit (Qiagen) protocol was used for bacterial genomic DNA extraction. Strains were cultured overnight at 30°C in 5 ml of LB broth while rolling. Cells were pelleted (16,000 x g, 3 min) from 1 ml of culture, re-suspended in 180 µl of lysis buffer (prepared in house), and incubated at 37°C for 45 min after which 25 µl of proteinase K (20 mg/mL) and 200 µl Buffer AL (Qiagen) was added. The samples were vortexed at maximum speed for 20 sec, incubated at 56°C for 30 min, and 200 µl of ethanol (96-100%) was added. The samples were vortexed at maximum speed for 30 sec, added to a DNeasy Mini spin column, centrifuged (16,000 x g, 1 min), and the supernatant was discarded. Buffer AW2 (Qiagen) was added (500 µl), followed by centrifugation (20,000 x g, 3 min). The DNeasy Mini spin column was placed into a sterile 2 ml microcentrifuge tube, and the gDNA was eluted in 100 µl of AE Buffer by centrifugation (20,000 x g, 1 min) following a 1 min incubation at room temperature. The gDNA concentration was quantified (1 µl sample volume) with a Thermo Scientific™ NanoDrop™ One Microvolume UV-Vis Spectrophotometer (840274100) and stored at -20°C.

#### **4.3.5 Bacterial genomic DNA extraction for PacBio whole-genome sequencing**

High molecular weight genomic DNA (20-160 kb) was extracted from 4 *Streptomyces* strains (DF, SFW, QF2, and JS) with the QIAGEN-Genomic-tip 500/G kit (10262) according to the manufacturer's protocol for bacteria.



#### **4.3.6 Bacterial whole-genome sequencing, assembly, and annotation**

The genome sequences of 4 *Streptomyces sp.* were generated using the Pacific Biosciences RS II (PacBio RS II) single molecule real-time (SMRT) sequencing platform at the IGM Genomics Center, University of California, San Diego, La Jolla, CA. Genome sequences were assembled using the HGAP protocol integrated in the PacBio RS II sequencer (smrt analysis v2.3.0/Patch5) resulting in a variable number (n = 1-95) of contigs per genome, and ranged in size from 5.42 to 7.79 Mb. The mauve contig mover was used to order the contigs of three draft genome sequences (genomes of strains SFW, QF2, and JS) relative to a closely related reference sequence (*S. pratensis* ATCC 33331, *S. globisporus* C-1027, and *S. parvulus* 2297 respectively). DNA sequencing of strain DF resulted in a single contig and did not require reordering to restore gene synteny. Gene prediction and annotation were made with the Rapid Annotations using Subsystems Technology (RASTtk)<sup>42</sup> platform.

#### **4.3.7 Genomic analysis of secondary metabolite biosynthetic potential**

Prediction of secondary metabolite biosynthetic gene clusters (smBGCs) was performed by antiSMASH v5.0<sup>23</sup>.

#### **4.3.8 Phage genomic DNA extraction**

5µl of RNAase A and 5 µl of DNase I were added to 10ml of lysate, incubated at 30°C for 30 minutes, and then precipitated overnight at 4°C by the addition of 4 ml of 20% polyethylene glycol 8,000. Samples were centrifuged at 10,000 g's for 30 minutes, and pellets resuspended in Qiagen PB buffer and DNA isolated using a Qiagen plasmid DNA isolation column as recommended by the manufacturer.

#### **4.3.9 Phage genome sequencing, assembly, and annotation**

Genomic DNA of 4 bacteriophages (TrvxScott, BartholomewSD, Shawty, and IceWarrior) was sequenced using the Illumina MiSeq platform at the Pittsburgh Bacteriophage Institute sequencing facility. The genomes were assembled with Newbler and checked for quality with Consed. The whole genome sequences were submitted to GenBank (Acc No. MH669016, MK460245, MK433266, and MK433259). DNA Master was used for annotation, and NCBI BLASTp was used to determine the potential function of gene products. Whole genome sequence comparisons were performed in Phamerator<sup>43</sup>.

#### **4.3.10 16S rRNA PCR amplification and sequencing**

16S ribosomal DNA templates (~1,465 bp) were amplified using Q5 high fidelity PCR (New England Biolabs) with the universal primer set 27F (5'-AGAGTTTGATCCTGGCTCAG-3') and 1492R (5'-GGTTACCTTGTTACGACTT-3')<sup>44</sup>. Each PCR mixture (50 µl) contained 100 ng of template gDNA, 500 pmol of each primer, and 200 µM dNTPs. PCR thermocycling conditions were as follows: 30 seconds of initial denaturation at 98°C, 30 cycles of denaturation at 98°C for 10 seconds, annealing for 15 seconds at 60°C, extension at 72°C for 1.5 minutes, and a final extension at 72°C for 5 minutes then held at 4°C. PCR products were purified with the oligonucleotide cleanup protocol as described in the Monarch PCR & DNA Cleanup Kit 5 µg user manual (NEB #T1030). Clean PCR products were sequenced using Sanger methods by Eton Biosciences (<https://www.etonbio.com/>) and trimmed for quality before analysis.

#### **4.3.11 CRISPR-Cas sequence analysis and predictions**

The sequences of all 4 *Streptomyces* were searched for CRISPR arrays (repeats and spacers) and potentially associated Cas genes using the following software tools; CRISPR-Cas++<sup>35,36</sup>, CRISPROne<sup>37</sup>, CRISPRMiner2<sup>38</sup>, and CRISPRDetect<sup>39</sup>.

#### **4.3.12 Phylogenetic analyses of bacterial isolates**

16S rRNA sequences were trimmed on both ends, (5' and 3') in Geneious Prime using the Trim Ends function with an error probability limit set at 0.05, which trims regions with more than a 5% chance of an error per base. Sequences were aligned using MUSCLE v3.8.425 with a maximum of 1,000 iterations, and profile-dependent parameters were set. Sequences were grouped based on similarity, and anchor optimized. Distance was measured using kmer4\_6 for the first iteration, and subsequent distance measurements were performed with pctid\_kimura. For all iterations the clustering method was neighbor joining, and the sequence weighting scheme was ClustalW with an spm objective score, 32 anchor spacing and -1 gap open score. Diagonals were set at a minimum length of 24 and a margin of 5, the minimum column anchor scores were set min best and min smoothed at 90. Finally, the hydrophobicity multiplier was set at 1.2 with a window size of 5. The tree was created with the Geneious Tree Builder using the Jukes-Cantor genetic distance model, Neighbor-Joining method to build the tree, with the outgroup set as *Chlorobium limicola* strain DSM 245. The resulting consensus tree was obtained after 1,000 iterations of bootstrap resampling, with a support threshold set at 70% and a random seed of 409,923.

#### **4.3.13 Cross-streak method for assessing antibacterial production potential**

From a single colony, using sterile Q-tips, *Streptomyces* isolates were streaked in a broad vertical line (2 inch) onto LB, and AIA, solid agar plates and incubated for one

week at 30°C. The day before the assay, test strains (*E. coli*  $\Delta tolC$ , *B. subtilis*, and *E. coli* MC4100) were grown in 5 ml of LB and incubated at 30°C overnight while rolling. On the day of the antibacterial screen, the overnight cultures of each test strain were diluted (1:100 in 5 ml LB) and grown to log phase OD600 0.15-0.2 (~1.5hr at 30°C while rolling). A volume of 10  $\mu$ l of each test strain was spotted in distinct lines almost to the edge of the *Streptomyces* line at a perpendicular angle. The plates were incubated overnight at 30°C, then investigated for the presence of zones of inhibition which were measured in millimeters.

#### **4.3.14 Bacterial cytological profiling (BCP) on plates**

Fluorescence microscopy and BCP on plates was performed as previously described by Nonejuie et. al.<sup>14</sup> Briefly, *Streptomyces sp.* strains (DF, SFW, QF2, and JS) were streaked in a vertical line down the center of LB, AIA, and ISP2 plates (for one liter: 4.0 g Difco yeast extract, 10.0 g Difco malt extract, 4.0 g dextrose, 20.0 g agar, pH 7.0), incubated for one week at 30°C. The test strain, *E. coli* JP313  $\Delta tolC$ , was prepared and spotted as described above in the cross-streak method. Following a 2 hr incubation at 30°C, a 1.5 x 1.5 cm square (~2.5 cm<sup>2</sup>) piece of agar containing the *E. coli* test strain was cut and prepared for high resolution fluorescence microscopy. The cut piece of agar was placed on a microscope slide, the *E. coli* cells were stained with fluorescent dyes, a coverslip was placed on top of the stained cells then imaged.

#### **4.3.15 Host range experiment**

The host range of 4 phages was determined against 10 different *Streptomyces* strains, including *S. platensis* and *S. coelicolor*. The experiment was blinded by assigning phages numbers i-iv and hosts letters A-J. A lawn of *Streptomyces* in LB top

agar was poured on LB CHX plates. After the top agar solidified, a grid was drawn on the bottom of the plate, and 5  $\mu$ l of pre-diluted phage ( $10^0$  to  $10^{-10}$  in phage buffer) was spotted in squares on the grid. Plaques were counted and used to calculate the titer which was compared to the detection limit to determine infectivity.

#### **4.3.16 Transmission electron microscopy**

10  $\mu$ l of lysate was applied to a copper grid, stained with 1% uranyl acetate, washed twice with phage buffer, and allowed to dry. Images were collected using a FEI Tecnai Spirit G2 BioTWIN Transmission Electron Microscope equipped with a bottom mount Eagle 4k camera.

#### **4.3.17 Strains used in this study**

*We used the following strains: S. platensis* JCM 4664 substr. MJ1A1<sup>41</sup>, *E. coli* MC4100<sup>45</sup>, *B. subtilis* PY79<sup>46</sup>, *P. aeruginosa* P4<sup>47</sup>, MRSA USA 300 TCH1516<sup>48</sup>, *S. coelicolor* A3(2) substr. M146<sup>49</sup>, *E. coli* JP313<sup>50-52</sup>  $\Delta tolC$  (MC4100 with a  $\Delta tolC5$  from strain CGSC5634), as well as two strains generously donated by Prof. Keith Poole at Queens University in Kingston, Canada – *P. aeruginosa* PA01 and *P. aeruginosa* K2733  $\Delta efflux$  ( $\Delta MexAB-OprM$ ,  $\Delta MexCD-OprJ$ ,  $\Delta MexEF-OprN$ ,  $\Delta MexXY-OprM$ ). The  $\Delta tolC5$  mutation is derived from strain PB3<sup>52</sup>, and was introduced into strain JP313<sup>50</sup> by P1 transduction. JP313<sup>50-52</sup> was transduced to tetracycline resistance with a lysate of strain CAG18475 (*metC162::Tn10*), and the methionine requirement of the transductants confirmed. This strain was then transduced to prototrophy with a lysate of PB3, and these transductants were screened on MacConkey agar for the presence of the  $\Delta tolC5$  mutation. Transductants that acquired  $\Delta tolC5$  were unable to grow on this medium owing to their sensitivity to bile salts such as deoxycholate<sup>51</sup>, and these arose at about 6%, in keeping

with previous linkage data<sup>52</sup>. PB3 and CAG18475 were obtained from the *Coli* Genetic Stock Center at Yale University.

## 4.4 Results and discussion

### 4.4.1 Isolation and antibacterial activities of *Streptomyces* sp.

To identify *Streptomyces* strains containing potentially novel BGCs, we collected 28 unique soil samples from sites across San Diego County. From these samples, we isolated a total of eight bacterial strains that were determined, based on 16S rRNA sequencing, to be members of the *Streptomyces* genus (Figure 1, Table 1). These strains and two known species (*Streptomyces coelicolor* A3(2) and *Streptomyces platensis* AB045882) were screened using the cross-streak method (as described in materials and methods) for their ability to inhibit the growth of wild type *E. coli* MC4100, an efflux defective mutant *E. coli* JP313  $\Delta tolC$ , and *B. subtilis* PY79. Since the production of bioactive secondary metabolites is highly dependent on growth conditions<sup>20</sup>, this screen was conducted on actinomycin isolation agar (AIA) as well as LB agar. Each of the 10 strains proved capable of inhibiting the growth of *E. coli* and/or *B. subtilis* on at least one of the tested media (Figure 2), suggesting that these strains likely produce antibiotics. As expected, however, the production of antibiotics often depended upon whether the strain was grown on AIA or LB agar. For example, strain ELW was incapable of inhibiting the growth of Gram-negative and Gram-positive bacteria when grown on AIA. However, when grown on LB agar, strain ELW inhibited the growth of both Gram-negative and Gram-positive bacteria. Conversely, strains JS and QF2 exhibited growth inhibition regardless of the media on which they were grown.

#### 4.4.2 Mechanistic analysis of natural products produced by four *Streptomyces* isolates

Strains QF2, JS, SFW and DF all inhibited the growth of *E. coli*  $\Delta tolC$  when grown on either AIA or LB agar, but in each case, the mechanism underlying inhibition was unknown. Thus, we utilized BCP to examine the mechanism of growth inhibition exhibited by the antibacterial natural products synthesized by these four *Streptomyces* isolates. Each of the four strains was grown on three different media (LB, AIA, or ISP2) for 5 days to allow for the synthesis and excretion of natural products into the surrounding agar. We then added exponentially growing *E. coli* cells adjacent to the *Streptomyces* lawn. After two hours of incubation at 30°C, the *E. coli* cells were stained with fluorescent dyes and imaged with high resolution fluorescence microscopy. *E. coli* cells incubated adjacent to each of the four *Streptomyces* isolates displayed characteristic cytological profiles that, in some cases, allowed for the classification of these strains according to the MOA of the natural products they produce (Figure 3).

