UC Berkeley
UC Berkeley Electronic Theses and Dissertations

Title

Identification of Systems Regulating Specialized Metabolism in the Bacterium Streptomyces coelicolor during Interspecies Interactions

Permalink

https://escholarship.org/uc/item/1q6162dn

Author

Bonet, Bailey

Publication Date

2020

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

Identification of Systems Regulating Specialized Metabolism in the Bacterium *Streptomyces coelicolor* during Interspecies Interactions

by

Bailey Bonet

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

In

Microbiology

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Assistant Professor Matthew Traxler, Chair
Associate Professor Michiko Taga
Associate Professor Kathleen Ryan
Professor Emeritus Jeremy Thorner

Fall 2020

Abstract

Identification of Systems Regulating Specialized Metabolism
in the Bacterium *Streptomyces coelicolor* during Interspecies Interactions

by

Bailey Bonet

Doctor of Philosophy in Microbiology

University of California, Berkeley

Assistant Professor Matthew Traxler, Chair

Streptomyces coelicolor is a soil-dwelling prokaryote representative of an order of Gram-positive filamentous bacteria designated Actinomycetales, the richest known bacterial source of bioactive natural products. These specialized / secondary metabolites (including macrolides, aminoglycosides, tetracyclines, cephalosporins and polyketides) often have potent anti-microbial activity and, hence, have been a staple of modern medicine. Genome sequencing has revealed that a given actinomycete genome encodes the potential biosynthetic capacity to produce from 10 to 30, or even more, distinct natural products. *S. coelicolor* is a model organism for studying actinomycete development, genetics, and specialized metabolism. However, in its natural environment, *S. coelicolor* must engage with a heterogenous and complex collection of other microorganisms, as well as with abiotic factors in soil. Therefore, in the research described in this dissertation, I explored the role of interspecies interactions in regulating the production of specialized metabolites by *S. coelicolor*. To obtain insight about how production of these compounds is induced, I exploited the fact that *S. coelicolor* produces two, readily visualized, red-colored antibiotics (actinorhodin and undecyl-prodigiosin), but only when placed in close contact to certain other bacterial genera. I used extensive analysis of gene expression (RNA-seq) during interspecies interactions and, thereby, identified involvement of conservon8 in controlling induction of the red-colored antibiotics. A conservon is an operon so named because it encodes a characteristic set of four proteins and the *S. coelicolor* genome encodes 13 such cassettes that are highly homologous to each other. I constructed single-gene deletions in conservon 8 and determined, using transcriptome analysis, that CvnA8 and CvnF8 seem to function together. On the basis of my data, combined with bioinformatic analysis, I propose that CvnA8 is a membrane-bound signaling protein and CvnF8, a GAF domain-containing protein (an element that typically binds ligands and allosterically regulates proteins), serves as a sensor for CvnA8. From the gene expression patterns of deletion mutants, I deduce that CvnC8 and CvnD8 likely function together and hypothesize that CvnC8 is a transcription factor that is negatively regulated by CvnD8, which resembles a small RAS-like GTPase. Similarly, I observed that a Δ*cvnB8* mutation has an intermediate effect on red pigment production and genome wide expression patterns, suggesting that it may connect CvnA8/F8 and CvnC8/D8, perhaps through regulation of the CvnD8 GTPase. Finally, my expression

data additionally demonstrate that these conservon 8 factors also control expression of the gene clusters for at least two other specialized metabolites aside from the two red pigments, including a cryptic cluster that is predicted to encode a lanthipeptide.

Table of Contents

List of Figures

List of Tables

Acknowledgements

This degree would not be possible without the support and help of so many people to whom I am extremely thankful.

I would like to first acknowledge Dr. Matthew Traxler who welcomed me into his lab and began supporting me from afar, before the Traxler lab was a physical place. My project is a continuation of Dr. Traxler's post-doctoral work. Thus, I feel very grateful for the work he contributed to the initiation of this project and throughout its entirety. I am also thankful for the exceptional writing and presenting skills he has passed on to me. I am undoubtedly a better science communicator because of him. Thank you for the time and effort spent with me to find a path through the vast unknown that was this project.

Secondly, I would like to thank the professors at Berkeley who provided a supportive, collaborative environment for me. Thank you to those who served on my qualifying exam (Dr. Arash Komeili, Dr. Michiko Taga, Dr. Steve Lindow, and Dr. Peter Quail), and those who served on my thesis committee (Dr. Matthew Traxler, Dr. Michiko Taga, Dr. Kathleen Ryan, and Dr. Jeremy Thorner). I am deeply grateful to Dr. Michiko Taga who served on both committees and provided me with crucial experimental advice and wisdom. Thank you for all your patience, time, and understanding. I am immensely thankful to Dr. Kathleen Ryan whose enthusiasm, encouragement, and scientific guidance was essential to these studies. Thank you to Dr. Jeremy Thorner for challenging me, encouraging me and lending his expertise to this project. I am grateful to Dr. Arash Komeili for providing his unending wisdom and inspiration during my time here.

I would like to thank the entire Traxler lab, both past and present, who contributed in so many ways to this project and my scientific trajectory. I am so thankful to have shared a bay with one of the most kind-hearted, genuinely good humans I have ever met, Dr. Vineetha Zacharia. Thank you for answering all my questions, teaching me so much about how to be a great scientist, and making me laugh every day. You helped me grow immensely as a scientist and a person. A special thanks to Neem Patel who helped me see the value in vulnerability, learn to reflect more deeply, and become more confident in myself. Thank you for understanding me and supporting me in the ways that I needed. I would like to thank Dr. Scott Behie for his friendship and all the smiles he brought to the lab. I am so appreciative to Dr. Rita Pessotti for her kindness and encouragement when I needed an extra boost of confidence. Thank you to the immensely talented Yein Ra who helped me with this work and who provided critical scientific input for much of this project. Thank you to Luis Cantu Morin who worked so hard to help with this project. I am glad I am leaving this project in the hands of such an extremely talented scientist. Lastly, I am thankful for the rest of the lab for their support and feedback, Jewel Reaso, Dr. Monika Fischer, Dylan McClung, Bridget Hansen, Mira Liu, Daniel Polyakov, Ryan Nguyen, Zoe Wildman, Hollie Osburn, Victor Chen and others.

I have infinite appreciation to the labs that I received my training in before coming to Berkeley, which gave me such a solid foundation of knowledge and skills to be successful in my graduate studies. Thank you to Dr. Paul Jensen and Dr. Bradley Moore for welcoming me into your groups and providing me with a space to learn so much about ecology, molecular

biology, natural products, and the wonderful world of actinomycetes. One of the most influential individuals on my scientific trajectory was my undergraduate mentor, Dr. Nadine Ziemert. I have endless gratitude to her for seeing potential in me as a young scientist and mentoring me with such dedication and care. I am here because of your unwavering support. Thank you so much.

Thank you to my friends and cohort members who became my greatest support system during graduate school; Cindy, Alyssa, Amanda, Morayma, Siwen, and Tess. Being surrounded by smart, strong, loving women like you sustained and motivated me throughout this adventure. I also would like to acknowledge my friends Taryn, Lindsey, Veronica, Arlene, Shelley, Myngoc, Laura, Noam, Annemarie, Sajuti, Tuesday, Gio, and Edwin. Each one of you impacted me and sustained me during this time. I would like to thank Javier Soto Bustos for supporting me and cheering me on during this endeavor along with helping write an elegant analysis program that made this body of work even better. I am so thankful for your love and encouragement.

Lastly, I would not be the person I am today without the constant love and support of my family. They are my biggest cheerleaders and I feel like this accomplishment belongs just as much to them as it does to. Thank you to my mom, Kathy, and my father, John, for sacrificing so much and working so hard to raise my siblings and me. Without all that you have provided me, I would absolutely not be here today. I feel so privileged to have you as parents. Thank you to Becca and Riley for being such loving and adventurous kids to grow up with. I would like to thank all the rest of my huge, amazing, loving family; Grammie, Tim, Emilee, Nancy, Mark, Sarah, Bryan, Bob, Christina, Monica, Chris, Dennis, Peter, Rick, Joanne, Sinclair, Colin, Terrie, Makena, Tyler, Donna, Greg, Barbara, Gary, Wendy, Aaron, Elli, Kiara, Sue, Brian, Grace, Lila, Michelle, Eric, Melissa, Mike, and Elin. Finally, I would like to thank my family who is no longer here to witness the end of this journey, but who shaped me in very significant ways; Grandpa Jack, Uncle Dennis, and Grandpa Domingo.

1 Chapter 1. An introduction to *Streptomyces* lifestyle, genome, and specialized metabolism

1.1 Life in the Soil as an Actinomycete

The soil is an extraordinarily complex and heterogenous environment from the kilometer scale all the way down to the nanometer scale. As a result of the great variation in biophysical and biochemical characteristics, microorganisms have evolved a myriad of growth and survival strategies in order to live in the soil (Young and Crawford, 2004). Pore structure, tortuosity, sorptive properties, moisture content, organic and inorganic carbon content, nitrogen content, trace metal availability, etc. all influence the habitability of soils and provide unique challenges for life at a small scale.

Some of the most abundant and prolific soil microorganisms are filamentous fungi, which can form massive networks of hyphae covering more than 2000 acres and surviving for more that 2,500 years (Smith et al., 1992; Anderson et al., 2018). Alongside bacteria, filamentous fungi are the main agents responsible for decomposing organic matter and unlocking the stored nutrients from dead organisms. The decomposition of nutrients in the soil plays an extremely crucial role in nutrient cycling and ecosystem function. So, what makes these fungi so stable and well adapted to the soils? Their unique lifestyle which involves filamentous growth, protection from harsh conditions by forming spores, and their ability to produce bioactive secondary metabolites are all factors that allow them to be successful in the soil environment.

Actinomycetes are saprophytes that are found ubiquitously in soils and decaying matter. More recently, they have been recognized as bona fide symbionts which can be found associated with plants, animals, and insects (Seipke et al., 2012). When the first actinomycetes were isolated, they were assumed to be filamentous fungi because of their filamentous growth, and their aerial spore formation that gave them a fuzzy colony morphology, similar to fungi. Though actinomycetes were subsequently proven to be bacteria through cellular morphological observations and the sequencing of their 16S gene, they still share many similar characteristics. *Streptomyces* are a genus of actinomycetes that grow in branching, filamentous networks of hyphae, some of which will branch and grow vertically from a surface to form aerial hyphae followed by differentiation into spore chains (Figure 1-1). The raising of aerial hyphae and formation of spores is a highly regulated process that typically occurs once nutrients become limited. The spores can then be dispersed to start the life cycle all over again in a more hospitable environment.

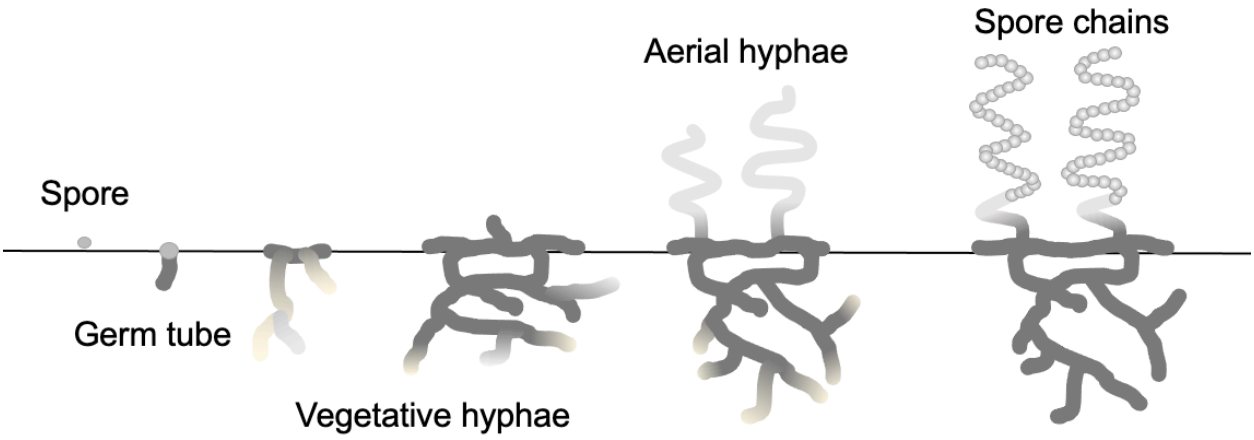


Figure 1-1 | The *Streptomyces* life cycle.
Streptomyces begin their life as spores that germinate when they find favorable conditions. They then grow by apical extension and form branching networks of vegetative hyphae. After depleting nutrients or sensing specific cues, they raise aerial hyphae that break the surface tension and grow vertically from the colony. The aerial hyphae will eventually differentiate into spore chains to begin the life cycle once more.

The similarities between the lifestyles of filamentous fungi and filamentous actinomycetes implies that filamentous growth and spore formation are useful traits for stability in soils. Indeed, this makes sense when considering the previously mentioned heterogeneity of soil environments. With networks of hyphae, filamentous organisms can move toward nutrients using hyphal extension while not sacrificing the space they are already occupying. The pockets of concentrated nutrients can also be shared across an entire network allowing for a community-wide acquisition of resources. Additionally, with a higher surface to volume ratio in filaments compared to rod-shape single cells, there is more space for the cells to contact and interact with their environment. If the environment becomes unfavorable due to lack of nutrients or desiccation, the microorganism can direct its resources toward spore differentiation in order to ensure its progeny's survival (Flärdh and Buttner, 2009). *Streptomyces* are also able to utilize a wide variety of sugars, alcohols, and organic acids (Practical Streptomyces Genetics | NHBS Academic & Professional Books; Dulaney, 1949; Basak and Majumdar, 1973) making them well adapted to the diverse nutrient sources in sediments. We would consider most *Streptomyces* to be generalists as they are capable of living in a diverse range of environments.

1.2 The *Streptomyces coelicolor* Genome

When the *Streptomyces coelicolor* genome was first sequenced using Sanger sequencing in 2002, it was the largest completely sequenced bacterial genome ever published (Bentley et al., 2002). The genome helped to expand our understanding of microbial life in the soil and unlocked new techniques for exploration of actinomycete development, gene regulation, and natural product production. This achievement was momentous in the actinomycete and natural products communities because it illuminated the complexity of the *Streptomyces* lifestyle through the variety of genes observed, the extensive dedication of genomic space to natural product production, the unprecedented amount of regulatory genes, and the structure of the genome itself (Bentley et al., 2002).

The 8,667,507 bp linear *S. coelicolor* genome is organized into a core region and variable arms flanking the core (Bentley et al., 2002). The central core region is well conserved across *Streptomyces* and contains almost all the genes likely to be unconditionally essential such as genes for transcription, translation, DNA replication, and cell division. In contrast, the variable arm regions house genes involved in non-essential functions such as hydrolytic enzymes, secondary metabolite biosynthesis, and gas vesicle proteins. The genome contains 7,825 predicted genes, allowing for a huge amount of coding potential and leading to expansions of known protein families along with identification of new families (Bentley et al., 2002). *S. coelicolor* had expansions in gene families related to transport, secreted proteins, and regulation. 7.8% of genes (614 total) and 10.5% of genes (819 total) are predicted to have a transport function or be secreted, respectively. Enrichment for transporters and secreted proteins underscore the investment made in order to interact with a complex environment, such as the soil.

The *S. coelicolor* genome also shows a clear emphasis on regulation with 12.3% of genes (965 total) predicted to have a regulatory function. It has been shown that larger genomes are typically enriched in regulatory genes and depleted in translation, DNA replication, and cell division proteins (Konstantinidis and Tiedje, 2004). This trend suggests an ecological strategy associated with genome expansion and implies that species with larger genomes might dominate environments where resources are scarce and slow growth does not have a huge negative selection pressure. Indeed, we think soil habitats are exactly this kind of environment.

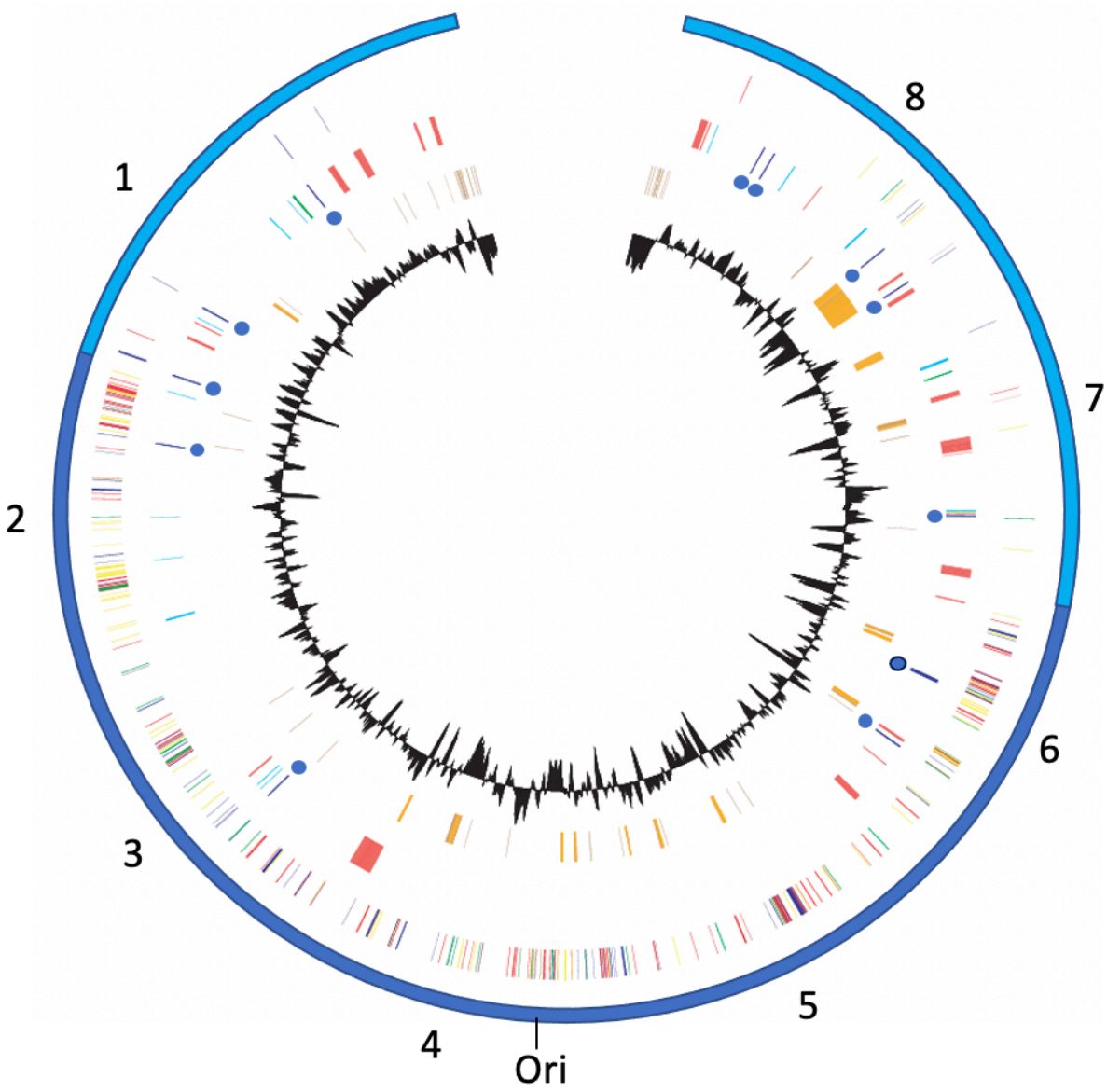


Figure 1-2 | The *Streptomyces* coelicolor genome.
Figure adapted from (Bentley et al., 2002).
The linear genome of *S. coelicolor* is shown, with dark blue indicating the core region and light blue indicating the variable arm regions. The length of the genome in mega bases is marked along the outside in a counterclockwise direction with the origin labeled along the outside as well. The innermost black circle represents GC content. The second ring indicates mobile elements, with brown and orange representing transposases and putative laterally acquired genes, respectively. The third ring from the inside indicates non-essential genes with specialized metabolite genes in red, exoenzymes in pale blue, conservons in dark blue with a dark blue circle adjacent to them, and gas vesicle proteins in

green. The fourth ring, closest to the genome, highlights essential genes for cell division, DNA replication, transcriptions, translation, and amino-acid biosynthesis.

S. coelicolor shows a strong preference for regulatory genes involved in stress response and extracellular sensing. The genome contains 65 sigma factors, 45 of which are ECF (extra-cytoplasmic function) sigma factors which are known to be involved in responding to external stimuli (Hong et al.; Bentley et al., 2002). *S. coelicolor* also has a predicted 100 sensor kinases and 87 response regulators making up a huge collection of two-component system genes (Rodríguez et al., 2013). Two-component systems consist of a kinase that senses an external signal and phosphorylates the response regulator which then can change gene expression to respond to an extracellular stimulus.

Along with sigma factors and two-component systems, a four gene operon, predicted to be involved in regulation, was found repeated throughout the genome 13 times (blue dots in Figure 1-2). This conserved operon was termed a conservon and thought to function as a different kind of regulatory system because of the predicted function of the four core genes. Subsequent studies on conservon 9 in *S. coelicolor*, and its homolog *rarA-E* in *S. griseus* showed that this particular conservon negatively regulates aerial mycelium formation and production of pigmented specialized metabolites in both species (Komatsu et al., 2003, 2006; Takano et al., 2011). Studies on conservon 9 showed that the proteins, CvnA-D9 interact with each other, CvnA9 is an ATPase, and CvnD9 is a GTPase (Komatsu et al., 2003). The authors conclude that the conservon resembles a eukaryotic G protein-coupled regulatory system that is typically involved in outside-in signaling.

One of the most illuminating discoveries that came from the sequenced *Streptomyces coelicolor* genome was the plethora of specialized metabolite biosynthetic genes. Prior to genome sequencing, just four natural products or specialized metabolites were known to be produced by *S. coelicolor*; actinorhodin, undecyl-prodiginine, calcium-dependent antibiotic (CDA), and the grey polyketide spore pigment (Wright and Hopwood, 1976; Hopwood and Merrick, 1977; Chong et al., 1998). After genome sequencing, another 18 natural product biosynthetic gene clusters were identified and today we are able to confidently predict a total of 27 biosynthetic gene clusters. This analysis revolutionized the natural product discovery field and thus began a new era of discovery aimed at waking up these silent, or unexpressed gene clusters.

1.3 Specialized Metabolism in *Streptomyces*

Actinomycetes are by far the richest source of bioactive, bacterial natural products, also called specialized or secondary metabolites. These natural products have been a staple of modern medicine including a subset which are used as frontline therapies for bacterial infections around the world. During the 'golden era' of natural product discovery in the 1940's and 1950's almost all important classes of antibiotics were discovered (macrolides, aminoglycosides,

tetracyclines, cephalosporins). Subsequently, the 1960's and 1970's saw an explosion in natural product discovery with more than 70% of all antibiotics discovered being derived from *Streptomyces*, a genus of actinobacteria. Once the 'gold rush' of the mid 20th century came to an end, scientists had to contend with issues of rediscovery from newly isolated, but closely related *Streptomyces* strains. It was not until the genome sequencing revolution in the late 20th and early 21st centuries that actinomycetes were thrust back into the natural products spotlight. This was due to the realization that each actinomycete genome had the potential to produce 10, 20, 30, or more natural products according. Currently, the field of natural product discovery has invested in discovering bioactive metabolites from silent or cryptic biosynthetic genes (not expressed or not connected to a known compound) to replace our current, increasingly less effective antibiotics.

Specialized metabolites are synthesized in enzyme complexes separate from the ribosome and can require many different core and tailoring enzymes to assemble the chemical structures. The biosynthesis genes of specialized metabolites are found as clusters of operons throughout the genome, making them relatively easy to identify and predict their structures. Despite this, the overwhelming majority of natural product biosynthetic gene clusters (BGCs) that have been discovered from genome sequencing are not expressed under laboratory conditions. While we know a bit about the regulation of these gene clusters by quorum sensing or nutrient availability, there are many more regulation mechanisms to explore (Takano et al., 2001; Sánchez et al., 2010; Du et al., 2011; Tsigkinopoulou et al., 2020).

Biosynthetic gene clusters in *Streptomyces coelicolor* are known to be regulated by a multitude of different nutrients, cellular stresses, and small molecules. Typically each gene cluster has a cluster-situated regulator (CSR) that is specifically regulated by these previously mentioned factors and controls the promoters of the operons of biosynthesis genes in response (Liu et al., 2013). Phosphate, N-Acetylglucosamine, glucose, xylose, nitrogen, zinc, amino acid availability, and more have been shown to regulate the expression of specialized metabolites in *S. coelicolor* (Liu et al., 2013). In addition, *S. coelicolor* produces an autoregulator, gamma-butyrolactone, that, after reaching a certain threshold concentration in the cell, binds to the cluster-situated repressor of coelimycin causing it to dissociate from the promoters (Takano et al., 2001, 2005). The *S. coelicolor* actinorhodin cluster-situated regulator, *actII*-ORF4, is a direct target of at least eight known regulatory proteins, highlighting the complexity of specialized metabolite regulation. Overall, there are a plethora of cues that feed into how each specialized metabolite biosynthetic gene cluster is regulated, and we have likely just begun to understand them.

Thinking about the ecology of natural product production, it is believed that these compounds confer a selective advantage to the producer, otherwise they would have been lost over evolutionary time. One prominent idea is that natural products are used to compete with other organisms for resources (Ryan and Dow, 2008; Cornforth and Foster, 2015) whereas other recent work has led to the hypothesis that secondary metabolites could be highly specialized signaling molecules that mediate interspecies interactions (Davies et al., 2006; Linares et al., 2006; Abrudan et al., 2015). The discrepancies of these findings emphasize that there remains

a fundamental gap in our understanding of the ecological role of natural products. Indeed, it could be that natural products function in both ways depending on the given situation. Recently, with the natural ecology of these soil dwelling bacteria in mind, researchers have begun looking towards co-culturing *Streptomyces* to elicit the production of cryptic natural products. Though there are many examples of the production of bioactive compounds in response to an interaction with another organism, little is known about how the compounds are induced (Straight et al., 2006; Onaka et al., 2011; Abrudan et al., 2015; Pishchany et al., 2018; Lee et al., 2020). The cues and molecular mechanisms used to sense and respond to interactions remain unidentified in actinomycetes making this a promising field for exploration.

In this work, I take an untargeted approach to explore how *Streptomyces coelicolor* senses and responds to interspecies actions. In chapter 2, I used transcriptomics to identify a conservon system that is necessary for induction of specialized metabolism during interspecies interactions and explore how the components of this system are influencing gene expression. Next, in chapter 3, I dive deep into the conservon components' protein sequences and predicted structures to hypothesize how these components are functioning. I also explore the biochemical activity of the CvnD8 protein. Though the conservon system identified is highly engaging, the complexity of the interaction environment led us to hypothesize that there are more factors involved in sensing and responding to interspecies interactions. Chapter 4 details a genetic screen performed to find additional genes involved in regulating specialized metabolism during interspecies interactions. Lastly, chapter 5 is a review paper that examined natural products and their functions *in situ*. It makes an argument for studying specialized metabolite production in natural environments in order to understand the ecological roles of these compounds. Altogether, this work makes significant progress towards identifying and characterizing systems that function in *S. coelicolor* to sense and respond to interspecies interactions.

2 Chapter 2. A conservon system regulates specialized metabolism in *S. coelicolor* during interspecies interactions

Bailey Bonet, Yein Ra, Luis Cantu Morin, Jonathan Livny, and Matthew F. Traxler.
Manuscript in preparation.

2.1 Abstract

Streptomyces are well-studied for their propensity to produce bioactive molecules. Recently, interspecies interaction assays have illuminated cryptic or silent biosynthetic gene clusters, which proved to be a promising avenue for novel drug discovery. Here we used transcriptomics to identify and study a conservon system that is upregulated and necessary for normal actinorhodin and undecyl-prodigiosin production during interactions with other actinomycetes. We determined that the *cvnA8* and *cvnF8* (SCO6939) genes function to positively regulate the biosynthesis genes of actinorhodin, and negatively regulate the coelimycin genes, likely through the same pathway during interactions. We observed that the *cvnC8* and *cvnD8* genes negatively regulate coelimycin and prodiginine biosynthesis genes and have similar overall effects on gene expression, implying they function together or in the same part of a regulatory pathway. Additionally, the CvnD8 protein is the first small bacterial GTPase shown to regulate gene expression. We found that all five conservon 8 genes strongly negatively regulate a cryptic lanthipeptide biosynthetic gene cluster, indicating that this could be the direct target of the conservon system. Together, we show that a conservon system is a broad regulator of specialized metabolism in *S. coelicolor*. Furthermore, the conservons form regulatory systems connected to many, sometimes cryptic, natural product biosynthesis genes, which opens avenues for discovery of new natural products from *Streptomyces*.

2.2 Introduction

Actinomycete bacteria produce an amazing array of natural products or specialized metabolites from which many clinically important antibiotics have been derived. The genus *Streptomyces* alone accounts for two-thirds of all naturally derived antibiotics. However, as they have been well studied for their applications in medicine, little is known about how these metabolites function in natural settings (Bérdy, 2005). It is thought that specialized metabolites confer an evolutionary advantage and might be used as signaling molecules or warfare agents during interspecies interactions. A growing body of work recognizes that natural product biosynthesis is often connect to social interactions, either intra- or inter-species (Abrudan et al., 2015).

There are many examples of interspecies interactions stimulating natural product biosynthesis of both known and novel metabolites in actinomycetes. In one such example, researchers showed that a coculture of *Myxococcus xanthus* and *Streptomyces coelicolor* stimulated the production of the antibiotic actinorhodin in the latter strain as a result of iron competition between the two organisms (Lee et al., 2020). Another study demonstrated that an interaction

between *Amycolatopsis* sp. AA4 and *S. coelicolor* stimulated the production of the production of the novel antibiotic, amycomycin, by *Amycolatopsis* sp. AA4 (Pishchany et al., 2018). All evidence indicates that determining the molecular mechanisms underlying the induction of secondary metabolism genes during interactions is a promising avenue to establish new approaches to finding novel bioactive secondary metabolites and understanding their function in microbial communities.

The conservons are an enigmatic family of operons found mostly in actinobacteria. The typical streptomycete has eight or more, while mycobacteria have at least one (Komatsu et al., 2003). Conservons were first identified when the first actinomycete genome was fully sequenced, that of *Streptomyces coelicolor* (Bentley et al., 2002). The researchers noticed multiple copies of an operon throughout the genome and termed the conserved operon a 'conservon'. The conservon consists of four genes encoding for a predicted membrane protein with a histidine kinase-like ATPase domain (*cvnA*), a protein with a roadblock/LC7 domain (*cvnB*), a protein with a conserved domain of unknown function (DUF 742) (*cvnC*), and a RAS-like GTPase (*cvnD*). Previous work has suggested that these operons are signaling systems similar to eukaryotic G protein-coupled receptors (Komatsu et al., 2006).

In this paper we examine interspecies interactions that stimulate production of two red pigmented antibiotics, actinorhodin and undecyl-prodigiosin, in the model streptomycete, *Streptomyces coelicolor*. Previous studies have identified that *S. coelicolor* interactions could stimulate the precocious synthesis of both of these molecules (Traxler et al., 2013). Here we set out to investigate the *S. coelicolor* molecular response to interactions by identifying genes involved in sensing or responding to the interactions and investigating how they are contributing to gene expression and pigmentation. We used transcriptomic analysis, bioinformatics, and pigmentation distribution analysis to determine that typical induction of genes involved in red pigmented antibiotic biosynthesis required the conservon 8 system. We show that the conservon 8 locus influences the expression of at least four biosynthetic gene clusters, thus it is a more global regulator of specialized metabolism during interactions. We found that *cvnA8* and *cvnF8* likely function together and that *cvnC8* and *cvnD8* likely function together as well. All together the results point to a model in which this conservon system is a broad regulator of specialized metabolism during interactions.

2.3 Methods

2.3.1 Bacterial Strains and Growth Conditions

All strains are listed in Table 2-2. Strains of *E. coli* were grown at 37°C in LB media with the appropriate antibiotics, if necessary. Strains of *Streptomyces coelicolor* were grown on ISP2 agar (10g malt extract, 4g glucose, 4g yeast extract, 15g agar, 1L ddH2O (distilled deionized water)). Streptomyces spores were generated and harvested using typical methods (Practical Streptomyces Genetics | NHBS Academic & Professional Books).

2.3.2 Plating interactions

S. coelicolor spore stocks were diluted to $2x10^8$ CFU/mL and 0.5ul was spotted onto a 60x15mm petri dish with 4ml ISP2 agar, resulting in agar approximately 2 mm thick. For interactions, 0.5 ul of an *Amycolatopsis* sp. AA4 stock ($8x10^8$ CFU/mL) was spotted 0.75 cm away from the *S. coelicolor* spot. Interactions were grown for 4 days at 30°C, then imaged or harvested for RNA isolation.

2.3.3 Generation of Deletion Mutants

Deletion mutants were generated in *S. coelicolor* M145 using the ReDirect PCR targeting system and the cosmid St1G8 (Gust et al., 2003). Deletions were verified using PCR spanning the deletion and sequencing of the PCR product. They were also confirmed with a PCR within the deleted gene, resulting in no PCR product. All primers used are listed in Table 2-1.

2.3.4 RNA Isolation

Biomass was collected after 4 days of growth at 30°C using a cell scraper. For the colonies growing alone the entire colony was collected and one sample consists of three whole colonies combined. For the colonies in interactions, only half of the colony closest to the initiator strain was collected and one sample consists of five half colonies. Precaution was taken to not scrape up any of the interacting strain.

Immediately after collection, the biomass was frozen in liquid nitrogen and ground up using a mortar and pestle. 200ul of trizol was added to the ground cells. The Zymogen Direct-zol RNA miniprep kit (R2051) was used to purify the RNA, with the optional DNase treatment step included. The purity of the RNA was verified using a NanoDrop One UV-Vis Spectrophotometer to measure the 260/280 and 260/230 absorbance ratios. The RNA was quantified on a Qubit fluorometer using an Invitrogen Qubit RNA-HS Assay Kit (Q32852).

2.3.5 RNA Sequencing and Analysis

RNA preparation was done by the UC Berkeley Functional Genomics Laboratory. Briefly, rRNA was removed using the Illumina Ribo-Zero rRNA removal kit (bacteria) and the remaining RNA was prepared with a Wafergen PrepX RNA library prep kit for Illumina. Indexing PCR was performed using in-house dual indexed primers and 14 cycles. Samples were sequenced on an Illumina HiSeq 4000 using 100bp paired end reads or an Illumina NovaSeq 600 using 150bp paired end reads.

Reads were cleaned and trimmed using trimmomatic version 0.33.0, mapped to the genome using bowtie2 version 2.2.4.1, counted using HTSeq version 0.11.3, and differential expression was tested using DESeq2 version 1.26.0 on R version 3.6.3. Each strain and environmental condition had three biological replicates.

2.3.6 Pigmentation detection and quantification

2.3.6.1 Pixel Color Distance Calculation (PCDC):
The PCDC method is used to determine how similar a pixel color is to a given target color.

First, we converted all the pixels in an image from RGB to CIELAB. We used the CIELAB color space because it represents our visual perception of colors more accurately than RGB. The CIELAB model expresses color as three values: L* for the lightness from black to white, a* from green to red, and b* from blue to yellow. Here we use the subscript λ to represent L*, α to represent a* and β to represent b*.
Given two colors in the CIELAB color space, color X (X_λ ,X_α, X_β) and a second color Y (Y_λ ,Y_α, Y_β), we calculate the distance (d) between them using equation 1 and determine how similar two colors are with 0 being the exact same color, and 1 being most different.

Equation 1: $$d = \sqrt{(X_\lambda - Y_\lambda)^2 + (X_\alpha - Y_\alpha)^2 + (X_\beta - Y_\beta)^2}$$

Since we would like the distances to be evenly distributed in the [0, 1] range, we normalized the distances for each pixel to the maximum distance d_{max} from the target color c. Given a target color *c,* when *c* is on the surface of a 3D sphere with diameter 1, that color has a maximum distance (d_{max}) of 1. If *c* is not on the surface of the sphere, but instead within the sphere, the maximum distance (d_{max}) from any given color to *c* will be less than 1. Given this, we can normalize the value d from equation 1 using the following steps:

- Calculate d_{max} of target color *c*:
  d_{max} is the distance from *c* to the center of the sphere (*L: 50, a: 0, b: 0*), plus the radius of the sphere (0.5)
- Calculate distance *d* between the color in a given pixel and the target color *c*.
- Normalize *d* to get d_s in the range [0, 1] using equation 2:

Equation 2: $$d_s = d / d_{max}$$

2.3.6.2 Colony masking:
In images of interactions, *Amycolatopsis* sp. AA4 colonies were carefully cropped out using Photoshop, ensuring only *S. coelicolor* colonies were used in the pigmentation detection workflow.

Using the PCDC method described above, we first averaged ten pixels from the background (agar) to determine that our target background color was #33393E. We then defined any pixel with a normalized distance of 0.15 or less from #33393E to be part of the background. With this, we convert the image to a binary mask, where every pixel is either black (part of the background) or white (part of the foreground).

To correct for any noisy white pixels that arise from small bubbles or imperfections in the agar, we counted the number of adjacent white pixels in the mask and if the count was less

than 2% of the total pixels in the image, we considered them noise and turned them black. Similarly, there were also missing black pixels that were within the colony (not part of the background) that had a d_s less than or equal to 0.15, which were removed from the mask. We used a similar counting method to identify these missing black pixels and if they were less than 20% of the total pixels in the image, we considered them part of the colony (not part of the background) and turned them white.

This method resulted in a black and white mask indicating where the *S. coelicolor* colony is (white pixels) and the background is (black pixels) in the image.

2.3.6.3 Pigmentation calculation:
We used the PCDC method to calculate the distance of each pixel within the mask to the target pigmentation color #803D33 that corresponds to red pigmentation. The target pigmentation color was calculated by averaging 20 pixels from highly pigmented colonies. We defined the pigmentation (p) using the distance to the pigmentation target color (d_s) with equation 3:

Equation 3: $p = 1 - d_s$

2.3.6.4 Baseline Normalization
We selected a region of interest (ROI) within the mask that is a rectangle centered vertically, with a height of 20% of the maximum colony height.

We wanted to differentiate red pigmentation from other components of the pixel color, such as brightness and glare, that otherwise would contribute to our pigmentation calculation. To do this, we first calculated the pigmentation in the ROI across twelve wild-type *S. coelicolor* single colonies which are not pigmented. We grouped the columns of pixels in the ROI into breakpoints of 200 per image across the horizontal axis. We then averaged these each of the 200 positional p values across all twelve images, resulting in an average pigmentation distribution of wild-type *S. coelicolor* single colonies. Wild-type single colonies were chosen because they have no visible pigmentation. This data set of spatial averages was named Pb, and Pb_{min} (0.342) was the minimum value in Pb distribution.

We then went back to our image of interest and adjusted the pigmentation value p in each pixel to give us evenly distributed values in the [0, 1] range, setting Pb_{min} to 0 . Given a pigmentation value p we define the effective pixel pigmentation P_e as:

$Pb_{min} = 0.342$

if $P < Pb_{min}$:
$P_e = 0$

if $P >= Pb_{min}$:
$P_e = (P - Pb_{min}) / (1 - Pb_{min})$

Edges were often shadowed because of the 3-dimensional aspect of the colony, and this created noise towards the edges of all the colonies. In order to account for these edge effects, we used the Pb dataset to apply a spatially specific pigmentation correction. Given the calculated P_e of each pixel, we determined which breakpoint the given pixel falls in (i) and then subtracted the difference between the baseline pigmentation of that breakpoint Pb_i and Pb_{min} to give the final pigmentation value, P_f, using equation 4:

Equation 4: $$P_f = P_e - (Pb_i - Pb_{min}) \qquad i = [1,200]$$

The P_f values for the pixels in a given breakpoint were then averaged and the standard deviation was determined.

Gene	Primer	Sequence
Primers for generating the PCR targeting cassette		
cvn8 targeting cassette	FW	CTCGCTTAGCCACCCACCCCCCTGACGAGGATTCCCATGATTCCGGGGATCCGTCGACC
	RV	GCGGCGTGCGGTAGGGGTCGCTGGGGTGCTGTGACATCATGTAGGCTGGAGCTGCTTC
cvnA8 targeting cassette	FW	CTCGCTTAGCCACCCACCCCCCTGACGAGGATTCCCATGATTCCGGGGATCCGTCGACC
	RV	GCCAGTGGATGTGCCGGGCCACCGGGAACGGACGGGTCATGTAGGCTGGAGCTGCTTC
cvnB8 targeting cassette	FW	CATCCACTGGCCCTTCCCCCTTCTGGAGCGATCCCTATGATTCCGGGGATCCGTCGACC
	RV	ACGCCGACTCCTCGGCCGCGCGGCCGTACGTGCTCACCGTGTAGGCTGGAGCTGCTTC
cvnC8 targeting cassette	FW	CCCCGTCCGCGTTGGCACGGGCGGCGAGGTCCGGTGAGCATTCCGGGGATCCGTCGACC
	RV	GCACCGGCCCGCGGGTGCGGCGAAACCGTGGACGGGTCATGTAGGCTGGAGCTGCTTC
cvnD8 targeting cassette	FW	CACCCGCGGGCCGGTGCGCCGACCGTGCTGAAGATCATGATTCCGGGGATCCGTCGACC
	RV	GCGGCGTGCGGTAGGGGTCGCTGGGGTGCTGTGACATCATGTAGGCTGGAGCTGCTTC
cvnF8 targeting cassette	FW	TGGTCACACCCCTTCTACCACCCTCCAGGACGCCTGATGATTCCGGGGATCCGTCGACC
	RV	ATCGCCTACGGCTCCCGGGTGGGGGTGTGGAACGGCTTATGTAGGCTGGAGCTGCTTC
Primers to verify gene deletion		
cvn8	FW	CCGACGACCTGGCTAGTGCTTC
	RV	GGTGATCACCCACTCGTTGCC
cvnA8	FW	CACATGGGCCTGATCCAGAG
	RV	GTGGGTACTCATAGGGATCGC
	RV2	GGATCAGCGACATGATGGCC
cvnB8	FW	CACATCCACTGGCCCTTCCC
	RV	GAGACGAGAACTTTGACGGCG
	RV2	ACGGGGATCACCATGTACGG
cvnC8	FW	CCTGTTCTTCGTCCAGAGCG
	RV	ATGTCCTCGAAAAAGCTGACGG
	RV2	GAGACGAGAACTTTGACGGCG
cvnD8	FW	AAGAGGAAGTGGCCTGATGC
	RV	AGGTCGTCGAAGTCCTCGTT
	RV2	ATGTCCTCGAAAAAGCTGACGG
cvnF8	FW	CTTGGAGAGCAGCTTCACGCC
	RV	GTCTCCCTCTCCCTGGATGAAG
	RV2	AGGTCGTCGAAGTCCTCGTT

Table 2-1 | List of primers used in this study.

Strain	Relevant Genotype	Use	Reference
E. coli ET12567/pUZ8002	*dam-*, *tra* genes	Streptomyces conjugation	MacNeil et al., 1992
E. coli/St1G8	S. coelicolor gDNA: 7674874-7712565 bp	gDNA cosmid with cvn8	Redenbach, et al., 1996
BW25113/plJ790	λ RED genes	Recombining the targeting cassette	Datsenko and Wanner, 2000
DH5α/plJ773	*aac(3)IV*, *oriT* (RK2), FRT sites	Apramycin cassette amplification	Gust, et al., 2006
S. coelicolor M145	Wild-type, SCP1-, SCP2-		Kieser et al., 2000
M145 *ΔcvnA8*	*cvnA8*::*acc(3)IV*		This study
M145 *ΔcvnB8*	*cvnB8*::*acc(3)IV*		This study
M145 *ΔcvnC8*	*cvnC8*::*acc(3)IV*		This study
M145 *ΔcvnD8*	*cvnD8*::*acc(3)IV*		This study
M145 *ΔcvnF8*	*cvnF8*::*acc(3)IV*		This study
M145 *Δcvn8*	*cvn8*::*acc(3)IV*		This study

Table 2-2 | List of bacterial strains used in this study.

2.4 Results

2. 4. 1 Conservon loci are upregulated and during interspecies interactions

Bacterial colony interactions have been previously shown to stimulate diversification of the *S. coelicolor* secreted metabolome (Traxler et al., 2013). We investigated how the transcriptome of *S. coelicolor* is changing during interactions compared to when it is grown as a single colony to identify *S. coelicolor* genes that may be involved in colony interactions, including the production of a diversified set of metabolites. We chose to look at genes differentially expressed in multiple interactions as this would lead us to genes more generally involved in sensing and responding to interactions. To set up the interactions, *S. coelicolor* was plated 0.5 cm away from the interacting strain on an ISP2 agar plate and grown for 4 days. We chose to use *Streptomyces* sp. SPB74, *Amycolatopsis* sp. AA4, *Streptomyces viridochromogenes*, and *Streptomyces albus* J1074 because the secreted metabolome has been explored in these interactions. Additionally, these strains stimulated a range of production of red pigmented metabolites in *S. coelicolor* (Figure 2-1).

Actinorhodin is a redox active, benzoisochromanequinone dimer polyketide antibiotic that is red below pH 8.5 or blue above that pH. In our studies, the pH is below 8.5, ensuring actinorhodin is in its protonated, red form. Actinorhodin has potent bacteriostatic activity against gram negative bacteria, including *B. subtilis*, *M. luteus*, *L. monocytogenes*, and *S. aureus* (Mak and Nodwell*, 2017).* Undecyl-prodigiosin is a linear tripyrrole molecule that is a member of the prodiginine family of molecules (Cerdeño et al., 2001). This family consist of red pigmented bioactive molecules that have a range of activities including antibiotic, antifungal, anticancer, and antimalarial (Stankovic et al., 2014). We used the induction of these two

pigmentated molecules as a proxy for the induction of specialized metabolism in general when studying how interactions were sensed and responded to by *S. coelicolor*.

After growing *S. coelicolor* in an interaction for 4 days, we isolated RNA from the half of the *S. coelicolor* colony closest to the interacting strain and from colonies of *S. coelicolor* growing alone. We compared the transcriptomes of *S. coelicolor* grown in each of the interactions to the control of *S. coelicolor* growing alone. When we investigated the genes that are differentially expressed in all four interactions, we saw that there were 39 genes shared between all four interactions and 93 genes shared between three or more interactions (Figure 2-2). We considered genes to be differentially expressed if they had a $\log_2$ fold change greater or less than two standard deviations away from the average $\log_2$ fold change of the entire set of genes and had a p-value of <0.05. Within the 39 genes that were differentially expressed in all four interactions, there were many natural product biosynthesis genes, genes predicted to be secreted or membrane bound, and a couple possible DNA binding genes (Figure 2-2). Many of the genes that we found to be mostly highly differentially expressed in the interactions were part of conserved operons, or conservons, with the most highly expressed conservon gene being more than 7.7 $\log_2$ fold (or 200-fold) upregulated during the interaction (Figure 2-1). These results indicate that conservons might be involved in *S. coelicolor's* process of responding to interactions and therefore warranted further study.

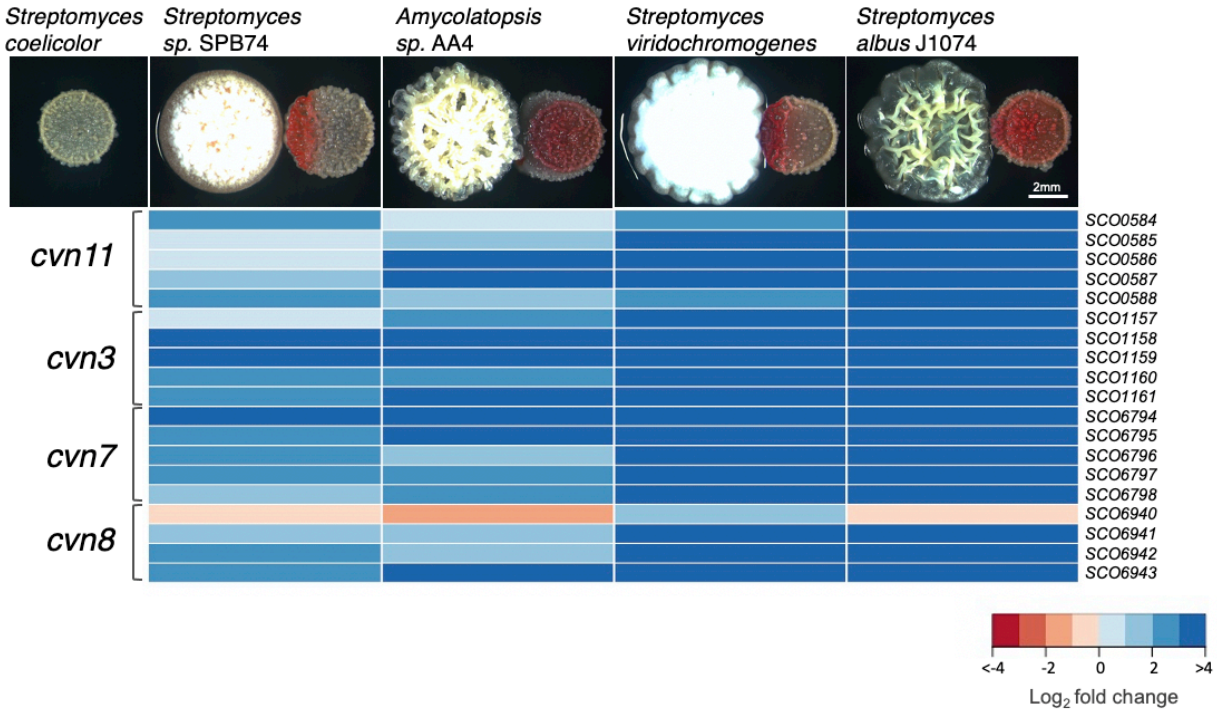
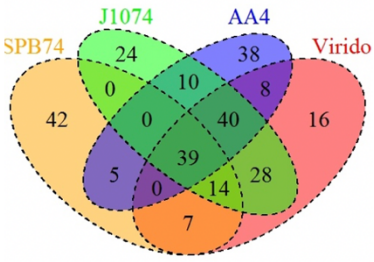


Figure 2-1 | Conservon loci are upregulated in *S. coelicolor* during interspecies interactions.
Phenotypes of *Streptomyces coelicolor* during growth as a single colony (top left) and during growth in interspecies interactions (right colony). Below each

interaction image is a heatmap of the $\log_2$ expression ratio of four conservon loci in *S. coelicolor* during that interspecies interaction compared to *S. coelicolor* grown as a single colony.



A

Localization or Function	Number of Genes
Secreted	8
Membrane associated	8
DNA binding	2
NP biosynthesis	10
Unknown	15

B

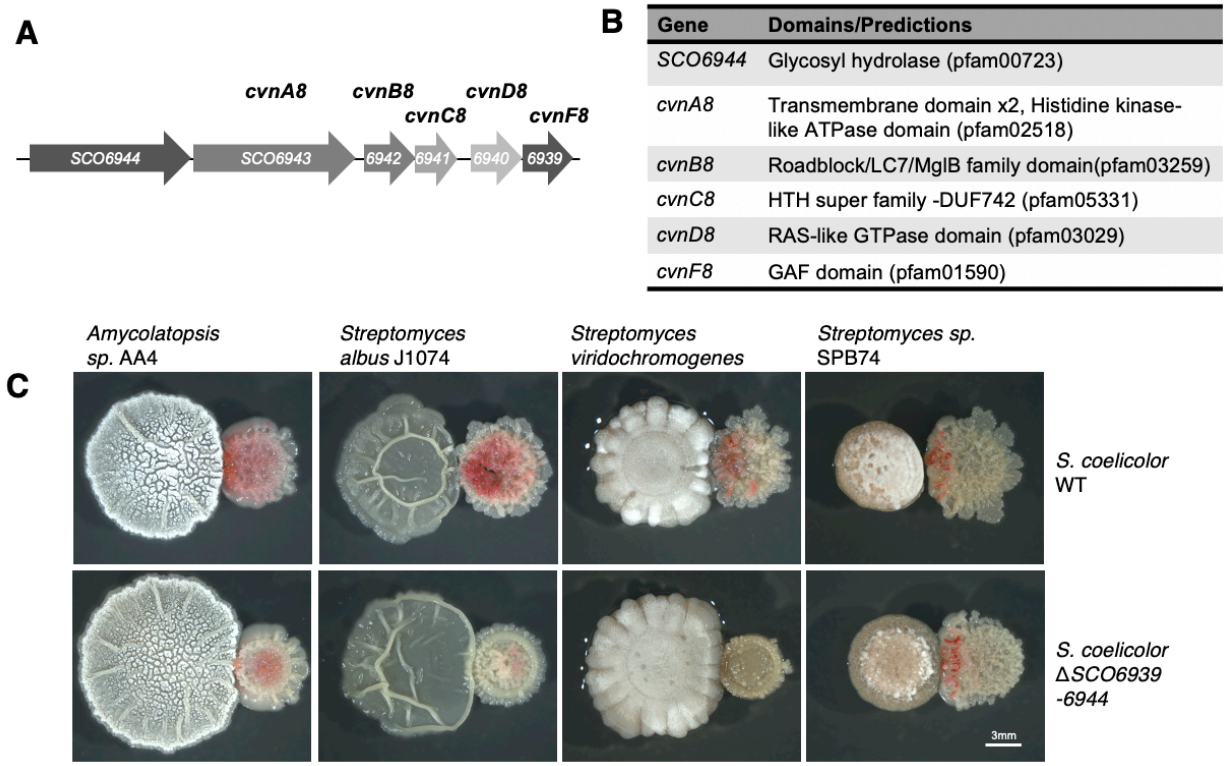
Genes of Interest Significantly Upregulated in all Four Interactions
ChpE
Seven CDA biosynthesis genes
Two Coelimycin biosynthesis genes
Two Conservons

C

Figure 2-2 | Genes differentially expression in *S. coelicolor* during interspecies interactions.
Four-way Venn diagram showing *S. coelicolor* genes significantly upregulated in the interactions (A), the predicted localization or function of the genes significantly upregulated in all four interactions (B), and some genes of note that were upregulated in all four interactions (C).

2. 4. 2 Conservon 8 is necessary for the timely induction of pigmented metabolites during interactions

In order to investigate the conservons further, we made deletions of the entire conservon clusters starting with conservon 7, which we saw no obvious change in the phenotype during interactions (Figure 2-4). When we made a deletion of conservon 8 we saw an obvious delay and reduction in pigmentation during three of the four interspecies interactions (Figure 2-3C). This means conservon 8 must be involved in timely induction of pigmentation during interspecies interactions. Conservon 8 is a typical conservon with the canonical four genes *cvnA,B,C,D* that are found within all the conservons (Figure 2-3B). We further characterized conservon 8 by analyzing the genomic structure and predicted protein domains for each gene.



Gene	Domains/Predictions
SCO6944	Glycosyl hydrolase (pfam00723)
cvnA8	Transmembrane domain x2, Histidine kinase-like ATPase domain (pfam02518)
cvnB8	Roadblock/LC7/MglB family domain(pfam03259)
cvnC8	HTH super family -DUF742 (pfam05331)
cvnD8	RAS-like GTPase domain (pfam03029)
cvnF8	GAF domain (pfam01590)

Figure 2-3 | Conservon 8 is involved in the induction of pigmented metabolites during interactions.
The conservon 8 genomic locus (A) and the predicted domains within each gene (B). Interspecies interactions of wild-type (WT) and Δ*cvn8* (ΔSCO6939-6944) *S. coelicolor* colonies when grown in interspecies interactions with four interacting strains (C). Micrographs were taken after 3 days of growth with *Amycolatopsis* sp. AA4 , after 4 days of growth with *S. albus* J1074, and after 5 days of growth with *S. viridochromogenes* and *Streptomyces* sp. SPB74.

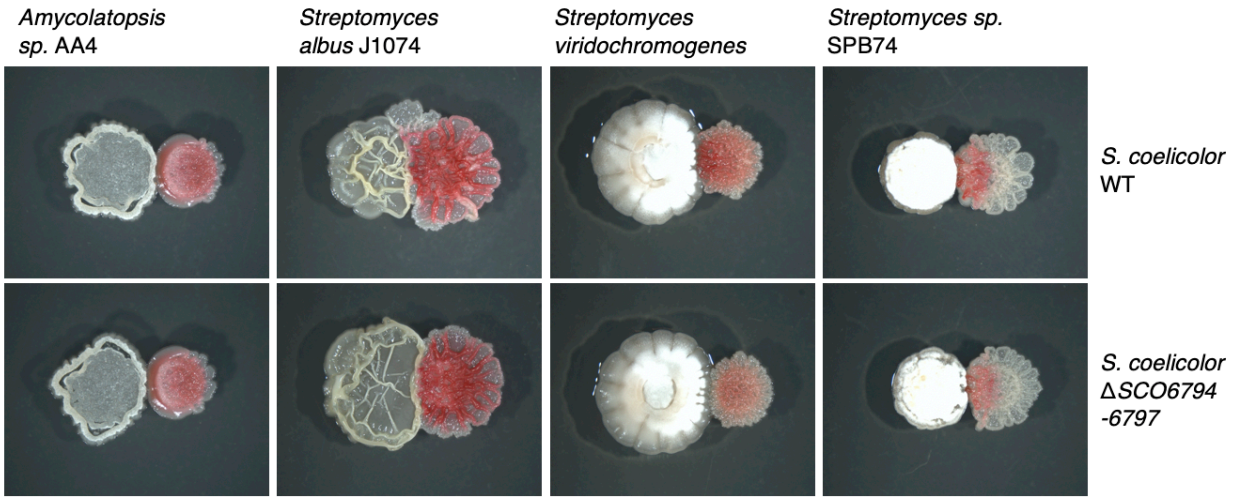


Figure 2-4 | *Δcvn7* interaction phenotypes.
Interspecies interactions of wild-type (WT) and Δ*cvn7* (Δ*SCO6794-6797*) *S. coelicolor* colonies when grown in interspecies interactions with four interacting strains. Micrographs were taken after 5 days of growth for every interaction shown.

We named the gene downstream of the core conservon 8 genes *cvnF* because gene expression evidence indicated that it functions with this conservon 8 system. The *cvnF* gene encodes a protein with a single predicted GAF domain. GAF domains are typically involved in sensing small molecules, with the most well studied GAF domains found in cyclic nucleotide phosphodiesterase (Zoraghi et al., 2004). 28bps upstream of the *cvnA* gene is a predicted glycosyl hydrolase, which we are unsure if it is involved in this system.

The CvnA8 protein is the second shortest CvnA protein out of the 13 total CvnA predicted proteins in the *S. coelicolor* genome. One reason for this is that it contains no predicted sensing domain. It has two transmembrane helixes (between residues 33-80) with eight predicted amino acids between them, indicating that CvnA8 does not have an extracellular sensor domain. Some of the other CvnA proteins in *S. coelicolor* have predicted sensing domains; for example, the CvnA1 and CvnA3 proteins both have an NIT superfamily, nitrate/nitrite sensing domain. Due to the lack of sensing a domain in the CvnA8 protein and the presence of the GAF domain in the CvnF8 protein, we wondered if the CvnF8 protein might function with the CvnA8 protein to sense a small molecule and regulate this system. The CvnA8 protein has a predicted histidine kinase-like ATPase domain (pfam02518) which includes the key residues present in the catalytic and ATP-binding (CA) domain, indicating that it is likely a functional ATPase. Previous studies have shown that the CvnA9 protein has ATPase activity (Komatsu et al., 2006).

The CvnB8 protein has a predicted roadblock/LC7 domain which is known to modulate different functions of NTPases (Koonin and Aravind, 2000). Interestingly, the predicted structure of the CvnB8 protein is very similar to the MglB protein crystal structure according to the protein structure prediction software, Phyre2. The MglB protein is a well characterized GTPase activating protein (GAP) and has recently been found to have guanine nucleotide exchange factor (GEF) activity as well (Baranwal et al., 2019). It exercises these activities to regulate the small-GTPase MglA in *Myxococcus xanthus* and help modulate cell polarity switches during motility. Bioinformatically, predicting the function of CvnB8 is challenging, but it likely plays a role in regulating a small-GTPase, possibly the CvnD8 protein.

The CvnD proteins are the most well conserved of all the conservon proteins in *S. coelicolor* with 54.1% pairwise identity between all 13 CvnD proteins in the genome. The CvnD8 protein is predicted to be a RAS-like GTPase. RAS proteins are small GTPases, typically found in eukaryotic systems, that participate in signal relay and regulation of critical cellular processes (Colicelli, 2004). The most well-studied bacterial RAS-like GTPase is the MglA protein in *M. xanthus* with which the CvnD8 protein shares 22% pairwise identity. The predicted catalytic and GTP binding domain of the CvnD8 protein has the p-loop, switch I, switch II, and other important residues for proper function, indicating that the CvnD8 protein is a functional GTPase.

It is somewhat challenging to predict the precise function of the CvnC8 protein. It contains a domain of unknown function (DUF742) that is part of the helix-turn-helix (HTH) superfamily. The specific HTH motif is typical for molecules that bind nucleic acids. Based on this, we hypothesize that CvnC8 binds DNA and possibly acts as a transcriptional regulator.

The observed phenotype of the conservon 8 deletion strain indicates that conservon 8 is necessary for the timely induction of actinorhodin and undecyl-prodigiosin production in *S. coelicolor* during interactions. Further analysis of predicted gene functions along with previous studies suggest that the conservons act as a sensing system. Therefore we hypothesize that genes in conservon 8 function to sense the interacting bacterium and regulate specialized metabolite production.

2. 4. 3 The genes in conservon 8 contribute differentially to the distribution of pigment production across interacting *S. coelicolor* colonies

After determining that conservon 8 is necessary for pigmented specialized metabolite production in *S. coelicolor* during interspecies interactions, we wanted to examine the role that each individual gene plays in this process. We constructed single gene deletions using a method involving the PCR targeting of a cosmid containing *S. coelicolor* genomic DNA, making sure to not disrupt the adjacent genes (Gust et al., 2003). We hypothesized that we would see a range of phenotypes in the single deletion strains indicating that the single genes contribute in different ways. We placed the *S. coelicolor* single gene deletion strains in interactions with

Amycolatopsis sp. AA4, spotting the spore stocks 0.75 cm apart and growing them for 4 days at 30°C before observing the phenotype (Figure 2-5D).

We developed in-house software to detect pigmentation in images of bacterial colonies to quantify the pigmentation intensity and distribution across *S. coelicolor* colonies in a robust and unbiased manner. Briefly, we calculated the distance of every pixel color in a given image to a target pigmentation color that was calculated from pixels selected in highly pigmented colony images. The distance value of each pixel was then converted to pigmentation intensities (Figure 2-5A). We normalized these values, set a baseline threshold for pigmentation, selected a region of interest across the diameter of the colony, and graphed the average pigmentation intensity across the colony diameter (Figure 2-5B). We performed this calculation for the *S. coelicolor* wild-type strain as a single colony and in an interaction, along with all the single gene deletion mutants (*⊿cvnA-F8*) in interactions. We had 20 replicate interactions for each strain we studied. We averaged the pigmentation intensity across all 20 of the replicate colonies, and chose the single replicate that best matched the average distribution as a representative image (Figure 2-5C). We determined that a pigmentation value of 0.5 or greater is consistent with intense pigmentation observed in the interacting wild-type strain. This 0.5 pigmentation threshold value is indicated as a dashed line on all graphs in Figure 2-5D.

We observed that the wild-type *S. coelicolor* strain, as a single colony, had almost no pigmentation. In contrast, the interacting wild-type strain displayed intense pigmentation in the first quarter of the colony, followed by a lighter pigmentation across the entire colony, decreasing in pigmentation at the right side of the colony. This is expected as pigmentation is induced in response to the interacting strain and the first quarter of the colony is closest to the interacting strain. We also observed a similar pattern in the *⊿cvnC8* and *⊿cvnD8* strains during interactions, with the most intense pigmentation in the first quarter of the colony followed by lighter pigmentation across the rest of the colony. Wild type differed from the *⊿cvnC8* and *⊿cvnD8* mutants in the first quarter peak, which was slightly wider in the deletion mutants than the peak in the wild-type strain. We wondered whether mutants *⊿cvnC8* and *⊿cvnD8* did in fact have more pigmentation there.

Interestingly, the *⊿cvnA8* and *⊿cvnF8* strains both had almost no pigmentation in the interactions and they showed very similar distributions with respect to each other across the diameter of the colonies (Figure 2-5D). *⊿cvnF8* did have slight pigmentation near the interaction edge of the colony and that can be seen in the pigmentation graph approaching the 0.5 pigmentation intensity threshold. *⊿cvnB8* showed a unique pigmentation pattern, where the most intense pigmentation is found around the center and third quarter of the colony in stark contrast to the wild-type *S. coelicolor*. This mislocalization of pigmentation seemed to be unique to *⊿cvnB8*, as we typically see the strongest pigmentation closest to the interaction edge.

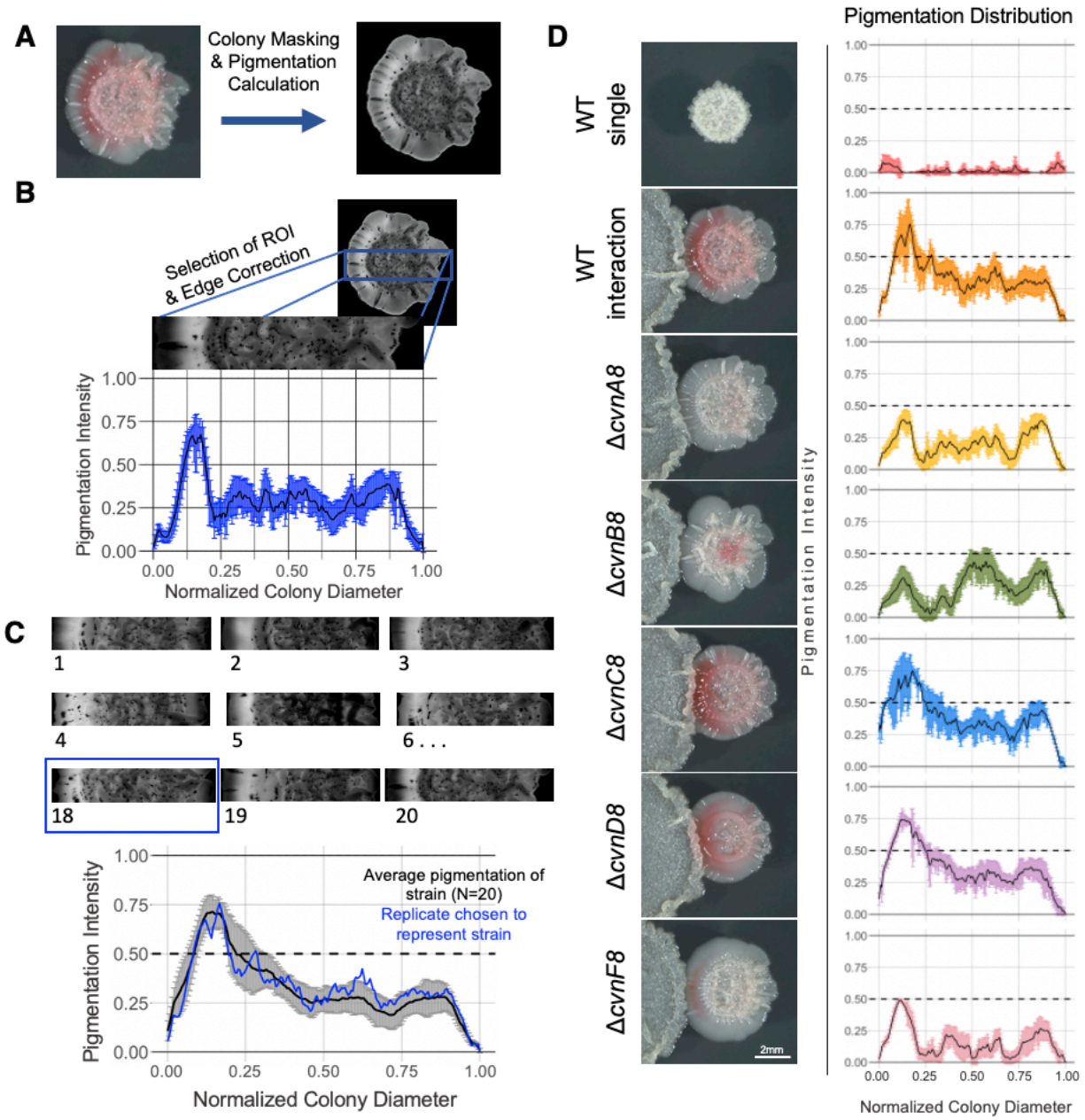


Figure 2-5 | The conservon 8 genes contribute distinctly to the distribution of pigmentation across interacting *S. coelicolor* colonies.
The input (left) and output (right) image from the colony masking and pigmentation detection method (A). Within the rectangular region of interest (ROI), the pixels that fall into each breakpoint are averaged, normalized, and graphed along with the standard deviations (in blue) (B). This calculation was done for 20 replicates of each *S. coelicolor* strain. One representative replicate was selected based on its similarity to the average pigmentation distribution across all 20 replicates (C). The representative replicates for each *S. coelicolor* strain are

shown in the micrographs along with the pigmentation profile across the width of that colony (D).

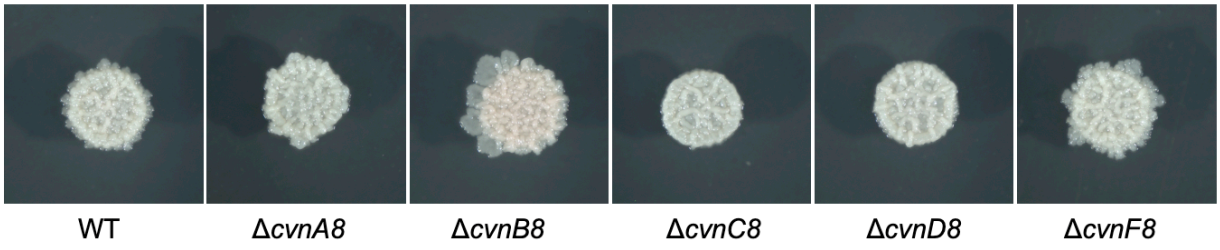


Figure 2-6 | Single colony phenotypes of *S. coelicolor* strains.
Wild-type *S. coelicolor* and five mutant strains were spotted on ISP2 agar and grown for 4 days to observe the phenotypes of the strains as single colonies.

In summary, the *ΔcvnC8* and *ΔcvnD8* mutants have a pigmentation pattern similar to wild-type strain, the *ΔcvnA8* and *ΔcvnF8* mutants have reduced pigmentation patterns, and the *ΔcvnB8* mutant showed a unique pattern of mislocalized pigmentation. Altogether, we can see differences in pigmentation distributions of the single gene deletion mutants, and we can conclude that the individual genes of conservon 8 are contributing differently to pigmentation production.

2.4.4 Expression patterns indicate some subsets of the conservon genes work together

We next wanted to understand how each of the conservon 8 genes contributes to gene expression in *S. coelicolor,* so we did RNAseq on the single gene deletion strains. We placed the *S. coelicolor* strains into interactions with *Amycolatopsis* sp. AA4, spotting the spore stocks 0.75 cm apart on ISP2 agar and growing them for 4 days before harvesting the entire single colonies or half of the interacting colonies for RNA isolation. We compared the transcriptome of each single gene deletion strain to the wild-type strain in both the single colony environment and the interaction environment. To broadly illustrate variation between these transcriptomes, we did a Principal Coordinate Analysis (PCA) on the $\log_2$ fold change expression values for each single gene mutant as compared to wild type, for both single colonies (Figure 2-7A) and for interacting colonies (Figure 2-7B). The first two principal components, or dimensions ('Dim') explained a substantial proportion of the variation for both the single (83.7%) and interacting colonies (82.4%), meaning these PCAs are very good representations of the entire data set.

We determined the threshold for significant differential expression to be at least 2 $\log_2$ fold (or 4-fold) change (either up or down) and a p-value < 0.05 (n=3) for each mutant compared to wild type. There were 122 significantly differentially expressed genes when the mutants

were in a single colony, and 329 genes that were significantly differentially expressed during interactions with *Amycolatopsis* sp. AA4 (Figure 2-7C,D).

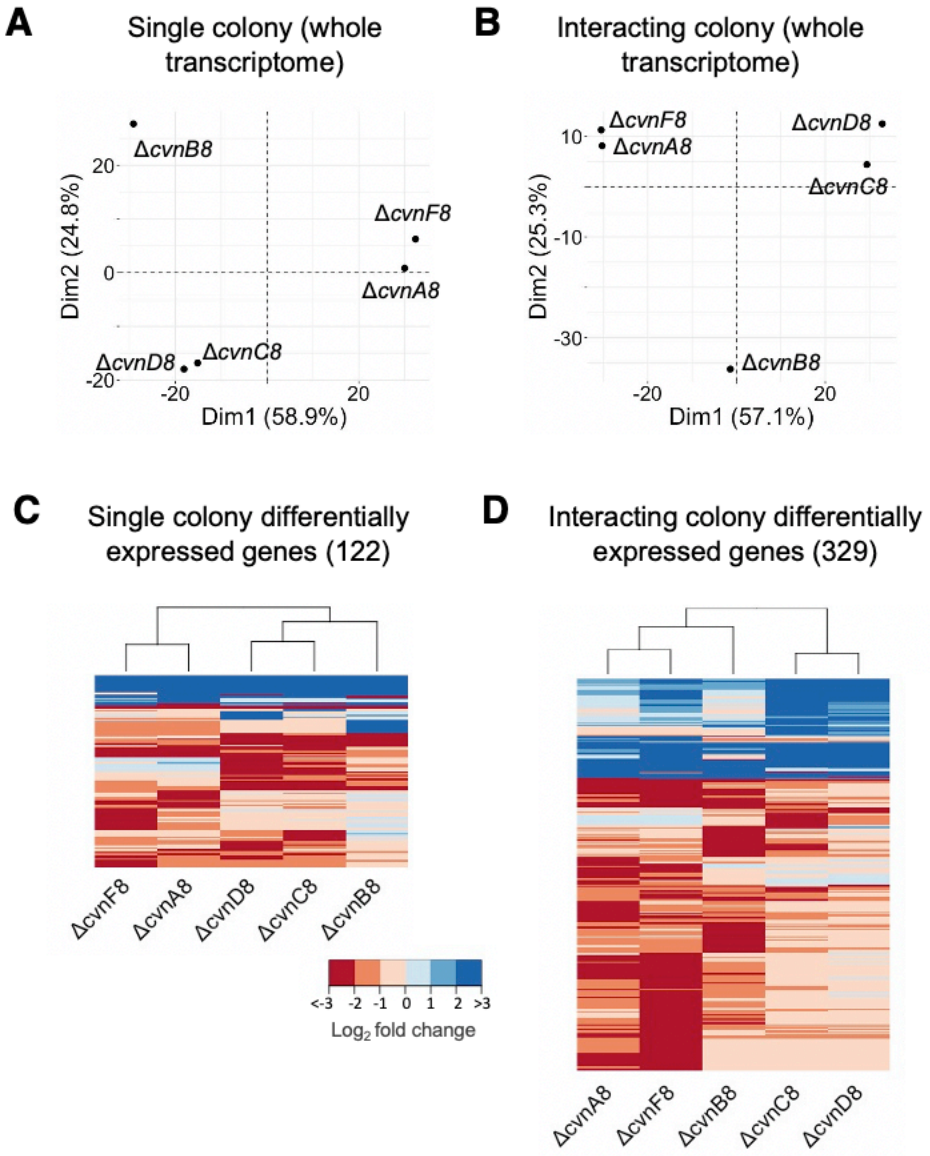


Figure 2-7 | General expression patterns in conservon 8 mutants.
Using the $\log_2$ fold change of all the genes in the transcriptome, the principal components of each strain was calculated and plotted for the strains in single colonies (A) and in interactions (B). Significantly differentially expressed genes ($p < 0.05$, absolute $\log_2$ fold change ≥ 2) are shown in the heatmaps (C,D). The gene had to be significantly differentially expressed in at least one of the mutant strains to be shown. The dissimilarity of the samples was calculated and used to hierarchically cluster the strains shown in the dendrogram. 122 genes and 329

genes were significantly differentially expressed in the single colonies and the interacting colonies, respectively.

Overall, our transcriptomic data mirrored the pigmentation phenotype for each mutant. The ⊿*cvnC8* and ⊿*cvnD8* mutants showed similar pigmentations patterns , and their transcriptomes clustered together (Figure 2-7). The ⊿*cvnA8* and ⊿*cvnF8* mutants also group together in our transcriptomic data (Figure 2-7), and these two mutants have similar pigmentation phenotypes. The ⊿*cvnB8* mutant stands out as unique in both our transcriptomic analysis and the pigmentation distribution analysis. Interestingly, when considering the subset of differentially expressed genes, the ⊿*cvnB8* strain is more similar to the ⊿*cvnC8* and ⊿*cvnD8* mutants when they are grown as single colonies and it is more similar to the ⊿*cvnA8* and ⊿*cvnF8* mutants when it is grown in an interaction.

In the PCA plot showing mutant transcriptomes during and interspecies interaction (Figure 2-7B), we note that the axis explained by Dim1 corresponds very nicely to the general amount of pigmentation observed. The ⊿*cvnA8* and ⊿*cvnF8* mutants showed little pigmentation (Dim1= ~-30), the ⊿*cvnB8* mutant showed a moderate level of pigmentation (Dim1 = ~0), and the ⊿*cvnC8* and ⊿*cvnD8* mutants showed the most pigmentation (Dim1 = ~30). From all of these data together, we can conclude that there are subsets of the conservon 8 genes that seem to function together or in a similar manner because their deletion causes similar changes in gene expression.

2. 4. 5 Conservon 8 controls the expression of multiple specialized metabolite biosynthetic gene clusters

2.4.5.1 Pigmented Metabolites

Since deletion of the entire conservon 8 locus caused *S. coelicolor* to produce a noticeably reduced amount of the two red pigmented specialized metabolites (actinorhodin and undecyl-prodigiosin), we used our single-gene mutant transcriptomic data to investigate how the conservon 8 genes regulate these biosynthesis pathways. We averaged the $\log_2$ fold expression values of all the genes within each red-pigment biosynthesis cluster for each mutant compared to wild type (Figure 2-8A,B). We also examined the $\log_2$ fold change in expression of each gene within the gene cluster (Figure 2-8C,D).

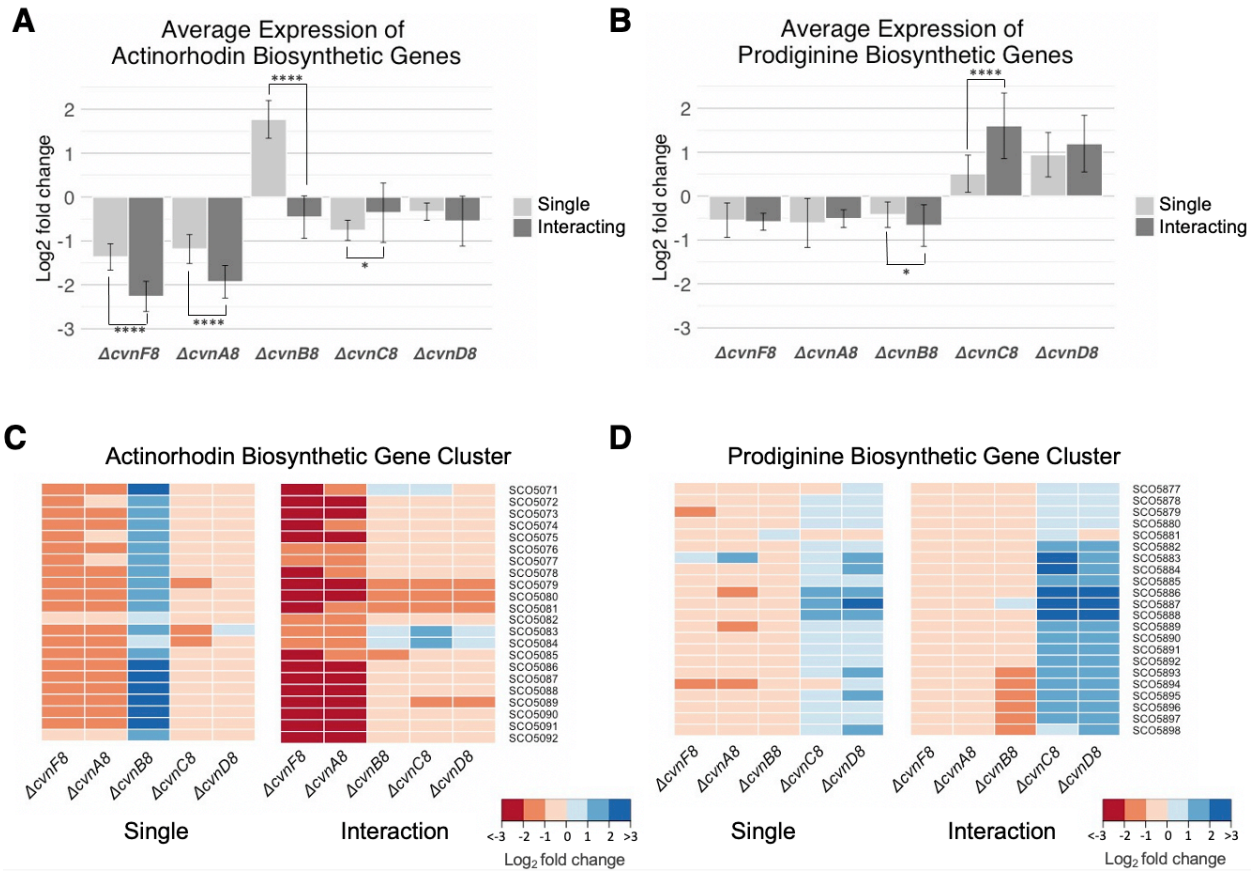


Figure 2-8 | Influence of conservon 8 genes on the expression of pigmented specialized metabolite biosynthetic genes.
The average $\log_2$ fold change of the actinorhodin (A) and prodiginine (B) biosynthetic genes from transcriptomic data collected during single colony growth and interaction growth. Error bars represent the standard deviation. Two-sample t-tests were done and significant differences between the single strain and interacting strain expression are shown (* $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$, **** $p \leq 0.0001$). Heatmaps of the individual genes in the biosynthetic gene clusters are shown with cool colors indicating up regulation and warm colors indicating down regulation (C,D).