When grown on either LB or ISP2, strain QF2 synthesized an antibiotic that caused the DNA of affected *E. coli* cells to assume a toroidal conformation (Figure 3). This phenotype is characteristic of bacteria treated with protein synthesis inhibitors such as chloramphenicol<sup>10,21</sup>, and thus, we concluded that strain QF2 is capable of synthesizing a translation-inhibiting natural product. QF2 also produced a membrane-active secondary metabolite, evidenced by Sytox Green permeability under all tested nutrient conditions (Figure 3). Strain JS appeared to induce similar phenotypes in *E. coli*, though under different growth conditions; protein synthesis phenotypes were observed on AIA and

ISP2 but not on LB. As was the case with strain QF2, Sytox Green permeability was observed in some cells regardless of the medium on which strain JS was grown.

Strain SFW induced distinct phenotypes in *E. coli* cells under each of the three nutrient conditions (Figure 3). On LB, a significant number of *E. coli* cells grown in the presence of strain SFW appeared to contain three chromosomes (white arrows), a phenotype that was not present in the untreated control cells. When strain SFW was grown on AIA, the *E. coli* cells became bent and lost their characteristic rod shape. Finally, strain SFW grown on ISP2 induced substantial swelling in *E. coli* cells that ultimately led to lysis. Notably, *E. coli* cells grown in the presence of strain DF exhibited nearly these same phenotypes under identical growth conditions, suggesting that these two strains produce compounds targeting similar pathways.

#### **4.4.3 Genomic analysis of four *Streptomyces* isolates**

The genomes of strains QF2, JS, SFW and DF were sequenced, and each was aligned to the most similar genome in the NCBI database (Figure 4A). Sequence reads for strain DF were assembled into a single contig that was most similar to the genome of *S. fulvissimus* DSM 40593. Sequencing of strains QF2, JS, and SFW yielded multiple contigs that were aligned to the genomes of *S. globisporus* C-1027, *S. parvulus* 12434, *S. pratensis* ATCC 33331, respectively.

In order to identify predicted gene clusters associated with secondary metabolism, the assembled genome sequences for strains QF2, JS, SFW and DF were annotated using RASTtk<sup>22</sup> and submitted to AntiSMASH 5.0<sup>23</sup> (Figure 4B). Each strain encoded between 18 and 34 BGCs, some of which were present in multiple strains (Figure 4C). Many of the encoded clusters closely resembled known BGCs in the MIBig repository (Tables 2-



5)<sup>24</sup>. For example, of the 23 putative BGCs identified in the genome of strain QF2 (Table 2), one of them (cluster 21) was similar to the viomycin BGC (Figure 5A). Viomycin inhibits protein synthesis by stabilizing tRNAs in the A site of the bacterial ribosome, inhibiting translocation<sup>25</sup>. According to AntiSMASH, 66% of the genes within the viomycin BGC were similar to genes within cluster 21. However, a global pairwise alignment of cluster 21 and the viomycin BGC revealed that the nucleotide sequence of cluster 21 is actually 98.5% identical over 32.5kb of the 36kb viomycin BGC (Figure 5A). This suggests that a viomycin related molecule is synthesized by strain QF2 and may account for strain's ability to inhibit protein synthesis in *E. coli* (Figure 3). While other clusters (Table 4: clusters 2, 4, 7, 8, 9, 12, 13, 22) have significant similarity to known BGCs, none of these clusters produce known antibiotics.

Strain JS contained 18 putative BGCs, six of which shared significant similarity (>60% of genes in common) with a known cluster (Table 3). Of these six, however, only cluster 7 was predicted to produce an antibiotic. All of the genes constituting a known terpene cluster that produces albaflavenone were present in cluster 7 (Figure 5B). Albaflavenone is capable of inhibiting the growth of *B. subtilis* by an unknown MOA<sup>26</sup>. Notably, albaflavenone has previously been isolated from *S. coelicolor* A3(2)<sup>27</sup>, a close relative of strain JS. Since the MOA of albaflavenone is unknown, it's not clear whether the products of cluster 7 or of a different cluster are responsible for the inhibition of protein synthesis and/or the membrane permeabilization observed in *E. coli* (Figure 3).

Of the 26 putative BGCs that were identified in the genome of strain SFW, only one cluster shared a high percentage of genes in common with a known antibiotic-producing cluster (Table 4). Cluster 1 shared similarity with 62% of the genes within a

known BGC that produces carbapenems (Figure 5C), a class of beta-lactam antibiotics that inhibit cell wall biogenesis<sup>28,29</sup>. Additionally, cluster 4 contained a low percentage of genes in common with a BGC involved in the synthesis of clavulanic acid, which inhibits beta-lactamase and consequently strengthens the bactericidal activity of beta-lactams<sup>30</sup>. Cluster 1 (and perhaps cluster 4) could, therefore, contribute to the synthesis of bioactive molecules that account for the inhibition of *E. coli* cell wall biogenesis on ISP2 media (Figure 3).

Strain DF encoded 34 BGCs (Table 5). Despite this rich supply of BGCs, however, we were only able to identify one cluster that participated in the synthesis of a known antibiotic. According to AntiSMASH v5.0, cluster 27 shared 85% identity with a known BGC that produces nonactin, a bioactive ionophore that disrupts membrane potential<sup>31</sup> (Figure 5D). The known clusters could not fully account for the antibacterial activity exhibited by strain DF, suggesting that antibiotics might be produced by novel clusters.

#### **4.4.4 Antimicrobial activity of four *Streptomyces* isolates against clinically relevant pathogens**

Strains JS, DF, SFW, and QF2 were screened for the ability to inhibit the growth of three clinically relevant pathogens using the cross-streak method (Table 6). Both strain QF2 and strain JS inhibited the growth of *E. coli*, *B. subtilis*, and MRSA, but not *P. aeruginosa* PA01 and *P. aeruginosa* P4, which were resistant to the antibiotics produced by all four *Streptomyces* isolates. Strain DF, though incapable of inhibiting the growth of wild-type *E. coli* MC4100, an efflux pump defective strain of *P. aeruginosa* K2733, *B. subtilis* PY79, and MRSA USA 300 TCH1516. Strain SFW was the least capable of

inhibiting the growth of pathogens, producing antibiotics only effective against *E. coli* MC4100 and *B. subtilis* PY79.

#### 4.4.5 Phage isolation and host range

Using *S. platensis* as a host, four bacteriophages (BartholomewSD, IceWarrior, Shawty, and TrvxScott) were isolated from the same soil samples from which we isolated the *Streptomyces* strains. These four phages were imaged using negative-stain transmission electron microscopy (Figure 6A) and were subsequently characterized as members of the family Siphoviridae due to their long filamentous tails and icosahedral capsids<sup>32,33</sup>. Genome sequencing revealed that BartholomewSD (52.1 kb) and TrvxScott (52.6 kb) are 89% identical (Figure 6B) and belong to the BD2 subcluster of *Streptomyces* phage, which currently contains 20 other members<sup>18</sup>. IceWarrior (55.5 kb) clustered in subcluster BII (19 members), and Shawty (40.7 kb) clustered in BB1 (6 members) (Table 7). The BLASTp-predicted functions of the gene products encoded by these phages are shown in Table 8.

To determine which of these four phages might be the most useful for genetically manipulating our newly isolated antibiotic-producing *Streptomyces* strains, we assessed the host ranges of the phages (Figure 6C). Five of our bacterial isolates (strains JS, DF, SFW, EDE, and ELW) and two laboratory strains (*S. platensis* and *S. coelicolor*) were capable of being infected by at least one of the four phages. However, the remaining three soil isolates (strains QF2, AH, and SK) were not successfully infected by any of the four phages. Despite a high degree of synteny between BartholomewSD and TrvxScott, these two phages exhibited distinct host ranges. While TrvxScott was capable of infecting strains ELW, DF, and JS in addition to *S. platensis*, BartholomewSD was only capable of

infecting *S. platensis*. This discrepancy may be accounted for by a DNA methylase encoded in the genome of TrvxScott but not BartholomewSD. Notably, methylation can protect the phage genome from digestion by host-encoded restriction enzymes<sup>34</sup>. Additionally, differences in the sequences of a minor tail protein (77% similarity) encoded by these two phages could contribute to the observed disparity in host ranges. Similar to BartholomewSD, Shawty exhibited a narrow host range, only infecting strain ELW and *S. platensis*. IceWarrior, however, displayed a relatively broad host range, as it was capable of infecting five of the soil isolates in addition to *S. platensis* and *S. coelicolor*.

#### **4.4.6 Characterization of CRISPR elements in the genomes of our bacteria**

Our bioinformatic analysis identified the presence of Cas enzymes and CRISPR arrays within the genomes of *Streptomyces* strains that could, in part, explain the results of our host range experiments. For example, CRISPRs were predicted in all four bacterial isolates, but the abundance of CRISPRs in each strain varied greatly (Table 9). QF2, which was uninfected by all four phages, contained the largest number of predicted CRISPRs – 39 in total, scattered around the chromosome, each containing between one and 25 spacers (Figure 4B, purple; Table 10). Predicted spacers within these arrays matched with 94-100% identity to sequences within TrvxScott (7 spacers), BartholomewSD (4 spacers), Shawty (2 spacers), and IceWarrior (5 spacers) (Table 10). Spacers targeted a variety of genes including those encoding minor tail proteins, tape measure proteins, deoxycytidylate deaminase, helix-turn-helix DNA binding proteins, endolysin, and capsid maturation protease (Figure 7). The large number of putative spacers in the QF2 genome targeting TrvxScott, BartholomewSD, Shawty, and

IceWarrior suggests that strain QF2 has previously encountered and acquired resistance to each of these phages. Moreover, strain QF2 was isolated from the same soil sample as BartholomewSD, providing support for these findings. Strain QF2 also encoded seven proteins of a Type IE CRISPR-Cas system<sup>35-39</sup>. The QF2 proteins were distantly related to the enzymes of the canonical Cas3 system in *E. coli* (Figure 8), but the operon in strain QF2 lacked two genes (Cas1 and Cas2) involved in spacer acquisition. This phenomenon, the absence of Cas1 and Cas2, was previously reported as common among Streptomycetaceae Type IE systems<sup>40</sup>. The presence, in the QF2 genome, of a Cas3 system and spacers mapping to essential proteins in each of the genomes of our phages could explain why this strain is resistant to infection by all four phages.

Additionally, the presence of specific spacers within the genomes of strains DF and SFW may explain their full or partial resistance to infection by some of our phage isolates. Strain DF, with one spacer mapping to a sequence within the genome of TrvxScott, was infected by TrvxScott at a low efficiency of plating that was approximately 4 orders of magnitude less than the efficiency of plating on *S. platensis* (Figure 6C). This strain also contained two spacers against BartholomewSD, a phage to which it was resistant (Figure 6C; Table 11, 12). Strain SFW contained two spacers that shared sequence similarity with genomic regions within Shawty and IceWarrior (Table 11, 13). Notably, strain SFW was resistant to infection by Shawty, while IceWarrior infected the strain at an efficiency of plating as low as  $10^{-7}$  (Figure 6C). Both strain DF and SFW encoded proteins containing regions with similarity to the RuvC and HNH endonuclease domains of known Cas enzymes. However, given the limited similarity of

these putative proteins to know Cas proteins, further study is necessary to determine if they constitute functional Cas systems that provide resistance against phage infection.

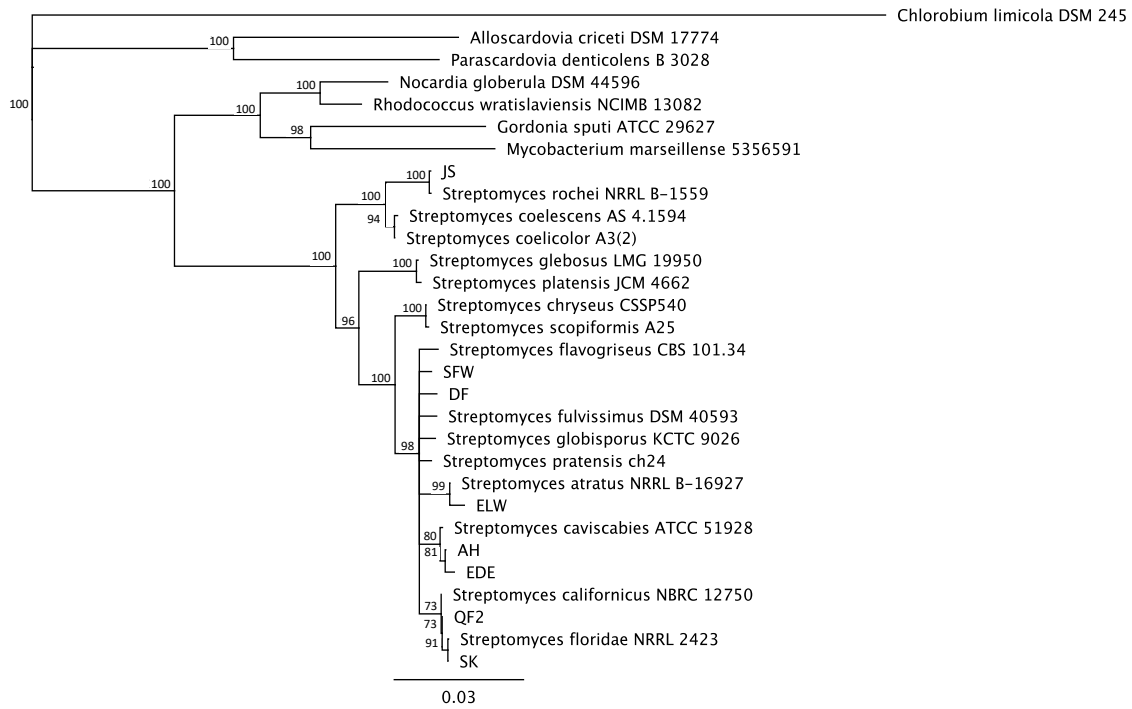
A curious feature emerged from our analysis of the CRISPRs within these strains. Among the 213 predicted spacers encoded by all four bacterial strains, 18 contained sequence similarity (95-100% identity) with at least one of the four phages (Table 9). The lengths of the matching sequences (100% identity) within bacterial spacers ranged from 14 to 18 nucleotides and accounted for approximately half the length of a typical spacer. Additionally, a single spacer occasionally appeared capable of targeting two distantly related phages. These spacers contained sequences mapping to two distinct genes encoded by different viral genomes. For example, spacer 106 in CRISPR 23 of strain QF2 is 32 nucleotides in length, and it contains 14 bases that share 100% identity with a region in the TrvxScott tape measure gene. These 14 bases overlap (by 8 nucleotides) with another sequence that is 14 base pairs in length and shares 100% identity with a noncoding region within the Shawty genome (Table 11). If these spacers functionally serve to resist infection, our analysis suggests that a single spacer may evolve to efficiently target more than one phage, thus providing broad immunity.