Looking at the overall expression of the actinorhodin biosynthetic gene cluster, termed *act* (Figure 2-8A,C), we observed reduced expression in the ⊿*cvnA8* and ⊿*cvnF8* mutants compared to wild type that correlated with the decrease in red pigmentation in these strains during interactions. Interestingly, we saw a lower expression in these mutants compared to wild-type as single colonies as well. However, in the interactions, the change in expression was significantly greater than the change in single colonies (t-test, $p < 0.0001$, n=3). The ⊿*cvnF8* and ⊿*cvnA8* mutants had an average $\log_2$ fold change of -1.36 and -1.18 (-4.8 and -3.81 fold change), respectively, across the 22 biosynthetic genes when in interactions, compared to wild type. When *S. coelicolor* is missing either the *cvnA8* or *cvnF8* genes, actinorhodin expression is

greatly reduced. This indicates that the *cvnA8* and *cvnF8* genes play a role in positively regulating the expression of the actinorhodin biosynthetic gene cluster, and this role is even more important during interspecies interaction. With mutants *⊿cvnC8* and *⊿cvnD8* we observed very slight down regulation of the actinorhodin biosynthetic genes, which was not consider significant (log_2 fold < 1 or > -1).

The *⊿cvnB8* mutant has a unique gene expression patter for the actinorhodin biosynthetic gene cluster. These genes are upregulated compared to wild type in a single colony environment by an average of 1.77 log_2 fold (~3.5-fold), and in contrast, these genes are not significantly differentially expressed when grown in an interspecies interaction (Figure 2-8A). It is notable that there are changes in expression in the absence of an interaction, indicating that the CvnB8, protein and possibly conservon 8 in general, has a role during single colony growth. With respect to the actinorhodin biosynthetic gene cluster, *act*, these data show that the CvnB8 protein is a negative regulator during growth as a single colony and the CvnA8 and CvnF8 proteins are positive regulators during growth as single or interacting colonies (Figure 2-10).

For the prodiginine biosynthetic gene cluster, the *red* genes, we observed slight, not significant down regulation in the *⊿cvnA8*, *⊿cvnF8*, and *⊿cvnB8* mutants as both single and interacting colonies (Figure 2-8B,D). Interestingly, SCO5893-5898 which are named *redF-K* are more down regulated than the rest of the biosynthetic gene cluster in the *⊿cvnB8* mutant during interactions (Figure 2-8D). The genes *redF-K* are tailoring enzymes involved in making modifications to the compound's core structure. Previous studies have hypothesized that *redF-K* are transcribed on an mRNA along with *redL-P* (SCO5888-58920), but our data indicates they may be transcribed separately or under the control of another form of regulation that results in a lower level of expression of *redF-K* compared to *redL-P*. Although, with only four nucleotides between the open reading frames of *redK* and *redL*, it is not likely to have a promoter between the two genes.

We observed increased expression of the prodiginine biosynthetic gene cluster in the *⊿cvnC8* and *⊿cvnD8* mutants as single and interacting colonies. For the *⊿cvnC8* mutant, the upregulation in the interaction was significant compared to the upregulation of the single colony. The *⊿cvnD8* strain displayed upregulation in both single and interacting colonies, a unique place where we observed a small difference between the *⊿cvnC8* and *⊿cvnD8* mutants. This increased expression supports our observation and hypothesis from the pigmentation data that the *⊿cvnC8* strain and the *⊿cvnD8* strain could be more pigmented than the wild-type strain in interactions and provides an explanation for this increase. The mostly highly upregulated genes were SCO5886-5888 named *redP,Q,R*, whose suggested role is to create a pool of fatty acid acyl carrier proteins (ACPs) from which dodecyl ACP can be specifically selected for the biosynthesis of 2-undecylpyrrole. It is this substrate that is combined with 4-methoxy-2,2'-bipyrrole-5-carboxaldehyde to produce undecyl-prodigiosin.

Taken together, these data indicate that the *cvnC8* gene is a negative regulator of the prodiginine biosynthesis genes, *red*, during interactions and the *cvnD8* gene is a negative regulator of the biosynthesis genes during both growth alone and in interactions (Figure 2-10). This points toward a dynamic conservon system that has differing effects on the gene expression of these two red-pigmented specialized metabolite biosynthetic gene clusters.

2.4.5.2 Other Specialized Metabolites

In addition to prodiginine and actinorhodin, we observed changes in the expression of other specialized metabolite biosynthetic gene clusters. In particular, we saw significant upregulation of the coelimycin biosynthetic gene cluster, *cpk*, in all five of the single gene deletion mutants during interactions. In contrast to the actinorhodin and prodiginine biosynthetic genes where the conservon 8 genes had differing effects, here we see each deletion has the same effect on *cpk* expression (Figure 2-9A). In this case, the whole conservon 8 system works together to negatively regulate the coelimycin biosynthetic gene cluster. Coelimycin is a highly reactive bis-epoxide specialized metabolite that has antibacterial properties and is synthesized by the *cpk* genes. It is not currently known what the effect of coelimycin is on *Amycolatopsis* sp. AA4, but we hypothesize that its highly reactive nature would have some disruptive effects.

Even though the change in gene expression the coelimycin cluster in each of the mutants is in the same direction, the magnitude of the change differs. The *⊿cvnC8* and *⊿cvnD8* strains exhibit a larger increase in average expression of approximately 3.46 and 2.98 $\log_2$ fold (11-fold and 7.9-fold), respectively, compared to the change in the *⊿cvnF8, ⊿cvnA8* and *⊿cvnB8* strains of approximately 2.21, 1.52, and 1.84 $\log_2$ fold (4.6, 2.9, and 3.6-fold) respectively when in interactions. Curiously, the only mutant that displays any change to the expression of the coelimycin biosynthetic gene cluster as a single colony is *⊿cvnF8*. This suggests that *cvnF8* plays a role in negatively regulating the *cpk* gene cluster in both the single and interacting environments.

Overall, the conservon 8 genes function to negatively regulate the coelimycin biosynthetic gene cluster during interspecies interactions specifically (Figure 2-10). We hypothesize that in order to do this, the conservon 8 system functions in a linear way to pass a signal through the proteins in this order: CvnA8/CvnF8, and CvnB8 to CvnD8/CvnC8. The CvnD8/CvnC8 proteins may also be receiving input from other sources and this could explain why we observed a greater increase in expression of the *cpk* genes in the *⊿cvnC8* and *⊿cvnD8* mutants compared to the *⊿cvnA8*, *⊿cvnF8*, and *⊿cvnB8*.

The most highly differentially expressed genes in our transcriptomic data are from a cryptic lanthipeptide biosynthetic gene cluster, we termed *lan*. Lanthipeptides are ribosomally synthesized, post-translationally modified specialized metabolites. They are known to have a broad range of bioactivities including antibacterial, antifungal, antiviral, and more (McAuliffe et al., 2001). *S. coelicolor* has one previously characterized lanthipeptide, SapB, that has biosurfactant activity and is essential for morphological differentiation on rich media (Capstick et al., 2007). No molecule has ever been described from this gene cluster, and conservon 8 is

directly adjacent to this lanthipeptide biosynthetic gene cluster, indicating that they may have an important, more direct connection.

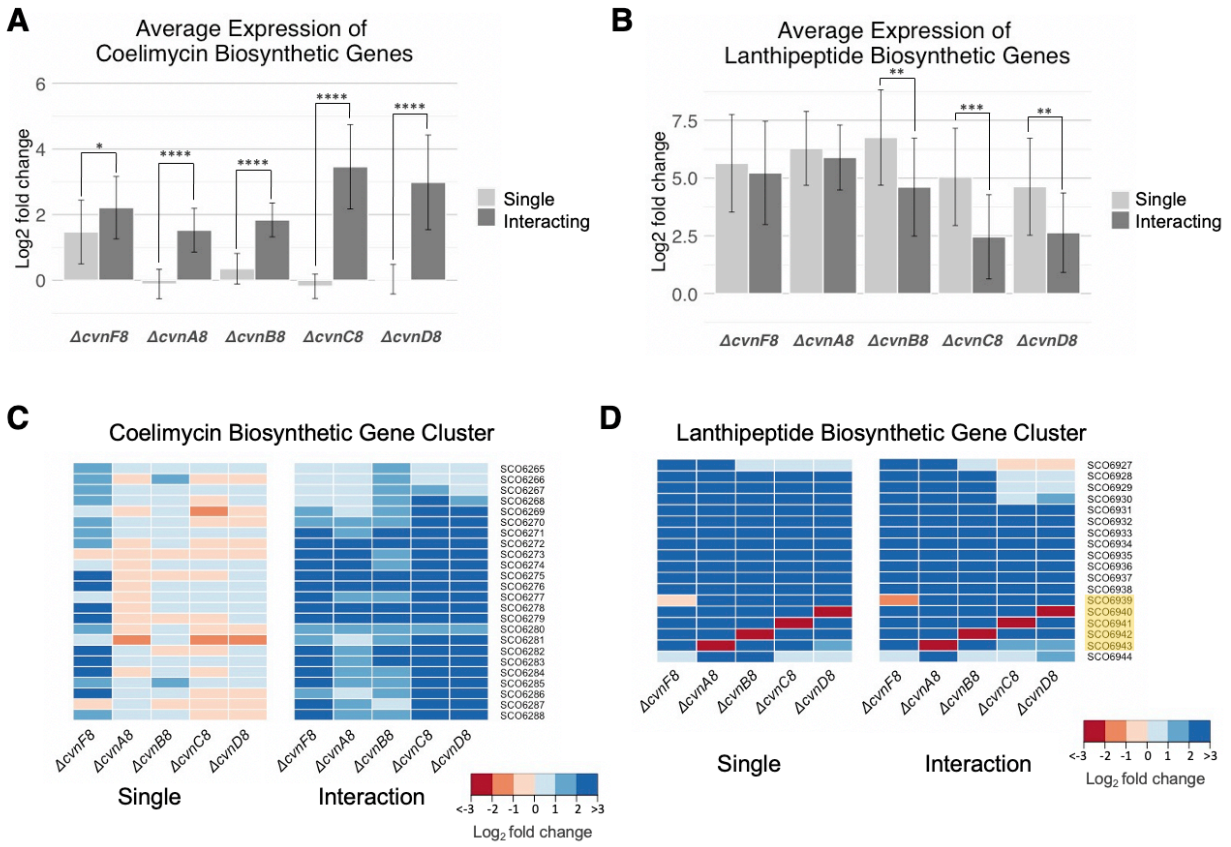


Figure 2-9 | Influence of conservon 8 genes on the expression of additional specialized metabolite biosynthetic genes.
The average $\log_2$ fold change of the coelimycin (A) and an unknown lanthipeptide (B) biosynthetic genes from transcriptomic data collected during single colony growth and interaction growth. Error bars represent the standard deviation. Two-sample t-tests were done and significant differences between the single strain and interacting strain expression are shown (* $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$, **** $p \leq 0.0001$). Heatmaps of the individual genes in the biosynthetic gene clusters are shown with cool colors indicating up regulation and warm colors indicating down regulation (C,D).

Two *lan* genes encoding the prepeptides, SCO6931 and SCO6932, were both strongly upregulated (> 9.27 $\log_2$ fold or 600-fold) in the ⊿*cvnB8* mutant compared to wild type during single colony growth. The *lan* gene cluster was upregulated on average 5.22 $\log_2$ fold (37-fold) or greater in the ⊿*cvnA8* and ⊿*cvnF8* mutants regardless of whether they were grown alone or in interactions. This implies that *cvnA8* and *cvnF8* significantly negatively regulate the cryptic

lanthipeptide biosynthetic gene cluster regardless of the presence of an interacting strain (Figure 2-10). Likewise, the *cvnB8*, *cvnC8*, and *cvnD8* genes also negatively regulate the *lan* genes in both the presence and absence of an interacting strain, but they exert more of an effect in the absence of the interaction (Figure 2-9B). This indicates that the conservon 8 system represses the expression of the *lan* genes without any interaction stimulus.

These data, coupled with the proximity of conservon 8 to this cryptic lanthipeptide biosynthetic gene cluster, led us to conclude that the cryptic lanthipeptide gene cluster is the direct target of conservon 8 regulation. In Figure 2-9D, we included the conservon 8 genes (highlighted in yellow) with the *lan* gene cluster. Overall, out data indicate that conservon 8 plays a role in regulating the cryptic lanthipeptide biosynthetic genes and suggests that the lanthipeptide is the direct target of conservon 8.

2.5 Discussion and Conclusions

We found that conservon 8 is important for *Streptomyces coelicolor* during interspecies interactions with other actinomycetes. When we deleted conservon 8 from the genome we saw a delay in the typical production of two red pigmented antibiotics indicating that conservon 8 functions to sense the interaction, regulate specialized metabolism during interactions, or possibly both (Figure 2-3C). Single gene deletion mutants allowed us to determine that there are pairs of conservon 8 genes that function together to regulate specialized metabolism. We found that conservon 8 is a global regulator of specialized metabolism during interspecies interactions, regulating the biosynthetic genes for actinorhodin, prodiginine, coelimycin, and a cryptic lanthipeptide (Figure 2-10). This points to a model where a histidine kinase-like protein functions with a GAF protein to sense an interaction cue, and relay that signal through a RAS-like GTPase, which alters gene expression in the bacterium *Streptomyces coelicolor*, which is unique and spectacular.

Previous studies on conservons have established that these gene clusters can effect both development and secondary metabolism in *Streptomyces* (Komatsu et al., 2003). The most in depth biochemical study, done by Komatsu et al., showed that the CvnA9 protein has ATPase activity and ATP-binding ability along with the ability to interact with the CvnB9 and CvnC9 proteins (Komatsu et al., 2006). Conservon 9 has the four typical conservon genes (*cvnA9-D9*) along with a cytochrome p450 directly downstream, *cvnE9*. The researchers hypothesized that the proteins of conservon 9 form a membrane-associated heterocomplex that parallels eukaryotic G-protein coupled regulatory systems involved in outside-in signaling. From our data, we agree that conservons are likely signaling systems, and we propose that there are different groups of conservons that function in distinct ways. With respect to conservon 8, we hypothesize that the CvnA8/F8 proteins function together as the sensor for this particular system.

Our data provides strong evidence for the cooperation of the CvnA8 and CvnF8 proteins. The first piece of evidence for this is a compelling reduction in red pigmentation of the deletion mutants when in interactions with *Amycolatopsis* sp. AA4 (Figure 2-5D). When looking

at the transcriptome of both mutants, we were surprised to see that their expression profiles phenocopy each other. This is powerful evidence that the CvnA8 and CvnF8 proteins work together because they have extremely similar effects on gene expression both during growth as a single colony and during interactions. The same similarities hold up when we focus on the four specialized metabolite biosynthetic gene clusters, *act*, *red*, *cpk*, and *lan*. Out of the four clusters under the two conditions, the *⊿cvnA8* and *⊿cvnF8* strains only show a significant difference ($p < 0.0001$) in the average expression of *cpk* in single colonies (Figure 2-9A).

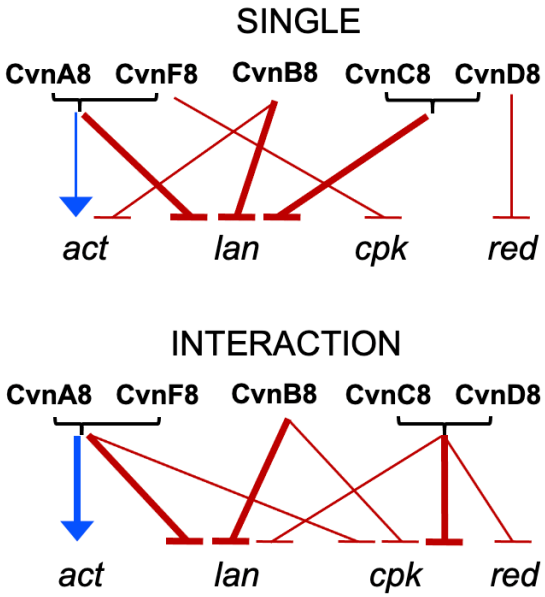


Figure 2-10 | Summary of regulation of specialized metabolite biosynthesis genes.
The effect of each conservon 8 gene on four specialized metabolite biosynthetic gene clusters determined from transcriptomic analysis: actinorhodin (*act*), prodiginine (*red*), coelimycin (*cpk*), and the cryptic lanthipeptide (*lan*). Blue arrows indicate positive regulation, red arrows indicate negative regulation, thicker arrows indicate > 2 $\log_2$ fold change and thinner arrows indicate between 1-2 $\log_2$ fold change.

Another hint that the CvnA8 and CvnF8 proteins might work together comes from the structure of the CvnA protein along with the genes downstream of the *cvnD* genes. We propose that these two traits can be used to define three different groups of conservons. The CvnA1-6 proteins have clear, extracellular, nitrate/nitrite sensing (NIT) domains (pfam08376) and these form conservon group 1 that seems to have a built in, extracellular sensing domain. The CvnA7-13 proteins have very short regions (<20 residues) between the two transmembrane domains, indicating they do not have an extracellular region that could act as a sensor domain nor do they have a predicted sensor domain anywhere else in the protein. This led us to

hypothesize that the CvnA7-13 proteins may interact with other proteins which fulfill the sensing role, which appears to be lacking in these versions of CvnA. Of those lacking a sensing domain, conservons 9 through 12 (what we termed conservon group 2) have a *cvnE* (cytochrome p450) gene following the *cvnD* gene which, for the conditions explored in Komatsu et al., 2006, had no effect on phenotype when deleted, but which we speculate may be a sensor for those conservon systems. The last set of conservons, group 3, contains conservons 7 and 8, which have a GAF domain containing gene downstream of the *cvnD* gene, which we have termed the *cvnF* gene. One common function of GAF domains is the binding of small molecules and regulation of enzyme activity. This fact, combined with our transcriptomic data showing the deletion of the *cvnA8* and *cvnF8* genes are phenocopies, led us to hypothesize the CvnF8 protein is acting alongside the CvnA8 protein as the sensor in this conservon system. We wonder if the CvnA protein is a different kind of sensor that has evolved from bacterial histidine kinases to function very similarly with an extracellular sensor domain (group 1) or to function using external sensing domains that interact with it (groups 2 and 3).

The data, taken together, illustrate that the CvnA8 and CvnF8 proteins function together or in very similar manners to affect red pigment production during interactions as well as gene expression in both alone and interactions conditions in *S. coelicolor*. It provides evidence that the CvnF8 protein (SCO6939), which contains a GAF domain, functions together with the conservon 8 system. It raises the question of whether other proteins in close proximity may also interact with the conservons. More work in needed to definitively prove that the CvnA8 and CvnF8 proteins function together, but our data strongly implies that this is the case.

The previous biochemical exploration of conservon 9 done by Komatsu et al., 2006 demonstrated that the CvnD9 protein has GTPase activity, can bind guanine nucleotides, and can interact with the CvnB9 and CvnC9 proteins in a GTP dependent manner. In particular, the CvnD protein caught our attention because small GTPases are used extensively in eukaryotic signaling systems and, although they are present in bacteria, we know very little about how they function in bacteria (Wuichet and Søgaard-Andersen, 2014).

The data strongly indicates that the CvnC8 and CvnD8 proteins function in a similar manner and possibly together to affect gene expression and pigmentation in *S. coelicolor* during interspecies interactions. Both the *⊿cvnC8* and *⊿cvnD8* mutants show no reduction in pigmentation intensity in interactions with *Amycolatopsis* sp. AA4 (Figure 2-5D). In fact, the peak in pigmentation intensity is wider, and the prodiginine biosynthesis gene cluster is strongly upregulated in these two mutants compared to the wild-type strain, indicating there might be an increase in red pigment production. We speculate that the wider pigmentation intensity peaks of these two mutants is due to the relieved repression of the *red* genes in Δ*cvnC8* and Δ*cvnD8* resulting in increased production of undecyl-prodigiosin. When looking at genome wide expression levels or specifically focusing on the specialized metabolite biosynthetic gene clusters, we see that the deletion of the *cvnC8* and *cvnD8* genes phenocopy each other, providing strong evidence for their connection. It is interesting to note that the expression of the *act* genes is not significantly changed, but the expression of the *red* genes

during interactions are upregulated more than 1 $\log_2$ fold or two-fold in both the *⊿cvnC8* and *⊿cvnD8* mutants.

When we analyzed the predicted structure of the CvnC8 and CvnD8 proteins and compared this with our transcriptomic data, which led us to hypothesize how these proteins function. The CvnC8 protein is a small, 136 residue protein with a conserved domain of unknown function (DUF742) that is in the HTH superfamily. Predicting the secondary structure shows that it does have a helix-turn-helix motif, which indicates the CvnC8 protein could bind nucleic acids. The CvnD8 protein indeed looks like a functional GTPase, with the correct predicted topology and almost all of the conserved residues important for guanine nucleotide binding. When we look at the secondary structure, we see that it has good homology to MglA, a small RAS-like GTPase in *Myxococcus xanthus* that is involved in gliding motility (Zhang et al., 2010; Nan et al., 2015; Galicia et al., 2019). The MglA protein and eukaryotic small GTPases have been shown to function as molecular switches where they are actively signaling to downstream proteins when bound to GTP (Colicelli, 2004; Zhang et al., 2010). They have a low intrinsic rate of GTP hydrolysis that eventually returns them to the inactive, GDP-bound form. The transition between active and inactive forms are regulated by other proteins in order to tightly control the timing of GTPase signaling. We speculate that the CvnD8 protein functions in the same way, but more work should be done to prove this hypothesis. Due to the similarity in the *cvnC8* and *cvnD8* deletion phenotypes, we hypothesize that the CvnD8 protein directly regulates the CvnC8 protein, possibly when it is in its active, GTP-bound form. Given the transcriptomic and bioinformatic data, we hypothesize that the CvnC8 protein is a transcriptional regulator that, when activated by its regulator, the CvnD8 protein, represses the *red* and *cpk* genes.

The most enigmatic member of conservon 8 is the CvnB8 protein. The phenotype of the *cvnB8* deletion shows an intermediate phenotype between the *⊿cvnA8/F8* and *⊿cvnC8/D8* mutants, because it still produces the red pigmented antibiotics, but at an intermediate level. Additionally, the *⊿cvnB8* mutant has mislocalized pigmentation, not near the interaction edge, but instead in the center and third quarter of the interacting colony (Figure 2-5D). This mislocalization might be a phenotype that results from misregulation of a factor that controls natural product production. When we analyze the collection of significantly differentially expressed genes, we also see that the *⊿cvnB8* mutant has more similarity to the *⊿cvnC8/D8* mutants when grown as a single colony, but is more similar to the *⊿cvnC8/D8* mutants when grown in interactions (Figure 2-7C,D). We think the CvnB8 protein might function to connect the CvnA8/F8 and CvnC8/D8 proteins, which can explain why we see an intermediate phenotype.

The predicted protein structure of CvnB8 looks very similar to the MglB protein crystal structure. MglB is a *M. xanthus* protein that regulates the small GTPase, MglA. The MglB protein has been shown to be a GTPase activating protein (GAP) (Zhang et al., 2010). These proteins bind to the small GTPases they regulate and stabilize the active site to increase the rate of GTP hydrolysis. Intriguingly, the MglB protein was recently shown to also act as a guanine nucleotide exchange factor (GEF) (Baranwal et al., 2019). GEFs promote the

exchange of GDP for GTP on the small GTPases they regulate to switch them from their inactive form to their active form. Combining this protein structure information with our data, the CvnB8 protein does not seem to be functioning strictly as a GAP for CvnD8. If this were the case, we would expect to see opposite single gene deletion phenotypes. If the CvnB8 protein has GEF activity specific for CvnD8, then we would expect the deletions to have the same phenotypes. We do not see the same phenotype for the *act* and *red* genes, but we do observer similar changes in expression levels with the *cpk* and *lan* genes. We hypothesize that the CvnB8 protein functions as a GEF to switch CvnD8 into its active state, but we do not think it is the only GEF that can activate the CvnD8 protein or that CvnD8 is the only GTPase that the CvnB8 protein may affect. CvnB8 may also function as a GAP under different conditions, not studied here.

Not all the transcriptomic regulation we observed can be explained by just the conservon 8 system. During interactions, the deletions of the *cvnA/F8* genes are the only mutants that show a significant change in expression of the *act* genes. This implies that the CvnA8 and CvnF8 proteins have an independent effect on the *act* genes, that seemingly does not involve the other members of the conservon 8 system. We hypothesize that this effect is the result of regulation flowing through other proteins, possibly other conservon proteins through crosstalk of the systems. Additionally, only the deletions of the *cvnC8* and *cvnD8* genes had an effect on the expression of the prodiginine biosynthetic genes, which indicates that the effect of the CvnC8 and CvnD8 proteins on the red genes is independent of the CvnA8,B8, and F8 proteins. Again, this is an area where we assume there must be other proteins, outside of the conservon 8 system, involved in passing a signal through CvnC8/D8.

With 13 copies of the conservons present throughout the *S. coelicolor* genome, it seems likely that there may be crosstalk between these systems. Aligning all 13 proteins from each member of the conservon showed that the CvnD protein is the most well conserved, and the CvnB8 protein is the second, while the CvnA8 protein was the least well conserved. CvnA protein length spanned from 400 to 1,329 amino acids. Further, there may be redundancy between some of conservon proteins and this may explain why we see different phenotypes with the different single gene deletions. The possibility of crosstalk and redundancy make this an extremely challenging system to work on and it will take a formidable effort to determine the details of how the conservons are functioning.

The conservon system is worth deeper study. *Streptomyces* are known to be very complex bacteria, with large genomes, an elaborate development life cycle, a huge proportion of regulatory genes, and many specialized metabolite synthesis genes. The conservon system is found 13 times throughout the genome, indicating that it serves an important function in order to be kept and propagated over evolutionary time. We think there might not be a universal way that all the conservons work, but conservon 8 seems to be a sensing and signaling system that is involved in regulating gene expression of specialized metabolite biosynthetic gene clusters during interspecies interactions.

2.6 Acknowledgements

This work was done in part by and under the supervision of Dr. Matthew Traxler. Some strain generation and RNA collection were done by Yein Ra. Pigmentation quantification script was written with the help of Javier Soto Bustos. Protein purification optimization was performed in part by Luis Cantu. Thank you to all of the above scientists for lending their talents to this project.

3 Chapter 3. Characterizing the conservon 8 system with a focus on CvnD8, a RAS-like GTPase involved in regulation of specialized metabolism

3.1 Introduction

We identified conservon 8 as being important for *Streptomyces coelicolor* during interspecies interactions and subsequent natural product production. Without the conservon 8 system, we observed abnormal interaction phenotypes and altered levels of natural product production. In the previous chapter we explored how conservon 8 is functioning through single gene deletions and transcriptomics. This allowed us to connect the conservon 8 system to downstream effects on gene expression and the visual natural product production. From this, we identified subsets of genes that function together and explored how they influence gene expression individually. This data is crucial to understanding the output of the conservon 8 system, but it is far from enough to give us an understanding of how it functions.

In this chapter, I explore connections within the conservon 8 system itself. We gathered information about each component of the system from bioinformatics data in order to generate hypotheses about how each component functions individually and eventually, how the system functions as a whole. This bioinformatics data has laid the groundwork for current and future work in the lab that is aimed at deciphering the molecular mechanisms by which this system functions. We dive deep into the sequences and structures of the conservon 8 proteins coupled with previous work on related systems to build a very speculative model of how the conservon 8 system is functioning.

From our analyses, we hypothesize that the CvnA8 protein is a membrane bound receptor that functions with the CvnF8 protein to sense a small, intercellular molecule. Once the signal is sensed, we think it travels through the CvnB8 protein, which goes on to regulate CvnD8 activity. When CvnD8 is active, we hypothesize that it passes this signal on to the CvnC8 protein to repress multiple biosynthetic gene clusters. Further, we hypothesize that there is crosstalk between other systems that can either influence or be influenced by conservon 8. We have begun biochemical experiments to characterize the RAS-like GTPase activity of CvnD8. This involves creating mutant proteins, purifying these mutants along with wild type, assaying for GTPase activity, and observing the *in vivo* effects of these mutations. Though these experiments are still in progress, the effort to characterize the CvnD8 protein is a substantial part of this body of work.

3.2 Methods

3.2.1 Strains and Plasmid Construction

E. coli BL21 (DE3) Codon Plus RIL cells were grown at 37°C shaking unless otherwise specified.

CvnD8 was cloned into the pETDuet plasmid using PCR amplification of both the genes and the plasmid followed by Gibson assembly. The 3x Strep Tag II along with the PreScission Protease cut site was ordered as a single gBlock from IDT. It was PCR amplified with the primers listed in Table 3-1 to add on the Gibson overhangs. The 9x His Tag and linker were also ordered as a single gBlock with the Gibson overhangs included. *cvnD8* and pETDuet were amplified with the primers listed in in Table 3-1. Since pETDuet is a roughly 5.4 kb plasmid we split it into two parts for smaller and easier amplification via PCR. A four-part Gibson was performed with the two pieces of pETDuet, *cvnD8*, and the 3x Strep Tag II gBlock. Once that plasmid was successfully generated, the 9x His Tag was added using Gibson assembly as well. For the second expression construct, the SUMO Tag was amplified from the TwinStrepSUMO3 plasmid provided by the Glaunsinger lab and Gibson assembled into the existing pETDuet construct.

The mutations in *cvnD8* were introduced by PCR using primers containing the base pair substitutions towards the center of an extra-long primer (30 bps). Forward and reverse overlapping primers containing the mutations were used to split the pETDuet-Strep-CvnD8-His plasmid at the site of the mutation and at another site on the plasmid backbone creating two roughly equal sized DNA fragments. The two fragments were then Gibson assembled and proper assembly and incorporation of the mutation was verified with Sanger sequencing of the PCR product.

3.2.2 Protein Sequence and Structure Analysis

Protein sequence alignments were done using Clustal Omega (version 1.2.4) (Sievers et al., 2011). An asterisk below the residue indicates it is fully conserved, a colon indicates conservation between groups of strongly similar properties, and a period indicates conservation between groups of weakly similar properties.

Protein structure prediction was done by Phyre2 in normal modelling mode (Kelley et al., 2015). Protein alignments and visualization was done in PyMOL (version 2.2.0).

Phylogenetic trees were made using a global alignment with free end gaps and the blosum62 cost matrix. The Jukes-Cantor genetic distance model and neighbor-joining tree building methods were used. There was no outgroup.

3.2.3 Protein expression and Isolation

For purification of CvnD8, CvnD8T11/12N, CvnD8Q76L, and CvnD8E67L, BL21 (DE3) Codon Plus RIL cells containing pETDuet with the appropriate expression construct were grown at 37°C in LB Broth containing 50 μg/μl ampicillin to an OD600 of 0.6. The cells were induced with 0.5 mM isopropyl-beta-D-thiogalactoside (IPTG) for 1 hour at 37°C. Cells were put on ice and resuspended in lysis buffer (100 mM Tris/HCl pH 7.5, 150 mM NaCl, 1 mM MgCl2, 100 μM GTP, 1% CHAPS, and protease inhibitor cocktail (Roche REF# 05892953001)). Cells were lysed using a probe tip sonicator at level 2 for 10 minutes with cycles of 10 seconds of sonication, 20 seconds of rest. The proteins were purified using dual affinity chromatography, beginning with Strep-Tactin® Superflow® resin (IBA Lifesciences). Columns were equilibrated with three column volumes (CVs) of lysis buffer, the soluble component of the lysed cells were flowed over the column three times, then the column was washed with two CVs of lysis buffer and two CVs of wash buffer (100 mM Tris/HCl pH 7.5, 150 mM NaCl, 1 mM MgCl2, 100 μM GTP). The proteins were eluted using one CVs of wash buffer containing 2.5 mM desthiobiotin. Subsequently, imidazole was added to the elution to a concentration of 5 mM. The elution was applied three times to the second column containing Ni-NTA Agarose (Qiagen) that had been equilibrated with wash buffer containing 5 mM imidazole. The column was washed with four CVs of wash buffer containing 15 mM imidazole and the proteins were eluted with one CV wash buffer containing 250 mM imidazole. The proteins were concentrated to about 50 μM using Amicon Ultra Centrifugal Filters (Millipore).

3.2.4 Fast Protein Liquid Chromatography

Protein was loaded on an equilibrated GE HiLoad 16/600 Superdex75 pg column and washed with SEC buffer (100 mM Tris/HCl pH 7.5, 150 mM NaCl, 1 mM MgCl2) at a flow rate of 2 mL per minute. Samples were collected in 1 mL fractions.

3.2.5 GTPase assays by LC/MS

The purified CvnD8, CvnD8T11/12N, and CvnD8Q76L, and CvnD8E67L were dialyzed into 1 L of assay buffer (50 mM Tris/HCl pH 7.5, 100 mM KCl, 10% glycerol, 1mM MgCl2, 1 mM DTT) overnight at 4°C. 100 μM GTP and 10 μM of protein was added to assay buffer and incubated at 30°C. 40 μl aliquots were removed and boiled at 95°C for 3 min to stop the reaction. 10ul of reaction was injected on a BioBasic AX HPLC 2.1 x 50 mm column (Thermo cat # 73105-052130). Buffer A was 5mM KH2PO4 at pH 3.2 and buffer B was 0.75M KH2PO4 at pH 3.2. The gradient method was as follows: 0-0.5 min 0% B, 0.5-9 min 30% B, 9-11.5 min 60% B, 14-14.5 min 100%, 14.5-18.5 min 100% B (to wash the column), 18.5-19 min 0% B, 19-23 min 0% B (to equilibrate the column).

Part	Primer	Sequence
Primers for building expression plasmid		
pET DUET part 1	FW	AACCTAGGCTGCTGCCACCG
	RV	GCTGCGGCGAGCGGTATC
pET Duet part 2	FW	GATACCGCTCGCCGCAGC
	RV	ACCAGACTCGAGGGTACCGA
cvnD8	FW	ATCGCCGGCGGTTTCGG
	RV	TAGAAGGGGTGTGACCAGCG
SUMO tag	FW	CGCACCCGCAGTTCGAGAAGTCCGGATCGCTGCAG GAGGA
	RV	CCCTGGAACAGGACCTCCAGGGCCGCCGCTGCCGC TGCCCCTCCCGTCTGCTGCTGG
3x Strep Tag II and percission cut site	FW	TCGGTACCCTCGAGTCTGGTATGTGGAGCCACCCGC AGTT
	RV	CCGCGCCGAAACCGCCGGCGATGCCCGGACCCTGG AACAGGA
G-block with 3x Strep Tag II and percission cut site	G-block	ATGTGGAGCCACCCGCAGTTCGAAAAGGGTGGTGG CTCGGGCGGTGGCAGCGGCGGAGGTAGCTGGTCC CACCCCCAGTTCGAAAAGGGCGGTGGAAGCGGAG GCGGCTCCGGAGGTGGAAGCTGGTCGCACCCGCA GTTCGAGAAGCTGGAGGTCCTGTTCCAGGGTCCGG GC
G-block with C-terminal linker and 9x His tag	G-block	GCTGGTCACACCCCTTCTACCACCCTCCAGGACGCC GGCCAGTTCTACCTGAACGAGGGCGGGTCCGGTGG CTCCGGCGGTTCCGGAGGCCACCATCATCACCACCA TCATCACCACTGAAACCTAGGCTGCTGCCACCGCTG AGCAATAACTAGC
Primers for introducing mutations in CvnD8		
T11/12N	FW	ATCGCCGGCGGTTTCGGCGCGGGCAAGAACAACGC AGTGGGC
	RV	CGCCGAAACCGCCGGCGAT
Q76L	FW	CGGCACGCCGGGCCTGGACCGGTTCGTCGA
	RV	TCGACGAACCGGTCCAGGCCCGGCGTGCCG
E67L	FW	GCGCCGATGCCTTTGCTGCTGTTCCTGTTC
	RV	GAACAGGAACAGCAGCAAAGGCATCGGCGC
Primers for validation of plasmid construction		
Expression construct	FW	GCCCAGTAGTAGGTTGAGG
	RV	GAGAAAGGAAGGGAAGAA

Table 3-1 | List of primers and gBlocks used in this study.

3.3 Results

3.3.1 Grouping the *Streptomyces coelicolor* conservons

Conservons are found across the actinobacterium phylum with 13 in the genome of *Streptomyces coelicolor* alone. We wondered if we could identify commonalities and differences in the *S. coelicolor* conservons to better understand how they may be functioning. Relatively little work has been done to understand the role of conservons in actinomycetes, so we decided to take a bioinformatic approach to generate testable hypotheses for future work.

Considering the previously generated hypothesis that the conservons are signaling systems that may operate similar to G-protein coupled regulatory systems (Komatsu et al., 2006), the sensing component was the first element we investigated for clues as to how this system may be functioning. The CvnA protein is a predicted membrane bound protein with two transmembrane helical domains, and a histidine kinase-like ATPase domain, making up what we consider to be the sensor of the conservon system. Upon comparison of all 13 *S. coelicolor* conservons, the CvnA protein is the most divergent of the core conservon proteins with just 16.6% pairwise identity, ranging in size from 400 to more than 1300 amino acids. Considering the *cvnB,C,D* genes of all 13 *S. coelicolor* conservons, we do not see obvious differences between them, and thus we focus on *cvnA* to find differences that might allow us to differentiate the 13 conservons.

Though the *cvnA* genes have a histidine kinase-like ATPase domain, they do not have the same structural features as typical histidine kinases. Histidine kinases (HKs) are typically transmembrane proteins that function in two-component signal transduction systems in microorganisms. The typical histidine kinase has two transmembrane helices flanking a sensing domain that is extracellular, a HAMP domain involved in signal propagation, a DHp (dimerization and histidine phosphorylation) domain, and an ATPase domain (Figure 3-2A) (Stock et al., 2000; Skerker et al., 2005). The sensing domains of histidine kinases are diverse and share very little primary sequence similarity due to the diversity of ligands they bind. There are some examples of non-canonical histidine kinases, like the chemotaxis kinase CheA or the nitrogen regulatory kinase NtrB, that are soluble, cytoplasmic histidine kinases (MacFarlane and Merrick, 1985; Stock et al., 1988). Cytosolic HKs, typically have PAS or GAF sensing domains that can monitor changes in light, redox potential, and small ligands (Taylor and Zhulin, 1999). CheA is a unique histidine kinase that does not contain any sensing domain and instead interacts with transmembrane chemoreceptors, known as methyl-accepting chemotaxis proteins (MCPs). These chemoreceptors have periplasmic ligand binding sites and conserved cytoplasmic signaling domains in order to act as the sensor for the chemotaxis system (Webre et al., 2003). Though it is certainly uncommon for histidine kinases to lack sensor domains, it is not unheard of in the bacterial world.

Conservons one through six all have a CvnA with a predicted nitrate/nitrite sensing domain (NIT domain) that is located between the two transmembrane helices, indicating it is

extracellular. Whether these domains are actually used to sense nitrate or nitrite is unclear, but we think that dedicating so much genomic space to sensing one nutrient is not likely. A more in-depth analysis paired with biochemical studies is needed to investigate what these NIT domains are sensing. We consider this group of conservons, with an extracellular sensing domain in their CvnA, group 1. We also analyzed the proximity of genes flanking the core conservons, to group the *Streptomyces coelicolor* conservon systems. Some of the conservons in group one have genes upstream or downstream that might also function with them, but with the lack of robust *Streptomyces* protomer prediction software, and enough space between the core conservon gene for it to be transcribed separately, it is hard to say if those genes function with the conservon system.

The second group of *S. coelicolor* conservons are those that do not have a sensing domain between the two transmembrane helices and are followed by an accessory conservon gene that has a cytochrome p450 domain, *cvnE*. This group includes conservons 9, 10, 11, and 12, with conservons 10 and 12 having two *cvnE* genes following the core conservon genes. Conservon 10, 11, and 12 have twelve or fewer nucleotides between their respective *cvnD* and *cvnE* genes, suggesting that those two genes are transcribed together. Conservon 9 has 76 nucleotides between *cvnD9* and *cvnE9*, providing enough room for a promoter in front of *cvnE9*. Interestingly, a *cvnE9* deletion was observed to have no phenotypic differences from a wild-type *S. coelicolor* strain in the conditions tested (Komatsu et al., 2006). Deletions of the other *cvnE* genes have not been studied. Cytochrome p450s are known to mediate a remarkable range of chemical transformations in bacteria through their heme cofactors.

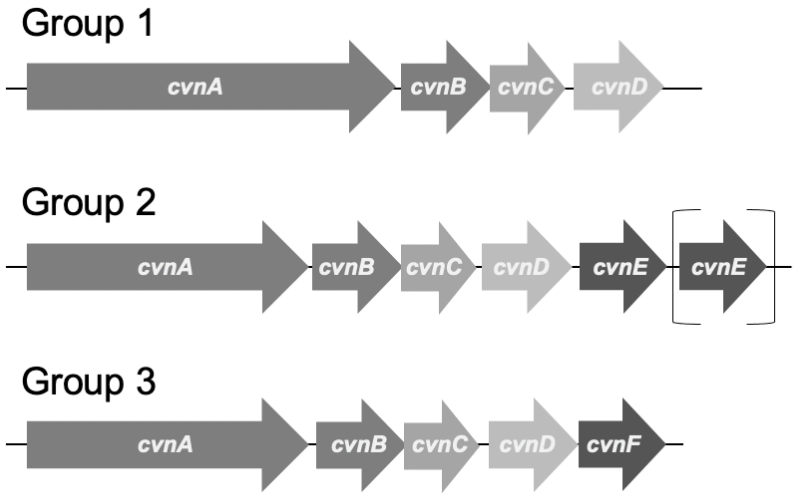


Figure 3-1 | *Streptomyces coelicolor* conservon structure.
The structure of conservon loci can be divided into three main groups. Group 1 has the four core conservon genes, including a larger *cvnA* gene with an extracellular NIT sensing domain. Group 2 has a cytochrome p450 containing gene, *cvnE*, or sometimes two *cvnE* genes following the four core genes. Group

3 has a GAF containing gene, *cvnF*, following the four core genes. Conservon 8 falls into group 3.

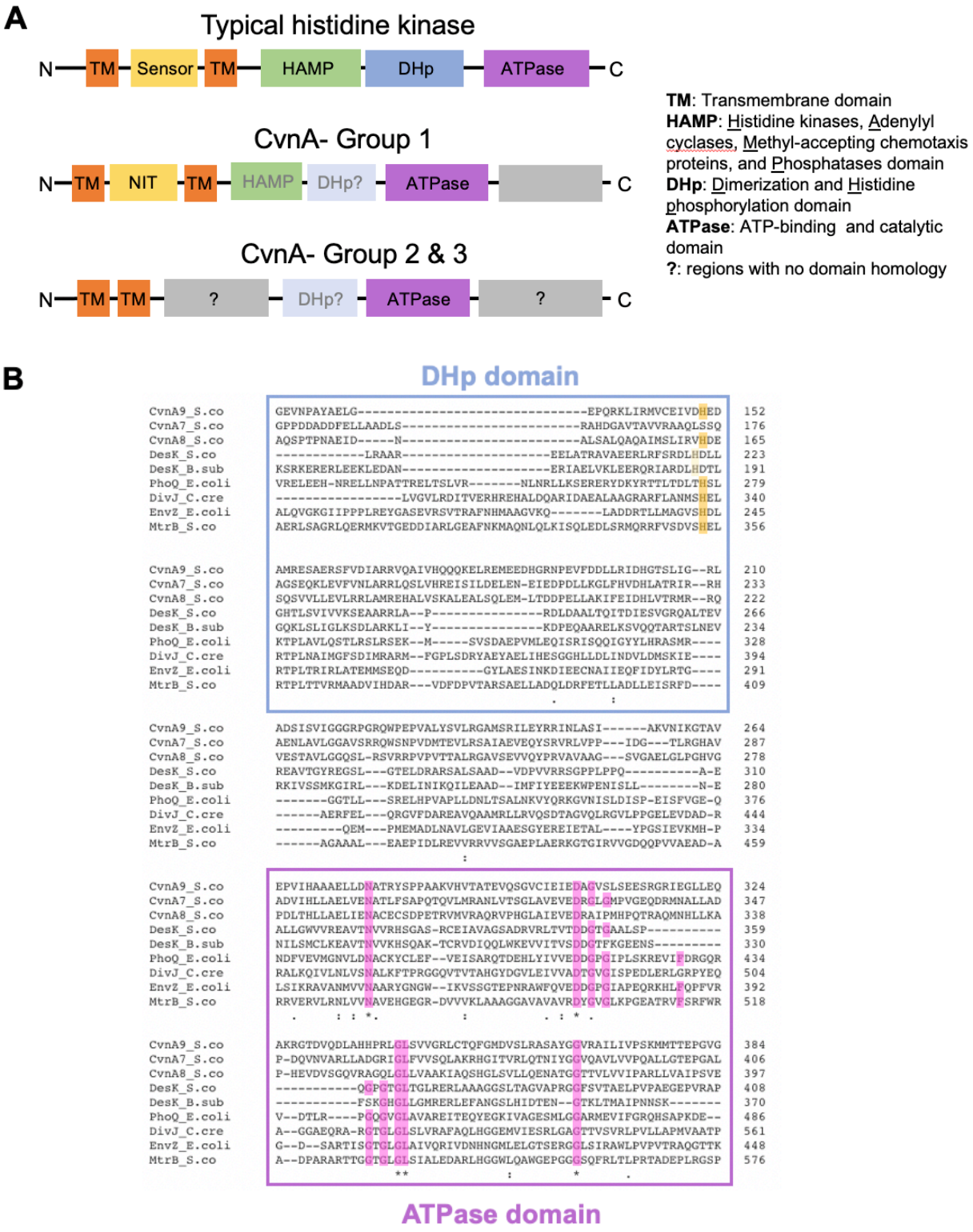


Figure 3-2 | CvnA structure.
A diagram of domains found in a typical histidine kinase along with domains found in CvnA (A). Protein sequence alignment of CvnA7, CvnA8, and CvnA9 from *S. coelicolor* along with histidine kinases from various microorganisms using Clustal Omega (B). The histidine kinases used in the alignment are DesK and MtrB from *S. coelicolor*, PhoQ and EnvZ from *E. coli*, DesK from *B. subtilis*, and DivJ from *C. crescentus*. The subset of sequence shown contains the DHp and ATPase domains.

The third and final group of conservons consist of conservon 7 and 8, which both have a *cvnA* gene without a predicted sensor, have no extracellular domain, and have *cvnF*, an accessory conservon gene with a GAF domain, immediately following the *cvnD* gene. The stop codon of the *cvnD8* gene is followed directly by that start codon of the *cvnF8* gene, and the *cvnF7* gene overlaps with the last 3 nucleotides of *cvnD7*, providing strong evidence that the *cvnF7/8* genes are transcribed with the *cvnD7/8* genes and likely functions with its respective conservon system. Previous studies outlined in chapter two showed that in a *cvnF8* deletion, the changes in the transcriptome phenocopy the changes associated with a *cvnA8* deletion providing additional strong evidence that these two gene products work together or in a very similar manner.

GAF domains are evolutionarily related to PAS domains and consist of a six-stranded anti-parallel β-sheet core. GAF domains have multiple known functions including binding cyclic nucleotides, binding small molecules, sensing redox states, and oligomerization (Zoraghi et al., 2004). One of the most well characterized GAF domains is found in phosphodiesterases (PDE) where it binds cGMP/cAMP and allosterically activates the catalytic domain of the enzyme. There are no invariant amino acids in GAF domains, making it challenging to predict what the role of a specific GAF domain is without experimental testing. To see if we could predict what the CvnF7/8 GAF domains are sensing, we built an evolutionary tree with the protein sequences of CvnF7, CvnF8, and previously studied GAF domains with known functions (Figure 3-3). We saw that the CvnF GAF domains clustered together and were most closely related to a branch of the tree containing GAF domain 2 from PDE2a from *H. sapiens* and GAF domain 1 from PDE5a from *M. musculus* which have both been shown to bind cGMP. The related branch also contains the GAF domain from NifA from *S. meliloti*, which binds glutamine. From this analysis we can simply speculate that the GAF domains in CvnF are involved in binding a small molecule, possibly a cyclic nucleotide.

Understanding the CvnA protein structure along with the presences of accessory conservon genes, *cvnE* and *cvnF*, we start to see patterns emerging in the 13 *S. coelicolor* conservons. Because of these differences we organized them into three groups, and we hypothesize that these groups function more similarly to one another (Figure 3-1).

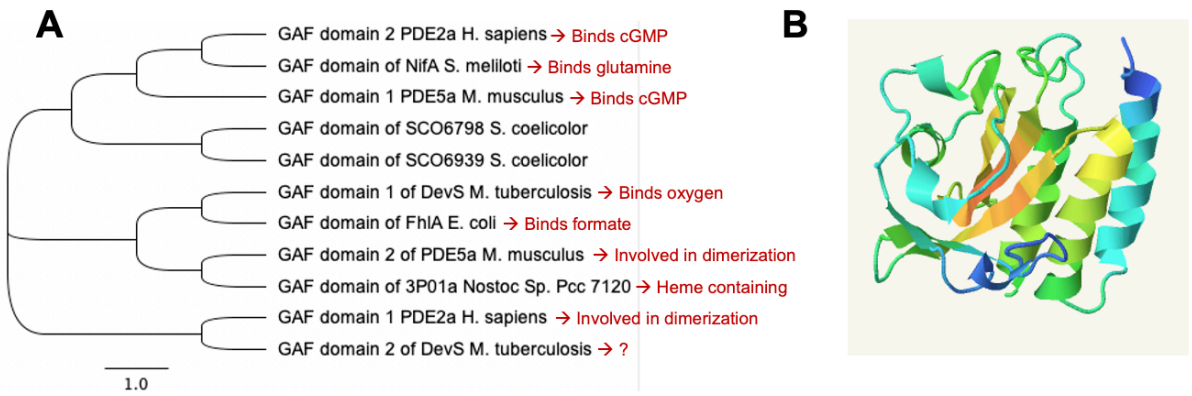


Figure 3-3 | GAF domain phylogeny and structure.
A phylogenetic tree of GAF domains with known function along with the GAF domain from CvnF7 (SCO6798) and CvnF8 (SCO6939). Known functions are labeled in red (A). The predicted secondary structure of CvnF8. Coloring is based on ProQ2 model quality assessment algorithm (B). The protein was compared with the GAF domain from a phosphoenolpyruvate-protein2 phosphotransferase (ptsp) from *Coxiella burnetiid*.

3.3.2 CvnC8 is a predicted transcriptional regulator

The *cvnC* genes contain a domain of unknown function (DUF742) that is conserved across the actinomycetes. This DUF742 falls into the helix-turn-helix (HTH) superfamily giving us a solid prediction for what this might do. HTH motifs are common to DNA and RNA binding proteins such as regulators or transcription factors. Looking at the predicted secondary structure of CvnC8, it most closely aligns with the HTH motif of the *Xylella fastidiosa* transcriptional repressor BigR. BigR represses an operon for genes promoting growth in biofilms (Barbosa and Benedetti, 2007; de Lira et al., 2018). The sequence and structural insights lead us to hypothesize that the CvnC8 protein is a transcriptional regulator that is acting at the DNA level. Data from chapter 2 show that expression of natural product biosynthetic genes in a Δ*cvnC8* mutant are upregulated in three specialized metabolite biosynthetic gene clusters, indicating that the CvnC8 protein negatively regulates these clusters. We hypothesize that CvnC8 is a transcriptional repressor of these gene clusters (Figure 3-10).

3.3.3 CvnD8 bioinformatic characterization

Another intriguing member of the conservon core genes is *cvnD* because it is predicted to be a small RAS-like GTPase. In eukaryotes, RAS GTPases are molecular switches that control essential processes such as proliferation, differentiation, cytoskeletal organization, vesicle trafficking, and gene expression (Colicelli, 2004). In fact, many cancers are caused by a

mutation in a RAS protein that decouples it from the cognate regulatory proteins leading to excessive or null signaling activity (Rajalingam et al., 2007). RAS GTPases are signal transduction components that are tightly regulated to be signaling at the appropriate time. They transition into their active signaling state by a guanine nucleotide exchange factor (GEF) protein that loads a GTP molecule onto the protein, shifting the conformation and activating the small GTPase. In their active, GTP bound form they interact with downstream regulatory proteins to relay the signal throughout the cell. RAS GTPases have a slow intrinsic rate of GTP hydrolysis so they can eventually release a free phosphate from the nucleotide and return to their inactive signaling state, bound to GDP (Colicelli, 2004). This process of turning off the signal can be induced by GTPase activating proteins (GAPs) that bind the small GTPase and stabilize the active site to promote GTP hydrolysis (Figure 3-4). These systems are widely used and are well-understood in eukaryotic organisms.

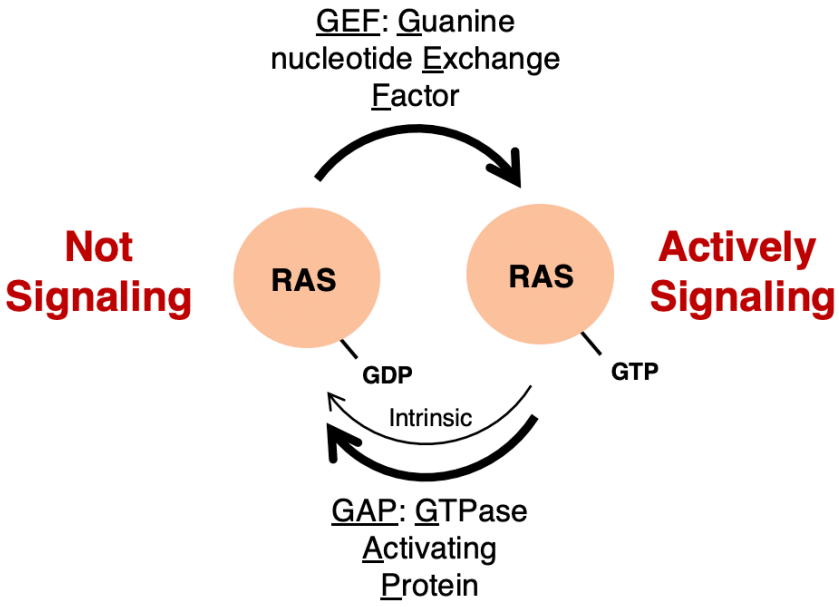


Figure 3-4 | Typical RAS GTPase systems.
Diagram of how RAS GTPases function with their cognate regulatory proteins. A guanine nucleotide exchange factors (GEF) promotes the loading of GTP onto the RAS protein to transition it to its active signaling state. The RAS GTPase has low intrinsic activity and can be regulated by a GTPase activating protein (GAP) to increases the rate of GTP hydrolysis and return the RAS to its off, or no-signaling state.

Comparably little is known about small GTPases in bacteria, even though they seem to be fairly widespread (Wuichet and Søgaard-Andersen, 2014). The most well studied bacterial small GTPase is MglA, a protein involved in switching cell polarity during gliding motility in

Myxococcus xanthus. The MglA protein has been shown to oscillate from pole to pole and this oscillation establishes the leading and lagging poles. Researchers have shown that MglA-GTP accumulates in the leading pole and is prevented from accumulating at the lagging pole by the cognate GTPase activating protein, MglB (Nan et al., 2015; Galicia et al., 2019). Sequence alignments of the CvnD8 protein with the RAS protein from *H. sapiens* and *S. cerevisiae*, the MglA protein from *M. xanthus* and *T. thermophilus*, and the MfpB protein from *M. smegmatis* show that CvnD8 has most of the key residues for the binding and hydrolysis of GTP (Figure 3-5A). This analysis, along with previous work showing the CvnD9 protein is able to hydrolyze GTP (Komatsu et al., 2006), gave us confidence to hypothesize that CvnD8 is a functional GTPase and it might function as a molecular switch in the same fashion as RAS GTPases.

Previous studies have examined the residues important for the function of the MglA protein; specifically, mutations were created to lock the MglA protein into its inactive or active form to study the effects of excessive or null signaling, respectively (Galicia et al., 2019). Mutations of the two threonine residues in positions 26 and 27 to asparagine residues correspond to the H-RAS mutation of the serine in position 17 to an asparagine. In H-RAS and MglA these mutations are predicted to create steric hinderance in the nucleotide binding pocket, specifically deeper in the pocket where GTP would occupy space when it is bound. These mutations have been shown to lower the binding affinity of the small GTPase for GTP and thus shift the GTPase to have a higher affinity for GDP, locking it into its inactive state (Zhang et al., 2010; Galicia et al., 2019). We aligned the CvnD8 predicted protein structure with the MglA protein crystal structure and saw very good structural overlap (Figure 3-5B). We used the structural and sequence alignments to identify two conserved threonine residues in CvnD8 at positions 11 and 12. I generated a mutant version, CvnD8$^{T11/12N}$, in *S. coelicolor* to study the effects of these residues on GTP/GDP binding affinity and GTP hydrolysis rate (Figure 3-5C).

An additional mutation was introduced in the MglA protein at the active site residue 82, changing it from a glutamine to a lysine. This mutation greatly reduces the GTPase activity (Zhang et al., 2010; Galicia et al., 2019). With this reduction in the MglA protein's ability to hydrolyze GTP, the small GTPase will be bound to GTP and in its active signaling state for much longer. In the structural and sequence alignments we see that this glutamine corresponds to residue 76 of the CvnD8 protein. We created a mutant version, CvnD8^{Q76L}, in which we would expect to see a similar reduction in GTPase activity if that residue is also involved in GTP hydrolysis. Lastly, we made a mutation in the glutamate residue in position 67 in the CvnD8 protein. We are not sure what this residue does as it is far from the active site, but we hypothesize that it might bind a magnesium cofactor. Nonetheless, we generated three mutant versions of the CvnD8 protein in order to begin to characterize residues important for CvnD8 function.

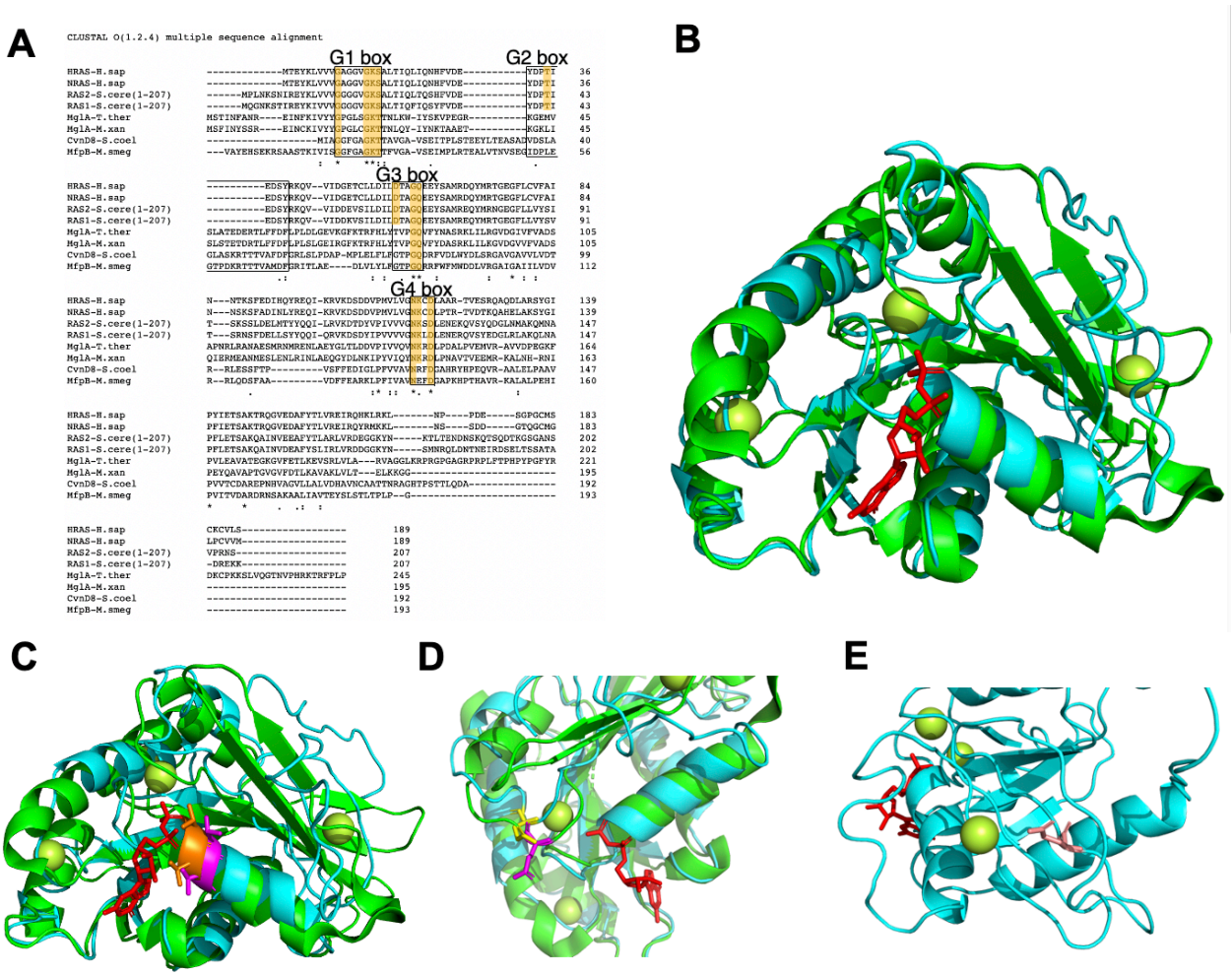


Figure 3-5 | CvnD8 sequence and structure.
Alignment of CvnD8 from *S. coelicolor* with HRAS and NRAS from *H. sapiens*, the first 207 residues of RAS1 and RAS2 from *S. cerevisiae*, MglA from *M. xanthus* and *T. thermophilus*, and MfpB from *M. smegmatis* using Clustal Omega. The G boxes are outlined with key residues highlighted (A). Phyre2 alignment of the predicted CvnD8 structure in blue with the MglA crystal structure in green bound to red GDP and coordinating Mg^{2+} ions represented as green spheres (B). The same alignment of CvnD8 and MglA is shown with MglA threonine 26 and 27 residues highlighted in magenta and the corresponding threonine residues in CvnD8 (T11, T12) highlighted in orange (C). The same alignment is shown with glutamine at position 82 in MglA highlighted in magenta and the glutamine at position 76 in CvnD8 highlighted in yellow (D). The predicted structure of CvnD8 with the glutamic acid residue in position 67 highlighted in pink (E).

3.3.4 CvnB8 is predicted to be a regulator of CvnD8

The CvnB8 protein has a roadblock/LC7 domain, which is not a well-studied protein domain. It is found in proteins associated with modulation of dynein function and researchers have speculated that this domain belongs to an ancient superfamily involved in NTPase regulation (Koonin and Aravind, 2000). We noticed that the *Myxococcus xanthus* protein, MglB has a roadblock/LC7 domain and though it only has 15.3% pairwise identity with CvnB8, they share many sites that are either strongly or weakly conserved between the two protein sequences (Figure 3-6A). The alignment of the CvnB8 predicted protein structure with the MglB protein crystal structure looked very similar (Figure 3-6B) and gave us confidence that we might be able to predict the function of the CvnB8 protein by understanding how the MglB protein functions.

The MglB protein has been shown to act as a GTPase activating protein (GAP) for the RAS-like GTPase, MglA (Baranwal et al., 2019). Unlike typical eukaryotic GAPs which contain an arginine finger that acts as an active site residue to hydrolyze GTP, MglB does not contain any active site residues. Instead it binds MglA, and a β-screw movement positions the active site residues in MglA favorably for GTP hydrolysis (Baranwal et al., 2019). The complex of MglA and MglB was recently crystalized and it showed that the C-terminal helix of one protomer of the dimeric MglB binds to a pocket far from the MglA active site and allosterically regulates the nucleotide exchange (Baranwal et al., 2019). This was surprising because it means MglB has both GAP and GEF activity, to accelerate the GTP hydrolysis rate during multiple turnovers. CvnB8 does have some extra C-terminal residues following the final fold of the roadblock/LC7 domain, but it does not have the same residues that comprise the MglB Ct-helix, which were shown to be necessary for the GEF activity of MglB (Figure 3-6A). Due to this dual functionality of MglB, it is difficult to conclude how CvnB8 might be functioning without further experiments. From the expression data of single gene deletions of *cvnB8* and *cvnD8* presented in chapter 2, we observe that sometimes they have opposite effects on specialized metabolite gene clusters as is the case for the *act* genes during single growth and the *red* genes during interactions. We also see that they can have the same effect on biosynthetic gene clusters when we look at the *cpk* genes during interactions and the *lan* genes in both environments. From the bioinformatic data, we hypothesize that CvnB8 is regulating CvnD8, but the mechanism and effect of that regulation is not clear at this time.

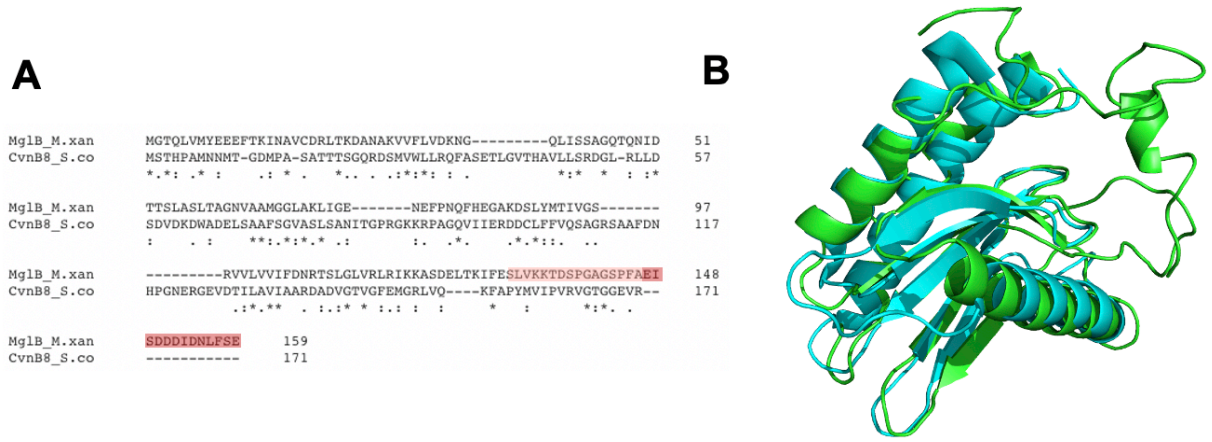


Figure 3-6 | CvnB8 and MglB alignments.
The alignment of MglB from *M. xanthus* and CvnB8 from *S. coelicolor* protein sequences. The C-terminal extension from MglB is highlighted in red with the Ct-helix highlighted in a darker red (A). The predicted structure of CvnB8 aligned with the MglB crystal structure (B).

3.3.5 Optimizing CvnD8 expression and purification

In order to assess whether CvnD8 is a functional GTPase and if it functions similarly to eukaryotic RAS GTPases, I set out to purify and test the activity of CvnD8 along with the three CvnD8 mutants that were made. We chose to use an affinity tag purification system utilizing the Strep-tag II, an eight-residue peptide sequence (Trp-Ser-His-Pro-Gln-Phe-Glu-Lys) that exhibits high affinity for streptavidin (Mao et al., 2017). We also used the polyhistidine-tag that has affinity to metal ions, such as nickel. We chose to use a tandem affinity purification method (TAP) in order to improve the purity and to ensure we purified full length protein. We used a Strep-tag II nucleotide sequence that was codon optimized for expression in *Streptomyces* and shown to be functional for use in *Streptomyces*, which will facilitate downstream, *in vivo* biochemical experiments (Mao et al., 2017). We ordered the 3xStrep tag II followed by a 6x alanine linker and a PreScission Protease cut site (to remove the tag if required for downstream studies) as a gBlock from IDT and made a transcriptional fusion to the N-terminus of the CvnD8 protein using Gibson assembly. The 9xhis-tag along with Gibson overhangs was also ordered as a gBlock with appropriate Gibson overhangs from IDT. *cvnD8* was PCR amplified from the genome. The pETDuet backbone was PCR amplified in two parts along with 20nt regions of homology for Gibson assembly. The 3xStrep-CvnD8-9xhis construct was cloned into the Multiple Cloning Site 2 (MCS2) of pETDuet. Proper assembly was confirmed via Sanger sequencing of the plasmid (Figure 3-7C). After assessing the solubility of the 3xStrep-CvnD8-9xhis construct, we decided to add a SUMO tag to improve the proportion of CvnD8 in the soluble fraction after lysis. The SUMO tag was PCR amplified from a TwinStrepSUMO3 plasmid provided by the Glaunsinger lab and cloned into pETDuet -Strep-CvnD8-His between the Strep tag and the PreScission Protease cut site via Gibson

assembly (Figure 3-7B). Both constructs were used to purify CvnD8, CvnD8$^{T11/12N}$, CvnD8^{Q76L}, and CvnD8^{E67L}.

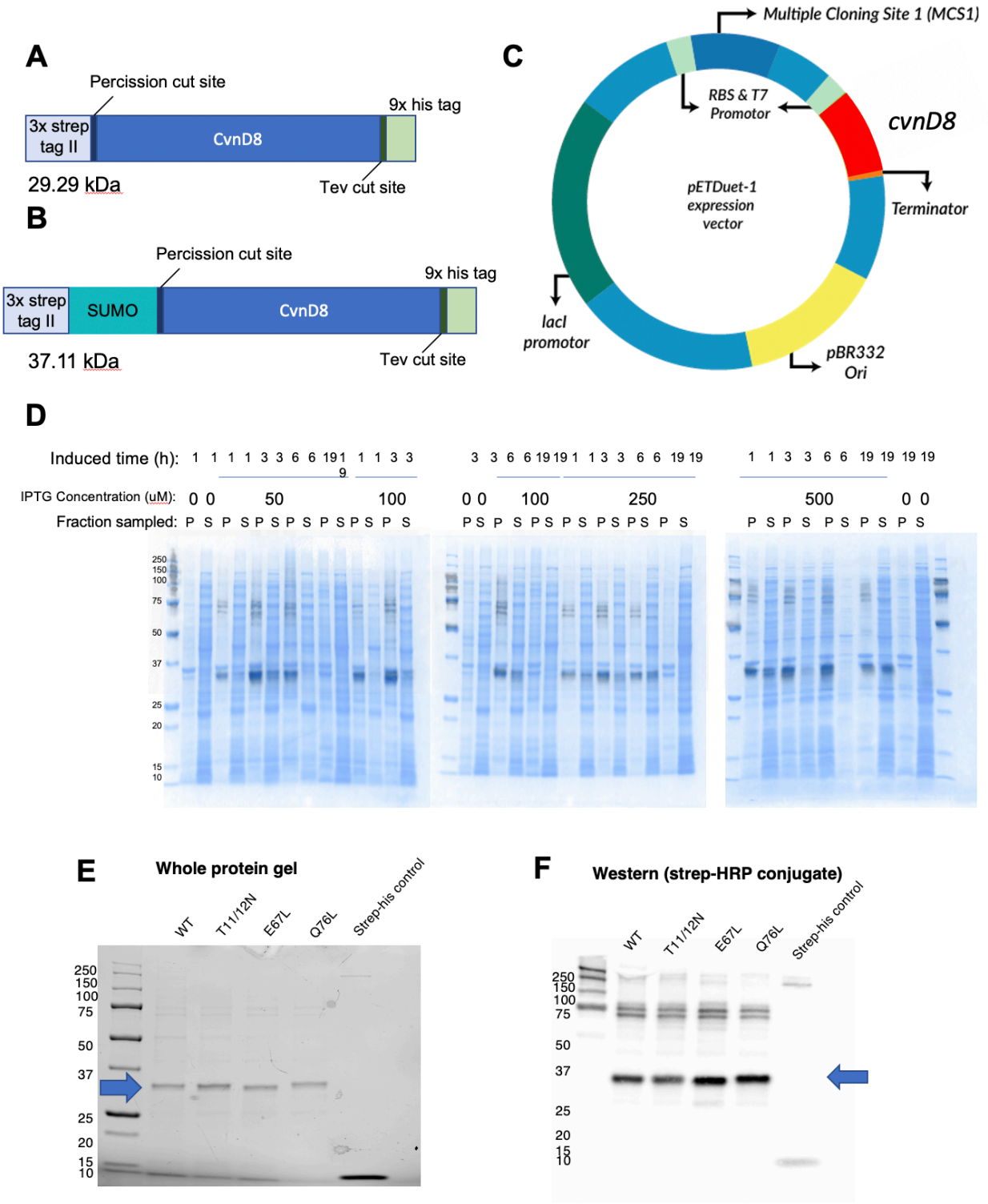


Figure 3-7 | CvnD8 expression constructs, plasmid, and purification results.
A diagram of expression constructs Strep tag II-CvnD8-His (A) and Strep tag II-SUMO-CvnD8-His (B). The pETDuet expression plasmid (C). Whole protein gel (blue) overlayed with the corresponding western blot (black) of tested expression conditions. For each condition, the samples were split into the insoluble fraction/pellet (P) or the soluble fraction (S)(D). The whole protein gel (E) of purified CvnD8 proteins and the corresponding western blot (F). The blue arrows indicate the bands containing the Strep tag II-CvnD8-His protein.

The CvnD8 expression construct in the pETDuet vector was transformed into BL21-CodonPlus(DE3)-RIL *E. coli*. Expression is driven by the *E. coli* T7 promoter and under control of the lac operator. LacI, the lac repressor, is encoded on the pETDuet vector backbone and driven by a constitutive promoter, keeping the expression of the CvnD8 construct off until the addition of the inducer, IPTG. To find the optimal expression conditions, we tested various lengths and strengths of induction. From this we found that the optimal induction strength is 500um of IPTG for 1 hour at 37°C, which resulted in the most CvnD8 protein in the soluble protein fraction as determined by western blotting with a Strep Tag II, HRP conjugated, monoclonal antibody (Figure 3-7D).

Following expression of the protein in the BL21 cells, the cells were cooled, pelleted and the pellets were stored in the -20°C freezer. Cell pellets were thawed, resuspended in lysis buffer and lysed via probe tip sonication. The lysed material was pelleted to separate the soluble and insoluble fractions. The soluble fraction was loaded onto a column packed with Strep-Tactin Superflow resin and flowed over the column three times to bind as much tagged protein as possible. The column was then washed with two column volumes of lysis buffer and two column volumes of wash buffer before being eluted with one column volumes of 2.5mM desthiobiotin. The protein was then loaded onto a column packed with Ni-NTA resin and washed with four column volumes of His wash buffer before being eluted with one column volume of buffer containing 250 mM of imidazole.

Once purification of the Strep-CvnD8-His constructs were complete, the protein was concentrated and dialyzed into assay buffer over night at 4°C. The resulting purified proteins were run on a gel and a western blot with a Strep Tag II, HRP conjugated, monoclonal antibody was performed to check the level of purity (Figure 3-7E,F).

3.3.6 CvnD8 can hydrolyze GTP

We wanted to test for GTPase activity of wild-type CvnD8 and compare the activity to that of the CvnD8 mutants to see if we could see a reduction in their ability to hydrolyze GTP, indicating the residues mutated are involved in this process. To do this, I developed an HPLC based GTPase assay that I performed with CvnD8 and the mutants.

The detection of GTP and GDP was done using an HPLC and the BioBasic anion exchange column, which has positively charged polyethyleneimine ligands bound to the silica resin. Anions interact with the resin depending on how negatively charged they are and GDP, being less negatively charged, elutes from the column earlier than GTP when we use a gradient of buffer with increasing ions to disrupt the ionic bonds. Using a set of standards to create a standard curve, we can quantify the amount of GDP and GTP present in samples that are eluted from the column. This was done with the proteins CvnD8, CvnD8$^{T11/12N}$, CvnD8^{Q76L}, and CvnD8^{E67L} that were purified previously (Figure 3-7E,F). We mixed 100uM GTP with 10uM of protein to observe the reaction over time. The GDP concentration was used to plot the activity of each enzyme over an 8-hour time window (Figure 3-8). We saw no GTP hydrolysis from the controls with no protein or the Strep-His tag construct lacking CvnD8. We did see a reduction in activity for CvnD8$^{T11/12N}$ and CvnD8^{E67L} at the first 0.5-hour time point, that grew as later timepoints were measured. We observed a reduction in CvnD8^{Q76L} beginning at the 1-hour timepoint and all timepoints beyond that.

From this experiment, we observed that CvnD8 could hydrolyze GTP and we also saw a reduction in the hydrolysis rate for the mutants. With these results, we can conclude that the CvnD8 protein is a functioning GTPase and the mutated residues seem to be partially involved in GTP hydrolysis. The reduction in activity in the mutants was not as dramatic as we were expecting, thus these conclusions come with some caveats which are addressed in the subsequent section.

Figure 3-8 | Hydrolysis of GTP by CvnD8 and CvnD8 mutants.

GTPase assay with wild-type CvnD8, CvnD8$^{T11/12N}$, CvnD8^{E67L}, and CvnD8^{Q76L} measuring the reaction product, GDP, concentration over time. The Strep tag II-His control is shown (black) along with the no protein control (grey). The assays contained 10mM of protein and 100mM GTP. Full details of the assay conditions can be found in the methods.

3.3.7 CvnD8 is a sticky protein

While optimizing the purification process, a myriad of avenues were explored to remove impurities and other unspecific proteins. It was observed that, although the proteins were passed through two affinity columns, other, larger bands could be seen in the protein gel and western blot which indicated the CvnD8 protein is not completely pure (Figure 3-7E,F). We employed many methods and techniques to try to remove these impurities, without success leading us to conclude that CvnD8 is a sticky protein. It seems that the CvnD8 protein specifically binds to proteins from the BL21 *E. coli* allowing for them to be co-purified with CvnD8. We investigated the nature of these proteins to determine if any might have GTPase activity and be influencing the GTPase assay results. This is possible, and more work is being done to disentangle the CvnD8 activity from any activity from unintentionally co-purified proteins.

We set out to determine the best way to increase the CvnD8 protein solubility and reduce the contaminants that co-purify with it. We first revisited our expression conditions and lowered the rate of expression by inducing the expression with lower concentrations of IPTG along with induction at 16°C overnight. We saw that overnight expression resulted in much less CvnD8 production, likely due to proteinase activity over the longer period of time. We saw that a slower induction led to an overall decreased amount of the CvnD8 protein with equivalent amounts of contaminants. We then tried increasing the concentration of detergent during our lysis and Strep-Tactin column wash steps in order to break up any protein-protein interactions that are happening at this stage. We saw that with the increase in detergent, more CvnD8 protein was in the soluble fractions, but the contaminants still eluted along with CvnD8. Additionally, we tried adding urea to our lysis conditions because it can disrupt protein-protein interactions and again we saw more soluble CvnD8, along with the contaminants. Next, we thought that the ratio of resin to soluble protein might not be optimal and it is possible that too much resin, with open spots for ligands to bind, might allow for more unspecific binding. To investigate this, we lowered the resin to supernatant ratio, which in turn lowered the amount of contaminants in elution, but also the total amount of the CvnD8 protein that was purified. We could observe some CvnD8 present in the flowthrough after loading the column with the supernatant, indicating we had more CvnD8 than spots on the resin for the protein to bind.

Considering that the contaminating proteins are much larger than CvnD8, we tried to separate them using size exclusion chromatography. Following the tandem affinity purification method with the Strep-Tactin and Ni-NTA columns, we loaded the proteins onto a GE HiLoad 16/600 Superdex75 pg column attached to a fast protein liquid chromatography (FPLC)

system. The resulting trace of the UV absorbance had three areas of interest which were collected, concentrated and run on a gel (Figure 3-9A,B). The first large peak is the void volume, which typically consists of very large proteins or protein aggregates that do not enter into the resin because they are larger than the resin openings. It is surprising that this is where we see some of the CvnD8 elute off the column, along with a large amount of a breakdown products. In the second region, we saw a very slight bump in the chromatogram and, given that this is where we would expect Strep-CvnD8-His to elute, we ran this region on a gel as well, but saw nothing here. Lastly, we saw a very large peak around the 120 min mark corresponding to the excess GTP that is added to the wash buffer to help stabilized CvnD8. GTP absorbs strongly in the 280nm region of UV, which is also the wavelength that is absorbed well by proteins and used to monitor proteins coming off the column. Nonetheless, we also concentrated this third peak and ran it on a gel to show there is no protein present here. Overall, this data suggests that CvnD8 is aggregating, coming off the column in the void volume, and possibly degrading by the third day post purification (Figure 3-9B).