#### **4.5 Acknowledgements**

These studies were supported by the National Institute of Health (R01-AI113295). KP and JP have an equity interest in Linnaeus Bioscience Incorporated, and receive consulting income from the company. The terms of this arrangement have been reviewed and approved by the University of California, San Diego in accordance with its conflict of interest policies.

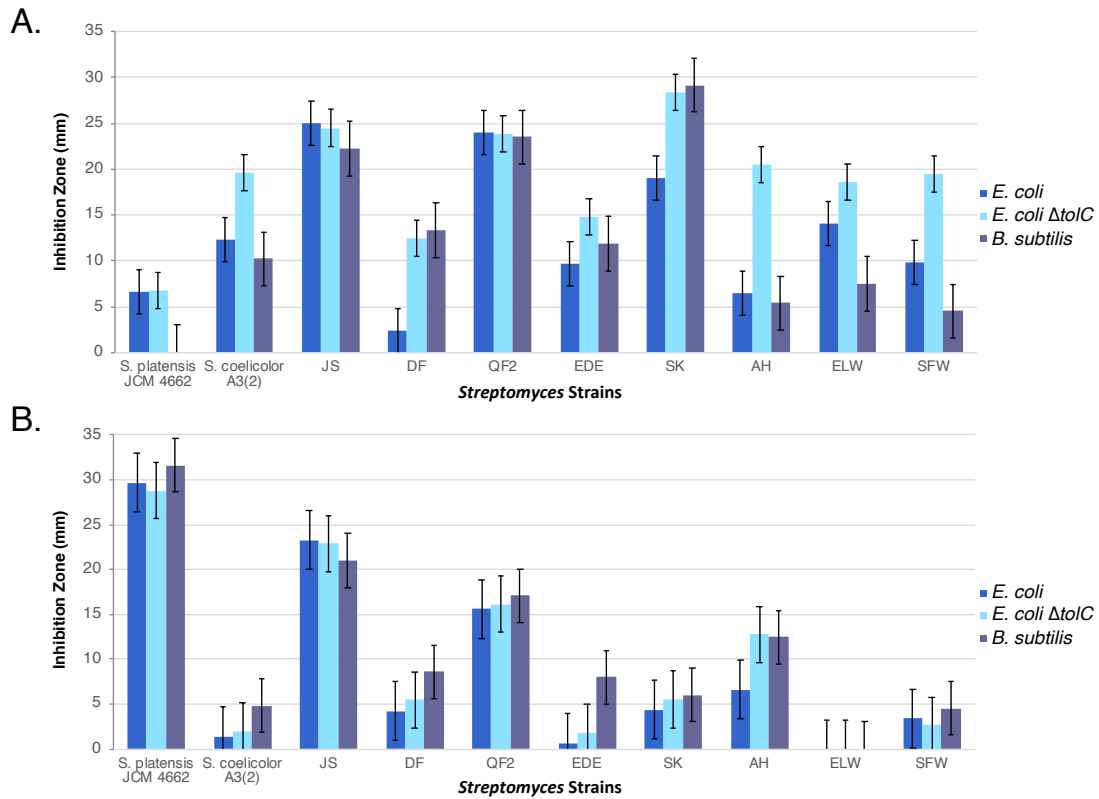
Chapter 4, in full, has been submitted for publication of the material as it may appear in Nature Science Reports, 2020. Elizabeth T. Montaña, Jason F. Nideffer, Lauren Brumage, Marcella Erb, Alan I. Derman, John Paul Davis, Elena Estrada, Sharon Fu, Aishwarya Vuppala, Cassidy Tran, Elaine Luterstein, Shivani Lakkaraju, Sriya Panchagnula, Caroline Ren, Jennifer Doan, Sharon Tran, Jamielyn Soriano, Yuya Fujita, Pranathi Gutala, Quinn Fujii, Minda Lee, Anthony Bui, Carleen Villarreal, Samuel R. Shing, Sean Kim, Danielle Freeman, Vipula Racha, Alicia Ho, Prianka Kumar, Kian Falah, Eray Enustun, Amy Prichard, Ana Gomez, Kanika Khanna, Shelly Trigg, Kit Pogliano, Joe Pogliano. The dissertation author was the primary investigator and author on these studies.

## 4.6 Figures

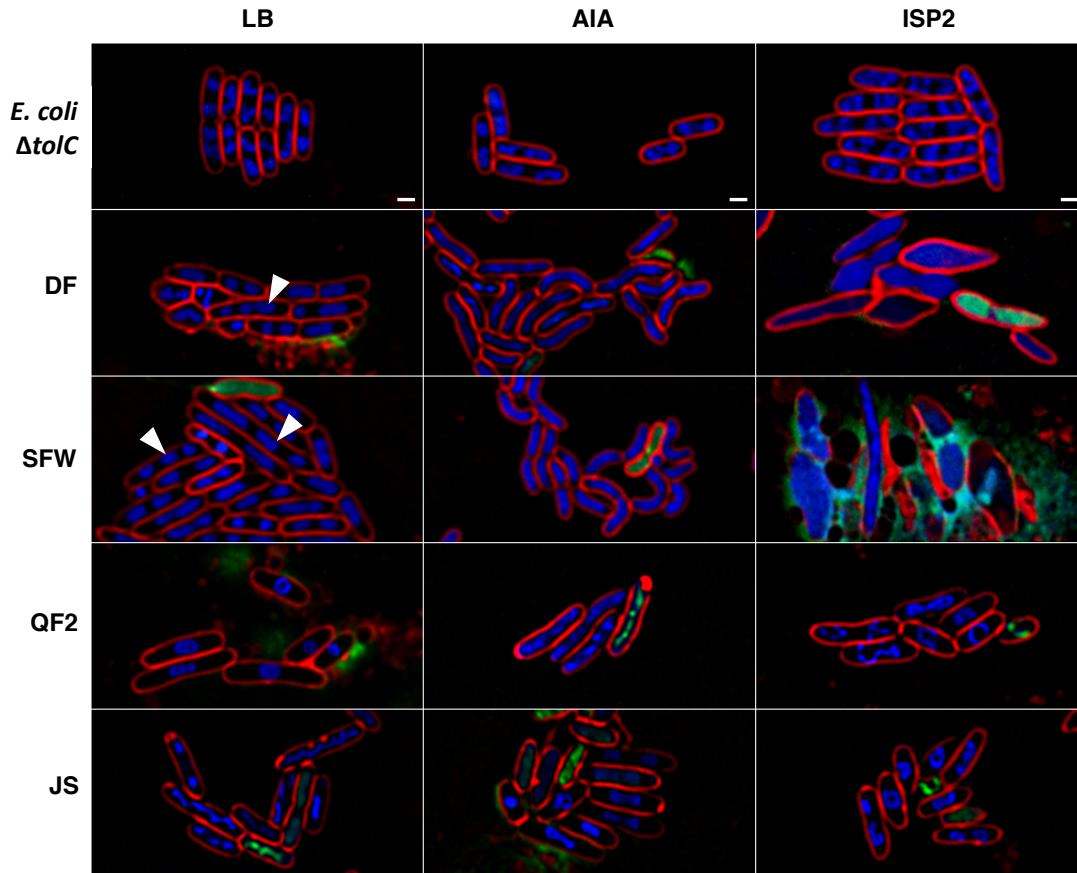


**Figure 4.1. Phylogenetic tree of *Streptomyces* bacteria isolated from soil samples.** This tree was constructed using PCR-amplified 16S rRNA sequences and was generated from a multiple sequence alignment using a Neighbor-Joining algorithm. The scale bar (0.03) indicates substitutions per nucleotide.

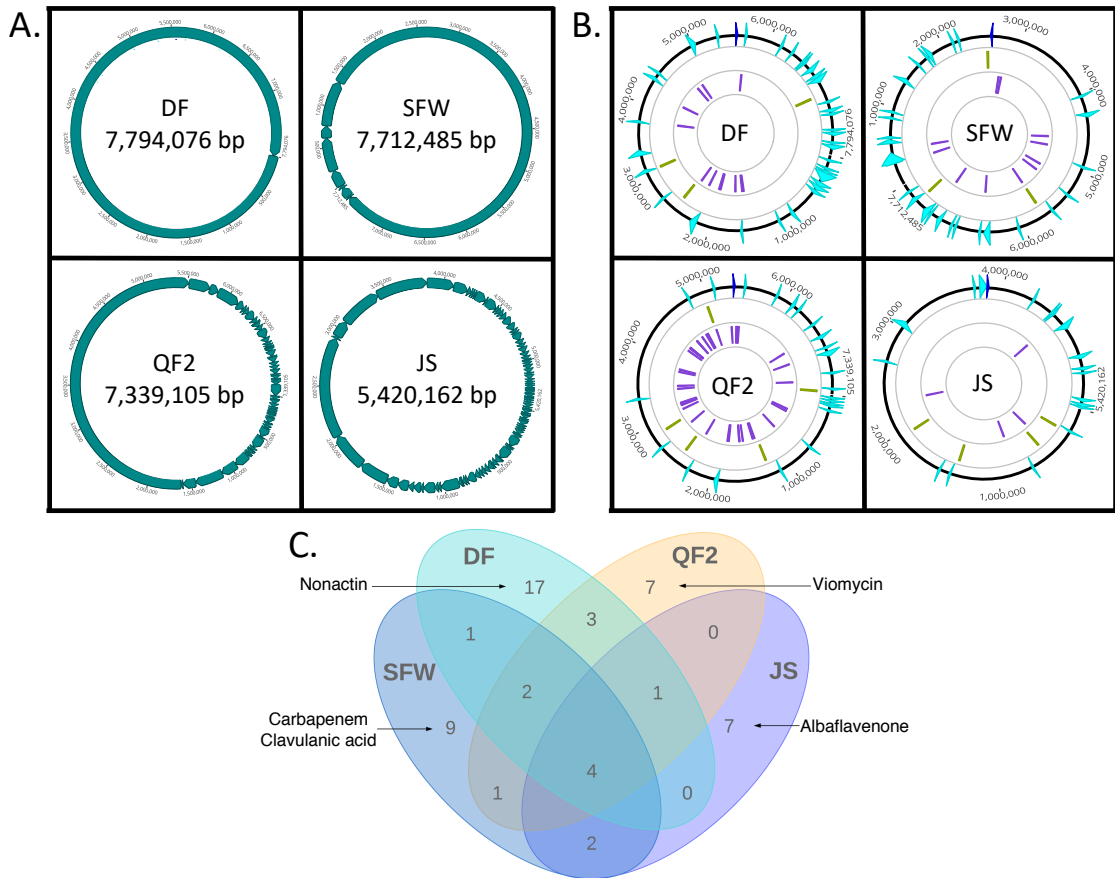




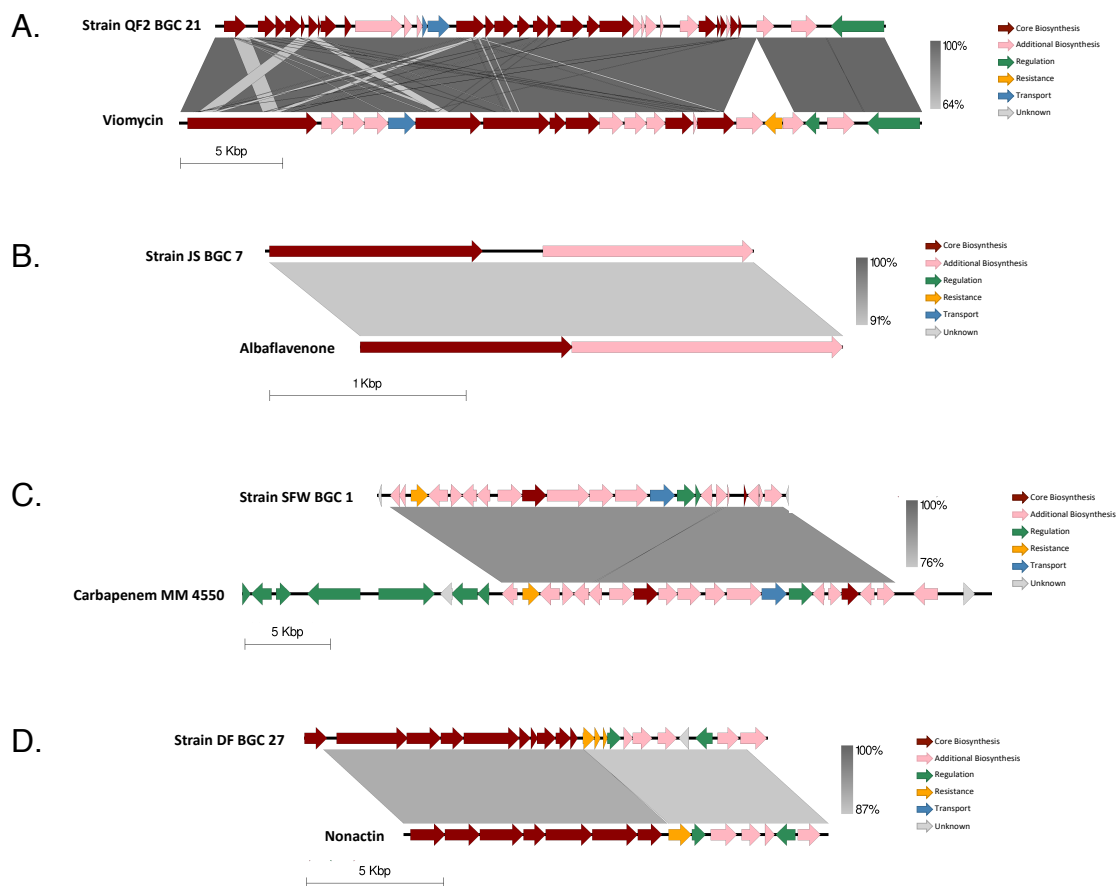
**Figure 4.2. Inhibition of bacterial growth by *Streptomyces* isolates.** The cross-streak method was used to measure the zone of inhibition among ten *Streptomyces* strains against two Gram-negative *E. coli* strains (MC4100 WT, JP313  $\Delta tolC$ ), and Gram-positive *B. subtilis* PY79 on (a) LB and (b) actinomycete isolation agar (AIA).