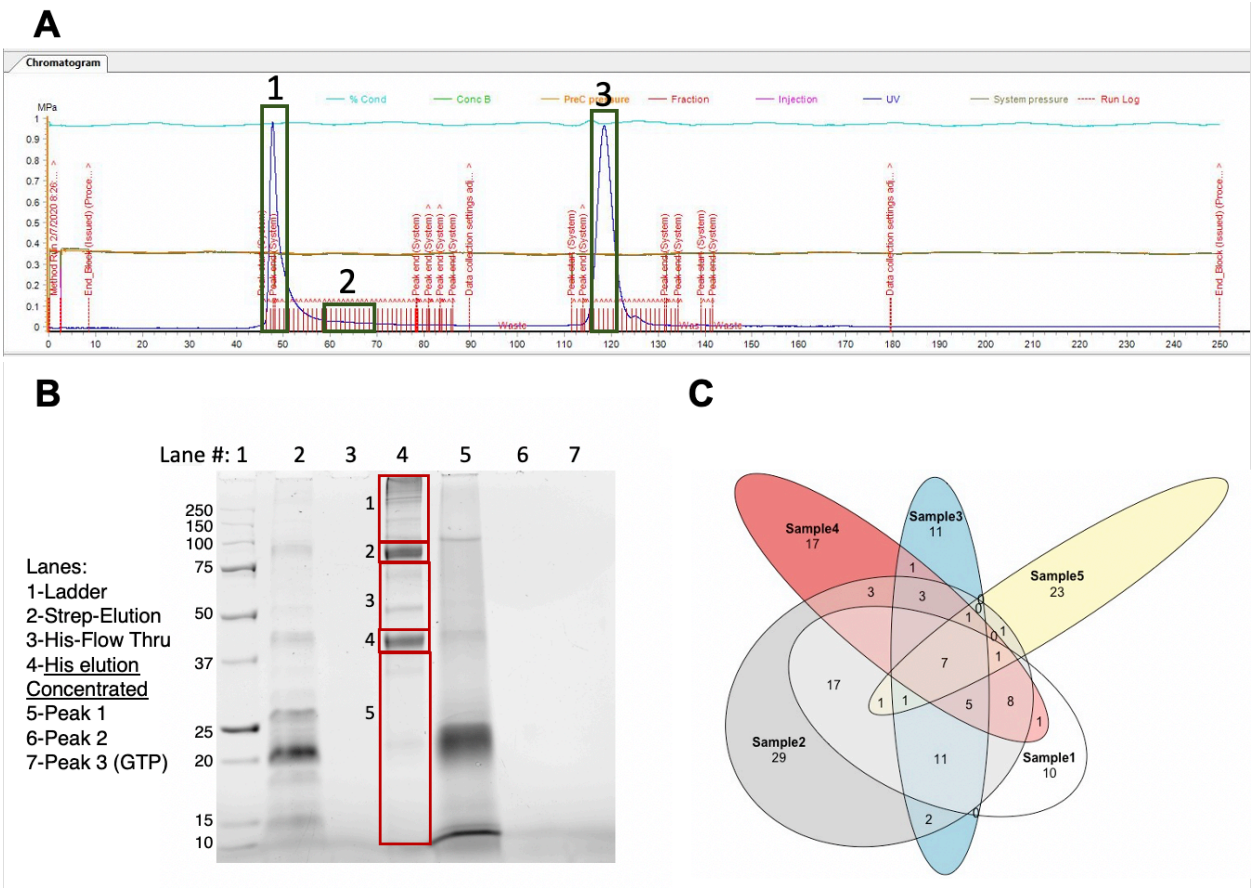


Figure 3-9 | CvnD8 is sticking to other *E. coli* proteins.
The UV chromatogram of protein separation of the Strep tag II-SUMO-CvnD8-His purification sample using a GE HiLoad 16/600 Superdex75 pg size exclusion column. Peaks of interest are boxed and labeled (A). The peaks from the size exclusion column were run on a gel (lanes 5-7) along with the elution

from the first affinity column (Strep-Tactin resin, lane 2) and the second affinity column (Ni-NTA, lane 4). The red boxes on lane 4 show the pieces of the gel that was cut out and sent for proteomics analysis (B). A Venn diagram showing the number of proteins found within and between samples (C).

Finally, in lieu of completely purifying the CvnD8 protein, we attempted to convince ourselves that the co-purifying proteins are not affecting the GTPase assay. To do this, we used proteomics to identify the contaminants that are co-purifying with Strep-SUMO-CvnD8-His. We took our purified protein sample post Strep-Tactin and Ni-NTA columns and ran the concentrated elution on a gel. From the protein gel, we cut out five regions from the sample lane to send for proteomics (Figure 3-9B red boxes). We saw a total of 147 *E. coli* K12 proteins in our sample along with the Strep-SUMO-CvnD8-His. Of those, only 60 were found in one or more samples and just 7 of these were found in all five samples. The list of the proteins found and which sample they were present in can be found in Table 2-1. Proteins of interest to us include nine proteins that have the ability to bind ATP or could have ATPase activity. These are of some concern because occasionally ATPases can also have a low level of GTPase activity, but we cannot say if they contribute to GTPase activity for sure. We also saw two elongation factor Tu (EF-Tu) proteins that were found in all five samples.

EF-Tu is a factor that delivers aminoacyl-tRNA to the A site of the ribosome in the elongation phase of protein synthesis. They function similarly to RAS GTPases and are regulated by GEFs and GAPs. As a result of the very slow release of GDP by EF-Tu, the formation of the active, GTP-bound form requires an association with its GEF, EF-Ts. EF-Tu is able to hydrolyze GTP in the absence of the ribosome but, the reaction is very slow (Hilgenfeld 2000, Maracci 2014, Rodnina 2000). We assume the EF-Tu proteins in our purified protein sample are able to hydrolyze a small amount of GTP, but at a very slow rate and with slow turnover as well. Further experimentation needs to be done in order to determine the CvnD8 GTP hydrolysis rate, excluding any contribution of EF-Tu.

Locus	Descriptive Name	sample1	sample2	sample3	sample4	sample5
SS-CvnD8-h	**SS-CvnD8-h**	1	1	1	1	1
gi|16129658|ref|NP_416217.1|	phosphoenolpyruvate synthase [Escherichia coli K12]	1	1	1	1	1
gi|16130606|ref|NP_417179.1|	recombinase A [Escherichia coli K12]	1	1	1	1	1
gi|16131218|ref|NP_417798.1|	elongation factor Tu [Escherichia coli K12]	1	1	1	1	1
gi|16131476|ref|NP_418062.1|	L-lactate dehydrogenase [Escherichia coli K12]	1	1	1	1	1
gi|16131764|ref|NP_418361.1|	glycerol kinase [Escherichia coli K12]	1	1	1	1	1
gi|16131810|ref|NP_418407.1|	elongation factor Tu [Escherichia coli K12]	1	1	1	1	1
gi|16128087|ref|NP_414636.1|	cell division protein [Escherichia coli K12]	1	1	1	1	0
gi|16128088|ref|NP_414637.1|	cell division protein FtsZ [Escherichia coli K12]	0	1	1	1	1
gi|16130124|ref|NP_416691.1|	nucleoid-associated protein NdpA [Escherichia coli K12]	1	1	0	1	1
gi|16130225|ref|NP_416793.1|	aspartate aminotransferase [Escherichia coli K12]	1	1	1	1	0
gi|16130843|ref|NP_417417.1|	S-adenosylmethionine synthetase [Escherichia coli K12]	1	1	1	1	0
gi|16131639|ref|NP_418230.1|	transcription termination factor Rho [Escherichia coli K12]	1	1	1	1	0
gi|17318569|ref|NP_006112.2|	keratin 1; Keratin-1; cytokeratin 1; hair alpha protein [Homo sapiens]	1	1	1	1	0
gi|90111564|ref|NP_417717.2|	regulator of ftsI, penicillin binding protein 3, septation function [Escherichia coli K12]	1	1	1	0	1
gi|16128009|ref|NP_414556.1|	chaperone with DnaK; heat shock protein [Escherichia coli K12]	1	1	0	1	0
gi|16128423|ref|NP_414972.1|	ATP-dependent protease ATP-binding subunit [Escherichia coli K12]	1	1	0	1	0
gi|16128698|ref|NP_415251.1|	succinate dehydrogenase catalytic subunit [Escherichia coli K12]	1	1	1	0	0
gi|16129058|ref|NP_415613.1|	3-oxoacyl-(acyl carrier protein) synthase [Escherichia coli K12]	1	1	0	1	0
gi|16129072|ref|NP_415627.1|	respiratory NADH dehydrogenase [Escherichia coli K12]	1	1	0	1	0
gi|16129170|ref|NP_415725.1|	ribose-phosphate pyrophosphokinase [Escherichia coli K12]	1	1	0	0	1
gi|16129426|ref|NP_415984.1|	cryptic nitrate reductase 2 beta subunit [Escherichia coli K12]	0	1	1	1	0
gi|16130041|ref|NP_416606.1|	phosphomethylpyrimidine kinase [Escherichia coli K12]	1	1	1	0	0
gi|16130166|ref|NP_416734.1|	DNA gyrase subunit A [Escherichia coli K12]	1	1	1	0	0
gi|16130176|ref|NP_416744.1|	sn-glycerol-3-phosphate dehydrogenase (anaerobic), large subunit [Escherichia coli K12]	1	1	1	0	0
gi|16130221|ref|NP_416789.1|	NADH dehydrogenase I chain C, D [Escherichia coli K12]	1	1	1	0	0
gi|16130442|ref|NP_417012.1|	putative pyruvate formate lyase activating enzyme 2 [Escherichia coli K12]	1	1	0	1	0

Locus	Descriptive Name	sample1	sample2	sample3	sample4	sample5
gi|16130445|ref|NP_417015.1|	hypothetical protein b2520 [Escherichia coli K12]	0	1	1	1	0
gi|16130733|ref|NP_417306.1|	PTS family enzyme I, transcriptional regulator (with NPR and NTR proteins) [Escherichia coli K12]	1	1	1	0	0
gi|16130894|ref|NP_417468.1|	hydrogenase-2, large subunit [Escherichia coli K12]	1	1	1	0	0
gi|16131066|ref|NP_417643.1|	phosphoglucosamine mutase [Escherichia coli K12]	1	1	0	1	0
gi|16131079|ref|NP_417656.1|	UDP-N-acetylglucosamine 1-carboxyvinyltransferase [Escherichia coli K12]	1	1	0	1	0
gi|16131600|ref|NP_418188.1|	ATP synthase subunit B [Escherichia coli K12]	1	0	1	1	0
gi|16131769|ref|NP_418366.1|	ATP-dependent protease ATP-binding subunit [Escherichia coli K12]	0	1	1	1	0
gi|16131845|ref|NP_418443.1|	B12-dependent methionine synthase [Escherichia coli K12]	1	1	1	0	0
gi|40354192|ref|NP_000412.2|	keratin 10; cytokeratin 10 [Homo sapiens]	1	1	1	0	0
gi|4557703|ref|NP_000414.1|	keratin 2a [Homo sapiens]	1	1	1	0	0
gi|4557705|ref|NP_000217.1|	keratin 9 [Homo sapiens]	1	0	1	1	0
gi|49176395|ref|YP_026241.1|	DNA gyrase subunit B [Escherichia coli K12]	1	1	1	0	0
gi|90111408|ref|NP_416756.4|	putative aminotransferase [Escherichia coli K12]	1	1	0	1	0
gi|16128175|ref|NP_414724.1|	lipid-A-disaccharide synthase [Escherichia coli K12]	0	1	0	1	0
gi|16128701|ref|NP_415254.1|	2-oxoglutarate decarboxylase, component of the 2-oxoglutarate dehydrogenase complex, thiamin-binding [Escherichia coli K12]	1	1	0	0	0
gi|16128703|ref|NP_415256.1|	succinyl-CoA synthetase subunit beta [Escherichia coli K12]	0	1	0	1	0
gi|16128891|ref|NP_415444.1|	condesin subunit B [Escherichia coli K12]	1	1	0	0	0
gi|16129138|ref|NP_415693.1|	cell division inhibitor, a membrane ATPase, activates minC [Escherichia coli K12]	0	1	0	0	1
gi|16129222|ref|NP_415777.1|	tryptophan synthase subunit beta [Escherichia coli K12]	1	1	0	0	0
gi|16129718|ref|NP_416278.1|	selenophosphate synthetase [Escherichia coli K12]	1	1	0	0	0
gi|16130032|ref|NP_416597.1|	PTS family enzyme IIA, galactitol-specific [Escherichia coli K12]	1	1	0	0	0
gi|16130033|ref|NP_416598.1|	putative tagatose 6-phosphate kinase 1 [Escherichia coli K12]	1	1	0	0	0
gi|16130191|ref|NP_416759.1|	hypothetical protein b2256 [Escherichia coli K12]	1	1	0	0	0
gi|16130440|ref|NP_417010.1|	4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase [Escherichia coli K12]	1	0	0	1	0
gi|16130441|ref|NP_417011.1|	hypothetical protein b2516 [Escherichia coli K12]	0	1	1	0	0

Table 3-2 | Top 52 most abundant proteins detected
The five protein gel slices from Figure 3-9B were analyzed with proteomics to determine what the contaminating proteins are. The locus, description, and samples each protein was detected in are shown. Greyed out rows are human derived proteins, assumed to be contamination during sample preparation.

From this long foray into improving the purity of CvnD8, we learned that CvnD8 is likely a sticky protein that is interacting with other proteins or protein complexes when it is expressed in *E. coli*. Currently we have created a mutant CvnD8 with residues 3-12 substituted with alanines to completely disrupt the active site of the GTPase. We are currently purifying this protein, testing the GTPase activity and we will use the assay results as a baseline measurement of GTPase activity of the co-purifying proteins. We hope to see a large reduction in activity when we assay the active site mutant, but any difference between the active site mutant and the wild-type CvnD8 can be attributed to the functioning active site of CvnD8.

3.4 Discussion and Conclusions

Considering the number of conservon systems in the *Streptomyces coelicolor* genome and the variability in some of the core and accessory conservon genes, characterizing the conservons will require considerable time and effort. Using information gathered from bioinformatic techniques, we can group the 13 conservons in three groups based on the organization of CvnA and the presence or absence of accessory genes, *cvnE* and *cvnF*. In group 1, the conservons have very large *cvnA* genes with a predicted extracellular sensing domain. The sensing domains are predicted to sense nitrate/nitrite, but dedicating six conservons to one nutrient seems excessive. We hypothesize that these sensing domains may have evolved from NIT domains as those are the most closely related characterized domains, however we think that these CvnA domains are involved in sensing different small molecules, signals, or nutrients.

In conservon group 2, there is no sensing domain in CvnA and they have a cytochrome p450 containing gene, *cvnE*, directly succeeding *cvnD*. Given the absence of a sensing domain in all the conservons with a *cvnE* (cytochrome p450), we wondered if somehow it is acting as a sensor for the conservons in group 2. Cytochrome p450s are a superfamily of enzymes that function as monooxygenases which used heme as a co-factor. They typically carry out reactions to transfer electrons between substrates (Kelly and Kelly, 2013). To our knowledge, there is no known case of cytochrome p450s acting as sensor domains in nature, but they have been employed in synthetic biosensors. Cytochrome p450s have been embedded in substrates, attached to electrodes, and used as electrochemical sensors to detect substrate presence and quantity in samples (Asturias-Arribas et al., 2014; Müller et al., 2016). Despite the fact that this has never been shown in nature, we hypothesize that the cytochrome may act as a sensor that can add an oxygen to a specific substrate and in doing so, become oxidized. This change in oxidation state may be sensed by another member of the conservon, possibly CvnA which can then relay the signal of the presence of the specific substrate that was reduced. Though we

have no experimental or concrete bioinformatic evidence for *cvnE* to function this way, we speculate that is it possible and it might be a very interesting concept to explore further in the future.

Conservon group 3 consists of just two conservons in *S. coelicolor* and are similar to group 2 in that they have a shorter CvnA protein, with no sensing domain, and they have a gene directly downstream of *cvnD*. However, this group contains *cvnF* as an additional gene. The CvnF protein has a GAF domain, which are typically involved in binding and sensing small molecules, leading us to hypothesize that CvnF acts as a sensing domain for the conservons in group 3. Based on transcriptomic data from single gene deletions, we see that the *cvnA8* and *cvnF8* deletions are phenocopies and likely function together or in a similar manner within the conservon 8 system. We think that CvnF resides in the cytoplasm, indicating it may be sensing intercellular small molecules, possible cyclic nucleotides. From this data, we confidently hypothesize that the CvnF protein acts as a sensor for CvnA in lieu of a sensing domain within CvnA itself.

We also hypothesize that the CvnC8 protein binds DNA and acts as a repressor for multiple biosynthetic gene clusters. Given the similarity in the expression patterns of the *⊿cvnC8* and *⊿cvnD8* mutants in chapter 2, we think that CvnC8 is positively regulated by the active form of CvnD8, bound to GTP (Figure 3-10). From our expression data, CvnB8 does not seem to be acting as a negative regulator of CvnD8 because if that were the case, we would expect to see opposite phenotypes of the deletions of *cvnB8* and *cvnD8*, which is not what we observe overall. Instead, in the case of the cryptic lanthipeptide and coelimycin, we see that CvnB8 and CvnD8 regulate the biosynthetic genes in the same manner. This leads us to think the CvnB8 protein is acting as a GEF in these situations to positively regulate CvnD8 and promote the nucleotide exchange to active CvnD8 (Figure 3-10). Due to the intermediate phenotype of the *⊿cvnB8* mutant between the *⊿cvnA/F8* and *⊿cvnC/D8* mutants, we hypothesize that it acts a connection between these two different portions of the conservon system. Though this model is very speculative and many of our hypotheses remain to be tested, it is the best description of the system we have for now, given our data.

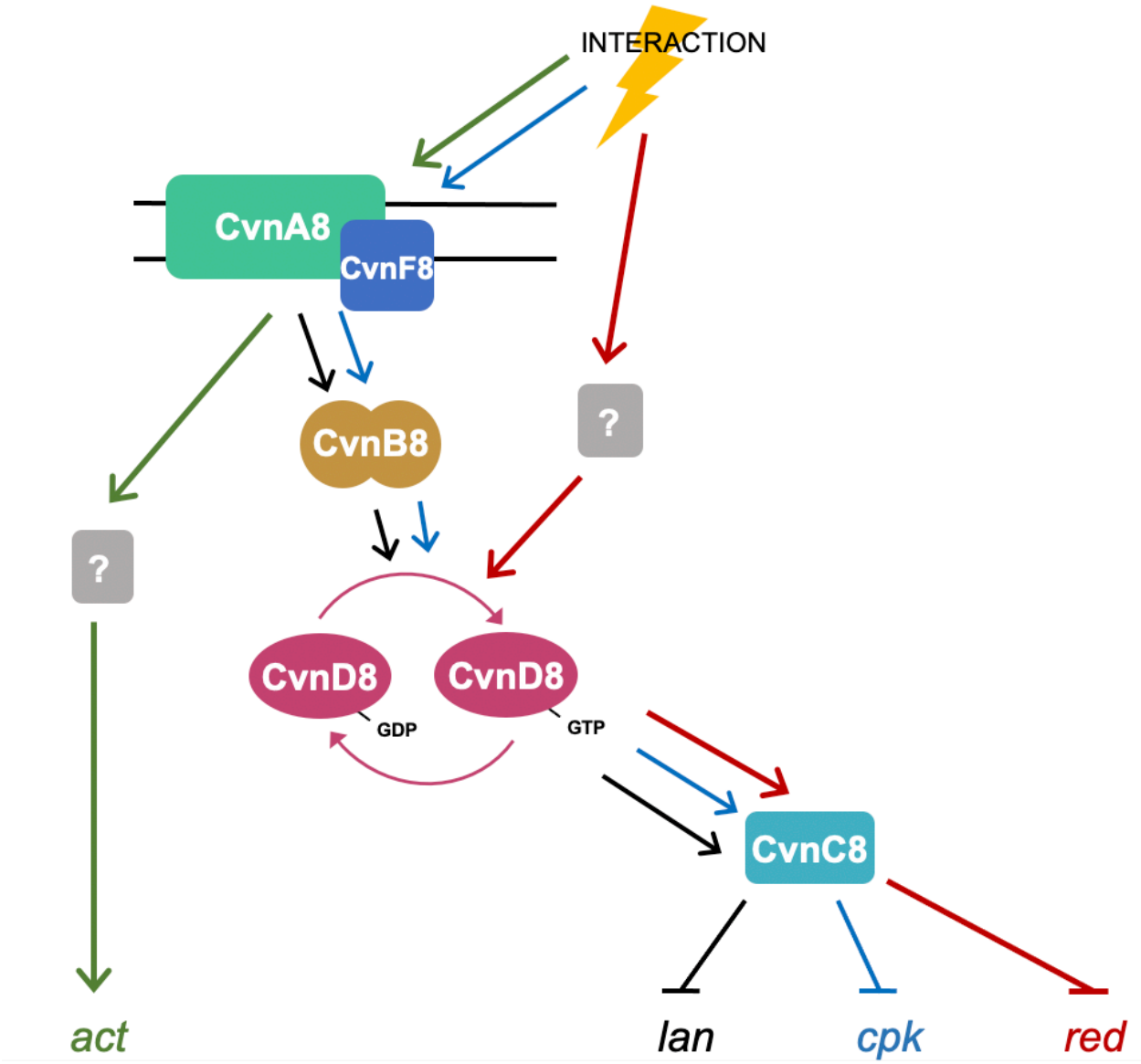


Figure 3-10 | A speculative model of conservon 8 regulation during interactions with *Amycolatopsis* sp. AA4.
The conservon 8 components are shown along with colored arrows for how the system seems to function for each of the four specialized metabolite biosynthetic gene clusters listed at the bottom. In green, CvnA8 and CvnF8 positively regulate the *act* genes through an unknown factor(s), and even more so when in an interaction. In black, all the conservon 8 genes function to negatively regulate the *lan* genes regardless of an interaction environment. In blue, during an interaction with *Amycolatopsis* sp. AA4, all the conservon 8 genes function to negatively regulate the *cpk* genes. We think the *lan* and *cpk* regulation happens by flowing first through CvnA/F8, which positively regulate CvnB8. We speculate that CvnB8 positively regulates CvnD8, possibly though GEF activity, which then stimulates CvnC8 to repress the *lan* and *cpk* genes. Lastly,

> we think the presence of an interspecies interaction stimulates another factor/sensor to activate CvnD8, which in turn positively regulates CvnC8 to repress the *red* genes.

Lastly, we ventured into protein purification and biochemistry to begin to characterize the CvnD8 protein, a predicted RAS-like GTPase. We learned that CvnD8 appears to be a sticky protein, bringing with it many other co-purifying proteins even after using a tandem affinity purification approach. We observed from our size exclusion chromatography data that CvnD8 is likely present as an aggregate or protein complex once purified. Characterizing this protein will take more work and making any concrete conclusions at this point is difficult. We did see GTPase activity of CvnD8 when it was mostly pure, and the activity was reduced with mutant versions lacking the optimal residues for GTP hydrolysis. This gives us confidence to say that CvnD8 is a functioning GTPase, but what residues are important for that and whether it functions similar to RAS GTPases remains to be explored.

4 Chapter 4. A forward genetics approach to identify additional genes involved in *S. coelicolor* interspecies interactions

4.1 Introduction

When thinking about how microbial genetics was done in the early days, researchers found incredibly simple and elegant ways to identify genes important for various microbial functions. None is more reliable and fruitful than the forward genetic screen to identify genes involved in a non-selectable phenotype or function. Nowadays, we have so many more genetic tools and technology to do very in-depth studies on cellular structure and function, but the classic techniques still hold a candle to the new ones, especially in more difficult to work with, or even genetically intractable organisms. The easiest, non-lethal, phenotypes to study are those that are visible because they make the screening process easier. Luckily, *Streptomyces coelicolor* has a clear visible phenotype when it is producing natural products as a response to interspecies interactions. We are interested in studying other genes involved in sensing and responding to interspecies interactions and can do so through forward genetics. The conservon 8 system identified in chapter 2 is involved in responding to interspecies interactions, but we expect that interactions are extremely complex. Thus multiple genes are likely involved in this process and we set out to identify additional genes using a genetic screen.

As we set out to search for additional genes involved in interspecies interactions in *Streptomyces coelicolor*, along with the reverse genetics approach outlined in chapter two, we set up a forward genetic screen. For mutagenesis, we attempted to use transposons due to their ease and the recent publication of one that was reported to work well in *Streptomyces*. We were unsuccessful with transposons in an acceptable timeframe, so we used a chemical mutagen to introduce point mutations in the *S. coelicolor* genome. We designed a genetic screen to identify mutants of interest that had altered interaction pigmentation intensities and distributions. We identified 27 unique mutations in 20 different genes contributing to normal pigmentation distribution during interspecies interactions. With this, we generated hypotheses about other regulators/proteins that may function in sensing and responding to interspecies interactions. This data set, combined with prior work, highlights some candidate genes that will be further studied.

4.2 Methods

4.2.1 Chemical mutagenesis and screen

Approximately 5.15x10^7 *S. coelicolor* spores in 10 ul of ddH_2O (distilled deionized water) were added to 750 ul of 0.01 M KPO_4 at pH 7.0 followed by centrifugation at 8000g for 3 min. The supernatant was removed and another 750 ul of 0.01 M KPO_4 at pH 7.0 was added. Next, the spores were resuspended by vortexing for 30 seconds. 25 ul of ethyl methanesulfonate (EMS,

Sigma M0880) was added to the spores, while 25 ul of sterile ddH_2O was added to the control. The spores containing the EMS or ddH_2O were shaken at 200 rpm, 30°C for the given amount of exposure time. The spores were centrifuged at 8000g for 3min and resuspended by vortexing in 1 mL, freshly made and filter-sterilized 5% w/v sodium thiosulfate solution to inactivate the EMS. Again, the spores were centrifuged at 8000g for 3min and resuspended by vortexing in 1 mL sterile ddH_2O. The water wash was repeated once more before a dilution series was made with each sample and plated on ISP2-agar or mixed with *Amycolatopsis* sp. AA4. The mixed cells were then plated on 50% strength ISP2-agar for the screen (Practical Streptomyces Genetics | NHBS Academic & Professional Books). Our *Amycolatopsis* sp. AA4 stock was approximately 1.6 x 10^{10} cells/mL, and we diluted it to 10^{-4} then diluted it again 1:50 before plating 100 ul of the dilution of *Amycolatopsis* sp. AA4 onto 150 mm diameter Petri dishes to give us approximately 300 CFU (colony forming units) per plate. For the 5 min EMS exposed *S. coelicolor* spores, 5 ul of spores were added to 1.396 mL of sterile ddH_2O and 20 ul was plates on each plate. For the 30 min EMS exposed *S. coelicolor* spores, 20 ul of spores were added to 780 ul of sterile ddH_2O and 20 ul was plates on each plate. We used only 50% strength ISP2-agar in order to lower the concentration of nutrients in the agar plates, resulting in quicker induction of pigmented metabolites.

Plates were incubated for 3-5 days at 30°C and checked on the third, fourth, and fifth days to observe interaction phenotypes. Strains of interest were picked and restreaked on ISP2-agar plates 2-3 times to purify the strain and observe the phenotype when grown in the absence of any other bacterium. Strains that looked pure and grew normally were streaked on MS-agar (Practical Streptomyces Genetics | NHBS Academic & Professional Books) in order to observe the developmental phenotype and collect spores of the mutant strain. Secondary screens were performed by quadrant streaking the strains on MS-agar, R2YE-agar, and ISP2-agar to observe growth in dense biomass and also as single colonies on a variety of media. Additionally, the mutant *S. coelicolor* strain were put into interactions again. The spore stocks were diluted to 2x10^8 CFU/mL and 0.5ul was spotted onto a 60x15 mm Petri dish with 4 ml ISP2-agar, approximately 2 mm thick and put into interactions. 0.5 ul of an *Amycolatopsis* sp. AA4 stock (8x10^8 CFU/mL) was spotted 0.75 cm away from the *S. coelicolor* spot. Interactions were grown for 4 days at 30°C, then imaged.

4.2.2 Whole-genome sequencing and analysis

High quality genomic DNA was isolated by standard procedure (Practical Streptomyces Genetics | NHBS Academic & Professional Books). Purity was confirmed using DNA gel electrophoresis, and a NanoDrop One UV-Vis Spectrophotometer to measure the 260/280 and 260/230 absorbance ratios. The concentration was accurately quantified on a Qubit fluorometer using an Invitrogen Qubit DNA-HS Assay Kit (Q32851). Samples were prepared by the UC Berkeley Functional Genomics Laboratory using the Illumina Nextera XT DNA sample preparation kit and sequenced on an Illumina HiSeq 4000 using 150 bp paired-end reads. Mutation identification and analysis was done using breseq (version 0.35.4) (Deatherage and Barrick, 2014).

4.2.3 Transposon plasmid construction and conjugation

The transposon plasmids were built with a system similar to the 'magic pools' method described in Liu et al., 2018. pJW16, a plasmid with a ColE1 origin of replication, the *cat* gene conferring resistance to chloramphenicol, and lacking the green fluorescent protein (GFP), adapted from pJW52 (Liu et al., 2018) was used as a holding vector for each part of the transposon plasmid. This allows for combinatorial construction of multiple different plasmids, switching out one or multiple parts using Golden-gate assembly with Type IIs restriction endonucleases. We used the BbsI (BpiI) enzyme for Golden-gate assembly, therefore we made sure to remove any BbsI sites from the parts of transposon plasmid. We added the BbsI sites on the ends of each part we cloned into pJW16, insuring the 4 bp sticky ends would match with the part it was meant to assembly specifically with. Each part of the transposon plasmid (Figure 4-3) was cloned into pJW16 using Gibson assembly.

For part 1, the codon optimized Tn5 or himar1 transposases and the first inverted repeat were ordered as a gBlocks with overlapping Gibson ends allowing for direct cloning into pBB2, along with the pUC18 origin of replication from pSET152 (Bierman et al., 1992). Part 2 and 3 were considered just one part for our purposes, because we had a promoter we knew would effectively express the antibiotic resistance gene successfully in *Streptomyces coelicolor*. Part 2/3 was amplified from pSET152 and Gibson overhangs were added via PCR. For part 4, the streptomycin resistance gene (*aadA*) was PCR amplified from pIJ778 (Gust et al., 2003) and the origin of transfer (*oriT*) was PCR amplified from pSET152 (Bierman et al., 1992). Both were Gibson assembled into pJW16. Part 5, the promoters driving the transposase gene, were ordered as primers or gBlocks with Gibson overhangs and assembled into holding vectors.

Once all the parts were made, we assembled them in different combinations into the transposon plasmids using Golden-gate assembly. We mixed one version of each of the four parts together to give a total of 0.1 pmols of each of the digested fragments. Then we mixed the four parts with 2 ul 10X T4 ligase buffer (NEB B0202S), 1 ul T4 DNA ligase (NEB M0202), and 1 ul BbsI (NEB R3539) to give a final reaction volume of 20ul. The reaction was cycled between 1 min at 37°C and 1 min at 16°C to alternate between the BbsI optimal temperature and the T4 ligase optimal temperature for 30 cycles. 1 ul of reaction was transformed into top10 *E. coli* and plated on apramycin (50 ug/ml) and spectinomycin (50 ug/ml) LB agar plates.

Conjugation was performed with *E. coli* ET12567 containing the mating plasmid pUZ8002 as described previously (Practical Streptomyces Genetics | NHBS Academic & Professional Books). *S. coelicolor* transposon mutants were plated on MS media containing apramycin (50 ug/ml) for selections of transposon mutants.

4.3 Results

4.3.1 Identification of genes involved in interspecies interactions

In order to perform a forward genetics screen to determine genes involved in *S. coelicolor*'s response to interspecies interactions, a collection of *S. coelicolor* mutants was generated. We used the chemical mutagen ethyl methanesulfonate (EMS), which introduces single nucleotide polymorphisms through guanine alkylation via interactions between the ethyl group and the guanine nucleotide (Sega, 1984). Our goal was to create *S. coelicolor* strains with as few mutations as possible so we would have an easier time identifying the causal mutations in our mutants of interest. Given this, a killing curve was generated to determine the optimal exposure times that would result in 50% and 90% killing. The optimal percentage of killing can be determined by knowing the number of essential genes in the genome. As this number is unknown for *S. coelicolor*, it was calculated that if 25% of all genes were essential in the genome, then at the 25% level of killing we would expect one mutation per spore and the 43.75% level of killing we would expect three mutations per spore. I exposed $5x10^7$ *S. coelicolor* spores to the ethyl methanesulfonate (EMS) and harvested the spores at different time points to generate killing curves (Figure 4-1B). From the killing curve we decided to generate mutations at the 63% and 92% killing levels that resulted from 5 min and 30 min of exposure respectively (Figure 4-1B).

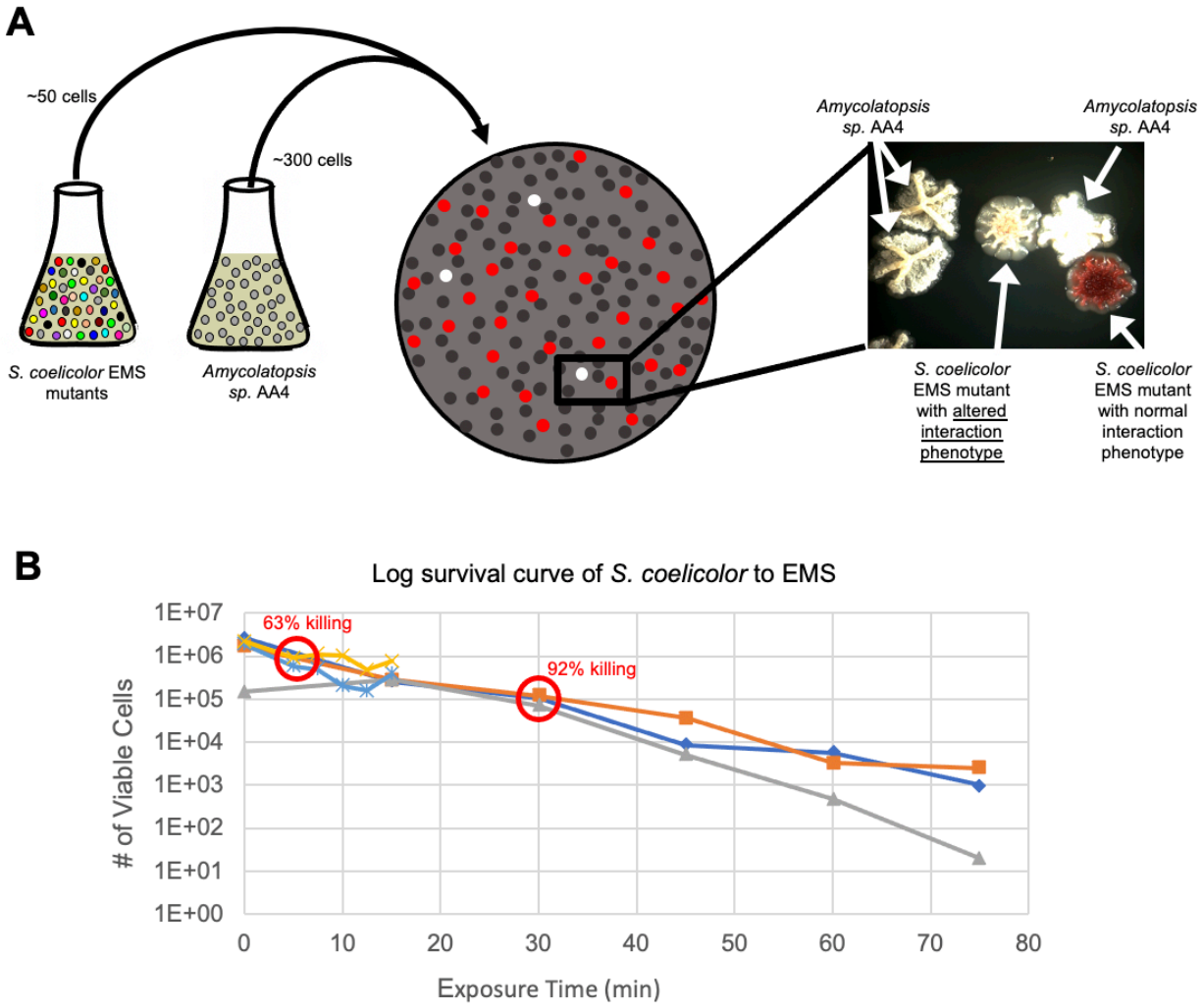


Figure 4-1 | **Genetic screen to identify genes involved in specialized metabolite induction during interspecies interactions.**
S. coelicolor germlings that have been exposed to ethyl methanesulfonate (EMS) are mixed with *Amycolatopsis* sp. AA4 in a ratio of 1:6 and plated on 50% ISP2 agar. The plates were visually screened for *S. coelicolor* colonies that had an altered interaction phenotype, with specific attention given to pigmentation intensity and distribution (A). The ethyl methanesulfonate (EMS) survival curve of *S. coelicolor* spores that were exposed for varying times (B).

To perform the screen for mutants of interest, the mutagenized *S. coelicolor* spores were mixed with *Amycolatopsis* sp. AA4 in liquid at a roughly 1:6 ratio. Once mixed, the strains were diluted and plated on large, 150 cm Petri dishes containing 50% ISP2-agar. The higher ratio of *Amycolatopsis* sp. AA4 to *S. coelicolor* and the density of colonies on the plates ensures that each *S. coelicolor* mutant will be able to interact with at least one colony of the *Amycolatopsis* initiator

strain (Figure 4-1A). In the screen, I looked for mutants that were lacking pigmentation, which means they fail to produce both pigmented natural products, actinorhodin and undecyl-prodigiosin. I also picked out mutants with more pigmentation or different pigmentation patterns than what we see during a wild-type *S. coelicolor* interaction with *Amycolatopsis* sp. AA4. These pigmented metabolites serve as a proxy for the induction of secondary metabolism. Furthermore, monitoring the loss or change of two pigmented natural products is advantageous as I can pick out genes involved in the global regulation of secondary metabolism rather than mutations in specific secondary metabolite biosynthetic pathways.

Multiple rounds of chemical mutagenesis and screening were done until a sufficient number and variety of mutants were collected. Overall, 250 mutants were selected for secondary screening, and 76 mutants of interest were identified subsequent to that. Secondary screening involved growth on sporulation media (R2YE and MS) to check for growth rate and production of aerial hyphae and spores. The mutants were also placed in interactions with normalized spore stocks to ensure they did in fact have an altered interaction phenotype in a controlled, isolated environment. The mutants were then grouped based on their phenotypes in the secondary screens and representatives from each group were selected for sequencing. In total, 22 mutants were selected for whole genome sequencing to determine the location of point mutations in the genome. The interaction phenotypes of the mutants varied greatly, from overly pigmented, to yellow, to greatly reduced in pigmentation (Figure 4-2A). We purposefully selected a high number of *S. coelicolor* mutants which had a considerable reduction in pigmentation during interactions (18 of 22) as we predicted these would have mutations in genes important for the sensing or the induction of pigmentation during interspecies interactions. When deciding on the mutants to select for genome sequencing, we also prioritized mutants that could raise aerial hyphae, sporulate, and produce pigmented antibiotics on the various media. This ensured that the mutants did not have more general problems with growth and development causing altered pigment production in interactions.