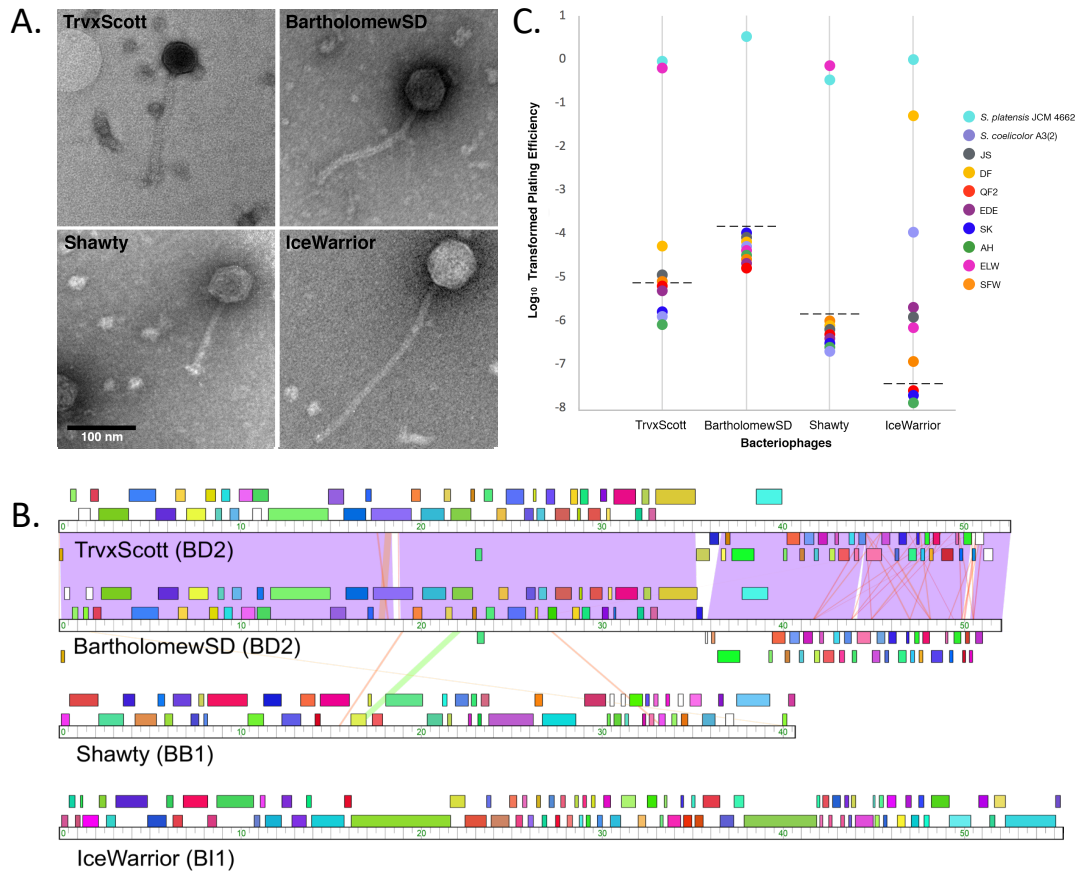
**Figure 4.3. Differential BCP phenotypes of *E. coli* JP313  $\Delta tolC$  exposed to natural products produced by four *Streptomyces* soil isolates grown on different solid media.** Also displayed, are *E. coli* JP313  $\Delta tolC$  untreated controls grown on the tested media (LB agar, AIA, and ISP2 agar). White arrows indicate cells with three chromosomes. BCP images were collected after staining the cells with FM4-64 (red), DAPI (blue), and SYTOX-green (green). The scale bar is 1 micron.



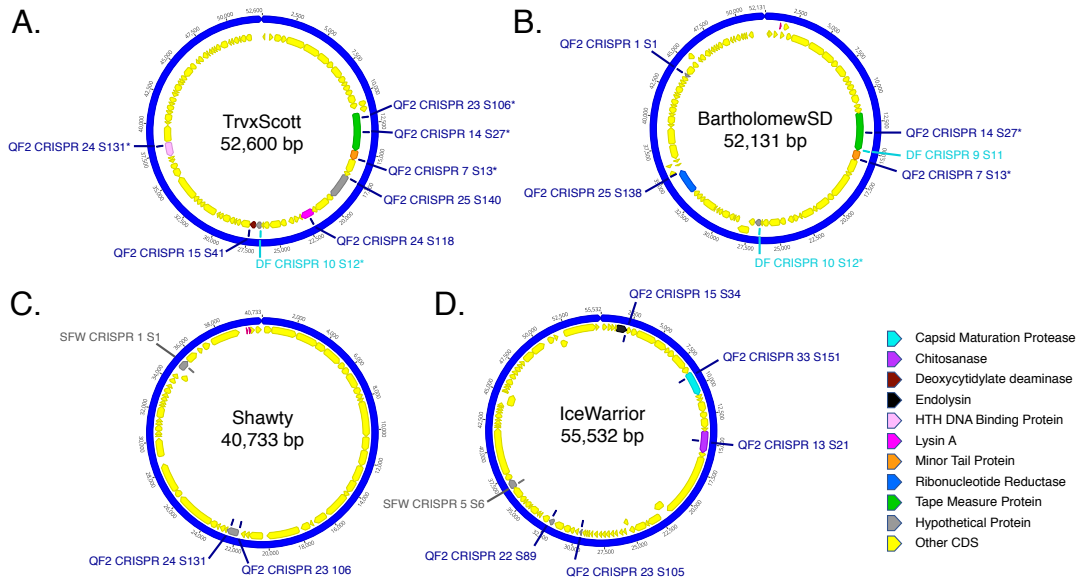
**Figure 4.4. Genome characteristics of *Streptomyces* sp. DF, SFW, QF2, and JS isolated from soil samples.** (a) Circular genomes of the four bacterial isolates represented by the assembled contigs obtained from genome sequencing. (b) Genomes oriented according to their threonine operons (dark blue). Predicted biosynthetic gene clusters (light blue), loci of Cas-associated protein-coding genes (green), and CRISPR arrays (purple) are shown. (c) A Venn diagram displaying the numbers of BGCs that are shared by and unique to the genomes of our isolates. Five clusters of particular importance are explicitly named.



**Figure 4.5. Comparison of BGCs encoded in the genomes of bacterial soil isolates and the predicted most similar, previously characterized BGC with an antibacterial product.** (a) Strain QF2, BGC 21, compared to the BGC previously described to produce the antibiotic viomycin (NCBI Acc No. AY263398.1), encoded in the WGS of *S. vinaceus* ATCC 11861. (b) Strain JS, BGC 7, compared to the BGC previously described to produce the antibacterial sesquiterpene Albaflavenone (NCBI Acc No. AL645882.2), encoded in the WGS of *S. coelicolor* A3(2). (c) Strain SFW, BGC 1, compared to the BGC previously described to produce the antibacterial beta-lactam Carbapenem MM 4550 (NCBI Acc No. KF042303.1), encoded in the WGS of *S. argenteolus* ATCC 11009. (d) Strain DF, BGC 27, compared to the BGC previously described to produce the ammonium ionophore antibiotic Nonactin (NCBI Acc No. AF074603.2), encoded in the WGS of *S. griseus subsp. griseus* ETH A7796. Cluster comparisons were constructed in Easyfig. Regions of nucleotide homology are indicated on a gray scale and genes are colored according to the putative function of the corresponding protein product.

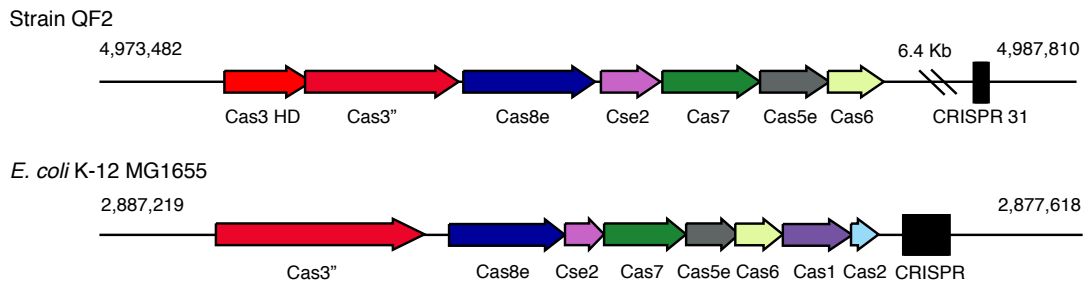


**Figure 4.6. Characterization of four Streptomyces phage isolated from soil samples.** (a) Electron micrographs of the four phages (IceWarrior, TrvxScott, BartholomewSD, and Shawty). Lysate samples were negatively stained and imaged with transmission electron microscopy (TEM). The scale bar is 100 nm. (b) A whole-genome sequence comparison of the four phages generated by Phamerator (top to bottom: TrvxScott, BartholomewSD, Shawty, IceWarrior). (c) Host ranges of our four phage isolates. The phages are listed on the horizontal axis, while the vertical axis indicates plating efficiency (log-transformed). Each circle represents one of ten Streptomyces bacteria that was tested for susceptibility to infection. Circles above the detection limit (dashed line) represent successfully infected strains of Streptomyces.



**Figure 4.7. Genomic maps of phages showing regions containing sequence similarity to spacers found within the CRISPRs of strains QF2, DF, and SFW. (a) TrvxScott, (b) BartholomewSD, (c) Shawty, and (d) IceWarrior. Key: Putative functions of CRISPR targeted genes.**

### Type I-E CRISPR-Cas System



**Figure 4.8. Class I, Type I-E CRISPR-Cas system encoded in the WGS of strain QF2.** (top) The Type I-E CRISPR-Cas operon encoded by strain QG is located from 4,973,482 to 4,987,810 and includes seven genes. The Type I-E cascade is followed by CRISPR 31, consisting of two repeats and a single spacer. (bottom) The canonical Type I-E CRISPR-Cas system encoded in the genome of *E. coli* K-12 MG1655 is located from 2,887,219 to 2,877,618 and includes eight genes. The Type I-E cascade is followed by a CRISPR 31, consisting of five repeats and four spacers.

## 4.7 Tables

**Table 4.1. Top NCBI BlastN hits of the 16S rRNA gene sequences used to create the phylogenetic tree containing eight *Streptomyces* isolates.**

Sample ID	NCBI BlastN Top 16S ribosomal RNA Hit Description	Max Score	Total Score	Query Cover	E value	Percent Identity	Accession No.
S. platensis JCM 4662	<i>Streptomyces platensis</i> strain JCM 4662	2748	2748	100%	0	100%	NR_024761.1
S. coelicolor A3(2)	<i>Streptomyces coelestis</i> strain AS 4.1594	2793	2793	98%	0	99.93%	NR_027222.1
JS	<i>Streptomyces rochei</i> strain NRRL B-1559	2741	2741	91%	0	99.93%	NR_116078.1
DF	<i>Streptomyces fulvissimus</i> strain DSM 40593	2691	2691	100%	0	99.93%	NR_103947.1
QF2	<i>Streptomyces californicus</i> strain NBRC 12750	2699	2699	100%	0	100%	NR_112257.1
EDE	<i>Streptomyces pratensis</i> strain ch24	1238	1238	99%	0	98.99%	NR_125616.1
SK	<i>Streptomyces californicus</i> strain NBRC 12750	1242	1242	100%	0	100%	NR_112257.1
AH	<i>Streptomyces pratensis</i> strain ch24	1194	1194	100%	0	100%	NR_125616.1
ELW	<i>Streptomyces atratus</i> strain NRRL B-16927	1138	1138	100%	0	99.84%	NR_043490.1
SFW	<i>Streptomyces caviscabies</i> strain ATCC 51928	2767	2767	100%	0	99.87%	NR_114493.1



**Table 4.2. BGCs encoded within the draft genome sequence of strain QF2.** The most similar BGCs in the MiBIG database are listed, as well as the percentage of genes in each MiBIG known cluster that have similarity to genes in the corresponding QF2 cluster. In cases where the most similar known BGC produces an antibiotic, the MOA was listed (Showdomycin<sup>50</sup>, Ficellomycin<sup>51</sup>, Skyllamycin<sup>52</sup>, Viomycin<sup>25</sup>, Alkylresorcinol<sup>53</sup>).