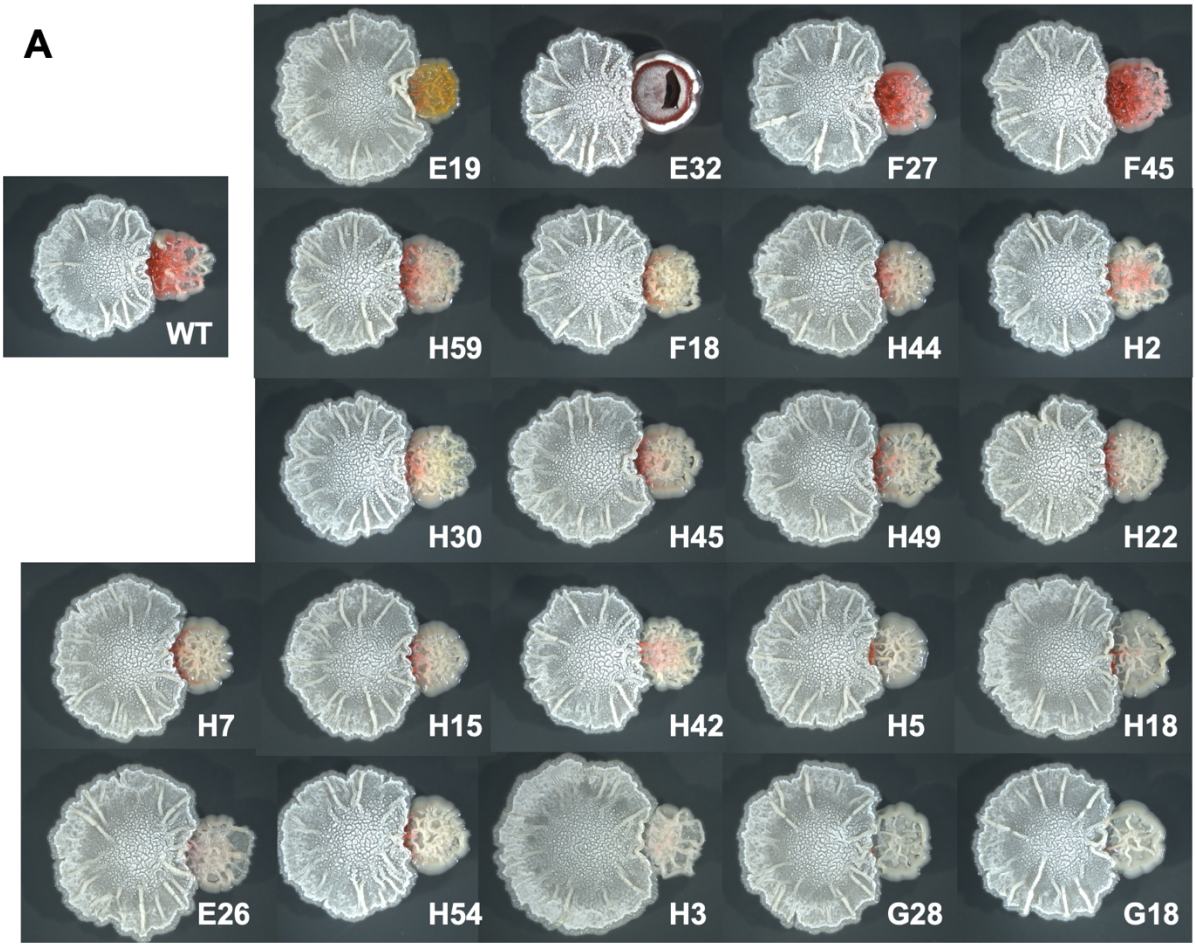


Figure 4-2 | **Sequenced *S. coelicolor* mutants.**
The subset of *S. coelicolor* EMS mutants whose genomes were sequenced with their labeled strain numbers (A). A bar plot shows the number of strains that had differing amounts of mutations (B).

4.3.2 Exploring and prioritizing genes involved in interspecies interactions

From the 22 *S. coelicolor* EMS mutants that were sequenced, we observed a total of 27 unique mutations in 20 genes or intergenic regions across the genome that are likely to disrupt a gene or promoter (Table 4-1). We sequenced 17 strains from the 5-minute EMS exposure time point and 5 strains from the 30-minute timepoint which had an average of 2.1 and 3.2 mutations per genome, respectively. There were seven strains with just a single mutation from the 5-minute exposure time point (Figure 4-2B). Two genes, *dnaA* (SCO3879) and SCO4215, were found to have multiple disruptions providing strong evidence that they are important for proper pigmentation induction during interspecies interactions. We also found that some of the mutants harbored only a single mutation in *dnaA* and SCO4215, providing even stronger evidence for their importance.

4.3.2.1 DnaA (SCO3879)

Strains of *S. coelicolor* with single mutations in the *dnaA* gene showed reduced pigmentation during interspecies interactions (strains H5, H7, H22 and more in Figure 4-2A). DnaA controls the initiation of bacterial chromosomal replication by forming an open initiation complex through interactions of DnaA bound to ATP with the DnaA boxes at the oriC. The DnaA proteins of *Streptomyces*, like those of other bacteria, contain four domains consisting of a divergent N-terminal sequence, a linker, an ATP-binding domain, and a DNA-binding region at the C terminus (Lee et al., 2002). The *dnaA* gene is essential in *Streptomyces*, just like in other bacteria, so it was surprising for us to find so many mutations in *dnaA* during our screen. First, we can conclude that the mutations that we observed in *dnaA* do not abolish its function. They do however affect the ability of *S. coelicolor* to produce actinorhodin and undecyl-prodiginine during interspecies interactions. Of the seven mutations in *dnaA*, five of them fall within the ATPase domain and two of them fall in the C-terminal DNA-binding region indicating that binding DNA, binding ATP, or hydrolyzing ATP is altered in these mutants.

4.3.2.2 SCO4215

S. coelicolor strains with a single mutation in SCO4215 showed reduced pigmentation during interspecies interactions (strains H3, and H42 in Figure 4-2A). SCO4215 is predicted to be a GntR transcriptional regulator. The GntR family of transcriptional regulators are named after the gluconate operon repressor in *Bacillus subtilis* and are known to regulate a variety of biological processes (Aravind and Anantharaman, 2003; Suvorova et al., 2015). They have a highly conserved, N-terminal, winged helix-turn-helix DNA binding domain and a diverse ligand binding C-terminal domain which is involved in effector binding or oligomerization. It is this second domain that gives rise to the sub families within the GntR superfamily. SCO4215 has a UTRA domain (pfam07702) which is predicted to bind small molecules. The sequence variance within the binding pocket of the UTRA ligand-binding domains is consistent with the capacity of these transcription regulators to respond to diverse small molecules, such as fatty acids (FarR), alkylphosphonate (PhnF), histidine (HutC), and sugars (TreR) (Aravind and Anantharaman, 2003). For now, we cannot predict what small molecule SCO4215 might be sensing, though we do know that it is important for normal pigment production during interspecies interactions.

Gene	Annotation	Type of mutations	Strains with (single) mutation	log2 interaction/ single
SCO3879 (dnaA)	chromosomal replication initiation protein	G618D (GGC→GAC), R589C (CGC→TGC), A511T (GCC→ACC), W473G (TGG→GGG), E464Q (GAG→CAG), T462S (ACC→AGC), I399V (ATC→GTC)	G18, G28, H18, H44, H45, (H5, H7, H15, H22, H49, H54)	-0.5934608
SCO4215	GntR family transcriptional regulator	L34R (CTC→CGC), coding (647/759 nt)	H2, (H3, H42)	-0.6537087
SCO4007	hypothetical protein	F220L (TTC→CTC)	H44, H45	-0.4408048
SCO0279	glycosyl hydrolase	coding (1076-1147/ 1272 nt)	E26	-0.0038362
SCO1051	hypothetical protein	coding (1040/1074 nt)	E26	-0.2262984
SCO1062	hydrolase	G557S (GGC→AGC)	H59	0.38153379
SCO1146, ***SCO1147***	lipoprotein/ABC transporter transmembrane subunit	intergenic (-166/-63)	H2	1.2426035
SCO1181	hypothetical protein	D108N (GAC→AAC)	F45	-0.4983441
SCO2121	two-component system sensor kinase	coding (31-41/1242 nt)	H30	-0.5260726
SCO2192	regulator	E443A (GAG→GCG)	F27	-1.5641943
SCO2263, ***SCO2264***	hypothetical protein/hypothetical protein	intergenic (+316/-84)	F45	0.47946618
SCO2868	hypothetical protein	coding (21-22/480 nt)	E32	-1.1196616
SCO3272	DNA-binding protein	W119* (TGG→TGA)	E32	0.07649952
SCO3671 (dnaK)	molecular chaperone DnaK	K55Q (AAG→CAG)	E26	0.75962608
SCO3883	inner membrane protein translocase component YidC	H148Y (CAC→TAC)	G18	-0.0419026
SCO4677, *SCO4678*	regulatory protein/hypothetical protein	intergenic (-52/-61)	H30	-2.6065546
SCO5800	hypothetical protein	T268I (ACC→ATC)	F27	1.15491847
SCO5892	polyketide synthase	coding (1842/6894 nt)	F18	1.47905807
SCO5955	hypothetical protein	N115Y (AAC→TAC)	E19	-0.317652
SCO7690	ABC transporter ATP-binding protein	R581Q (CGA→CAA)	H30	3.12061213

Table 4-1 | Mutations identified in strains with *S. coelicolor* EMS mutants with altered interaction phenotypes.

Genes identified as having mutations in the sequenced *S. coelicolor* mutant strains. In the 'Type of mutations' column, the missense mutations are shown in blue with the nucleotide polymorphism highlighted in red. An intergenic mutation's location relative to each gene is indicated to be upstream (-) or downstream (+) the specified number of nucleotides. Coding mutations are insertions or deletions of base pairs at the location specified. The '$\log_2$ interaction/single' column indicates the $\log_2$ fold change of each gene during wild-type *S. coelicolor* in interactions compared to when grown as a single colony. Red indicates lower expression in the interaction and blue indicates higher expression in the interaction. The intensity of red or blue reflects the amount of increased or decreased expression. The bolded cells in the '$\log_2$ interaction/single' column indicate which genes were significantly different from wild type (t-test, $p < 0.05$, $n=3$).

4.3.2.3 SCO5892

We also found mutations in genes from natural product biosynthetic gene clusters. We observed one mutation in the core polyketide synthase gene in the prodiginine biosynthetic gene cluster, SCO5892. A strain with multiple mutations, including one in SCO5892 (F18 in Figure 4-2A) showed a reduction in pigmentation during interspecies interactions. Our screen relied on the production of two pigmented antibiotics, actinorhodin and undecyl-prodigiosin, to try to minimize hits in the biosynthetic genes producing those antibiotics and this was the only mutation found in either of the gene clusters, indicating that our screening method was successful.

4.3.2.4 SCO5800

We also observed a mutation in a cryptic siderophore synthase component, SCO5800, which is not associated with a known compound. The strain with multiple mutations, including one in SCO5800 (F27 in Figure 4-2A) showed an increase in pigmentation during interspecies interactions. Siderophores are iron scavenging molecules that are secreted by bacteria. They can bind to Fe^{3+}, solubilizing it and allowing for uptake into the cell. The strain harboring this mutation had another missense mutation so we cannot say for sure if this mutation is affecting pigmentation in *S. coelicolor* during interactions. It has been shown that siderophores play important role is *Streptomyces* during interactions so this is a gene that should be investigated further (Yamanaka et al., 2005; Traxler et al., 2013; Lee et al., 2020).

4.3.2.5 SCO7690

Additionally, there was a mutation in an ABC transporter in the coelibactin gene cluster, predicted to be involved in the cellular export of coelibactin, SCO7690. A strain with multiple mutations, including one in SCO7690 (H30 in Figure 4-2A) showed a reduction in pigmentation during interspecies interactions. Coelibactin is a non-ribosomal peptide whose

structure was predicted computationally from the non-ribosomal peptide synthetase gene (Bentley et al., 2002). Coelibactin itself has never been identified, but it is predicted to have zincophore activity. The expression of coelibactin genes has been shown to abolish specialized metabolite production, indicating it can regulate specialized metabolism (Hesketh et al., 2009; Kallifidas et al., 2010; Gomez-Escribano et al., 2012). Both SCO5800 and SCO7690, found in specialized metabolite biosynthetic gene clusters, are also upregulated during interspecies interactions compared to when they are grown alone. This indicates that they are involved in the response to the interspecies interactions (Table 4-1). Specifically, the coelibactin gene cluster (SCO7676-7692) has an average $\log_2$ fold change of 1.71 during interactions compared to single colony growth. Unlike what was seen in previous experiments, where expression of coelibactin abolished production of the pigmented antibiotics, here we see higher production of these antibiotics and increased expression of coelibactin, indicating the negative regulation is overcome during interspecies interactions (Kallifidas et al., 2010).

4.3.2.6 SCO4677

The same strain with a mutation in SCO7690 (H30 in Figure 4-2A) also had a mutation in SCO4677 and had an interaction phenotype with a reduction in pigmentation. SCO4677 is predicted to be a small, 144 amino acid, protein with a histidine kinase-like ATPase domain (pfam 13581) that spans nearly the entire protein. It is predicted to be similar to the histidine kinase found in RsbW, a sigma-B anti sigma factor in *B. subtilis*. The gene appears to be transcribed on its own so we cannot derive information from the surrounding genes. We also saw that the expression of SCO4677 was much lower in *S. coelicolor* colonies in interactions compared to when grown alone. Therefore, we hypothesize that it might play a role in regulating a sigma factor in *S. coelicolor*.

4.3.2.7 SCO2192

Another gene of interest identified during this screen that was also down regulated during interspecies interactions was SCO2192. A strain with multiple mutations, including one in SCO2192 (F27 in Figure 4-2A) showed an increase in pigmentation during interspecies interactions. The predicted protein has no detectable conserved domains preventing us from predicting how this might be functioning. Nonetheless, hypothetical genes with no predicted domains are an alluring avenue for discovery of novel proteins.

4.3.2.8 SCO1147

Lastly, SCO1147 is an ABC-type multidrug transport system, which was found mutated in one of the *S. coelicolor* strains (H2 in Figure 4-2A) identified during the screen and upregulated during interspecies interactions. The strain showed a reduction in pigmentation during interspecies interactions. The identification of this gene in the screen indicates that *S. coelicolor* might be pumping out its own drugs or pumping out drugs from the interacting strains during interactions. This process might be crucial for *S. coelicolor* to compete and effectively use the natural products it is producing.

4.3.2.9 Strain E19

The E19 *S. coelicolor* mutant strain displayed a unique phenotype among the mutants we observed in our screen (Figure 4-2A). This mutant was very yellow and did not make aerial hyphae or spores on sporulation media (R2YE). What we found in this strain is that it was missing the first 78kb of DNA from its linear genome and 650kb from the end of its genome along with a point mutation in SCO5955. This was astonishing to see that *S. coelicolor* E19 was missing around 8.5% of its entire genome. *S. coelicolor* is known to chop off large regions on the ends of its genome, which has been shown to provide an advantage when present in a heterogenous population (Zhang et al., 2020). This yellow pigmentation is likely a spontaneously formed derivative of coelimycin that has been described to have a yellow color. Interestingly, E19 still produced some red pigments in response to interspecies interactions, indicating the system to sense and respond to these interactions in still intact and functioning despite missing roughly 728kb of its genome.

4.3.3 Transposons remain a challenging tool in *Streptomyces*

Transposons were first discovered in maize by the talented scientist Barbara McClintock in the 1940's and have since become a staple in the microbial genetics tool box (McClintock, 1950). Transposons are DNA sequences that can undergo transposition events and move location from a chromosome, phage genome, or plasmid in the absence of any complementary DNA sequence. They occur naturally and are often found to transfer antibiotic resistance, pathogenicity genes, or genes that can confer some advantage to the recipient (Muñoz-López and García-Pérez, 2010). Transposons are flanked by inverted repeats (IRs) which define the boundaries. Transposition is catalyzed by a transposase gene that recognizes the inverted repeats and excises the transposon from the DNA and integrates it into a new location. Researchers have figured out how to engineer transposons, introduce them into microbes, and get them to transpose into the genome of that microbe to create insertions randomly throughout the genome.

Transposons are an ideal method to create loss-of-function mutations in single genes because it is possible to engineer them to have antibiotic resistance, which we can select for, and we can remove the transposase gene from the transposon to ensure that it cannot hop anywhere else in the genome. We provide the transposase on a non-replicative plasmid in order to do this. Early expression of the transposase allows for the transposition event to occur, and during subsequent outgrowth, the plasmid is lost. If this is done with a transposon that has no sequence preference for insertion, we can introduce mutations randomly across the genome of a microbial strain and generate a collection of single transposon mutants, called a transposon library. Additionally, linking the transposon mutant strain with the transposon insertion site can be done using a primer off the end of the transposon and arbitrary primers to hopefully create a product with the transposon boarder and the genomic DNA region flanking the transposon. Altogether, transposons can generate single-gene, loss-of-function mutants, the mutants can be selected for, and the site of mutation can be easily identified making them a very powerful tool for generating a mutant library for genetic screens and selections.

We attempted to generate a *Streptomyces coelicolor* transposon mutant library with which we could use to screen for mutants that are impaired in interspecies interactions described previously. Earlier attempts to use transposon mutagenesis in *Streptomyces* have proven difficult because many native transposons destabilize the *S. coelicolor* genome or have trouble inserting properly and incorporate the entire plasmid backbone, including the transposase, into the genomic DNA (Widenbrant and Kao, 2007; Zhang et al., 2012). These problems were found when using native actinomycete transposons, therefore many researchers have shifted to using transposons from distantly related organisms that may be more stable, efficient, and predictable.

Due to the high GC content of *Streptomyces* genomes (*S. coelicolor* has 72% GC content) the transposases and transposons from other organisms must be codon optimized in order to be expressed efficiently. Other research groups have only recently systematically characterized strong *Streptomyces* promoters, synthetic promoters, and ribosomal binding sites to drive the optimal expression of the transposase gene (Bai et al., 2015). Recent work using a mariner-based or Tn5-based transposon systems seemed to have more success in efficiency of transposition, stability of the insertion, and proper incorporation (Petzke and Luzhetskyy, 2009; Bilyk et al., 2013). I first attempted to use the published himar1 transposon system optimized for *Streptomyces* called pHAM (Bilyk et al., 2013). After repeated attempts following the published protocol, I was unsuccessful in generating transposon mutants and the authors proved to be unavailable when we reached out for help. Therefore, I decided to build my own transposon system and optimize it for use in *S. coelicolor*.

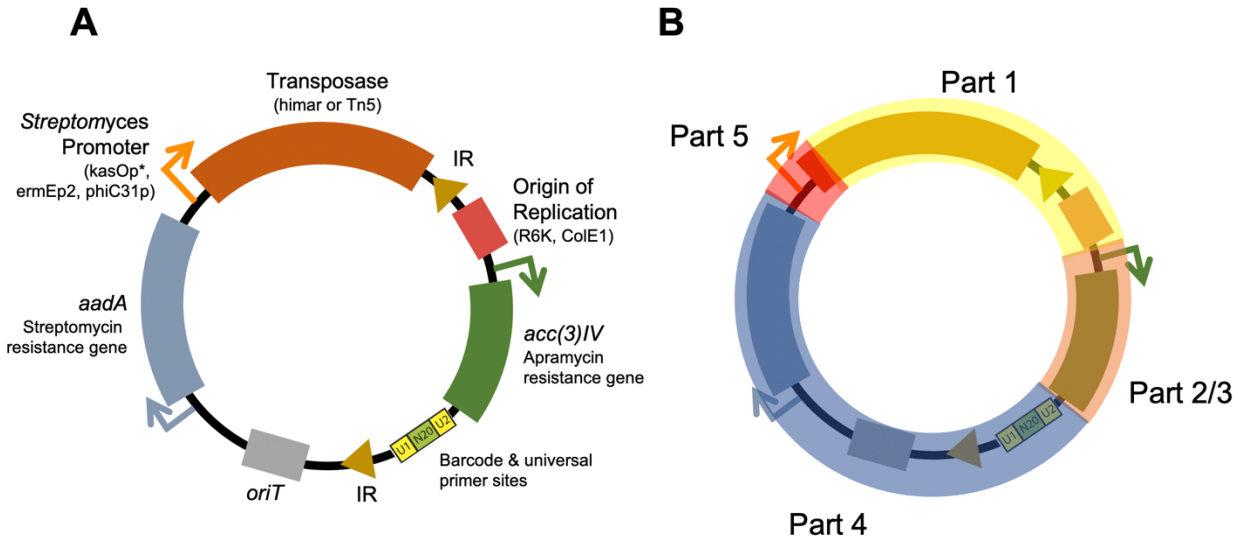


Figure 4-3 | Transposon mutagenesis plasmid construction.
Plasmid design for transposon mutagenesis in *S. coelicolor* (A). The elements and each version of the element that was used are shown around the plasmid.

Elements are not drawn to scale. The parts of the plasmid used for Golden Gate assembly of the constructs (B).

Altogether I built more than 20 transposon plasmid constructs using two different transposon systems (Tn5 and himar1), three different *Streptomyces* promoters driving the transposase expression (kasOp*, ermEp2, and phiC31p), two E. coli origins of replication for delivery by different E. coli strains (R6K and ColE1), and two versions of *Streptomyces* origins (pSG5 rep or none) (Figure 4-3A). The plasmid was designed to take advantage of Golden Gate cloning with Type IIs restriction endonucleases. These endonucleases recognize asymmetric sequences and cleave those sequences a couple nucleotides from the recognition site. The plasmid was designed in four parts with DNA ends that were flanked by a Type IIs restriction site, so that digestion of the fragments removes the enzyme recognition sites and generates complementary overhangs (Figure 4-3B) (Engler et al., 2008). The ends of each part were then ligated to create a junction that lacks the original site or scars. Unfortunately none of these plasmids proved to be successful and other mutagenesis techniques were thus explored.

4.4 Discussion and Conclusions

The forward genetics screen employed for discovery of *Streptomyces coelicolor* genes involved in production of red pigmented antibiotics during interspecies interactions produced many compelling leads. The gene with the most mutations from our genetic screen was *dnaA* (SCO3879)(Table 4-1). Half of all the *S. coelicolor* EMS mutants we sequenced were found to have a mutation in *dnaA*, with six of them having just a singular mutation. Of those six, three mutants are genetically identical, indicating that they probably arose as a singular mutant that replicated and fragmented before being plated to give rise to multiple colonies on our screening plates. This is an area of the EMS mutagenesis protocol that could be optimized to provide less time for *Streptomyces* replication or more careful handling of the filaments before plating. Nonetheless, there is very strong evidence that *dnaA* is important for the proper induction of pigmented specialized metabolite production in *S. coelicolor* during interactions with *Amycolatopsis* sp. AA4.

DnaA activity is best studied and understood in the rod shaped, proteobacteria, *E. coli*. To initiate DNA replication, DnaA-ATP binds to DnaA-boxes at the origin and promotes the unwinding of DNA. The level of DnaA-ATP increases just before initiation. After replication has been initiated the level of DnaA-ATP drops due to hydrolysis, yielding inactive DnaA-ADP. In *E. coli*, DnaA-ATP hydrolysis is crucial to regulation of initiation. Replication initiation in *Streptomyces* has distinct differences from initiation in *E. coli*; the replication origin is much larger (~1,000bp) (Zawilak-Pawlik et al., 2005) compared to E. coli (250bp), the sequences for strong DnaA-boxes differ (Majka et al., 2001), and the replication is initiated and proceeds bidirectionally to the ends of the linear chromosome (Musialowski et al., 1994). In *E. coli*, initiation from a newly formed oriC region is prevented by the SeqA protein, which bind to the hemimethylated DNA resulting from the newly synthesized strand. The *E. coli*

origin contains a high number of GATC sites recognized by the dam methyltransferase that eventually methylates those sites on the newly synthesized strand, releasing SeqA and allowing for DnaA to rebind and initiate replication again. Relatively little is known about the regulation in *Streptomyces*, but *Streptomyces* strains which have been studied so far do not have a dam methyltransferase or a system comparable to that.

Our finding using the forward genetic screen also indicated that the function of the ATPase domain and the C-terminal DNA binding domain are important for pigment production during interactions. I hypothesize that all these mutations decrease the amount or affinity of DnaA-ATP for the DnaA-boxes, resulting in a decreased amount of chromosome replication. Any mutation that reduces the ability of DnaA to bind ATP or bind DNA would reduce the pool of DnaA-ATP or the affinity for DnaA-ATP to bind to DNA and initiate replication, respectively. There is a possibility that these mutations actually increase binding to ATP, decrease the hydrolysis rate, or increase the affinity of DnaA for binding the DnaA-boxes, which would all lead to increased replication initiation. Deciphering how many genomes are in each cellular compartment can distinguish between these two hypotheses.

Streptomyces are filamentous multiploidy organisms, which grow by apical extension and branching networks of hyphae. During sporulation, *Streptomyces* go through coordinated cell divisions along with chromosome compaction and ParAB facilitated chromosomal segregation. However, during vegetative growth cell compartments have multiple copies of the chromosome and it was shown that ParA tethers just one of those chromosomes to the apical polarisome to ensure proper elongation and chromosome distribution. The segregation machinery targets the chromosomes to new branches or germ tubes, reminiscent of filamentous fungi (Kois-Ostrowska et al., 2016). Vegetative septa are then laid down at regularly spaced intervals, creating compartments that do not fully separate the cells. We hypothesize that a reduction in chromosome replication initiation in *S. coelicolor* would lead to fewer genomes in the cellular compartments or even compartments without a chromosome. This may explain the large reduction in pigmentation produced during interactions. It could be that there are simply less genomes to transcribe specialized metabolite biosynthesis genes from. The next steps to explore this phenomenon further is to observe the structure of the filament networks and the distribution and copy number of genomes in each cellular compartment.

The forward genetic screen also identified that mutations in SCO4215, a GntR family transcriptional regulator with a UTRA ligand binding domain, affected pigmented antibiotic production during interspecies interactions. From a bioinformatic exploration, SCO4215 is likely a transcriptional repressor that binds small molecules. This transcriptional regulator is a good candidate for further study during interspecies interactions. Connecting what genes it is regulating and what proteins it is interacting with will shed light on what role it plays during interactions.

We also identified mutations in genes found in specialized metabolite biosynthetic gene clusters. The most interesting of these mutations was that of SCO7690, a coelibactin

transporter. Disruption of the coelibactin biosynthesis genes has been shown to suppress a Zur repressor deletion mutant which is conditionally defective in sporulation (Kallifidas et al., 2010). In *Streptomyces*, development and pigment production are genetically linked and we wonder if the disruption of the coelibactin transport system could be responsible for the reduction of pigmented antibiotic production in *S. coelicolor* during interactions with *Amycolatopsis* sp. AA4 due to build-up of intercellular bioactive molecules. We also identified a mutation in a cryptic siderophore synthesis gene, SCO5800. Iron acquisition systems have been shown to influence morphology and specialized metabolism in *Streptomyces* during interactions (Seyedsayamdost et al., 2011; Lee et al., 2020). In fact, *Amycolatopsis* sp. AA4 has been shown to trigger the production of desferrioxamines in *S. coelicolor* during interspecies interactions (Traxler et al., 2013). We hypothesize that the ability to produce this cryptic siderophore might be important for *S. coelicolor* to be able to successfully compete for the limited iron that is available in the media, which is used as a co-factor for enzymes synthesizing specialized metabolites.

The regulatory networks controlling the expression of specialized metabolites are likely extremely complex with many systems playing a role in regulating the numerous biosynthetic gene clusters. The genes identified from the forward genetic screen are good avenues to investigate their role and this genetic screen provides a foundation for many exploratory studies of the genes identified here.

5 Chapter 5. Molecules to ecosystems: actinomycete natural products *in situ*

Scott W. Behie*, Bailey Bonet*, Vineetha M. Zacharia, Dylan J. McClung, Matthew F. Traxler
(2016) *Frontiers in Microbiology*
*both authors contributed equally

5.1 Abstract

Actinomycetes, filamentous actinobacteria found in numerous ecosystems around the globe, produce a wide range of clinically useful natural products. In natural environments, actinomycetes live in dynamic communities where environmental cues and ecological interactions likely influence natural product biosynthesis. Our current understating of these cues, and the ecological roles of natural products, is in its infancy. We postulate that understanding the ecological context in which actinomycete metabolites are made is fundamental to advancing the discovery of novel natural products. In this review we explore the ecological relevance of actinomycetes and their secondary metabolites from varying ecosystems, and suggest that investigating the ecology of actinomycete interactions warrants particular attention with respect to metabolite discovery. Furthermore, we focus on the chemical ecology and *in situ* analysis of actinomycete natural products and consider the implications for natural product biosynthesis at ecosystem scales.

5.2 Introduction

Of the top ten leading causes of death in 1900, three were bacterial infections, and they accounted for greater than one third of total deaths in the US (Jones et al., 2012). Today, bacterial infections are absent from the top ten list due to the discovery and development of antibiotics from fungi and bacteria. In 1943 the first actinomycete-produced therapeutic, streptomycin, was discovered by Albert Schatz, Elizabeth Bugie, and Selman Waksman and was rapidly put into use to treat tuberculosis (Schatz and Waksman, 1944). Actinomycetes are remarkable Gram-positive, filamentous, bacteria responsible for producing an estimated 70% of the antibiotics used in human therapy (Bérdy, 2005), making them the most robust natural source of antibiotics. Antibiotics and other unique compounds produced by actinomycetes are known as natural products (NP), or secondary or specialized metabolites (SM).

Historically, most discoveries of natural products from actinomycetes have involved their growth in rich media, in monoculture, which is strikingly different from their natural environment. During the "golden era" of antibiotics discovery in the 1950s and 1960s, this method worked remarkably well, and lead to a rapid increase in the number of medically relevant natural products found from bacteria, including anti-cancer therapeutics, immunosuppressants, and anthelminthics (Bérdy, 2005). From the 1980s and onward, the discovery of novel natural products from actinomycetes began dropping precipitously as

compound re-discovery emerged as a major impediment (Van Middlesworth and Cannell, 1998). The 1990s and 2000s witnessed the large-scale shutdown of most commercial efforts to isolate novel compounds from actinomycetes. This situation was, and continues to be, exacerbated by the rise of antibiotic resistance in pathogenic microbes.

Now, in the post-genomic era, with thousands of actinomycete genome sequences available, we have come the realization that actinomycetes possess the genetic capability to produce a multitude of natural products never before observed in the laboratory. This realization, coupled with advances in genetic tools, has rekindled interest in exploring actinomycete natural products in new and creative ways. Some of these strategies include physiological or genetic manipulation, sampling actinomycetes from new sources, using methods to isolate rare actinomycetes, and co-culturing (Bode et al., 2002; Traxler et al., 2013; Yamanaka et al., 2014; Harvey et al., 2015; Pimentel-Elardo et al., 2015). Beyond these, efforts that capitalize on genome mining supported by massive sequencing efforts (Doroghazi et al., 2014; Ju et al., 2015; Goering et al., 2016), high-throughput heterologous expression of NP gene clusters from environmental DNA (Subramani and Aalbersberg, 2013; Milshteyn et al., 2014; Iqbal et al., 2016), NP pathway re-factoring and heterologous expression, *in situ* cultivation of 'unculturable' bacterial strains (eg. the iChip, (Ling et al., 2015)), and the MS-guided genome-mining (termed peptidogenomics) (Kersten et al., 2011), have been undertaken. Many of these efforts have yielded exciting payoffs in terms of novel compounds discoveries in just the past few years, hinting that a new era of discovery may be at hand.

While developing new ways to discover natural products has been, and will certainly continue to be, an important goal, our understanding of the ecological role of these natural products is relatively primitive. We suggest that understanding the ecological relevance of these molecules for the microbes that make them is of great importance to inform future strategies for natural products discovery. For example, if we can fully understand the environmental cues that stimulate natural product production *in situ,* we can exploit these cues in the laboratory to induce normally silent natural product gene clusters. Our nascent understanding has lead researchers to propose that natural products may play a role in competition for resources, communication and signaling with other organisms, or defense of a symbiotic partner (Davies et al., 2006; Linares et al., 2006; Ryan and Dow, 2008; Cornforth and Foster, 2013; Abrudan et al., 2015). While the initial experiments underpinning these hypotheses are important first steps towards understanding natural product ecology, relatively few studies have actually detected the effects of natural products *in situ*.

This review focuses on studies where actinomycete natural product biosynthesis has been shown *in situ*. We briefly review these systems and consider the possible implications of natural product activity within their respective ecological contexts, both at a proximate and indirect level. Actinomycete natural products from soil or marine sediments have been addressed in recent reviews (Subramani and Aalbersberg, 2012; Traxler and Kolter, 2015), and therefore we do not address them here. The role of actinomycetes in symbiotic relationships has also been recently reviewed, and thus we emphasize studies post-2012 (Seipke et al., 2012). We conclude

by considering new tools and experimental paradigms that hold promise for exploring the chemical ecology of actinomycete natural products in the future.

5.3 Notable systems and examples

5.3.1 Use of actinomycete natural products by fungus-farming ants

One of the most extensively characterized natural systems that harbor actinomycetes is that of the leaf-cutter ants. Actinobacteria associated with fungus-growing leaf-cutter ants produce natural products that have been shown to have key roles in shaping this ecosystem (Figure 5-1) (Currie et al., 1999). Primarily found in the Americas, leaf-cutter ants (including the genera *Atta* and *Acromyrmex*) maintain a delicately balanced network that includes their fungal food source, *Leucoagaricus gongylophorus*, and actinomycetes, which produce antimicrobials to protect the fungal gardens against pathogen invasion (Schultz and Brady, 2008; Hölldobler and Wilson, 2010; Andersen et al., 2015). If the balance of the microbial community is perturbed, fungal pathogens belonging to the *Escovopsis* genus can invade the fungal garden, resulting in reduced garden biomass and the eventual destruction of entire ant colonies (Currie et al., 2003; Reynolds and Currie, 2004). Ultimately, the natural products synthesized by the symbiotic actinobacteria play a critical role in preserving the overall fitness of leaf-cutter ant colonies by protecting *L. gongylophorus* from pathogen infection.

Actinomycetes predominantly belonging to the S*treptomyces* and *Pseudonocardia* genera are also directly associated with the cuticle of leaf-cutter ants. These actinomycetes were studied in assays against invading fungal species and were found to produce a suite of antifungals including antimycins, candicidin, dentigerumycin, and nystatin variants which suppress *Escovopsis* and other potential pathogens while leaving *L. gongylophorus* unharmed (Haeder et al., 2009; Seipke et al., 2011; Dângelo et al., 2016). More recently, Sit and colleagues discovered, and chemically characterized, new antifungals related to dentigerumycin, called gerumycins, synthesized by *Pseudonocardia spp.* associated with both *Apterostigma dentigerum*, and the *Trachymyrmex cornetzi* ant (Sit et al., 2015). To add another level of complexity to this ecological system, the natural products synthesized by leaf-cutter ant actinomycetes may also serve to fend off closely-related bacteria that may displace the resident strain (Poulsen et al., 2007). This antagonism between different species of ant-associated *Pseudonocardia* was pronounced in more distantly-related species, suggesting that competition between *Pseudonocardia* have shaped this association from its evolutionary origin (Poulsen et al., 2007). Consistent with this idea, Van Arnam et al. recently found that one *Pseudonocardia* strain produced a novel rebeccamycin analog that inhibited the growth of competing *Pseudonocardia* (Figure 5-1). PacBio sequencing of these strains revealed variations in the rebeccamycin BGCs which were located on plasmids, suggesting plasmid-encoded niche defense (Van Arnam et al., 2015). Furthermore, rebeccamycins are of interest to human health since they possess anti-tumor activities and are currently being tested in clinical trials (Xu and Her, 2015).

Interestingly, recent evidence suggests that the direct interaction of actinomycetes with their leaf-cutter ant hosts is beneficial for worker ant protection from pathogenic fungi and bacteria (Schoenian et al., 2011; Mattoso et al., 2012; de Souza et al., 2013). With high resolution mass spectrometry (LC-ESI-HR-MS), and matrix-assisted laser desorption ionization (MALDI) imaging on the bodies of the ants, Schoenian et al. identified valinomycins, actinomycins, and antimycins from *Streptomyces* isolates associated with the integument of *Acromyrmex echinatior* workers. In their bioassays, Schoenian et al. demonstrated the inhibitory effects of these natural products on insect pathogens such as *Metarhizium anisopliae*, and *Cordyceps militaris*, and on *E. weberi* and *F. decemcellulare*, which are fungal pathogens of the fungus garden. Strikingly, the MALDI-imaging data clearly showed the localization of valinomycin secreted by *Streptomyces* on the ant cuticle at varying concentrations. Taken together, these data suggest that the presence of integumental biofilms and their secreted natural products work in concert to protect fungal gardens from invading bacteria and fungi, while simultaneously contributing to worker ant immunity by inhibiting the growth of entomopathogens (Figure 5-1)(Schoenian et al., 2011). To corroborate the importance of actinomycete natural products in their protective roles against ant-specific pathogens, cuticle-associated biofilms were removed with antibiotic treatment, thereby exposing the ants to attack by the entomopathogenic fungus *Metarhizium anisopliae*, which resulted in increased mortality of *Acromyrmex* ants (Mattoso et al., 2012). Furthermore, the immunity of *Acromyrmex subterraneus* ants with and without their symbiotic actinomycetes was evaluated. In their study, de Souza and colleagues observed that young worker ants lose their bacterial coating and mature into older, external worker ants. Furthermore, they found that in the absence of their associated bacteria, the young ants had an altered innate immune response, indicating that the actinomycetes confer protection to young internal workers until their immune systems have matured (de Souza et al., 2013). These works collectively stand as a rare concrete example in which the actinomycetes have been removed from the system and their influence ascertained, with the overall result being that their hosts became vulnerable to disease. The authors of these works logically conclude that the most likely reason for this is the activity of the natural products made by the actinomycetes *in situ*. Ultimately this hypothesis might be tested in the future by re-colonizing ants with mutant actinomycete strains lacking the ability to make specific natural products. The study of Seipke et al. takes a step toward conducting this type of experiment as they generated *Streptomyces* mutant strains unable to synthesize candicidin and antimycin (verified by PCR and LC-MS) and tested its bioactivity towards fungi *in vitro* (Seipke et al., 2011). The results, however, were inconclusive in terms of mutant effect, therefore it was not tested *in situ*. This highlights the challenge of executing complex, *in situ* microbial ecology.

As they are the dominant herbivores in much of the Neotropics (Costa et al., 2008), the foraging activity of leaf-cutter ants has a large impact in the overall function of their ecosystems (Figure 5-1, Figure 5-2). These ants forage up to 17% of the foliar biomass and contribute substantially to carbon turnover in their forest habitats (Costa et al., 2008). Their fungal crop, *L. gongylophorus*, produces lignocellulases capable of breaking down plant polymers (Aylward et al., 2013), which contribute significantly to organic material degradation and re-introduction into the neotropical terrestrial environment. As described above, actinomycete natural products likely play a role in protecting the ant's fungal gardens from invasion, and

also play a role in protecting the ants themselves from parasites and diseases. While equivocally assessing the direct contribution of these natural products is extremely challenging at an ecosystem scale, this body of evidence supports the notion that actinomycete natural products are intimately involved in maintaining one of the largest herbivory cycles on earth.

We also note that actinomycetes are associated with non-attine ants such as the Japanese carpenter ant (*Camponotus japonicas*). The actinomycetes *Nocardia camponoti* sp. and *Promicromonospora alba* sp. were recently isolated from the heads of these ants, and their genomes were sequenced (Guo et al., 2016; Liu et al., 2016). While the natural product repertoire of these strains has not yet been explored, further study of these newly reported interactions might lead to the discovery of novel bioactive compounds that play a role in shaping the Japanese carpenter ant's ecosystem.