Strain - QF2							
Cluster	Type	Most Similar MiBIG Cluster and Predicted Percent Similarity	Antibacterial Activity	MiBig BGC-ID	Minimum	Maximum	Length (nt.)
1	butyrolactone	Coelmycin P1 T1PKS (8%)	-	BGC0000038	28795	39496	10702
2	terpene	Geosmin Terpene (100%)	-	BGC0001181	60346	82527	22182
3	NRPS	Griseobactin NRPS (35%)	-	BGC0000368	101394	123548	22155
4	NRPS	Coelichelin NRPS (72%)	-	BGC0000325	123549	145157	21609
5	T3PKS	Herboxidiene T1PKS, T3PKS (6%)	-	BGC0001065	184513	202853	18341
6	terpene	Isorenieratene Terpene (57%)	-	BGC0000664	681001	697369	16369
7	ectoine	Ectoine Other (75%)	-	BGC0000853	1125565	1134445	8881
8	T2PKS	Griseorhodin T2PKS (69%)	-	BGC0000230	1908342	1950906	42565
9	siderophore	Desferrioxamine B Siderophore (80%)	-	BGC0000941	2278919	2290697	11779
10	LAP,thiopeptide	-	-	-	2690041	2722548	32508
11	ectoine,butyrolactone	Showdomycin Other (47%)	Nucleic Acid and Protein Synthesis	BGC0001778	3344033	3359401	15369
12	melanin	Melanin Other (100%)	-	BGC0000911	4777563	4787988	10426
13	lanthipeptide	AmfS Lanthipeptide (100%)	-	BGC0000496	5105643	5127766	22124
14	terpene	-	-	-	5476141	5497007	20867
15	siderophore	Ficellomycin NRPS (3%)	DNA Replication	BGC0001593	5881146	5896034	14889
16	NRPS	Vioprolide A NRPS (25%)	-	BGC0001822	6093895	6137611	43717
17	bacteriocin	-	-	-	6194201	6205608	11408
18	NRPS-like	-	-	-	6430967	6442494	11528
19	NRPS-like,ladderane,arylpolyene	Skylamycin NRPS (14%)	Unknown MOA	BGC0000429	6608620	6645073	36454
20	terpene	Hopene Terpene (46%)	-	BGC0000663	6755130	6764031	8902
21	NRPS,T1PKS	Viomycin NRPS (66%)	Protein Synthesis	BGC0000458	6806327	6870853	64527
22	T3PKS	Alkylresorcinol T3PKS (66%)	Unknown MOA	BGC0000282	7058484	7080991	22508
23	lassopeptide	-	-	-	7260793	7282889	22097

**Table 4.3. BGCs encoded within the draft genome sequence of strain JS.** The most similar BGCs in the MiBIG database are listed, as well as the percentage of genes in each MiBIG known cluster that have similarity to genes in the corresponding JS cluster. In cases where the most similar known BGC produces an antibiotic, the MOA was listed (Methylenomycin<sup>54</sup>, Albaflavenone<sup>27</sup>, Lipopeptide 8D1-1 & 8D1-2<sup>55</sup>, Lysolipin<sup>56</sup>, Kinamycin<sup>57</sup>).

Strain - JS							
Cluster	Type	Most Similar MiBIG Cluster and Predicted Percent Similarity	Antibacterial Activity	MiBig BGC-ID	Minimum	Maximum	Length (nt.)
1	T3PKS	Herboxidiene T1PKS, T3PKS (7%)		BGC0001065	214109	229085	14977
2	ectoine	Ectoine Other (100%)		BGC0000853	672467	682865	10399
3	melanin	Melanin Other (60%)		BGC0000911	1414668	1425276	10609
4	siderophore	Desferrioxamine B Siderophore (66%)		BGC0000941	1507497	1519344	11848
5	furan	Methylenomycin Other (9%)	Potentially Inhibits Cell Wall Biosynthesis	BGC0000914	2667706	2688702	20997
6	NRPS	Ansamitocin P-3 T1PKS (7%)		BGC0001511	3021294	3076234	54941
7	terpene	Albaflavenone Terpene (100%)	Unknown MOA	BGC0000660	3743916	3764845	20930
8	T2PKS	Spore pigment T2PKS (66%)		BGC0000271	3800219	3857023	56805
9	siderophore	-	-	-	4274698	4286154	11457
10	bacteriocin	-	-	-	4471015	4482384	11370
11	terpene	-	-	-	4488880	4508472	19593
12	NRPS	Lipopeptide 8D1-1 & 8D1-2 NRPS (25%)	PMF Collapse	BGC0001370	4616241	4669719	53479
13	NRPS	Lipopeptide 8D1-1 & 8D1-2 NRPS (15%)	PMF Collapse	BGC0001370	4929292	4967094	37803
14	terpene	Hopene Terpene (76%)		BGC0000663	5042481	5069167	26687
15	terpene	Lysolipin T2PKS (4%)	Cell Wall Biosynthesis	BGC0000242	5088579	5103665	15087
16	T1PKS	Candicidin T1PKS (28%)		BGC0000034	5331251	5350185	18935
17	T2PKS, butyrolactone	Kinamycin T2PKS (25%)	DNA Synthesis	BGC0000236	5370217	5395391	25175
18	T1PKS	FR-008/Levorin A3 T1PKS (28%)		BGC0000061	5395392	5413565	18174

**Table 4.4. BGCs encoded within the draft genome sequence of strain SFW.** The most similar BGCs in the MiBIG database are listed, as well as the percentage of genes in each MiBIG known cluster that have similarity to genes in the corresponding SFW cluster. In cases where the most similar known BGC produces an antibiotic, the MOA was listed (Carbapenem<sup>29</sup>, Clavulanic acid<sup>58</sup>, Kanamycin<sup>59</sup>, Lipopeptide 8D1-1 & 8D1-2<sup>55</sup>, Ficellomycin<sup>51</sup>, Lactonamycin<sup>60</sup>, Istamycin<sup>61</sup>, Cinerubin<sup>62</sup>, Tetronasin<sup>63</sup>).

Strain - SFW							
Cluster	Type	Most Similar MiBIG Cluster and Predicted Percent Similarity	Antibacterial Activity	MiBig BGC-ID	Minimum	Maximum	Length (nt.)
1	NRPS, blactam	Carbapenem MM 4550 Other (62%)	Cell Wall Biosynthesis	BGC00000842	280243	420495	140253
2	NRPS	Coelichelin NRPS (90%)		BGC00000325	537068	587954	50887
3	terpene	Isorenieratene Terpene (28%)		BGC00000664	601924	615201	13278
4	blactam	Clavulanic acid Other (20%)	Beta-lactamase Inhibition	BGC00000845	839861	863248	23388
5	terpene	Hopene Terpene (69%)		BGC00000663	920123	946635	26513
6	T1PKS	Sceliphrolactam T1PKS (72%)		BGC0001770	1325302	1388343	63042
7	bacteriocin	-	-	-	1598693	1609196	10504
8	lanthipeptide	Kanamycin Saccharide (1%)	Protein Synthesis	BGC0000703	1734703	1760347	25645
9	NRPS	Lipopeptide 8D1-1 & 8D1-2 NRPS (6%)	PMF Collapse	BGC0001370	1773501	1830128	56628
10	siderophore	Ficellomycin NRPS (3%)	DNA replication	BGC0001593	2107372	2120491	13120
11	terpene	-	-	-	2186197	2205874	19678
12	butyrolactone	Lactonamycin T2PKS (3%)	Protein Synthesis	BGC0000238	4018281	4029096	10816
13	NRPS	Istamycin Saccharide (11%)	Protein Synthesis	BGC0000700	4251191	4307412	56222
14	siderophore	Desferrioxamine B Siderophore (83%)		BGC0000941	4924798	4936579	11782
15	lanthipeptide	-	-	-	5321361	5346350	24990
16	terpene	-	-	-	5588840	5608518	19679
17	ectoine	Ectoine Other (100%)		BGC00000853	6072388	6080990	8603
18	T2PKS, PKS-like	Cinerubin B T2PKS (25%)	DNA Intercalation	BGC0000212	6443938	6515214	71277
19	terpene	Steffimycin T2PKS-Saccharide (16%)		BGC0000273	6560271	6580717	20447
20	terpene, ectoine	Ectoine Other (100%)		BGC00000853	6860752	6881669	20918
21	bacteriocin	-	-	-	6909859	6920014	10156
22	T3PKS	Tetronasin T1PKS (11%)	PMF Collapse	BGC0000163	7071589	7112647	41059
23	melanin	Melanin Other (100%)		BGC0000911	7208921	7219385	10465
24	T2PKS, terpene	Spore pigment T2PKS (83%)		BGC0000271	7244899	7317424	72526
25	NRPS	Rimosamide NRPS (21%)		BGC0001760	7458615	7511513	52899
26	butyrolactone	-	-	-	7592249	7602533	10285

**Table 4.5. BGCs encoded within the closed genome sequence of strain DF.** The most similar BGCs in the MiBIG database are listed, as well as the percentage of genes in each MiBIG known cluster that have similarity to genes in the corresponding DF cluster. In cases where the most similar known BGC produces an antibiotic, the MOA was listed (Arsenopolyketides<sup>64</sup>, Naringenin<sup>65</sup>, Laspartomycin<sup>66</sup>, Asukamycin<sup>67</sup>, Tetrapetalone<sup>68</sup>, Nonactin<sup>31</sup>, SGR PTMs NRPS<sup>69</sup>).

Strain - DF							
Cluster	Type	Most Similar MiBIG Cluster and Predicted Percent Similarity	Antibacterial Activity	MiBIG BGC-ID	Minimum	Maximum	Length (nt.)
1	ectoine	-	-	-	55194	65595	10402
2	butyrolactone	Gamma-butyrolactone (100%)	-	BGC0000038	156789	167693	10905
3	terpene	Geosmin Terpene (100%)	-	BGC0001181	196683	218891	22209
4	transAT-PKS, T1PKS, NRPS	Griseobactin NRPS (82%)	-	BGC0000368	227963	343592	115630
5	NRPS	Coelichelin NRPS (81%)	-	BGC0000325	363422	414309	50888
6	NRPS	Arsenopolyketides Other (41%)	Unknown MOA	BGC0001283	421648	466618	44971
7	T3PKS	Naringenin Terpene (100%)	Potentially Inhibits Fatty Acid Biosynthesis	BGC0001310	481558	522683	41126
8	T2PKS	Granaticin T2PKS (29%)	-	BGC0000227	874325	916838	42514
9	terpene	Steffimycin T2PKS-Saccharide (19%)	-	BGC0000273	1079402	1100700	21299
10	ectoine	Ectoine Other (100%)	-	BGC0000853	1578303	1588701	10399
11	NRPS	Laspartomycin NRPS (9%)	Cell Wall Biosynthesis	BGC0000379	2182777	2259643	76867
12	siderophore	Desferrioxamine B Siderophore (100%)	-	BGC0000941	2696558	2708336	11779
13	LAP, thiopeptide	-	-	-	3100604	3132306	31703
14	NRPS	Microsclerodermins NRPS, T1PKS (14%)	-	BGC0001019	3327941	3392550	64610
15	betalactone	Divergolide T1PKS (6%)	-	BGC0001119	3740002	3763908	23907
16	T2PKS	Rabelomycin T2PKS-Saccharide (25%)	-	BGC0000262	4318297	4360617	42321
17	lassopeptide	Keywimycin RiPP (100%)	-	BGC0001634	4419581	4442206	22626
18	T1PKS	Streplazone E T1PKS (91%)	-	BGC0001296	4973731	5038740	65010
19	lanthipeptide	AmS Lanthipeptide (100%)	-	BGC0000496	5317536	5340199	22664
20	terpene	-	-	-	5677735	5698600	20866
21	siderophore	-	-	-	6150075	6164335	14261
22	butyrolactone	4-formylaminoxyvinylglycine Other (21%)	-	BGC0001488	6317636	6328541	10906
23	bacteriocin	-	-	-	6488518	6499849	11332
24	terpene	2-methylisoborneol Terpene (100%)	-	BGC0000658	6514280	6535368	21089
25	NRPS	Asukamycin T2PKS (12%)	Unknown MOA	BGC0000187	6618679	6679113	60435
26	arylpolyene	Tetrapetalone A-D PKS (26%)	Unknown MOA	BGC0001798	6726255	6767369	41115
27	NRPS	Nonactin T2PKS (85%)	Dissipates Transmembrane Electric Potential	BGC0000252	6774441	6840334	65894
28	terpene	Hopene Terpene (69%)	-	BGC0000663	7093211	7119781	26571
29	NRPS	SGR PTMs NRPS, T1PKS (100%)	Unknown MOA	BGC0001043	7204325	7247340	43016
30	bacteriocin	-	-	-	7260810	7271608	10799
31	melanin	Melanin Other (100%)	-	BGC0000911	7452699	7463166	10468
32	T3PKS	Herboxidiene T1PKS, T3PKS (7%)	-	BGC0001065	7495580	7536629	41050
33	terpene	Isorenieratene Terpene (100%)	-	BGC0000664	7626626	7652190	25565
34	thiopeptide, LAP, NRPS	Lactazole Thiopeptide (33%)	-	BGC0000606	7666072	7717669	51598

**Table 4.6. Inhibition of growth of clinically relevant pathogens by *Streptomyces* strains DF, SFW, QF2, and JS.** Plus signs indicate growth inhibition, while minus signs indicate pathogen growth.