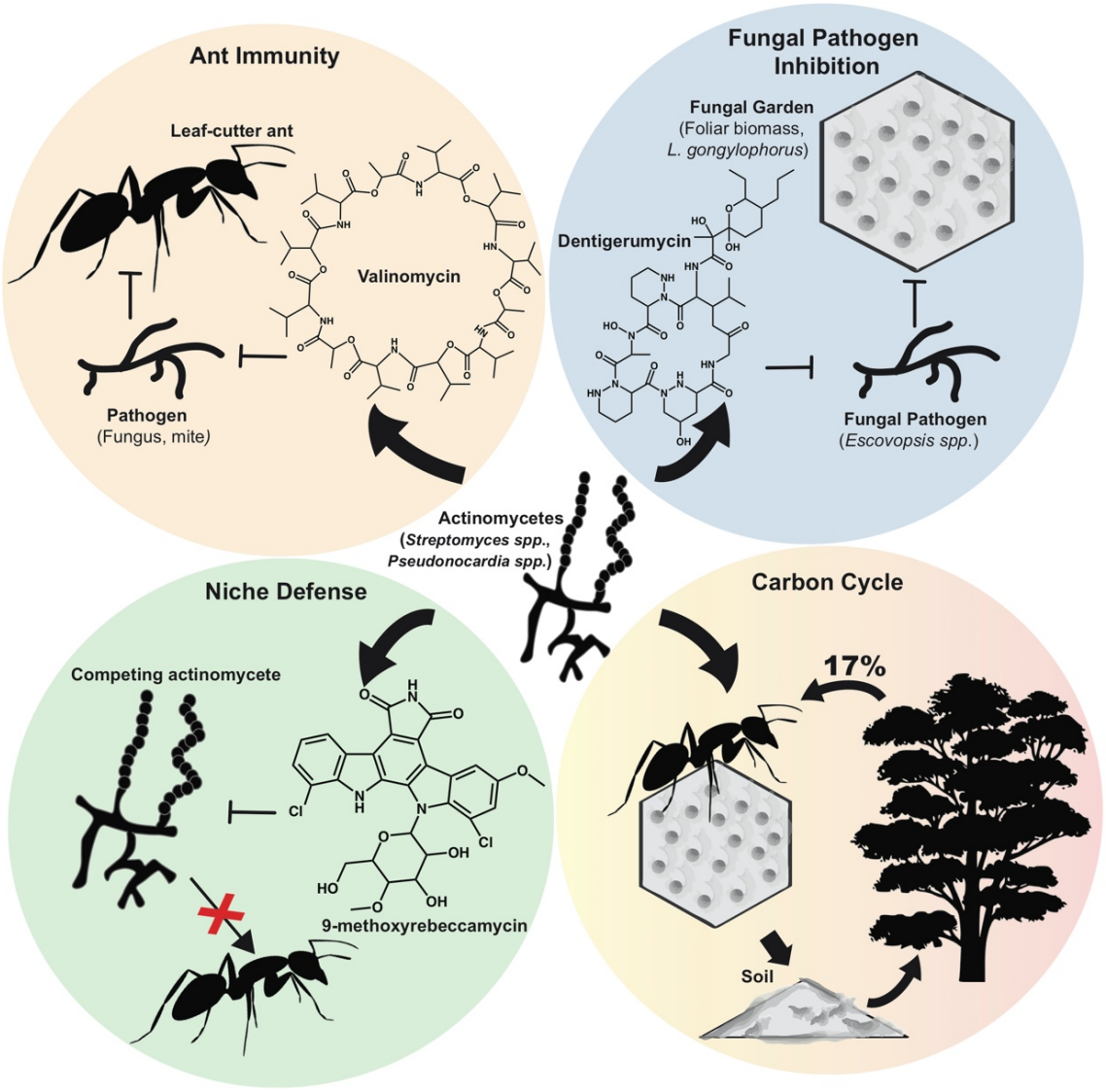


Figure 5-1 | Roles of actinomycete natural products in the leaf-cutter ant ecosystem.
Different actinomycete species isolated from attine ant integument (*Streptomyces* spp., *Pseudonocardia* spp., etc.) secrete natural products with various functions and roles. (Top left) *Streptomyces* spp. confer ant immunity by producing valinomycin, antimycins, and actinomycins, which inhibit pathogens and parasites. (Top Right) Actinomycetes produce a range of antifungals including dentigerumycins, candicidin, and nystatin variants, which inhibit fungal pathogens (e.g. *Escovopsis* spp.) but are not detrimental to the attine ant fungal symbiont, *L. gongylophorus*. (Bottom Left) *Pseudonocardia* competing for residency on the ant integument use plasmid-encoded niche defense mechanisms by

synthesizing 9-methoxyrebeccamycin. (Bottom Right) Natural products from actinomycetes help preserve the attine ant community, which in turn, affects the larger ecosystem. Leaf-cutter ants forage up to 17% of the foliar biomass and the *L. gongylophorus* symbiont helps degrade organic material, which collectively contributes to carbon turnover.

5.3.2 Offspring protection via actinomycete natural products by Beewolf wasps

The beewolf digger wasps (*Philanthus* spp., *Hymenoptera, Crabronidae*) live in a symbiosis with the streptomycete *Candidatus Streptomyces philanthi*, which is housed within the antennal glands of the female beewolves (Kaltenpoth et al., 2005). Female wasps deposit this bacterial symbiont onto the inner walls of the protective underground burrows where they lay their eggs. These streptomycetes are incorporated into the silk as the larvae begin spinning their cocoons (Kaltenpoth et al., 2006). The cocoons remain immobile in the humid burrows, and are thus easy targets for fungal and bacterial pathogens that live in the soil or are introduced by the honeybees brought into the burrow as a food source for the larvae (Kaltenpoth et al., 2005). However, the streptomycete symbionts protect the cocoons by secreting nine antimicrobial or antifungal compounds onto the outer surface of the cocoon (Kroiss et al., 2010; Seipke et al., 2012; Koehler et al., 2013).

Kroiss et al. (2010) were able to detect three of these compounds, the antibiotics piericidin A1, piericidin B1, and streptochlorin *in situ*. They further elucidated that there was a uniform distribution of all three compounds on the outer surface of the cocoon as compared to the inner surface (Kroiss et al., 2010). This distribution of antimicrobials suggests that the metabolites protect the cocoon from invading pathogens on the outside, but are at lower concentrations inside to prevent disruption of larval development. Four of the nine antimicrobials were shown to be inhibitory to the growth of 10 soil fungi and bacteria that would likely pose a threat of infection to the cocoons such as *Aspergillus* and *Metarhizium* (Strohm and Linsenmair, 2001; Kroiss et al., 2010). *In situ* analysis with MALDI-TOF/MS in conjunction with fluorescence *in situ* hybridization (FISH), showed that the insect associated streptomycetes co-localize with piericidin A1 and B1 (Kaltenpoth et al., 2016). This data illustrates the role of *Streptomyces* produced antimicrobials *in situ* that protect the beewolf cocoon from potential microbial infections, ensuring offspring survival.

At an ecosystem level, it is challenging to imagine what the overall effect of these actinomycete-produced compounds might be. However, this system is notable because it is one of the few examples in which natural products from actinomycetes have been visualized *in situ*. To the extent that this system may be manipulated at the genetic and chemical level, it may represent one of the best opportunities for assessing the direct impact, and therefore the *bona fide* role, of a natural product made by an actinomycete in a natural setting.

5.3.3 Actinomycete specialized metabolism in the rhizosphere

Actinomycetes that inhabit the rhizosphere have been shown to secrete a wide range of natural products that can contribute to defense against pathogenic bacteria and fungi, and aid in plant nutrient scavenging (El-Tarabily et al., 2009; Solans et al., 2011; Rungin et al., 2012; Harikrishnan et al., 2014). Rhizosphere soils contain relatively high levels of actinobacteria (Hirsch and Mauchline, 2012), and the relationship between soil actinomycetes and plant roots can have dramatic effects on plant health.

Root-colonizing *Streptomyces* have been found to produce a number of antifungal and antimicrobial compounds *in situ* such as staurosporine, 3-acetonylidene-7-prenylindolin-2-one, diastaphenazine, and antimycin A18 (Yan et al., 2010; Li et al., 2014; Zhang et al., 2014). These *in situ* analyses offer a critical lens for beginning to understand bioactive compound production in the natural context of the rhizosphere. Studies examining the large-scale effect of rhizosphere actinomycetes have been done with respect to plant health and disease resistance. For example, when cucumber plants (*Cucumis sativus*) were inoculated with one or a combination of three endophytic actinomycetes, *Actinoplanes campanulatus*, *Micromonospora chalcea*, or *Streptomyces spiralis*, the overall effects of *Pythium aphanidermatum*, a soil-borne, fungal phytopathogen, were mitigated. Specifically, host plants showed significantly less root and crown rot, and were healthier overall (El-Tarabily et al., 2009). A number of other studies have shown similar results, in various plant systems, where root-inoculated actinomycetes were able to protect the plant from harmful pathogen invasion (Xue et al., 2013; Nabti et al., 2014; Solans et al., 2016). In a notable example, *Streptomyces lividans* was shown to produce prodiginines on the roots of *Arabidopsis thaliana*, which antagonized the particularly damaging fungal pathogen *Verticillium dahlia* (Meschke et al., 2012). The ability of rhizosphere-inhabiting actinomycetes to produce a large range of metabolites antagonistic to phytopathogens, and subsequently provide protection, point to an important role for soil actinomycetes in ensuring plant health and primary production (Figure 5-2).

Soil actinomycetes are also known to enhance plant health by stimulating plant growth and development through the production of phytohormones (Solans et al., 2011; Harikrishnan et al., 2014). Endophytic actinomycetes such as *Nocardiopsis*, *Actinoplanes* spp., and *Micromonospora* have been found to produce a number of important phytohormones such as indole-3-acetic acid (IAA) (Figure 5-2) and indole-3-pyruvic acid (IPYA), compounds required for fundamental plant functions including coordinated cell growth, and gene regulation (Subramaniam et al., 2016). This suite of compounds has been specifically shown to promote growth and development in wheat, lettuce, rye and tomato (Merzaeva and Shirokikh, 2010; Bonaldi et al., 2015; Abbamondi et al., 2016; Toumatia et al., 2016). Directly delineating microbially produced phytohormones, as opposed to phytohormones made by the host, is technically challenging. However, the *in vitro* production of these compounds by root colonizing, or rhizosphere competent, actinomycetes suggests a potential means by which the soil microbiome is able to impact plant health (Figure 5-2).

Soil *Streptomyces* are also capable of enhancing the overall nutrient acquisition efficiency of plants by promoting the growth of critical symbiotic nitrogen fixers such as Rhizobia, and by aiding in root nutrient scavenging. Leguminous plants enter into a symbiosis with bacteria of the *Rhizobiales* clade in which the bacterium stimulates the plant to produce a root nodule that house the bacteria. In turn the bacteria fix atmospheric nitrogen, making it available to the plant. When chickpea plants were inoculated with *Streptomyces*, the result was increased nodulation as well as an increase in nodule size suggesting an overall stimulating effect on the nitrogen-fixing capability of the plant (Gopalakrishnan et al., 2015). This is especially notable given that fixed nitrogen is a limiting resource for plants in all systems. The production of siderophores by actinomycetes within the rhizosphere, or directly on the root, can specifically aid the plant in scavenging iron from surrounding soils (Rungin et al., 2012). Through the production of their own siderophores, plants are able to scavenge the soil for soluble iron (Ma, 2005). However, plants are also able to uptake iron-bound, actinomycete-produced, siderophores, increasing their overall iron uptake efficiency.

The example studies highlighted above point to a ubiquitous association between actinomycetes and plant roots in the rhizosphere. By in large, this association is beneficial for plants as they may receive multiple chemical inputs from actinobacteria, including siderophores, phytohormones, antibiotics and antifungals that inhibit potential pathogens (Solans et al., 2011; Rungin et al., 2012; Harikrishnan et al., 2014; Li et al., 2014; Zhang et al., 2014). Most studies have examined actinomycete/plant relationships in the context of food crops, some of which are grasses such as wheat, sorghum, rye, maize, and rice (Bressan, 2003; Merzaeva and Shirokikh, 2010; Harikrishnan et al., 2014; Toumatia et al., 2016). Beyond this, metagenomic and culture-based studies of rhizosphere soil from grasslands have revealed a wide array of plant-associated actinomycetes (Lugo et al., 2007; Delmont et al., 2012). Based on this collective evidence, we suggest that chemical interactions between actinomycetes and plants (especially non-woody plants) are likely to influence plant health at an ecosystem scale, especially in grasslands (Figure 5-2). Though the exact *in situ* role of each actinomycete produced natural product is not fully understood, it is clear that the presence of actinomycetes alter the plant's susceptibility to pathogens, nutrient uptake, and growth. Of course, other microbes also likely to play a major role in these complex interaction webs, but the large number of natural products produced by actinomycetes may make their role unique.

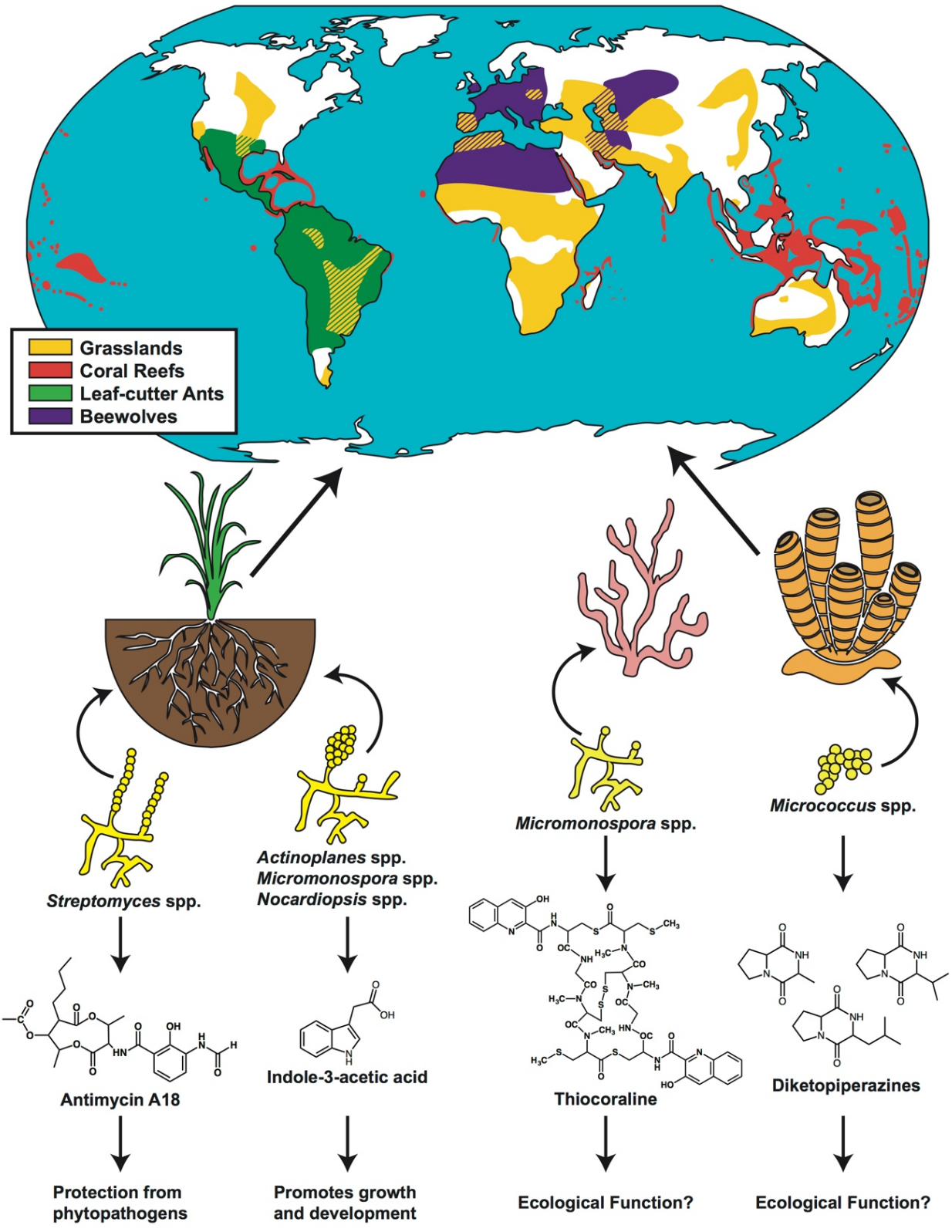

Grasslands
Coral Reefs
Leaf-cutter Ants
Beewolves
Streptomyces spp.
Actinoplanes spp.
Micromonospora spp.
Nocardiopsis spp.
Micromonospora spp.
Micrococcus spp.
Antimycin A18
Indole-3-acetic acid
Thiocoraline
Diketopiperazines
Protection from phytopathogens
Promotes growth and development
Ecological Function?
Ecological Function?

Figure 5-2 | Actinomycete-produced natural products aid in ecosystem function across the globe.
Distributions of grasslands (yellow), coral reefs (red), and leaf-cutter ants (green) are denoted on a map of the globe. Representative actinomycetes and the natural products that they produce in rhizosphere, coral and sponge communities are shown. Distribution maps adapted from: www.thinglink.com/scene/655045986153922561 (grasslands); www.antweb.org/antblog/2011/03/geographic-range-of-leaf-cutter-ants-don-indianapolis-in-usa.html (leaf-cutter ants); H. Ahlenius, UNEP/GRID-Arendal www.grida.no/graphicslib/detail/distribution-of-coldwater-and-tropical-coral-reefs_1153 (coral reefs).

5.3.4 Actinomycete natural products in marine organisms

The ocean covers 70% of the earth's surface and is responsible for 50% of the earth's daily primary productivity, which amounts to an estimated 140 million tons of carbon production per day (Field et al., 1998). Actinomycetes have been found in many marine habitats including marine sediments, estuaries, fish, mollusks, mangroves, and seaweeds, and marine actinomycetes have long been a productive source of prospecting for novel natural products. Isolating actinomycetes from marine sediments has proven to be a successful strategy for discovery of potent bioactive molecules (Fenical et al., 2009) and coupling this with sequencing efforts has shown that there is great potential still to be discovered (Udwary et al., 2007; Ziemert et al., 2014). For now we will focus on coral reefs and marine sponges as these are the best studied marine actinomycete environments. Studies have measured the diversity and abundance of actinobacterial symbionts, finding that they are ubiquitously and stably associated within these systems (Abdelmohsen et al., 2014; Braña et al., 2015; Mahmoud and Kalendar, 2016). Subsequently, actinomycetes have been isolated from various sponges and corals to investigate their antimicrobial activities and potential to produce novel bioactive compounds (Schneemann et al., 2010; Cheng et al., 2015; Kuang et al., 2015; Pham et al., 2015; Mahmoud and Kalendar, 2016). The isolates have proved to be a rich source of natural products with antifungal, antimicrobial, anti-cancer, and anti-HIV compounds (Abdelmohsen et al., 2014).

The majority of these studies involved isolating the actinomycetes and culturing them in the laboratory setting. Comprehensive laboratory analyses done on sponge associated actinomycetes has found that a number of different species are culturable from marine sponges, and that they are capable of producing bioactive secondary metabolites (Schneemann et al., 2010; Xi et al., 2012). With respect to the sponge *Halichondria panacea*, for example, five actinomycete genera were isolated and the bacterial extracts were found to contain secondary metabolites with bioactivites (Schneemann et al., 2010). However, ecological analysis of marine environments remains challenging, and there are few studies that have reported *in situ* data. Recently, however, Yarnold and colleagues used MALDI-MS imaging to map the spatial distribution of brominated pyrrole-2-aminoimidazole alkaloids, which have cytotoxic and anti-

fouling properties, on cross sections of the marine sponge *Stylissa flabellate* (Yarnold et al., 2012). These compounds were found in discrete locations within the sponges and were present in high amounts near sites most susceptible to predation and biofouling. There is yet no clear delineation between the suite of metabolites produced by the sponge, and those produced by associated microorganisms, but gradients and pockets of metabolite production suggest that these are likely to be microbially produced. In one specific example, three diketopiperazines, one of which has antibiotic properties, that were isolated and characterized from the sponge *Tedania ignis*, were later found to be produced by associated *Micrococcus* (Figure 5-2) (Stierle et al., 1988).

Due to the overall paucity of information about specific microbially produced metabolites on marine sponges, there is little information on the roles they may play for their sponge hosts. Recently, however, the genomes of three marine sponge associated actinomycetes (*Micromonospora* sp. RV43, *Rubrobacter* sp. RV113, and *Nocardiopsis* sp. RV163) were sequenced, and subsequently mined for the presence of secondary metabolite producing gene clusters (Horn et al., 2015). All three species under investigation contained gene clusters required to produce natural products of known antibiotic and antifungal activity. To study how bacterial symbionts affect the ecology of the sponge, Mehbub and colleagues performed controlled aquarium experiments designed to test the effects of increased levels of a sponge-associated actinomycete, *Streptomyces* ACT-52A, on the sponge's microbial community, metabolite profile, and bioactivity (Mehbub et al., 2016). When *Streptomyces* ACT-52A was exogenously added to the water in the aquarium, they observed shifts in the microbial community, accompanied by a change in metabolite profile composition and the bioactivity of sponge extracts. Though it is still not clear if the changes in metabolite production were of microbial or sponge origin, it is clear that interactions between the sponge and its associated bacteria can impact natural product biosynthesis and microbial diversity. We can envision this as a system where the effect of mutant actinomycetes, deficient in the production of a specific natural product, may be tested. A possible next step would be to investigate how the health and function of the sponge, and its associated microbiome, may impact the marine ecosystem as a whole.

Coral reefs themselves have also been of interest with respect to bioprospecting as they house an immense diversity of secondary metabolites (Leal et al., 2013). A large number of these metabolites, however, are produced by symbiotic bacteria living in the corals (Leal et al., 2013). In one instance, thiocoraline, a thiodepsipeptide, was isolated from a coral associated marine actinomycete, *Micromonospora marina* (Romero et al., 1997). Though thiocoraline was shown to have anti-tumor and antibiotic activity *in vitro*, no *in situ* experiments have been done, thus the ecological impacts of this metabolite on the corals remains unexplored (Figure 5-2). These and other bacterial symbionts potentially have a number of functions with regard to reef health, including protection from invading pathogens. However, the culturing of the majority of marine bacteria remains a major impediment, and thus the analysis of the majority of reef associated actinomycetes remains elusive. Beyond this, large amounts of reef biomass are required for the isolation of coral-associated metabolites, and while the aquaculturing of reefs represents a potentially viable option to produce reef associated metabolites (Leal et al., 2013),

the complexity of marine environments is difficult to replicate, and thus the chemical ecology, and symbiotic microbiome, may be drastically different in laboratory experiments (Hay, 1996).

Collectively reefs occupy <0.2% of the world's oceans (Figure 5-2), yet are estimated to be one of the most biodiverse ecosystems on the planet (Knowlton et al., 2010), that includes a panoply of coral and sponge species. Taken together, the studies described above point to a long-lasting symbiosis between corals, sponges and actinomycetes. As for the other systems addressed in this review, assessing the contribution of actinomycete natural products to the overall health and stability of reef ecosystems is challenging. However, any advantage conferred by actinomycete symbiosis is likely to have ecosystem-wide implications. Further studies aimed at understanding the role of these compounds in simplified, controlled aquarium settings may provide a window into the chemical underpinnings of these ecosystems.

5.4 Concluding remarks

Accurately assessing the role of a given bacterial species in its natural environment is a daunting task. However, as we transition to the 'age of the microbiome', questions about microbial community function, and community impact on their respective ecosystems, will become increasingly important. Actinomycetes are interesting to consider in this regard since the role played by their extensive natural product repertoires has long been the source of speculation. Beyond this, at a practical level, understanding the roles and ecological drivers that stimulate production of these compounds may open new avenues to natural products discovery. Only by conducting these kinds of studies, will we begin to understand how these molecules function in natural settings, and what advantages might be gained for the organisms that make these compounds. We suggest that knowledge about the ecological and chemical contexts in which these molecules are used by the producing organisms may provide insights into how we can minimize the development of pathogen resistance and extend the working life of antibiotics in the clinic.

What does it take to effectively evaluate the impact of a natural product *in situ*? The framework for considering this question could be adapted from the fundamental principles of Koch's postulates. Namely, the natural product should be detectable *in situ*, removing the producing organism from the system should eliminate the natural product and result in a measurable effect on the system, and replacing the producing organism should lead to detection of the compound *in situ* and reverse the effect of removing it. Undoubtedly, the most rigorous approach would be to replace the producing organism with a mutant defective in its ability to produce the natural product in question, followed by verifying the absence of the natural product and assessment of the effect on the system overall. These stringent criteria are obviously challenging to implement in practice, especially in natural settings.

Determining the ecological drivers of natural product biosynthesis *in situ* may be ultimately prove even more challenging than assessing the impact of these compounds. However, we suggest that understanding the effect that a given natural product has on the surrounding

microbial community may provide valuable context for considering what cues may be connected to its production. Another starting point will be to understand cues that drive actinomycete specialized metabolism in simple, binary interactions. Surprisingly, while it has been shown many times that binary interspecies interactions can lead to activation of natural product biosynthesis (Ueda et al., 2000; Yamanaka et al., 2005; Onaka et al., 2011; Vetsigian et al., 2011; Traxler et al., 2013; Kinkel et al., 2014; Abrudan et al., 2015), a molecular/physiological understanding of this phenomenon remains largely elusive. Only in one case, described by Yamanaka et al., 2005, was a molecule identified as triggering antibiotic production in another nearby strain. In that instance, the stimulatory molecule revealed to be the common siderophore desferrioxamine E. Accordingly, understanding the chemical and mechanistic nature of these interactions is an active area of research being undertaken in multiple labs. Beyond binary interactions, simplified synthetic communities that can be experimentally manipulated will also offer a tractable starting point for assessing connections between cues, natural products, and ecosystem effects.

However difficult it may be to carry out studies on complex natural systems that include microbial communities, the development of new tools provides a reason to be optimistic about future efforts aimed at understanding the ecology of natural products. Specifically, advances in mass spectrometry imaging and direct chemical sampling of microhabitats via technologies like NanoDESI MS are making detection of these fascinating molecules *in situ* a much more tractable prospect. One predictable outcome of removing a molecule with antibiotic or antifungal activity from a microbial community might be a shift in the overall community composition. Fortunately, the best method for assessing these changes, eg. metagenomic sequencing, has also become radically more accessible. Moreover, as the study of microbiomes advances with the development of model microbial communities and more portable technologies to enable measurement of key microbiome parameters in the field (Alivisatos et al., 2015; Blaser et al., 2016) the questions posed here will become more tractable.

5.5 Acknowledgements

This review was a lab wide effort. Thanks to Dr. Matthew Traxler, Dr. Scott Behie, Dr. Vineetha Zacharia, and Dylan McClung for their efforts and contributions to this work. All contributing authors have granted permission for this work to be used in this dissertation.

6 Chapter 6. Discussion and Conclusions

This body of work is the first, to our knowledge, to use an untargeted approach to identify systems involved in globally regulating the production of specialized metabolites during interspecies interactions. Because the genus *Streptomyces* is a rich source of antibacterial, antifungals, and many other biologically useful compounds, understanding when and how these specialized metabolites are produced is fundamental to discovery efforts (Bérdy, 2005). Regulation of specialized metabolite expression has been studied for specific cues or for individual biosynthetic gene clusters, but never at a more universal scale (Liu et al., 2013). Identifying more general regulatory systems controlling natural product production can highlight avenues for natural product drug discovery. Here we employed a forward genetic screen, transcriptomics analysis, and mutational characterization to identify genes involved in sensing and responding to interspecies interactions in the model actinomycete, *Streptomyces coelicolor*.

From our transcriptional analysis of *S. coelicolor* in interactions with four other actinomycetes, we identified four conservons as being highly upregulated in this environment. Subsequent deletion of the conservon 8 locus and each individual member illuminated how these members relate to one another and how each member affects the gene expression of multiple specialized metabolite biosynthetic gene clusters. Very little information is known about the conservons and how they function. Only one *S. coelicolor* conservon, conservon 9, has ever been studied. The deletion of *cvnA9* and *cvnD9* resulted in *S. coelicolor* strains that conditionally overproduced actinorhodin and were able to form aerial mycelium and spores on 2% glucose medium, which typically inhibits those processes (Komatsu et al., 2006). This body of work linked the conservons to natural product production and development, though only one of the 13 systems present in the *S. coelicolor* genome and numerous conservon systems throughout the actinobacteria phylum was studied (Wuichet and Søgaard-Andersen, 2014). In this work, we link another conservon system, conservon 8, to natural product production during interspecies interactions and growth as a single colony, adding significantly to our growing knowledge of these systems.

We determined that the CvnA8 and CvnF8 proteins likely function together and we propose that CvnA8 is a membrane bound signaling component while CvnF8, a GAF containing protein, acts as the small molecule sensor. We found that the CvnC8 and CvnD8 proteins likely function together and we hypothesize that CvnC8 is a transcriptional regulator that, together with CvnD8, a molecular switch similar to RAS GTPases, regulates the expression of multiple specialized metabolite biosynthetic gene clusters. We observed that CvnB8 has an intermediate effect on pigmentation and gene expression leading us to propose that it acts as a connection between CvnA/F8 and CvnC/D8 likely through the regulation of the GTPase, CvnD8.

There is much more work to be completed in the future in order to fully understand the conservons systems. First and foremost, a thorough exploration of the CvnA protein should

be performed to determine if it has autophosphorylation activity, if it interacts with CvnF, CvnE, or other proteins, and where in the cell it is located. The determination of what conservon 8 is sensing from the interacting bacterium is of utmost interest to us, as well. We think that determining the role of the CvnF8 protein and whether it binds a small molecule will be a good start towards resolving what the signal(s) or cue(s) may be. Another important question to answer is can the CvnC8 protein bind DNA? Answering this question can help us understand how conservon 8 is exerting its influence and will surely generate many new questions. Additionally, how is the CvnD8 protein functioning? What do the CvnD8 on/off mutants do *in vivo*? Can the CvnB8 protein alter CvnD8's activity? These are all outstanding questions that will need to be addressed in order to understand the molecular mechanisms of the conservon system. Lastly, determining how the conservons integrate into previously characterized regulatory networks and how they may interact with different conservon systems will be crucial. Though the exploration of the conservons has copious unanswered questions, this work has laid the foundation for future studies of this underexplored, intriguing system.

We also used a forward genetic screen to identified multiple genes that are related to pigmentation production during interspecies interactions. Notably, *dnaA* (SCO3879) was identified using this screen. We hypothesize that mutations in *dnaA* cause a reduction in chromosome replication initiation, leading to fewer chromosomes in each cellular compartment. Additionally, SCO4215, a predicted GntR transcriptional regulator with a UTRA ligand-binding domain, was identified in our screen. GntR regulators with UTRA domains typically are regulated by the binding of small molecules. Multiple other genes were implicated in interaction induced specialized metabolite production from our genetic screen. This is an area of active exploration and highlights the complex regulation that exists in *Streptomyces*.

For the first time, we have shown that a conservon system regulates the expression of multiple natural product biosynthetic gene clusters, including a cryptic cluster in *Streptomyces coelicolor*. This discovery highlights the potential for discovery of new compounds through manipulation of conservon signaling. We also conclude that this system must be intertwined with other regulatory networks which will take more work to unravel. It may be the case that regulators identified in the forward genetic screen might link to the conservon system. Additionally, CvnD8 is the first ever bacterial small GTPase shown to alter gene expression, indicating that these types of proteins play an important role in bacterial cell processes not just in eukaryotic signaling systems. Thus, they should be given more thought. Overall, we excitedly hypothesize that the conservons are another type of signaling system in actinomycetes that interfaces with the environment in various ways. We assert that the conservons should be considered alongside two-component systems and ECF (extra-cytoplasmic function) sigma factors when considering bacterial systems employed by actinomycetes to interact with their environments.

7 References

Abbamondi, G. R., Tommonaro, G., Weyens, N., Thijs, S., Sillen, W., Gkorezis, P., et al. (2016). Plant growth-promoting effects of rhizospheric and endophytic bacteria associated with different tomato cultivars and new tomato hybrids. *Chem. Biol. Technol. Agric.* 3, 1. doi:10.1186/s40538-015-0051-3.

Abdelmohsen, U. R., Bayer, K., and Hentschel, U. (2014). Diversity, abundance and natural products of marine sponge-associated actinomycetes. *Nat. Prod. Rep.* 31, 381–399. doi:10.1039/C3NP70111E.

Abrudan, M. I., Smakman, F., Grimbergen, A. J., Westhoff, S., Miller, E. L., van Wezel, G. P., et al. (2015). Socially mediated induction and suppression of antibiosis during bacterial coexistence. *Proc. Natl. Acad. Sci. U. S. A.* 112, 11054–11059. doi:10.1073/pnas.1504076112.

Alivisatos, A. P., Blaser, M. J., Brodie, E. L., Chun, M., Dangl, J. L., Donohue, T. J., et al. (2015). A unified initiative to harness Earth's microbiomes. *Science* 350, 507–508. doi:10.1126/science.aac8480.

Andersen, S. B., Yek, S. H., Nash, D. R., and Boomsma, J. J. (2015). Interaction specificity between leaf-cutting ants and vertically transmitted Pseudonocardia bacteria. *BMC Evol. Biol.* 15, 27. doi:10.1186/s12862-015-0308-2.

Anderson, J. B., Bruhn, J. N., Kasimer, D., Wang, H., Rodrigue, N., and Smith, M. L. (2018). Clonal evolution and genome stability in a 2500-year-old fungal individual. *Proc. Biol. Sci.* 285, 20182233. doi:10.1098/rspb.2018.2233.

Aravind, L., and Anantharaman, V. (2003). HutC/FarR-like bacterial transcription factors of the GntR family contain a small molecule-binding domain of the chorismate lyase fold. *FEMS Microbiol. Lett.* 222, 17–23. doi:10.1016/S0378-1097(03)00242-8.

Asturias-Arribas, L., Alonso-Lomillo, M. A., Domínguez-Renedo, O., and Arcos-Martínez, M. J. (2014). Cytochrome P450 2D6 based electrochemical sensor for the determination of codeine. *Talanta* 129, 315–319. doi:10.1016/j.talanta.2014.05.053.

Aylward, F. O., Burnum-Johnson, K. E., Tringe, S. G., Teiling, C., Tremmel, D. M., Moeller, J. A., et al. (2013). Leucoagaricus gongylophorus Produces Diverse Enzymes for the Degradation of Recalcitrant Plant Polymers in Leaf-Cutter Ant Fungus Gardens. *Appl. Environ. Microbiol.* 79, 3770–3778. doi:10.1128/AEM.03833-12.

Bai, C., Zhang, Y., Zhao, X., Hu, Y., Xiang, S., Miao, J., et al. (2015). Exploiting a precise design of universal synthetic modular regulatory elements to unlock the microbial natural products in Streptomyces. *Proc. Natl. Acad. Sci. U. S. A.* 112, 12181–12186. doi:10.1073/pnas.1511027112.

Baranwal, J., Lhospice, S., Kanade, M., Chakraborty, S., Gade, P. R., Harne, S., et al. (2019). Allosteric regulation of a prokaryotic small Ras-like GTPase contributes to cell polarity oscillations in bacterial motility. *PLoS Biol.* 17, e3000459. doi:10.1371/journal.pbio.3000459.

Barbosa, R. L., and Benedetti, C. E. (2007). BigR, a Transcriptional Repressor from Plant-Associated Bacteria, Regulates an Operon Implicated in Biofilm Growth. *J. Bacteriol.* 189, 6185–6194. doi:10.1128/JB.00331-07.

Basak, K., and Majumdar, S. K. (1973). Utilization of Carbon and Nitrogen Sources by Streptomyces kanamyceticus for Kanamycin Production. *Antimicrob. Agents Chemother.* 4, 6–10.

Bentley, S. D., Chater, K. F., Cerdeño-Tárraga, A.-M., Challis, G. L., Thomson, N. R., James, K. D., et al. (2002). Complete genome sequence of the model actinomycete Streptomyces coelicolor A3(2). *Nature* 417, 141–147. doi:10.1038/417141a.

Bérdy, J. (2005a). Bioactive microbial metabolites. *J. Antibiot. (Tokyo)* 58, 1–26. doi:10.1038/ja.2005.1.

Bérdy, J. (2005b). Bioactive Microbial Metabolites. *J. Antibiot. (Tokyo)* 58, 1–26. doi:10.1038/ja.2005.1.

Bierman, M., Logan, R., O'Brien, K., Seno, E. T., Rao, R. N., and Schoner, B. E. (1992). Plasmid cloning vectors for the conjugal transfer of DNA from Escherichia coli to Streptomyces spp. *Gene* 116, 43–49. doi:10.1016/0378-1119(92)90627-2.

Bilyk, B., Weber, S., Myronovskyi, M., Bilyk, O., Petzke, L., and Luzhetskyy, A. (2013). In vivo random mutagenesis of streptomycetes using mariner-based transposon Himar1. *Appl. Microbiol. Biotechnol.* 97, 351–359. doi:10.1007/s00253-012-4550-x.

Blaser, M. J., Cardon, Z. G., Cho, M. K., Dangl, J. L., Donohue, T. J., Green, J. L., et al. (2016). Toward a Predictive Understanding of Earth's Microbiomes to Address 21st Century Challenges. *mBio* 7, e00714-16. doi:10.1128/mBio.00714-16.

Bode, H. B., Bethe, B., Höfs, R., and Zeeck, A. (2002). Big Effects from Small Changes: Possible Ways to Explore Nature's Chemical Diversity. *ChemBioChem* 3, 619–627. doi:10.1002/1439-7633(20020703)3:7<619::AID-CBIC619>3.0.CO;2-9.

Bonaldi, M., Chen, X., Kunova, A., Pizzatti, C., Saracchi, M., and Cortesi, P. (2015). Colonization of lettuce rhizosphere and roots by tagged Streptomyces. *Plant Biot. Interact.* 6, 25. doi:10.3389/fmicb.2015.00025.

Braña, A. F., Braña, A. F., Fiedler, H.-P., Nava, H., González, V., Sarmiento-Vizcaíno, A., et al. (2015). Two Streptomyces species producing antibiotic, antitumor, and anti-inflammatory compounds are widespread among intertidal macroalgae and deep-sea

coral reef invertebrates from the central Cantabrian Sea. *Microb. Ecol.* 69, 512–524. doi:10.1007/s00248-014-0508-0.

Bressan, W. (2003). Biological control of maize seed pathogenic fungi by use of actinomycetes. *BioControl* 48, 233–240. doi:10.1023/A:1022673226324.

Capstick, D. S., Willey, J. M., Buttner, M. J., and Elliot, M. A. (2007). SapB and the chaplins: connections between morphogenetic proteins in Streptomyces coelicolor. *Mol. Microbiol.* 64, 602–613. doi:10.1111/j.1365-2958.2007.05674.x.

Cerdeño, A. M., Bibb, M. J., and Challis, G. L. (2001). Analysis of the prodiginine biosynthesis gene cluster of Streptomyces coelicolor A3(2): new mechanisms for chain initiation and termination in modular multienzymes. *Chem. Biol.* 8, 817–829. doi:10.1016/s1074-5521(01)00054-0.

Cheng, C., MacIntyre, L., Abdelmohsen, U. R., Horn, H., Polymenakou, P. N., Edrada-Ebel, R., et al. (2015). Biodiversity, Anti-Trypanosomal Activity Screening, and Metabolomic Profiling of Actinomycetes Isolated from Mediterranean Sponges. *PloS One* 10, e0138528. doi:10.1371/journal.pone.0138528.

Chong, P. P., Podmore, S. M., Kieser, H. M., Redenbach, M., Turgay, K., Marahiel, M., et al. (1998). Physical identification of a chromosomal locus encoding biosynthetic genes for the lipopeptide calcium-dependent antibiotic (CDA) of Streptomyces coelicolor A3(2). *Microbiol. Read. Engl.* 144 (Pt 1), 193–199. doi:10.1099/00221287-144-1-193.

Colicelli, J. (2004). Human RAS Superfamily Proteins and Related GTPases. *Sci. STKE Signal Transduct. Knowl. Environ.* 2004, RE13. doi:10.1126/stke.2502004re13.

Cornforth, D. M., and Foster, K. R. (2013). Competition sensing: the social side of bacterial stress responses. *Nat. Rev. Microbiol.* 11, 285–293. doi:10.1038/nrmicro2977.

Cornforth, D. M., and Foster, K. R. (2015). Antibiotics and the art of bacterial war. *Proc. Natl. Acad. Sci.* 112, 10827–10828. doi:10.1073/pnas.1513608112.

Costa, A. N., Vasconcelos, H. L., Vieira-Neto, E. H. M., and Bruna, E. M. (2008). Do herbivores exert top-down effects in Neotropical savannas? Estimates of biomass consumption by leaf-cutter ants. *J. Veg. Sci.* 19, 849–854. doi:10.3170/2008-8-18461.

Currie, C. R., Scott, J. A., Summerbell, R. C., and Malloch, D. (1999). Fungus-growing ants use antibiotic-producing bacteria to control garden parasites. *Nature* 398, 701–704. doi:10.1038/19519.

Currie, C. R., Wong, B., Stuart, A. E., Schultz, T. R., Rehner, S. A., Mueller, U. G., et al. (2003). Ancient Tripartite Coevolution in the Attine Ant-Microbe Symbiosis. *Science* 299, 386–388. doi:10.1126/science.1078155.

Dângelo, R. A. C., de Souza, D. J., Mendes, T. D., Couceiro, J. da C., and Lucia, T. M. C. D. (2016). Actinomycetes inhibit filamentous fungi from the cuticle of Acromyrmex leafcutter ants. *J. Basic Microbiol.* 56, 229–237. doi:10.1002/jobm.201500593.

Davies, J., Spiegelman, G. B., and Yim, G. (2006). The world of subinhibitory antibiotic concentrations. *Curr. Opin. Microbiol.* 9, 445–453. doi:10.1016/j.mib.2006.08.006.

de Lira, N. P. V., Pauletti, B. A., Marques, A. C., Perez, C. A., Caserta, R., de Souza, A. A., et al. (2018). BigR is a sulfide sensor that regulates a sulfur transferase/dioxygenase required for aerobic respiration of plant bacteria under sulfide stress. *Sci. Rep.* 8, 3508. doi:10.1038/s41598-018-21974-x.

de Souza, D. J., Lenoir, A., Kasuya, M. C. M., Ribeiro, M. M. R., Devers, S., Couceiro, J. da C., et al. (2013). Ectosymbionts and immunity in the leaf-cutting ant Acromyrmex subterraneus subterraneus. *Brain. Behav. Immun.* 28, 182–187. doi:10.1016/j.bbi.2012.11.014.

Deatherage, D. E., and Barrick, J. E. (2014). Identification of mutations in laboratory evolved microbes from next-generation sequencing data using breseq. *Methods Mol. Biol. Clifton NJ* 1151, 165–188. doi:10.1007/978-1-4939-0554-6_12.

Delmont, T. O., Prestat, E., Keegan, K. P., Faubladier, M., Robe, P., Clark, I. M., et al. (2012). Structure, fluctuation and magnitude of a natural grassland soil metagenome. *ISME J.* 6, 1677–1687. doi:10.1038/ismej.2011.197.

Doroghazi, J. R., Albright, J. C., Goering, A. W., Ju, K.-S., Haines, R. R., Tchalukov, K. A., et al. (2014). A roadmap for natural product discovery based on large-scale genomics and metabolomics. *Nat. Chem. Biol.* 10, 963–968. doi:10.1038/nchembio.1659.

Du, Y.-L., Shen, X.-L., Yu, P., Bai, L.-Q., and Li, Y.-Q. (2011). Gamma-Butyrolactone Regulatory System of Streptomyces chattanoogensis Links Nutrient Utilization, Metabolism, and Development ▿. *Appl. Environ. Microbiol.* 77, 8415–8426. doi:10.1128/AEM.05898-11.

Dulaney, E. L. (1949). Observations on Streptomyces Griseus. III. Carbon Sources for Growth and Streptomycin Production. *Mycologia* 41, 1–10. doi:10.2307/3755267.

El-Tarabily, K. a., Nassar, A. h., Hardy, G. E. St. J., and Sivasithamparam, K. (2009). Plant growth promotion and biological control of Pythium aphanidermatum, a pathogen of cucumber, by endophytic actinomycetes. *J. Appl. Microbiol.* 106, 13–26. doi:10.1111/j.1365-2672.2008.03926.x.

Engler, C., Kandzia, R., and Marillonnet, S. (2008). A One Pot, One Step, Precision Cloning Method with High Throughput Capability. *PLoS ONE* 3. doi:10.1371/journal.pone.0003647.

Fenical, W., Jensen, P. R., Palladino, M. A., Lam, K. S., Lloyd, G. K., and Potts, B. C. (2009). Discovery and Development of the Anticancer Agent Salinosporamide A (NPI-0052). *Bioorg. Med. Chem.* 17, 2175. doi:10.1016/j.bmc.2008.10.075.

Field, C. B., Behrenfeld, M. J., Randerson, J. T., and Falkowski, P. (1998). Primary Production of the Biosphere: Integrating Terrestrial and Oceanic Components. *Science* 281, 237–240. doi:10.1126/science.281.5374.237.

Flärdh, K., and Buttner, M. J. (2009). Streptomyces morphogenetics: dissecting differentiation in a filamentous bacterium. *Nat. Rev. Microbiol.* 7, 36–49. doi:10.1038/nrmicro1968.

Galicia, C., Lhospice, S., Varela, P. F., Trapani, S., Zhang, W., Navaza, J., et al. (2019). MglA functions as a three-state GTPase to control movement reversals of Myxococcus xanthus. *Nat. Commun.* 10, 5300. doi:10.1038/s41467-019-13274-3.

Goering, A. W., McClure, R. A., Doroghazi, J. R., Albright, J. C., Haverland, N. A., Zhang, Y., et al. (2016). Metabologenomics: Correlation of Microbial Gene Clusters with Metabolites Drives Discovery of a Nonribosomal Peptide with an Unusual Amino Acid Monomer. *ACS Cent. Sci.* 2, 99–108. doi:10.1021/acscentsci.5b00331.

Gomez-Escribano, J. P., Song, L., Fox, D. J., Yeo, V., Bibb, M. J., and Challis, G. L. (2012). Structure and biosynthesis of the unusual polyketide alkaloid coelimycin P1, a metabolic product of the cpk gene cluster of Streptomyces coelicolor M145. *Chem. Sci.* 3, 2716–2720. doi:10.1039/C2SC20410J.

Gopalakrishnan, S., Srinivas, V., Alekhya, G., Prakash, B., Kudapa, H., Rathore, A., et al. (2015). The extent of grain yield and plant growth enhancement by plant growth-promoting broad-spectrum Streptomyces sp. in chickpea. *SpringerPlus* 4, 31. doi:10.1186/s40064-015-0811-3.

Guo, L., Liu, C., Zhao, J., Li, C., Guo, S., Fan, J., et al. (2016). Promicromonospora alba sp. nov., a novel actinomycete isolated from the cuticle of Camponotus japonicas Mayr. *Int. J. Syst. Evol. Microbiol.* doi:10.1099/ijsem.0.000885.

Gust, B., Challis, G. L., Fowler, K., Kieser, T., and Chater, K. F. (2003). PCR-targeted Streptomyces gene replacement identifies a protein domain needed for biosynthesis of the sesquiterpene soil odor geosmin. *Proc. Natl. Acad. Sci. U. S. A.* 100, 1541–1546. doi:10.1073/pnas.0337542100.

Haeder, S., Wirth, R., Herz, H., and Spiteller, D. (2009). Candicidin-producing Streptomyces support leaf-cutting ants to protect their fungus garden against the pathogenic fungus Escovopsis. *Proc. Natl. Acad. Sci.* 106, 4742–4746. doi:10.1073/pnas.0812082106.

Harikrishnan, H., Shanmugaiah, V., and Balasubramanian, N. (2014). Optimization for production on Indole acetic acid (IAA) by plant growth promoting Streptomyces sp VSMGT1014 isolated from rice rhizosphere. *Int. J. Curr. Microbiol. Appl. Sci.* 3, 158–171.

Harvey, A. L., Edrada-Ebel, R., and Quinn, R. J. (2015). The re-emergence of natural products for drug discovery in the genomics era. *Nat. Rev. Drug Discov.* 14, 111–129. doi:10.1038/nrd4510.

Hay, M. E. (1996). Marine chemical ecology: what's known and what's next? *J. Exp. Mar. Biol. Ecol.* 200, 103–134. doi:10.1016/S0022-0981(96)02659-7.

Hesketh, A., Kock, H., Mootien, S., and Bibb, M. (2009). The role of absC, a novel regulatory gene for secondary metabolism, in zinc-dependent antibiotic production in Streptomyces coelicolor A3(2). *Mol. Microbiol.* 74, 1427–1444. doi:10.1111/j.1365-2958.2009.06941.x.

Hirsch, P. R., and Mauchline, T. H. (2012). Who's who in the plant root microbiome? *Nat. Biotechnol.* 30, 961–962. doi:10.1038/nbt.2387.

Hölldobler, B., and Wilson, E. O. (2010). *The Leafcutter Ants: Civilization by Instinct*. Original edition. New York: W. W. Norton & Company.

Hong, H., Bibb, M. J., and Buttner, M. J. *The ECF sigma factors of Streptomyces coelicolor A3(2).*

Hopwood, D. A., and Merrick, M. J. (1977). Genetics of antibiotic production. *Bacteriol. Rev.* 41, 595–635.

Horn, H., Hentschel, U., and Abdelmohsen, U. R. (2015). Mining Genomes of Three Marine Sponge-Associated Actinobacterial Isolates for Secondary Metabolism. *Genome Announc.* 3, e01106-15. doi:10.1128/genomeA.01106-15.

Iqbal, H. A., Low-Beinart, L., Obiajulu, J. U., and Brady, S. F. (2016). Natural Product Discovery through Improved Functional Metagenomics in Streptomyces. *J. Am. Chem. Soc.* 138, 9341–9344. doi:10.1021/jacs.6b02921.

Ju, K.-S., Gao, J., Doroghazi, J. R., Wang, K.-K. A., Thibodeaux, C. J., Li, S., et al. (2015). Discovery of phosphonic acid natural products by mining the genomes of 10,000 actinomycetes. *Proc. Natl. Acad. Sci.* 112, 12175–12180. doi:10.1073/pnas.1500873112.

Kallifidas, D., Pascoe, B., Owen, G. A., Strain-Damerell, C. M., Hong, H.-J., and Paget, M. S. B. (2010). The zinc-responsive regulator Zur controls expression of the coelibactin gene cluster in Streptomyces coelicolor. *J. Bacteriol.* 192, 608–611. doi:10.1128/JB.01022-09.

Kaltenpoth, M., Goettler, W., Dale, C., Stubblefield, J. W., Herzner, G., Roeser-Mueller, K., et al. (2006). 'Candidatus Streptomyces philanthi', an endosymbiotic streptomycete in the antennae of Philanthus digger wasps. *Int. J. Syst. Evol. Microbiol.* 56, 1403–1411. doi:10.1099/ijs.0.64117-0.

Kaltenpoth, M., Göttler, W., Herzner, G., and Strohm, E. (2005). Symbiotic Bacteria Protect Wasp Larvae from Fungal Infestation. *Curr. Biol.* 15, 475–479. doi:10.1016/j.cub.2004.12.084.

Kaltenpoth, M., Strupat, K., and Svatoš, A. (2016). Linking metabolite production to taxonomic identity in environmental samples by (MA)LDI-FISH. *ISME J.* 10, 527–531. doi:10.1038/ismej.2015.122.

Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., and Sternberg, M. J. E. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* 10, 845–858. doi:10.1038/nprot.2015.053.

Kelly, S. L., and Kelly, D. E. (2013). Microbial cytochromes P450: biodiversity and biotechnology. Where do cytochromes P450 come from, what do they do and what can they do for us? *Philos. Trans. R. Soc. B Biol. Sci.* 368. doi:10.1098/rstb.2012.0476.

Kersten, R. D., Yang, Y.-L., Xu, Y., Cimermancic, P., Nam, S.-J., Fenical, W., et al. (2011). A mass spectrometry–guided genome mining approach for natural product peptidogenomics. *Nat. Chem. Biol.* 7, 794–802. doi:10.1038/nchembio.684.

Kinkel, L. L., Schlatter, D. C., Xiao, K., and Baines, A. D. (2014). Sympatric inhibition and niche differentiation suggest alternative coevolutionary trajectories among Streptomycetes. *ISME J.* 8, 249–256. doi:10.1038/ismej.2013.175.

Knowlton, N., Brainard, R. E., Fisher, R., Moews, M., Plaisance, L., and Caley, M. J. (2010). "Coral Reef Biodiversity," in *Life in the World's Oceans*, ed. A. D. McIntyre (Wiley-Blackwell), 65–78. Available at: http://onlinelibrary.wiley.com/doi/10.1002/9781444325508.ch4/summary [Accessed October 13, 2016].

Koehler, S., Doubský, J., and Kaltenpoth, M. (2013). Dynamics of symbiont-mediated antibiotic production reveal efficient long-term protection for beewolf offspring. *Front. Zool.* 10, 3. doi:10.1186/1742-9994-10-3.

Kois-Ostrowska, A., Strzałka, A., Lipietta, N., Tilley, E., Zakrzewska-Czerwińska, J., Herron, P., et al. (2016). Unique Function of the Bacterial Chromosome Segregation Machinery in Apically Growing Streptomyces - Targeting the Chromosome to New Hyphal Tubes and its Anchorage at the Tips. *PLOS Genet.* 12, e1006488. doi:10.1371/journal.pgen.1006488.

Komatsu, M., Kuwahara, Y., Hiroishi, A., Hosono, K., Beppu, T., and Ueda, K. (2003). Cloning of the conserved regulatory operon by its aerial mycelium-inducing activity in an amfR mutant of Streptomyces griseus. *Gene* 306, 79–89. doi:10.1016/S0378-1119(03)00405-0.

Komatsu, M., Takano, H., Hiratsuka, T., Ishigaki, Y., Shimada, K., Beppu, T., et al. (2006). Proteins encoded by the conservon of Streptomyces coelicolor A3(2) comprise a membrane-associated heterocomplex that resembles eukaryotic G protein-coupled regulatory system. *Mol. Microbiol.* 62, 1534–1546. doi:10.1111/j.1365-2958.2006.05461.x.

Konstantinidis, K. T., and Tiedje, J. M. (2004). Trends between gene content and genome size in prokaryotic species with larger genomes. *Proc. Natl. Acad. Sci.* 101, 3160–3165. doi:10.1073/pnas.0308653100.

Koonin, E. V., and Aravind, L. (2000). Dynein light chains of the Roadblock/LC7 group belong to an ancient protein superfamily implicated in NTPase regulation. *Curr. Biol. CB* 10, R774-776. doi:10.1016/s0960-9822(00)00774-0.

Kroiss, J., Kaltenpoth, M., Schneider, B., Schwinger, M.-G., Hertweck, C., Maddula, R. K., et al. (2010). Symbiotic streptomycetes provide antibiotic combination prophylaxis for wasp offspring. *Nat. Chem. Biol.* 6, 261–263. doi:10.1038/nchembio.331.

Kuang, W., Li, J., Zhang, S., and Long, L. (2015). Diversity and distribution of Actinobacteria associated with reef coral Porites lutea. *Extreme Microbiol.*, 1094. doi:10.3389/fmicb.2015.01094.

Leal, M. C., Calado, R., Sheridan, C., Alimonti, A., and Osinga, R. (2013). Coral aquaculture to support drug discovery. *Trends Biotechnol.* 31, 555–561. doi:10.1016/j.tibtech.2013.06.004.

Lee, L.-F., Yeh, S.-H., and Chen, C. W. (2002). Construction and Synchronization of dnaA Temperature-Sensitive Mutants of Streptomyces. *J. Bacteriol.* 184, 1214–1218. doi:10.1128/jb.184.4.1214-1218.2002.

Lee, N., Kim, W., Chung, J., Lee, Y., Cho, S., Jang, K.-S., et al. (2020). Iron competition triggers antibiotic biosynthesis in Streptomyces coelicolor during coculture with Myxococcus xanthus. *ISME J.* doi:10.1038/s41396-020-0594-6.

Li, X., Huang, P., Wang, Q., Xiao, L., Liu, M., Bolla, K., et al. (2014). Staurosporine from the endophytic Streptomyces sp. strain CNS-42 acts as a potential biocontrol agent and growth elicitor in cucumber. *Antonie Van Leeuwenhoek* 106, 515–525. doi:10.1007/s10482-014-0220-6.

Linares, J. F., Gustafsson, I., Baquero, F., and Martinez, J. L. (2006). Antibiotics as intermicrobial signaling agents instead of weapons. *Proc. Natl. Acad. Sci. U. S. A.* 103, 19484–19489. doi:10.1073/pnas.0608949103.

Ling, L. L., Schneider, T., Peoples, A. J., Spoering, A. L., Engels, I., Conlon, B. P., et al. (2015). A new antibiotic kills pathogens without detectable resistance. *Nature* 517, 455–459. doi:10.1038/nature14098.

Liu, C., Guan, X., Li, Y., Li, W., Ye, L., Kong, X., et al. (2016). Nocardia camponoti sp. nov., an actinomycete isolated from the head of an ant (Camponotus japonicas Mayr). *Int. J. Syst. Evol. Microbiol.* 66, 1900–1905. doi:10.1099/ijsem.0.000963.

Liu, G., Chater, K. F., Chandra, G., Niu, G., and Tan, H. (2013). Molecular Regulation of Antibiotic Biosynthesis in Streptomyces. *Microbiol. Mol. Biol. Rev. MMBR* 77, 112. doi:10.1128/MMBR.00054-12.

Liu, H., Price, M. N., Waters, R. J., Ray, J., Carlson, H. K., Lamson, J. S., et al. (2018). Magic Pools: Parallel Assessment of Transposon Delivery Vectors in Bacteria. *mSystems* 3. doi:10.1128/mSystems.00143-17.

Lugo, M. A., Ferrero, M., Menoyo, E., Estévez, M. C., Siñeriz, F., and Anton, A. (2007). Arbuscular Mycorrhizal Fungi and Rhizospheric Bacteria Diversity Along an Altitudinal Gradient in South American Puna Grassland. *Microb. Ecol.* 55, 705. doi:10.1007/s00248-007-9313-3.

Ma, J. F. (2005). Plant Root Responses to Three Abundant Soil Minerals: Silicon, Aluminum and Iron. *Crit. Rev. Plant Sci.* 24, 267–281. doi:10.1080/07352680500196017.

MacFarlane, S. A., and Merrick, M. (1985). The nucleotide sequence of the nitrogen regulation gene ntrB and the glnA-ntrBC intergenic region of Klebsiella pneumoniae. *Nucleic Acids Res.* 13, 7591–7606. doi:10.1093/nar/13.21.7591.

Mahmoud, H. M., and Kalendar, A. A. (2016). Coral-Associated Actinobacteria: Diversity, Abundance, and Biotechnological Potentials. *Extreme Microbiol.*, 204. doi:10.3389/fmicb.2016.00204.

Majka, J., Zakrzewska-Czerwińska, J., and Messer, W. (2001). Sequence Recognition, Cooperative Interaction, and Dimerization of the Initiator Protein DnaA of Streptomyces. *J. Biol. Chem.* 276, 6243–6252. doi:10.1074/jbc.M007876200.

Mak, S., and Nodwell, J. R. (2017). Actinorhodin is a redox-active antibiotic with a complex mode of action against Gram-positive cells. *Mol. Microbiol.* 106, 597–613. doi:10.1111/mmi.13837.

Mao, X.-M., Sun, N., Zheng, Y., and Li, Y.-Q. (2017). Development of Series of Affinity Tags in Streptomyces. *Sci. Rep.* 7, 6854. doi:10.1038/s41598-017-07377-4.

Mattoso, T. C., Moreira, D. D. O., and Samuels, R. I. (2012). Symbiotic bacteria on the cuticle of the leaf-cutting ant Acromyrmex subterraneus subterraneus protect workers from attack by entomopathogenic fungi. *Biol. Lett.* 8, 461–464. doi:10.1098/rsbl.2011.0963.

McAuliffe, O., Ross, R. P., and Hill, C. (2001). Lantibiotics: structure, biosynthesis and mode of action. *FEMS Microbiol. Rev.* 25, 285–308. doi:10.1111/j.1574-6976.2001.tb00579.x.

McClintock, B. (1950). The origin and behavior of mutable loci in maize. *Proc. Natl. Acad. Sci.* 36, 344–355. doi:10.1073/pnas.36.6.344.

Mehbub, M. F., Tanner, J. E., Barnett, S. J., Franco, C. M. M., and Zhang, W. (2016). The role of sponge-bacteria interactions: the sponge Aplysilla rosea challenged by its associated bacterium Streptomyces ACT-52A in a controlled aquarium system. *Appl. Microbiol. Biotechnol.*, 1–18. doi:10.1007/s00253-016-7878-9.

Merzaeva, O. V., and Shirokikh, I. G. (2010). The production of auxins by the endophytic bacteria of winter rye. *Appl. Biochem. Microbiol.* 46, 44–50. doi:10.1134/S0003683810010072.

Meschke, H., Walter, S., and Schrempf, H. (2012). Characterization and localization of prodiginines from Streptomyces lividans suppressing Verticillium dahliae in the absence or presence of Arabidopsis thaliana. *Environ. Microbiol.* 14, 940–952. doi:10.1111/j.1462-2920.2011.02665.x.

Milshteyn, A., Schneider, J. S., and Brady, S. F. (2014). Mining the metabiome: identifying novel natural products from microbial communities. *Chem. Biol.* 21, 1211–1223. doi:10.1016/j.chembiol.2014.08.006.

Müller, M., Agarwal, N., and Kim, J. (2016). A Cytochrome P450 3A4 Biosensor Based on Generation 4.0 PAMAM Dendrimers for the Detection of Caffeine. *Biosensors* 6. doi:10.3390/bios6030044.

Muñoz-López, M., and García-Pérez, J. L. (2010). DNA Transposons: Nature and Applications in Genomics. *Curr. Genomics* 11, 115–128. doi:10.2174/138920210790886871.

Musialowski, M. S., Flett, F., Scott, G. B., Hobbs, G., Smith, C. P., and Oliver, S. G. (1994). Functional evidence that the principal DNA replication origin of the Streptomyces coelicolor chromosome is close to the dnaA-gyrB region. *J. Bacteriol.* 176, 5123–5125. doi:10.1128/jb.176.16.5123-5125.1994.

Nabti, E., Bensidhoum, L., Tabli, N., Dahel, D., Weiss, A., Rothballer, M., et al. (2014). Growth stimulation of barley and biocontrol effect on plant pathogenic fungi by a Cellulosimicrobium sp. strain isolated from salt-affected rhizosphere soil in northwestern Algeria. *Eur. J. Soil Biol.* 61, 20–26. doi:10.1016/j.ejsobi.2013.12.008.

Nan, B., Bandaria, J. N., Guo, K. Y., Fan, X., Moghtaderi, A., Yildiz, A., et al. (2015). The polarity of myxobacterial gliding is regulated by direct interactions between the gliding motors and the Ras homolog MglA. *Proc. Natl. Acad. Sci.* 112, E186–E193. doi:10.1073/pnas.1421073112.

Onaka, H., Mori, Y., Igarashi, Y., and Furumai, T. (2011). Mycolic acid-containing bacteria induce natural-product biosynthesis in Streptomyces species. *Appl. Environ. Microbiol.* 77, 400–406. doi:10.1128/AEM.01337-10.

Petzke, L., and Luzhetskyy, A. (2009). In vivo Tn5-based transposon mutagenesis of Streptomycetes. *Appl. Microbiol. Biotechnol.* 83, 979–986. doi:10.1007/s00253-009-2047-z.

Pham, T. M., Wiese, J., Wenzel-Storjohann, A., and Imhoff, J. F. (2015). Diversity and antimicrobial potential of bacterial isolates associated with the soft coral Alcyoniumdigitatum from the Baltic Sea. *Antonie Van Leeuwenhoek* 109, 105–119. doi:10.1007/s10482-015-0613-1.

Pimentel-Elardo, S. M., Sørensen, D., Ho, L., Ziko, M., Bueler, S. A., Lu, S., et al. (2015). Activity-Independent Discovery of Secondary Metabolites Using Chemical Elicitation and Cheminformatic Inference. *ACS Chem. Biol.* 10, 2616–2623. doi:10.1021/acschembio.5b00612.

Pishchany, G., Mevers, E., Ndousse-Fetter, S., Horvath, D. J., Paludo, C. R., Silva-Junior, E. A., et al. (2018). Amycomicin is a potent and specific antibiotic discovered with a targeted interaction screen. *Proc. Natl. Acad. Sci.* 115, 10124–10129. doi:10.1073/pnas.1807613115.

Poulsen, M., Erhardt, D. P., Molinaro, D. J., Lin, T.-L., and Currie, C. R. (2007). Antagonistic Bacterial Interactions Help Shape Host-Symbiont Dynamics within the Fungus-Growing Ant-Microbe Mutualism. *PLoS ONE* 2. doi:10.1371/journal.pone.0000960.

Practical Streptomyces Genetics | NHBS Academic & Professional Books Available at: https://www.nhbs.com/practical-streptomyces-genetics-book [Accessed October 8, 2020].

Rajalingam, K., Schreck, R., Rapp, U. R., and Albert, Š. (2007). Ras oncogenes and their downstream targets. *Biochim. Biophys. Acta BBA - Mol. Cell Res.* 1773, 1177–1195. doi:10.1016/j.bbamcr.2007.01.012.

Reynolds, H. T., and Currie, C. R. (2004). Pathogenicity of Escovopsis weberi: The parasite of the attine ant-microbe symbiosis directly consumes the ant-cultivated fungus. *Mycologia* 96, 955–959.

Rodríguez, H., Rico, S., Díaz, M., and Santamaría, R. I. (2013). Two-component systems in Streptomyces: key regulators of antibiotic complex pathways. *Microb. Cell Factories* 12, 127. doi:10.1186/1475-2859-12-127.

Romero, F., Espliego, F., Pérez Baz, J., García de Quesada, T., Grávalos, D., de la Calle, F., et al. (1997). Thiocoraline, a new depsipeptide with antitumor activity produced by a

marine Micromonospora. I. Taxonomy, fermentation, isolation, and biological activities. *J. Antibiot. (Tokyo)* 50, 734–737.

Rungin, S., Indananda, C., Suttiviriya, P., Kruasuwan, W., Jaemsaeng, R., and Thamchaipenet, A. (2012). Plant growth enhancing effects by a siderophore-producing endophytic streptomycete isolated from a Thai jasmine rice plant (Oryza sativa L. cv. KDML105). *Antonie Van Leeuwenhoek* 102, 463–472. doi:10.1007/s10482-012-9778-z.

Ryan, R. P., and Dow, J. M. (2008). Diffusible signals and interspecies communication in bacteria. *Microbiol. Read. Engl.* 154, 1845–1858. doi:10.1099/mic.0.2008/017871-0.

Sánchez, S., Chávez, A., Forero, A., García-Huante, Y., Romero, A., Sánchez, M., et al. (2010). Carbon source regulation of antibiotic production. *J. Antibiot. (Tokyo)* 63, 442–459. doi:10.1038/ja.2010.78.

Schatz, A., and Waksman, S. A. (1944). Effect of Streptomycin and Other Antibiotic Substances upon Mycobacterium tuberculosis and Related Organisms. *Exp. Biol. Med.* 57, 244–248. doi:10.3181/00379727-57-14769.

Schneemann, I., Nagel, K., Kajahn, I., Labes, A., Wiese, J., and Imhoff, J. F. (2010). Comprehensive Investigation of Marine Actinobacteria Associated with the Sponge Halichondria panicea. *Appl. Environ. Microbiol.* 76, 3702–3714. doi:10.1128/AEM.00780-10.

Schoenian, I., Spiteller, M., Ghaste, M., Wirth, R., Herz, H., and Spiteller, D. (2011). Chemical basis of the synergism and antagonism in microbial communities in the nests of leaf-cutting ants. *Proc. Natl. Acad. Sci.* 108, 1955–1960. doi:10.1073/pnas.1008441108.

Schultz, T. R., and Brady, S. G. (2008). Major evolutionary transitions in ant agriculture. *Proc. Natl. Acad. Sci.* 105, 5435–5440. doi:10.1073/pnas.0711024105.

Sega, G. A. (1984). A review of the genetic effects of ethyl methanesulfonate. *Mutat. Res.* 134, 113–142. doi:10.1016/0165-1110(84)90007-1.

Seipke, R. F., Barke, J., Brearley, C., Hill, L., Yu, D. W., Goss, R. J. M., et al. (2011). A Single Streptomyces Symbiont Makes Multiple Antifungals to Support the Fungus Farming Ant Acromyrmex octospinosus. *PLOS ONE* 6, e22028. doi:10.1371/journal.pone.0022028.

Seipke, R. F., Kaltenpoth, M., and Hutchings, M. I. (2012). Streptomyces as symbionts: an emerging and widespread theme? *FEMS Microbiol. Rev.* 36, 862–876. doi:10.1111/j.1574-6976.2011.00313.x.

Seyedsayamdost, M. R., Traxler, M. F., Zheng, S.-L., Kolter, R., and Clardy, J. (2011). Structure and Biosynthesis of Amychelin, an Unusual Mixed-Ligand Siderophore from Amycolatopsis sp. AA4. *J. Am. Chem. Soc.* 133, 11434–11437. doi:10.1021/ja203577e.

Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., et al. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* 7, 539. doi:10.1038/msb.2011.75.

Sit, C. S., Ruzzini, A. C., Arnam, E. B. V., Ramadhar, T. R., Currie, C. R., and Clardy, J. (2015). Variable genetic architectures produce virtually identical molecules in bacterial symbionts of fungus-growing ants. *Proc. Natl. Acad. Sci.* 112, 13150–13154. doi:10.1073/pnas.1515348112.

Skerker, J. M., Prasol, M. S., Perchuk, B. S., Biondi, E. G., and Laub, M. T. (2005). Two-component signal transduction pathways regulating growth and cell cycle progression in a bacterium: a system-level analysis. *PLoS Biol.* 3, e334. doi:10.1371/journal.pbio.0030334.

Smith, M. L., Bruhn, J. N., and Anderson, J. B. (1992). The fungus Armillaria bulbosa is among the largest and oldest living organisms. *Nature* 356, 428–431. doi:10.1038/356428a0.

Solans, M., Scervino, J. M., Messuti, M. I., Vobis, G., and Wall, L. G. (2016). Potential biocontrol actinobacteria: Rhizospheric isolates from the Argentine Pampas lowlands legumes. *J. Basic Microbiol.*, n/a-n/a. doi:10.1002/jobm.201600323.

Solans, M., Vobis, G., Cassán, F., Luna, V., and Wall, L. G. (2011). Production of phytohormones by root-associated saprophytic actinomycetes isolated from the actinorhizal plant Ochetophila trinervis. *World J. Microbiol. Biotechnol.* 27, 2195–2202. doi:10.1007/s11274-011-0685-7.

Stankovic, N., Senerovic, L., Ilic-Tomic, T., Vasiljevic, B., and Nikodinovic-Runic, J. (2014). Properties and applications of undecylprodigiosin and other bacterial prodigiosins. *Appl. Microbiol. Biotechnol.* 98, 3841–3858. doi:10.1007/s00253-014-5590-1.

Stierle, A. C., Cardellina, J. H., and Singleton, F. L. (1988). A marine Micrococcus produces metabolites ascribed to the sponge Tedania ignis. *Experientia* 44, 1021.

Stock, A., Chen, T., Welsh, D., and Stock, J. (1988). CheA protein, a central regulator of bacterial chemotaxis, belongs to a family of proteins that control gene expression in response to changing environmental conditions. *Proc. Natl. Acad. Sci.* 85, 1403–1407. doi:10.1073/pnas.85.5.1403.

Stock, A. M., Robinson, V. L., and Goudreau, P. N. (2000). Two-component signal transduction. *Annu. Rev. Biochem.* 69, 183–215. doi:10.1146/annurev.biochem.69.1.183.

Straight, P. D., Willey, J. M., and Kolter, R. (2006). Interactions between Streptomyces coelicolor and Bacillus subtilis: Role of Surfactants in Raising Aerial Structures. *J. Bacteriol.* 188, 4918–4925. doi:10.1128/JB.00162-06.

Strohm, E., and Linsenmair, K. E. (2001). Females of the European beewolf preserve their honeybee prey against competing fungi. *Ecol. Entomol.* 26, 198–203. doi:10.1046/j.1365-2311.2001.00300.x.

Subramani, R., and Aalbersberg, W. (2012). Marine actinomycetes: An ongoing source of novel bioactive metabolites. *Microbiol. Res.* 167, 571–580. doi:10.1016/j.micres.2012.06.005.

Subramani, R., and Aalbersberg, W. (2013). Culturable rare Actinomycetes: diversity, isolation and marine natural product discovery. *Appl. Microbiol. Biotechnol.* 97, 9291–9321. doi:10.1007/s00253-013-5229-7.

Subramaniam, G., Arumugam, S., and Rajendran, V. (2016). *Plant Growth Promoting Actinobacteria: A New Avenue for Enhancing the Productivity and Soil Fertility of Grain Legumes*. Springer.

Suvorova, I. A., Korostelev, Y. D., and Gelfand, M. S. (2015). GntR Family of Bacterial Transcription Factors and Their DNA Binding Motifs: Structure, Positioning and Co-Evolution. *PLoS ONE* 10. doi:10.1371/journal.pone.0132618.

Takano, E., Chakraburtty, R., Nihira, T., Yamada, Y., and Bibb, M. J. (2001). A complex role for the gamma-butyrolactone SCB1 in regulating antibiotic production in Streptomyces coelicolor A3(2). *Mol. Microbiol.* 41, 1015–1028. doi:10.1046/j.1365-2958.2001.02562.x.

Takano, E., Kinoshita, H., Mersinias, V., Bucca, G., Hotchkiss, G., Nihira, T., et al. (2005). A bacterial hormone (the SCB1) directly controls the expression of a pathway-specific regulatory gene in the cryptic type I polyketide biosynthetic gene cluster of Streptomyces coelicolor. *Mol. Microbiol.* 56, 465–479. doi:10.1111/j.1365-2958.2005.04543.x.

Takano, H., Hashimoto, K., Yamamoto, Y., Beppu, T., and Ueda, K. (2011). Pleiotropic effect of a null mutation in the cvn1 conservon of Streptomyces coelicolor A3(2). *Gene* 477, 12–18. doi:10.1016/j.gene.2011.01.005.

Taylor, B. L., and Zhulin, I. B. (1999). PAS Domains: Internal Sensors of Oxygen, Redox Potential, and Light. *Microbiol. Mol. Biol. Rev.* 63, 479–506.

Toumatia, O., Compant, S., Yekkour, A., Goudjal, Y., Sabaou, N., Mathieu, F., et al. (2016). Biocontrol and plant growth promoting properties of Streptomyces mutabilis strain IA1 isolated from a Saharan soil on wheat seedlings and visualization of its niches of colonization. *South Afr. J. Bot.* 105, 234–239. doi:10.1016/j.sajb.2016.03.020.

Traxler, M. F., and Kolter, R. (2015). Natural products in soil microbe interactions and evolution. *Nat. Prod. Rep.* 32, 956–970. doi:10.1039/c5np00013k.

Traxler, M. F., Watrous, J. D., Alexandrov, T., Dorrestein, P. C., and Kolter, R. (2013). Interspecies interactions stimulate diversification of the Streptomyces coelicolor secreted metabolome. *mBio* 4. doi:10.1128/mBio.00459-13.

Tsigkinopoulou, A., Takano, E., and Breitling, R. (2020). Unravelling the γ-butyrolactone network in Streptomyces coelicolor by computational ensemble modelling. *PLOS Comput. Biol.* 16, e1008039. doi:10.1371/journal.pcbi.1008039.

Udwary, D. W., Zeigler, L., Asolkar, R. N., Singan, V., Lapidus, A., Fenical, W., et al. (2007). Genome sequencing reveals complex secondary metabolome in the marine actinomycete Salinispora tropica. *Proc. Natl. Acad. Sci. U. S. A.* 104, 10376–10381. doi:10.1073/pnas.0700962104.

Ueda, K., Kawai, S., Ogawa, H., Kiyama, A., Kubota, T., Kawanobe, H., et al. (2000). Wide distribution of interspecific stimulatory events on antibiotic production and sporulation among Streptomyces species. *J. Antibiot. (Tokyo)* 53, 979–982.

Van Arnam, E. B., Ruzzini, A. C., Sit, C. S., Currie, C. R., and Clardy, J. (2015). A Rebeccamycin Analog Provides Plasmid-Encoded Niche Defense. *J. Am. Chem. Soc.* 137, 14272–14274. doi:10.1021/jacs.5b09794.

Van Middlesworth, F., and Cannell, RichardJ. P. (1998). "Dereplication and Partial Identification of Natural Products," in *Natural Products Isolation* Methods in Biotechnology., ed. RichardJ. P. Cannell (Humana Press), 279–327. doi:10.1007/978-1-59259-256-2_10.

Vetsigian, K., Jajoo, R., and Kishony, R. (2011). Structure and Evolution of Streptomyces Interaction Networks in Soil and In Silico. *PLOS Biol.* 9, e1001184. doi:10.1371/journal.pbio.1001184.

Webre, D. J., Wolanin, P. M., and Stock, J. B. (2003). Bacterial chemotaxis. *Curr. Biol. CB* 13, R47-49. doi:10.1016/s0960-9822(02)01424-0.

Widenbrant, E. M., and Kao, C. M. (2007). Introduction of the foreign transposon Tn4560 in Streptomyces coelicolor leads to genetic instability near the native insertion sequence IS1649. *J. Bacteriol.* 189, 9108–9116. doi:10.1128/JB.00983-07.

Wright, L. F., and Hopwood, D. A. (1976). Actinorhodin is a chromosomally-determined antibiotic in Streptomyces coelicolar A3(2). *J. Gen. Microbiol.* 96, 289–297. doi:10.1099/00221287-96-2-289.

Wuichet, K., and Søgaard-Andersen, L. (2014). Evolution and Diversity of the Ras Superfamily of Small GTPases in Prokaryotes. *Genome Biol. Evol.* 7, 57–70. doi:10.1093/gbe/evu264.

Xi, L., Ruan, J., and Huang, Y. (2012). Diversity and Biosynthetic Potential of Culturable Actinomycetes Associated with Marine Sponges in the China Seas. *Int. J. Mol. Sci.* 13, 5917–5932. doi:10.3390/ijms13055917.

Xu, Y., and Her, C. (2015). Inhibition of Topoisomerase (DNA) I (TOP1): DNA Damage Repair and Anticancer Therapy. *Biomolecules* 5, 1652–1670. doi:10.3390/biom5031652.

Xue, L., Xue, Q., Chen, Q., Lin, C., Shen, G., and Zhao, J. (2013). Isolation and evaluation of rhizosphere actinomycetes with potential application for biocontrol of Verticillium wilt of cotton. *Crop Prot.* 43, 231–240. doi:10.1016/j.cropro.2012.10.002.

Yamanaka, K., Oikawa, H., Ogawa, H., Hosono, K., Shinmachi, F., Takano, H., et al. (2005). Desferrioxamine E produced by Streptomyces griseus stimulates growth and development of Streptomyces tanashiensis. *Microbiol. Read. Engl.* 151, 2899–2905. doi:10.1099/mic.0.28139-0.

Yamanaka, K., Reynolds, K. A., Kersten, R. D., Ryan, K. S., Gonzalez, D. J., Nizet, V., et al. (2014). 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. doi:10.1073/pnas.1319584111.

Yan, L.-L., Han, N.-N., Zhang, Y.-Q., Yu, L.-Y., Chen, J., Wei, Y.-Z., et al. (2010). Antimycin A18 produced by an endophytic Streptomyces albidoflavus isolated from a mangrove plant. *J. Antibiot. (Tokyo)* 63, 259–261. doi:10.1038/ja.2010.21.

Yarnold, J. E., Hamilton, B. R., Welsh, D. T., Pool, G. F., Venter, D. J., and Carroll, A. R. (2012). High resolution spatial mapping of brominated pyrrole-2-aminoimidazole alkaloids distributions in the marine sponge Stylissa flabellata via MALDI-mass spectrometry imaging. *Mol. Biosyst.* 8, 2249–2259. doi:10.1039/c2mb25152c.

Young, I. M., and Crawford, J. W. (2004). Interactions and Self-Organization in the Soil-Microbe Complex. *Science* 304, 1634–1637. doi:10.1126/science.1097394.

Zawilak-Pawlik, A., Kois, A., Majka, J., Jakimowicz, D., Smulczyk-Krawczyszyn, A., Messer, W., et al. (2005). Architecture of bacterial replication initiation complexes: orisomes from four unrelated bacteria. *Biochem. J.* 389, 471–481. doi:10.1042/BJ20050143.

Zhang, J., Wang, J.-D., Liu, C.-X., Yuan, J.-H., Wang, X.-J., and Xiang, W.-S. (2014). A new prenylated indole derivative from endophytic actinobacteria Streptomyces sp. neau-D50. *Nat. Prod. Res.* 28, 431–437. doi:10.1080/14786419.2013.871546.

Zhang, X., Bao, Y., Shi, X., Ou, X., Zhou, P., and Ding, X. (2012). Efficient transposition of IS204-derived plasmids in Streptomyces coelicolor. *J. Microbiol. Methods* 88, 67–72. doi:10.1016/j.mimet.2011.10.018.

Zhang, Y., Franco, M., Ducret, A., and Mignot, T. (2010). A Bacterial Ras-Like Small GTP-Binding Protein and Its Cognate GAP Establish a Dynamic Spatial Polarity Axis to Control Directed Motility. *PLoS Biol.* 8. doi:10.1371/journal.pbio.1000430.

Zhang, Z., Du, C., de Barsy, F., Liem, M., Liakopoulos, A., van Wezel, G. P., et al. (2020). Antibiotic production in Streptomyces is organized by a division of labor through terminal genomic differentiation. *Sci. Adv.* 6. doi:10.1126/sciadv.aay5781.

Ziemert, N., Lechner, A., Wietz, M., Millán-Aguiñaga, N., Chavarria, K. L., and Jensen, P. R. (2014). Diversity and evolution of secondary metabolism in the marine actinomycete genus Salinispora. *Proc. Natl. Acad. Sci. U. S. A.* 111, E1130-1139. doi:10.1073/pnas.1324161111.

Zoraghi, R., Corbin, J. D., and Francis, S. H. (2004). Properties and Functions of GAF Domains in Cyclic Nucleotide Phosphodiesterases and Other Proteins. *Mol. Pharmacol.* 65, 267–278. doi:10.1124/mol.65.2.267.