	Gram-Negative Bacteria					Gram-Positive Bacteria		
	<i>E. coli</i>		<i>P. aeruginosa</i>			<i>B. subtilis</i>	MRSA	
	JP313 $\Delta tolC$	MC4100	PA01	P4	PA01 $\Delta efflux$	PY79	USA 300	TCH1516
<b>DF</b>	+	-	-	-	+	+		+
<b>SFW</b>	+	+	-	-	-	+		-
<b>QF2</b>	+	+	-	-	+	+		+
<b>JS</b>	+	+	-	-	+	+		+

**Table 4.7. Summary of the NCBI WGS annotations of four phage isolates.**

<b>Bacteriophage</b>	<b>Taxa</b>	<b>Genome (bp)</b>	<b>GC%</b>	<b>Genes</b>	<b>Cluster</b>	<b>Subcluster</b>	<b>Genbank Acc. No.</b>
TrvxScott	unclass. Aequatrovirus	52600	67.8	81	BD	BD2	MH669016
IceWarrior	unclass. Rimavirus	55532	59.5	86	BI	BI1	MK433259
BartholomewSD	unclass. Aequatrovirus	52131	67.6	88	BD	BD2	MK460245
Shawty	unclass. Lomovskayavirus	40733	63.2	58	BB	BB1	MK433266

**Table 4.8. Functions of the putative proteins encoded within the genomes of four phage isolates.**

Streptomyces Phage TrvxScott taxon:2301575		Streptomyces phage Shawly taxon:2510521		Streptomyces phage IceWarrior taxon:2510515		Streptomyces phage BartholomewSD taxon:2510587	
CDS No.	Product	CDS No.	Product	CDS No.	Product	CDS No.	Product
1	hypothetical protein	1	terminase small subunit	1	hypothetical protein	1	hypothetical protein
2	HNH endonuclease	2	terminase large subunit	2	HNH endonuclease	2	hypothetical protein
3	thioredoxin	3	portal protein	3	hypothetical protein	3	HNH endonuclease
4	terminase small subunit	4	capsid maturation protease	4	hypothetical protein	4	tRNA-Phe
5	terminase large subunit	5	major capsid protein	5	endolysin	5	hypothetical protein
6	portal protein	6	head-to-tail adaptor	6	head-to-tail connector complex protein	6	thioredoxin
7	capsid maturation protease	7	hypothetical protein	7	hypothetical protein	7	hypothetical protein
8	scaffolding protein	8	major tail protein	8	terminase large subunit	8	terminase
9	major capsid protein	9	hypothetical protein	9	hypothetical protein	9	portal protein
10	head-to-tail connector complex protein	10	tail assembly chaperone	10	hypothetical protein	10	MuF-like minor capsid protein
11	head-to-tail connector complex protein	11	tail assembly chaperone	11	hypothetical protein	11	scaffolding protein
12	hypothetical protein	12	tape measure protein	12	portal protein	12	major capsid protein
13	head-to-tail connector complex protein	13	minor tail protein	13	hypothetical protein	13	head-to-tail adaptor
14	major tail protein	14	minor tail protein	14	capsid maturation protease	14	head-to-tail stopper
15	tail assembly chaperone	15	minor tail protein	15	hypothetical protein	15	hypothetical protein
16	tail assembly chaperone	16	minor tail protein	16	hypothetical protein	16	tail terminator
17	tape measure protein	17	hypothetical protein	17	major tail protein	17	major tail protein
18	minor tail protein	18	tail fiber	18	hypothetical protein	18	tail assembly chaperone
19	minor tail protein	19	lysin A	19	major tail protein	19	tail assembly chaperone
20	hypothetical protein	20	hypothetical protein	20	hypothetical protein	20	tape measure protein
21	hypothetical protein	21	deoxynucleoside monophosphate kinase	21	chitinase	21	minor tail protein
22	hypothetical protein	22	immunity repressor	22	hypothetical protein	22	minor tail protein
23	minor tail protein	23	Cas4 family exonuclease	23	tape measure protein	23	hypothetical protein
24	hypothetical protein	24	hypothetical protein	24	minor tail protein	24	minor tail protein
25	lysin A	25	hypothetical protein	25	minor tail protein	25	hypothetical protein
26	hypothetical protein	26	hypothetical protein	26	hypothetical protein	26	hypothetical protein
27	hypothetical protein	27	hypothetical protein	27	hypothetical protein	27	hypothetical protein
28	hypothetical protein	28	hypothetical protein	28	hypothetical protein	28	lysin A
29	hypothetical protein	29	hypothetical protein	29	holin	29	hypothetical protein
30	exonuclease	30	HNH endonuclease	30	hypothetical protein	30	hypothetical protein
31	hypothetical protein	31	DNA primase	31	hypothetical protein	31	immunity repressor
32	hypothetical protein	32	restriction endonuclease	32	hypothetical protein	32	hypothetical protein
33	hypothetical protein	33	DNA polymerase I	33	hypothetical protein	33	Cas4 family exonuclease
34	deoxycytidylate deaminase	34	RNA polymerase sigma factor	34	hypothetical protein	34	hypothetical protein
35	DNA helicase	35	hypothetical protein	35	hypothetical protein	35	hypothetical protein
36	holliday junction resolvase	36	hypothetical protein	36	hypothetical protein	36	hypothetical protein
37	hypothetical protein	37	hypothetical protein	37	hypothetical protein	37	deoxycytidylate deaminase
38	DNA primase	38	hypothetical protein	38	hypothetical protein	38	DnaB-like helicase
39	DNA primase	39	hypothetical protein	39	hypothetical protein	39	holliday junction resolvase
40	hypothetical protein	40	ThyX-like thymidylate synthase	40	hypothetical protein	40	hypothetical protein
41	hypothetical protein	41	hypothetical protein	41	hypothetical protein	41	DNA primase
42	exonuclease	42	hypothetical protein	42	hypothetical protein	42	DNA primase
43	hypothetical protein	43	thioredoxin	43	hypothetical protein	43	hypothetical protein
44	HTH DNA binding protein	44	hypothetical protein	44	hypothetical protein	44	hypothetical protein
45	hypothetical protein	45	deoxycytidylate deaminase	45	hypothetical protein	45	hypothetical protein
46	ribonucleotide reductase	46	hypothetical protein	46	hypothetical protein	46	Mre11 family dsDNA break repair endo/exonuclease
47	DNA methylase	47	hypothetical protein	47	hypothetical protein	47	hypothetical protein
48	hypothetical protein	48	hypothetical protein	48	hypothetical protein	48	helix-turn-helix DNA binding protein
49	hypothetical protein	49	hypothetical protein	49	hypothetical protein	49	hypothetical protein
50	hypothetical protein	50	hypothetical protein	50	hypothetical protein	50	ribonucleotide reductase
51	HTH DNA binding protein	51	hypothetical protein	51	hypothetical protein	51	hypothetical protein
52	integrase	52	hypothetical protein	52	hypothetical protein	52	hypothetical protein
53	hypothetical protein	53	protein kinase	53	hypothetical protein	53	hypothetical protein
54	thymidylate synthase	54	integrase	54	hypothetical protein	54	hypothetical protein
55	hypothetical protein	55	tRNA-Asp	55	hypothetical protein	55	helix-turn-helix DNA binding protein
56	hypothetical protein	56	tRNA-Thr	56	hypothetical protein	56	integrase
57	hypothetical protein	57	hypothetical protein	57	hypothetical protein	57	hypothetical protein
58	hypothetical protein	58	HNH endonuclease	58	hypothetical protein	58	ThyX-like thymidylate synthase
59	hypothetical protein			59	DNA primase-polymerase	59	hypothetical protein
60	hypothetical protein			60	hypothetical protein	60	hypothetical protein
61	deoxynucleoside monophosphate kinase			61	hypothetical protein	61	hypothetical protein
62	hypothetical protein			62	hypothetical protein	62	hypothetical protein
63	hypothetical protein			63	hypothetical protein	63	hypothetical protein
64	hypothetical protein			64	hypothetical protein	64	hypothetical protein
65	hypothetical protein			65	hypothetical protein	65	deoxymononucleoside kinase
66	hypothetical protein			66	hypothetical protein	66	hypothetical protein
67	hypothetical protein			67	hypothetical protein	67	hypothetical protein
68	hypothetical protein			68	hypothetical protein	68	hypothetical protein
69	hypothetical protein			69	hypothetical protein	69	hypothetical protein
70	hypothetical protein			70	hypothetical protein	70	hypothetical protein
71	hypothetical protein			71	hypothetical protein	71	hypothetical protein
72	hypothetical protein			72	hypothetical protein	72	hypothetical protein
73	hypothetical protein			73	hypothetical protein	73	hypothetical protein
74	hypothetical protein			74	hypothetical protein	74	hypothetical protein
75	hypothetical protein			75	hypothetical protein	75	hypothetical protein
76	hypothetical protein			76	hypothetical protein	76	hypothetical protein
77	hypothetical protein			77	hypothetical protein	77	hypothetical protein
78	acetyltransferase			78	hypothetical protein	78	hypothetical protein
79	hypothetical protein			79	hypothetical protein	79	hypothetical protein
80	hypothetical protein			80	hypothetical protein	80	hypothetical protein
81	hypothetical protein			81	DNA helicase	81	hypothetical protein
				82	HNH endonuclease	82	hypothetical protein
				83	hydrolase	83	hypothetical protein
				84	DNA helicase	84	hypothetical protein
				85	helix-turn-helix DNA binding domain protein	85	hypothetical protein
				86	hypothetical protein	86	hypothetical protein
						87	hypothetical protein
						88	hypothetical protein

**Table 4.9. General characteristics of predicted CRISPR-Cas systems within the genomes of strains DF, SFW, QF2, and JS.** Included in this Table 4.is the number of spacers within the genome of each bacterial strain with sequence similarity to regions within any of the four phage isolates (IceWarrior, TrvxScott, BartholomewSD, or Shawty).

<b>Strains</b>	<b>CRISPR</b>	<b>Spacers</b>	<b>Repeats</b>	<b>Spacers with Blastn Hits to Host Range Phage</b>	<b>Cas Loci</b>	<b>Cas- Associated Genes</b>
<b>DF</b>	12	14	26	2	3	9
<b>SFW</b>	11	23	34	2	3	17
<b>QF2</b>	39	168	207	14	5	22
<b>JS</b>	4	8	12	0	4	20



**Table 4.10. Characteristics of the 38 CRISPRs predicted in the draft genome sequence of strain QF2.** Spacers with sequence similarity to any of the four phages in this study are listed next to their corresponding CRISPR and are identified according to their position relative to all other spacers within the QF2 genome.

Strain QF2						
CRISPRs	Min	Max	Length (nt.)	No. Repeats	No. Spacers	Spacer Blastn Hit to Host Range Phage
1	384313	384428	116	2	1	S1 [BartholomewSD]
2	437656	437755	100	2	1	
3	835820	835922	103	2	1	
4	1307659	1307752	94	2	1	
5	1344898	1345205	308	6	5	
6	1543053	1543144	92	2	1	
7	1616159	1616542	384	5	4	S13 [BartholomewSD, TrvxScott]
8	1618208	1618291	84	2	1	
9	1833873	1833977	105	2	1	
10	1861097	1861197	101	2	1	
11	2316101	2316187	87	2	1	
12	2704894	2704990	97	2	1	
13	3015981	3016376	396	7	6	S21 [IceWarrior]
14	3106815	3107262	448	8	7	S27 [BartholomewSD, TrvxScott]
15	3112452	3113433	982	17	16	S34 [IceWarrior], S41 [TrvxScott]
16	3138610	3140161	1,552	26	25	
17	3145439	3145830	392	7	6	
18	3444685	3444779	95	2	1	
19	3507440	3507598	159	3	2	
20	3791739	3792012	274	5	4	
21	3838730	3838804	75	2	1	
22	4257871	4258080	210	4	3	S89 [IceWarrior]
23	4327550	4328916	1,367	23	22	S105 [IceWarrior], S106 [Shawty, TrvxScott]
24	4333365	4334666	1,302	22	21	S118 [TrvxScott], S131 [Shawty, TrvxScott]
25	4335904	4336481	578	10	9	S138 [BartholomewSD], S140 [TrvxScott]
26	4522773	4522879	107	2	1	
27	4528080	4528148	69	2	1	
28	4657265	4657374	110	2	1	
29	4754273	4754356	84	2	1	
30	4787509	4787642	134	3	2	
31	4987714	4987810	97	2	1	
32	5305650	5305745	96	2	1	
33	5400452	5400541	90	2	1	
34	5417083	5417162	80	2	1	S151 [IceWarrior]
35	5441923	5442032	110	2	1	
36	6552625	6552699	75	2	1	
37	6798173	6798309	137	3	2	
38	7177566	7177797	232	6	5	

**Table 4.11. Characteristics of the spacers within the genomes of strain DF, SFW, and QF2 that have sequence similarity to at least one of the four phage isolates. The bold portion of each spacer shares high sequence similarity with a region in the genome of the listed host range phage. In cases where a single spacer mapped to two phages, bold and underlined are used to sequences distinguish the two.**

Strain DF									
CRISPR	Spacer	Blastn Hit to Host Range Phage	Score (bits) / E Val.	No. Identities (%ID)	Strand	Minimum	Maximum	Length (nt.)	Sequence
9	S11	BartholomewSD	30.2 (15) / 2.4	15/15 (100%)	Plus / Minus	4171019	4171050	32	<b>TGCCCA<b>CCGGCCGAGCC</b>CCCTCCCGCAGGCAG</b>
		TrvxScott	30.2 (15) / 4.9	15/15 (100%)	Plus / Plus				
10	S12	BartholomewSD	28.2 (14) / 9.5	14/14 (100%)	Plus / Minus	4684554	4684602	49	GGGTCCCCGCGGTGCGTGCAT <b>GTCTTCGGCTTGA</b> GC <b>GGGCTGCCG</b>
Strain SFW									
CRISPR	Spacer	Blastn Hit to Host Range Phage	Score (bits) / E Val.	No. Identities (%ID)	Strand	Minimum	Maximum	Length (nt.)	Sequence
1	S1	Shawly	28.2 (14) / 5.9	14/14 (100%)	Plus / Plus	250625	250650	26	GAGTCA <b>CCAGCCGGCGAAAGGCACGC</b>
5	S6	IceWarrior	32.2 (16) / 0.99	19/20 (95%)	Plus / Plus	4553831	4553872	42	CGGGCGTCCAGCGGTGACGAGCG <b>TCGCCTGCTACTTCTCCTT</b> G
Strain QF2									
CRISPR	Spacer	Blastn Hit to Host Range Phage	Score (bits) / E Val.	No. Identities (%ID)	Strand	Minimum	Maximum	Length (nt.)	Sequence
1	S1	BartholomewSD	32.2 (16) / 1.2	16/16 (100%)	Plus / Plus	384347	384394	48	GCGGACGGCG <b>CGCGCCGGTACCC</b> CCGGTGTCCACGACGGCGCGCGG
7	S13	BartholomewSD	32.2 (16) / 1.9	16/16 (100%)	Plus / Plus	1616356	1616423	68	CGACCTGCGGTACCCTCGATCCGGGGCGGTCCCATCT <b>ACAAGGGCACGGTGC</b> TCCAGCGGACCGAG
		TrvxScott	30.2 (15) / 1.9	16/16 (100%)	Plus / Plus				
13	S21	IceWarrior	30.2 (15) / 2.5	15/15 (100%)	Plus / Minus	3016077	3016109	33	CGCGGAACCTC <b>AAGGAGGAGAACGGCGCGGG</b>
14	S27	BartholomewSD	30.2 (15) / 3.1	18/19 (94%)	Plus / Minus	3106902	3106938	37	<b>AGGGCTGGCCGTCCGGGTGCGGGTGAGTGTGT</b>
		TrvxScott	30.2 (15) / 3.1	18/19 (94%)	Plus / Minus				
15	S34	IceWarrior	30.2 (15) / 2.4	15/15 (100%)	Plus / Plus	3112541	3112572	32	ACAGCGACGTGCCT <b>TACAACCTACGCCGCCTGG</b>
		TrvxScott	28.2 (14) / 9.5	14/14 (100%)	Plus / Plus	3112961	3112992	32	GGTGTGAACCGTCC <b>GGCGCGTGA</b> ACTTGT
22	S89	IceWarrior	30.2 (15) / 2.5	15/15 (100%)	Plus / Plus	4257958	4257990	33	CCCGGGCGTCC <b>TCCCGGAGGACCTG</b> CCCC
		TrvxScott	30.2 (15) / 2.5	15/15 (100%)	Plus / Plus	4257958	4257990	33	CATCAGC <b>GTCTGAAGCAGCACGCC</b> CCATCGCCTT
23	S106	Shawly	28.2 (14) / 9.5	14/14 (100%)	Plus / Plus	4328491	4328522	32	TG <b>ATCGAGCCGGACGGCAC</b> ATCAGCGGCC
		TrvxScott	28.2 (14) / 9.5	14/14 (100%)	Plus / Minus	4328491	4328522	32	
	S118	TrvxScott	28.2 (14) / 9.5	14/14 (100%)	Plus / Plus	4333692	4333723	32	GCCGCGTCCGGCT <b>ACGGCTACGGCTCCG</b> CCCC
24	S131	Shawly	28.2 (14) / 9.5	14/14 (100%)	Plus / Plus	4334483	4334514	32	AACG <b>CCCTCCATGAGGGCTG</b> CGGTTGCGTC
		TrvxScott	28.2 (14) / 9.5	14/14 (100%)	Plus / Minus	4334483	4334514	32	
	S138	BartholomewSD	28.2 (14) / 9.5	14/14 (100%)	Plus / Plus	4336177	4336208	32	AACGCGCAGCGATGG <b>CCCGTACGAGCGGGG</b>
		TrvxScott	28.2 (14) / 9.5	14/14 (100%)	Plus / Minus	4336177	4336208	32	
25	S140	TrvxScott	28.2 (14) / 9.5	14/14 (100%)	Plus / Minus	4336299	4336330	32	ATCCTCGCC <b>TCCAGACCGCTCAGCGC</b> AGAT
33	S151	IceWarrior	36.2 (18) / 0.063	18/18 (100%)	Plus / Minus	5400476	5400517	42	GTGGTGGCCTCGCCGACCA <b>GTGCTCGGACGCCTGGCGGCC</b>

**Table 4.12. Characteristics of the 12 CRISPRs predicted in the complete genome sequence of strain DF.** Spacers with sequence similarity to any of the four phages in this study are listed next to their corresponding CRISPR and are identified according to their position relative to all other spacers within the DF genome.

Strain DF						
CRISPRs	Min	Max	Length (nt.)	No. Repeats	No. Spacers	Spacer Blastn Hit to Host Range Phage
1	1491701	1491797	97	2	1	
2	1517080	1517203	124	2	1	
3	1669449	1669522	74	2	1	
4	1995371	1995463	93	2	1	
5	2013671	2013864	194	3	2	
6	2241854	2241958	105	2	1	
7	2443114	2443225	112	2	1	
8	3767692	3767796	105	2	1	
9	4170935	4171076	142	3	2	S11 [BartholomewSD]
10	4684531	4684625	95	2	1	S12 [BartholomewSD, TrvxScott]
11	4802187	4802274	88	2	1	
12	5682554	5682639	86	2	1	

**Table 4.13. Characteristics of the 11 CRISPRs predicted in the draft genome sequence of strain SFW.** Spacers with sequence similarity to any of the four phages in this study are listed next to their corresponding CRISPR and are identified according to their position relative to all other spacers within the SFW genome.

Strain SFW						
CRISPRs	Min	Max	Length (nt.)	No. Repeats	No. Spacers	Spacer Blastn Hit to Host Range Phage
1	250601	250674	74	2	1	S1 [Shawty]
2	452415	452685	271	3	2	
3	2779718	2779796	79	2	1	
4	2824197	2824275	79	2	1	
5	4553801	4554168	368	5	4	S6 [IceWarrior]
6	4731148	4731651	504	10	9	
7	5166339	5166424	86	2	1	
8	5317268	5317371	104	2	1	
9	5722922	5723016	95	2	1	
10	6532966	6533033	68	2	1	
11	7183989	7184080	92	2	1	

## 4.8 References

1. Wencewicz, T. A. Crossroads of Antibiotic Resistance and Biosynthesis. *J. Mol. Biol.* **431**, 3370–3399 (2019).
2. Pogue, J. M., Kaye, K. S., Cohen, D. A. & Marchaim, D. Appropriate antimicrobial therapy in the era of multidrug-resistant human pathogens. *Clin. Microbiol. Infect.* **21**, 302–312 (2015).
3. Brown, E. D. & Wright, G. D. Antibacterial drug discovery in the resistance era. *Nature* **529**, 336–343 (2016).
4. Newman, D. J. & Cragg, G. M. Natural Products as Sources of New Drugs from 1981 to 2014. *J. Nat. Prod.* **79**, 629–661 (2016).
5. Fischbach, M. A. & Walsh, C. T. Antibiotics for emerging pathogens. *Science (80-. ).* **325**, 1089–1093 (2009).
6. Hover, B. M., Kim, S. H., Katz, M., Charlop-Powers, Z., Owen, J. G., Ternei, M. A., Maniko, J., Estrela, A. B., Molina, H., Park, S., Perlin, D. S. & Brady, S. F. Culture-independent discovery of the malacidins as calcium-dependent antibiotics with activity against multidrug-resistant Gram-positive pathogens. *Nat. Microbiol.* **3**, 415–422 (2018).
7. Ling, L. L., Schneider, T., Peoples, A. J., Spoering, A. L., Engels, I., Conlon, B. P., Mueller, A., Schäberle, T. F., Hughes, D. E., Epstein, S., Jones, M., Lazarides, L., Steadman, V. A., Cohen, D. R., Felix, C. R., Fetterman, K. A., Millett, W. P., Nitti, A. G., Zullo, A. M., Chen, C. & Lewis, K. A new antibiotic kills pathogens without detectable resistance. *Nature* **517**, 455–459 (2015).
8. Chen, R., Wong, H. & Burns, B. New Approaches to Detect Biosynthetic Gene Clusters in the Environment. *Medicines* **6**, 32 (2019).
9. Pye, C. R., Bertin, M. J., Lokey, R. S., Gerwick, W. H. & Linington, R. G. Retrospective analysis of natural products provides insights for future discovery trends. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 5601–5606 (2017).
10. Nonejuie, P., Burkart, M., Pogliano, K. & Pogliano, J. Bacterial cytological profiling rapidly identifies the cellular pathways targeted by antibacterial molecules. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 16169–16174 (2013).
11. Lamsa, A., Lopez-Garrido, J., Quach, D., Riley, E. P., Pogliano, J. & Pogliano, K. Rapid Inhibition Profiling in *Bacillus subtilis* to Identify the Mechanism of Action of New Antimicrobials. *ACS Chem. Biol.* **11**, 2222–2231 (2016).
12. Lamsa, A., Liu, W. T., Dorrestein, P. C. & Pogliano, K. The *Bacillus subtilis*

- cannibalism toxin SDP collapses the proton motive force and induces autolysis. *Mol. Microbiol.* **84**, 486–500 (2012).
13. Peters, C. E., Lamsa, A., Liu, R. B., Quach, D., Sugie, J., Brumage, L., Pogliano, J., Lopez-Garrido, J. & Pogliano, K. Rapid Inhibition Profiling Identifies a Keystone Target in the Nucleotide Biosynthesis Pathway. *ACS Chem. Biol.* (2018). doi:10.1021/acscchembio.8b00273
  14. Nonejuie, P., Trial, R. M., Newton, G. L., Lamsa, A., Ranmali Perera, V., Aguilar, J., Liu, W. T., Dorrestein, P. C., Pogliano, J. & Pogliano, K. Application of bacterial cytological profiling to crude natural product extracts reveals the antibacterial arsenal of *Bacillus subtilis*. *J. Antibiot. (Tokyo)*. **69**, 353–361 (2016).
  15. Quach, D. T., Sakoulas, G., Nizet, V., Pogliano, J. & Pogliano, K. Bacterial Cytological Profiling (BCP) as a Rapid and Accurate Antimicrobial Susceptibility Testing Method for *Staphylococcus aureus*. *EBioMedicine* **4**, 95–103 (2016).
  16. Alberti, F. & Corre, C. Editing streptomycete genomes in the CRISPR/Cas9 age. *Nat. Prod. Rep.* **36**, 1237–1248 (2019).
  17. de Jonge, P. A., Nobrega, F. L., Brouns, S. J. J. & Dutilh, B. E. Molecular and Evolutionary Determinants of Bacteriophage Host Range. *Trends Microbiol.* **27**, 51–63 (2019).
  18. Russell, D. A. & Hatfull, G. F. PhagesDB: The actinobacteriophage database. *Bioinformatics* **33**, 784–786 (2017).
  19. Genilloud, O. Actinomycetes: Still a source of novel antibiotics. *Nat. Prod. Rep.* **34**, 1203–1232 (2017).
  20. Thakur, D., Bora, T. C., Bordoloi, G. N. & Mazumdar, S. Influence of nutrition and culturing conditions for optimum growth and antimicrobial metabolite production by *Streptomyces* sp. 201. *J. Mycol. Med.* **19**, 161–167 (2009).
  21. Zimmerman, S. B. Toroidal nucleoids in *Escherichia coli* exposed to chloramphenicol. *J. Struct. Biol.* **138**, 199–206 (2002).
  22. Aziz, R. K., Bartels, D., Best, A., DeJongh, M., Disz, T., Edwards, R. A., Formsma, K., Gerdes, S., Glass, E. M., Kubal, M., Meyer, F., Olsen, G. J., Olson, R., Osterman, A. L., Overbeek, R. A., McNeil, L. K., Paarmann, D., Paczian, T., Parrello, B., Pusch, G. D., Reich, C., Stevens, R., Vassieva, O., Vonstein, V., Wilke, A. & Zagnitko, O. The RAST Server: Rapid annotations using subsystems technology. *BMC Genomics* **9**, 1–15 (2008).

23. Blin, K., Shaw, S., Steinke, K., Villebro, R., Ziemert, N., Lee, S. Y., Medema, M. H. & Weber, T. antiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. *Nucleic Acids Res.* **47**, W81–W87 (2019).
24. Kautsar, S. A., Blin, K., Shaw, S., Navarro-mu, J. C., Terlouw, R., Hooft, J. J. J. Van Der, Santen, J. A. Van, Tracanna, V., Duran, H. G. S., Andreu, P., Selemojica, N., Alanjary, M., Robinson, S. L., Lund, G., Epstein, S. C., Sisto, A. C., Charkoudian, L. K., Linington, R. G., Weber, T. & Medema, M. H. MIBiG 2.0: a repository for biosynthetic gene clusters. 1–5 (2019). doi:10.1093/nar/gkz882
25. Holm, M., Borg, A., Ehrenberg, M. & Sanyal, S. Molecular mechanism of viomycin inhibition of peptide elongation in bacteria. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 978–983 (2016).
26. Gürtler, H., Pedersen, R., Anthoni, U., Christophersen, C., Nielsen, P. H., Wellington, E. M. H., Pedersen, C. & Bock, K. Albaflavenone, a sesquiterpene ketone with a zizaene skeleton produced by a streptomycete with a new rope morphology. *J. Antibiot. (Tokyo)*. **47**, 434–439 (1994).
27. Zhao, B., Lin, X., Lei, L., Lamb, D. C., Kelly, S. L., Waterman, M. R. & Cane, D. E. Biosynthesis of the sesquiterpene antibiotic albaflavenone in *Streptomyces coelicolor* A3(2). *J. Biol. Chem.* **283**, 8183–8189 (2008).
28. Nagarajan, R., Boeck, L. D., Gorman, M., Hamill, R. L., Higgins, C. E., Hoehn, M. M., Stark, W. M. & Whitney, J. G.  $\beta$ -Lactam Antibiotics from *Streptomyces*. *J. Am. Chem. Soc.* **93**, 2308–2310 (1971).
29. Papp-Wallace, K. M., Endimiani, A., Taracila, M. A. & Bonomo, R. A. Carbapenems: Past, present, and future. *Antimicrob. Agents Chemother.* **55**, 4943–4960 (2011).
30. Neu, H. C. & Fu, K. P. Clavulanic acid, a novel inhibitor of  $\beta$ -lactamases. *Antimicrob. Agents Chemother.* **14**, 650–655 (1978).
31. Kusche, B. R., Smith, A. E., McGuirl, M. A. & Priestley, N. D. Alternating pattern of stereochemistry in the nonactin macrocycle is required for antibacterial activity and efficient ion binding. *J. Am. Chem. Soc.* **131**, 17155–17165 (2009).
32. Ackermann, H. W. 5500 Phages examined in the electron microscope. *Methods Mol. Biol.* **394**, 213–234 (2007).
33. Jin, H., Jiang, Y.-L., Yang, F., Zhang, J.-T., Li, W.-F., Zhou, K., Ju, J., Chen, Y. & Zhou, C.-Z. Capsid Structure of a Freshwater Cyanophage Siphoviridae Mic1. *Structure* **27**, 1508-1516.e3 (2019).
34. Murphy, J., Mahony, J., Ainsworth, S., Nauta, A. & van Sinderen, D.

- Bacteriophage orphan DNA methyltransferases: Insights from their bacterial origin, function, and occurrence. *Appl. Environ. Microbiol.* **79**, 7547–7555 (2013).
35. Couvin, D., Bernheim, A., Toffano-Nioche, C., Touchon, M., Michalik, J., Néron, B., Rocha, E. P. C., Vergnaud, G., Gautheret, D. & Pourcel, C. CRISPRCasFinder, an update of CRISRFinder, includes a portable version, enhanced performance and integrates search for Cas proteins. *Nucleic Acids Res.* **46**, W246–W251 (2018).
  36. Abby, S. S., Néron, B., Ménager, H., Touchon, M. & Rocha, E. P. C. MacSyFinder: A program to mine genomes for molecular systems with an application to CRISPR-Cas systems. *PLoS One* **9**, (2014).
  37. Zhang, Q. & Ye, Y. Not all predicted CRISPR-Cas systems are equal: Isolated cas genes and classes of CRISPR like elements. *BMC Bioinformatics* **18**, 1–12 (2017).
  38. Zhang, F., Zhao, S., Ren, C., Zhu, Y., Zhou, H., Lai, Y., Zhou, F., Jia, Y., Zheng, K. & Huang, Z. CRISPRminer is a knowledge base for exploring CRISPR-Cas systems in microbe and phage interactions. *Commun. Biol.* **1**, (2018).
  39. Biswas, A., Staals, R. H. J., Morales, S. E., Fineran, P. C. & Brown, C. M. CRISPRDetect: A flexible algorithm to define CRISPR arrays. *BMC Genomics* **17**, 1–14 (2016).
  40. Faure, G., Shmakov, S. A., Yan, W. X., Cheng, D. R., Scott, D. A., Peters, J. E., Makarova, K. S. & Koonin, E. V. CRISPR–Cas in mobile genetic elements: counter-defence and beyond. *Nat. Rev. Microbiol.* **17**, (2019).
  41. Kang, M. J., Jones, B. D., Mandel, A. L., Hammons, J. C., DiPasquale, A. G., Rheingold, A. L., La Clair, J. J. & Burkart, M. D. Isolation, structure elucidation, and antitumor activity of spirohexenolides A and B. *J. Org. Chem.* **74**, 9054–9061 (2009).
  42. Brettin, T., Davis, J. J., Disz, T., Edwards, R. A., Gerdes, S., Olsen, G. J., Olson, R., Overbeek, R., Parrello, B., Pusch, G. D., Shukla, M., Thomason, J. A., Stevens, R., Vonstein, V., Wattam, A. R. & Xia, F. RASTtk: A modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. *Sci. Rep.* **5**, (2015).
  43. Cresawn, S. G., Bogel, M., Day, N., Jacobs-Sera, D., Hendrix, R. W. & Hatfull, G. F. Phamerator: A bioinformatic tool for comparative bacteriophage genomics. *BMC Bioinformatics* **12**, (2011).
  44. Lane, D. J. *16S/23S rRNA Sequencing*. In: *Stackebrandt, E. and Goodfellow, M., Eds., Nucleic Acid Techniques in Bacterial Systematics*. John Wiley & Sons, Print. (1991).



45. Casadaban, M. J. Transposition and fusion of the lac genes to selected promoters in *Escherichia coli* using bacteriophage lambda and Mu. *J. Mol. Biol.* **104**, 541–555 (1976).
46. Youngman, P., Perkins, J. B. & Losick, R. A novel method for the rapid cloning in *Escherichia coli* of *Bacillus subtilis* chromosomal DNA adjacent to Tn 917 insertions. *MGG Mol. Gen. Genet.* **195**, 424–433 (1984).
47. Buch, A., Archana, G. & Naresh Kumar, G. Metabolic channeling of glucose towards gluconate in phosphate-solubilizing *Pseudomonas aeruginosa* P4 under phosphorus deficiency. *Res. Microbiol.* **159**, 635–642 (2008).
48. Highlander, S. K., Hultén, K. G., Qin, X., Jiang, H., Yerrapragada, S., Mason, E. O., Shang, Y., Williams, T. M., Fortunov, R. M., Liu, Y., Igboeli, O., Petrosino, J., Tirumalai, M., Uzman, A., Fox, G. E., Cardenas, A. M., Muzny, D. M., Hemphill, L., Ding, Y., Dugan, S., Blyth, P. R., Buhay, C. J., Dinh, H. H., Hawes, A. C., Holder, M., Kovar, C. L., Lee, S. L., Liu, W., Nazareth, L. V., Wang, Q., Zhou, J., Kaplan, S. L. & Weinstock, G. M. Subtle genetic changes enhance virulence of methicillin resistant and sensitive *Staphylococcus aureus*. *BMC Microbiol.* **7**, 1–14 (2007).
49. Gomez-Escribano, J. P. & Bibb, M. J. Engineering *Streptomyces coelicolor* for heterologous expression of secondary metabolite gene clusters. *Microb. Biotechnol.* **4**, 207–215 (2011).
50. Economou, A., Pogliano, J. A., Beckwith, J., Oliver, D. B. & Wickner, W. SecA membrane cycling at SecYEG is driven by distinct ATP binding and hydrolysis events and is regulated by SecD and SecF. *Cell* **83**, 1171–1181 (1995).
51. Morona, R. & Reeves, P. Molecular cloning of the tolC locus of *Escherichia coli* K-12 with the use of transposon Tn 10. *MGG Mol. Gen. Genet.* **184**, 430–433 (1981).
52. Whitney, E. N. The tolC locus in *Escherichia coli* K12. *Genetics* **67**, 39–53 (1971).
53. Komatsu, Y. & Tanaka, K. Mechanism of Action of Showdomycin: Part II. Effect of Showdomycin on the Synthesis of Deoxyribonucleic Acid in *Escherichia coli*. *Agric. Biol. Chem.* **34**, 891–899 (1970).
54. Reusser, F. Ficellomycin and Feldamycin; Inhibitors of Bacterial Semiconservative DNA Replication. *Biochemistry* **16**, 3406–3412 (1977).
55. Pohle, S., Appelt, C., Roux, M., Fiedler, H. P. & Süßmuth, R. D. Biosynthetic gene cluster of the non-ribosomally synthesized cyclodepsipeptide skyllamycin: Deciphering unprecedented ways of unusual hydroxylation reactions. *J. Am. Chem. Soc.* **133**, 6194–6205 (2011).

56. Kanda, N., Ishizaki, N., Inoue, N., Oshima, M., Handa, A. & Kitahara, T. Db-2073, a New Alkylresorcinol Antibiotic. I. Taxonomy, Isolation and Characterization. *J. Antibiot. (Tokyo)*. **28**, 935–942 (1975).
57. Hobbs, G., Obanye, A. I. C., Petty, J., Mason, J. C., Barratt, E., Gardner, D. C. J., Flett, F., Smith, C. P., Broda, P. & Oliver, S. G. An integrated approach to studying regulation of production of the antibiotic methylenomycin by *Streptomyces coelicolor* A3(2). *J. Bacteriol.* **174**, 1487–1494 (1992).
58. Liu, Y., Ding, S., Shen, J. & Zhu, K. Nonribosomal antibacterial peptides that target multidrug-resistant bacteria. *Nat. Prod. Rep.* **36**, 573–592 (2019).
59. Drautz, H., Keller-Schierlein, W. & Zähler, H. Stoffwechselprodukte von Mikroorganismen - 149. Mitteilung. Lysolipin I, ein neues Antibioticum aus *Streptomyces violaceoniger*. *Arch. Microbiol.* **106**, 175–190 (1975).
60. Ballard, T. E. & Melander, C. Kinamycin-mediated DNA cleavage under biomimetic conditions. *Tetrahedron Lett.* **49**, 3157–3161 (2008).
61. Reading, C. & Cole, M. Clavulanic acid: a beta lactamase inhibiting beta lactam from *Streptomyces clavuligerus*. *Antimicrob. Agents Chemother.* **11**, 852–857 (1977).
62. Suzuki, J., Kunimoto, T. & Hori, M. Effects of kanamycin on protein synthesis : Inhibition of elongation of peptide chains. *J. Antibiot. (Tokyo)*. **23**, 99–101 (1970).
63. Höltzel, A., Dieter, A., Schmid, D. G., Brown, R., Goodfellow, M., Beil, W., Jung, G. & Fiedler, H. P. Lactonamycin Z, an antibiotic and antitumor compound produced by *Streptomyces sanglieri* strain AK 623. *J. Antibiot. (Tokyo)*. **56**, 1058–1061 (2003).
64. Wehmeier, U. F. & Piepersberg, W. *Chapter 19 Enzymology of Aminoglycoside Biosynthesis-Deduction from Gene Clusters. Methods in Enzymology* vol. 459 (Elsevier Inc., 2009).
65. Kersten, W., Kersten, H. & Szybalski, W. Physicochemical Properties of Complexes between Deoxyribonucleic Acid and Antibiotics Which Affect Ribonucleic Acid Synthesis (Actinomycin, Daunomycin, Cinerubin, Nogalamycin, Chromomycin, Mithramycin, and Olivomycin). *Biochemistry* **5**, 236–244 (1966).
66. Newbold, C. J., Wallace, R. J., Watt, N. D. & Richardson, A. J. Effect of the novel ionophore tetronasin (ICI 139603) on ruminal microorganisms. *Appl. Environ. Microbiol.* **54**, 544–547 (1988).
67. Cruz-Morales, P., Kopp, J. F., Martínez-Guerrero, C., Yáñez-Guerra, L. A., Selem-Mojica, N., Ramos-Aboites, H., Feldmann, J. & Barona-Gómez, F. Phylogenomic

analysis of natural products biosynthetic gene clusters allows discovery of arseno-organic metabolites in model streptomycetes. *Genome Biol. Evol.* **8**, 1906–1916 (2016).

68. Wang, L. H., Wang, M. S., Zeng, X. A., Xu, X. M. & Brennan, C. S. Membrane and genomic DNA dual-targeting of citrus flavonoid naringenin against: *Staphylococcus aureus*. *Integr. Biol. (United Kingdom)* **9**, 820–829 (2017).
69. Kleijn, L. H. J., Oppedijk, S. F., T Hart, P., Van Harten, R. M., Martin-Visscher, L. A., Kemmink, J., Breukink, E. & Martin, N. I. Total Synthesis of Laspartomycin C and Characterization of Its Antibacterial Mechanism of Action. *J. Med. Chem.* **59**, 3569–3574 (2016).
70. Rui, Z., Petříčková, K., Škanta, F., Pospíšil, S., Yang, Y., Chen, C. Y., Tsai, S. F., Floss, H. G., Petříček, M. & Yu, T. W. Biochemical and genetic insights into asukamycin biosynthesis. *J. Biol. Chem.* **285**, 24915–24924 (2010).
71. Cheng, L., Xlnyu, L. I. A. & Ran, H. Synthetic study on tetrapetalones: Stereoselective cyclization of N-acyliminium ion to construct substituted 1-benzazepines. *Org. Lett.* **11**, 4036–4039 (2009).
72. Luo, Y., Huang, H., Liang, J., Wang, M., Lu, L., Shao, Z., Cobb, R. E. & Zhao, H. Activation and characterization of a cryptic polycyclic tetramate macrolactam biosynthetic gene cluster. *Nat. Commun.* **4**, 1–8 (2013).

## **Chapter 5: Concluding remarks**

This study began by collecting samples from four subterranean limestone caves in New Mexico (Lechuguilla Cave, Spider Cave, Carlsbad Cavern, and Backcountry Cave). I isolated and characterized bacteria from these samples and in order to determine the culturable bacterial diversity from these four caves. Comparative 16S rRNA analysis revealed the dominant group was Actinobacteria, accounting for 45% of all cave isolates. Members of the Actinobacteria including *Streptomyces sp.* were selectively cultured for this study because they are a rich source of secondary metabolic natural products. I used Bacterial Cytological Profiling to identify strains capable of producing metabolites with bioactivity against Gram-negative *E. coli*. Cytological profiling revealed strains within this collection killed *E. coli* by a variety of different mechanisms, the most prominent targeting the cell envelope. The genomes of 9 isolates (4 Actinobacteria and 5 *Bacillus*) were sequenced and mined for biosynthetic gene clusters associate with secondary metabolism. Some of the gene clusters were predicted to produce antibiotics that would produce phenotypes identified by BCP, suggesting that these clusters might be responsible for the observed inhibitory activity. To further identify novel molecules capable of killing Gram negative bacteria, in my second chapter I describe the identification of small synthetic molecules that inhibit growth of *E. coli tolC*. I characterized one molecule in detail and demonstrated that it represents a chemical series that inhibit the *E. coli* thymidylate kinase gene in vivo and in vitro that is required for DNA replication. However I also showed that some members of this family inhibit both DNA replication and cell wall synthesis *in vivo*. This particular family of molecules belongs to a group known as “PAINS” for pan-assay-inhibitory screen, which are notoriously difficult to study because they are chemically reactive, have nonspecific

binding and frequently hit multiple targets within the cell. I was able to show that BCP is a powerful tool for characterizing molecules such as these that have multiple mechanisms of action and are traditionally very difficult to study. Finally, in my third chapter I describe the characterization of *Streptomyces* bacteria and characterization of phage that infect them. These phage might one be used to perform genetic experiments with *Streptomyces* strains in order to improve their ability to produce antibiotics. Taken together, these chapters provide additional insight into the cave microbiome and methods for identifying and characterizing new antibiotics.