THE PENNSYLVANIA STATE UNIVERSITY
SCHREYER HONORS COLLEGE

DEPARTMENT OF BIOCHEMISTRY AND MOLECULAR BIOLOGY

Activation of Silent Gene clusters in Streptomyces spp.

JOAO RAMOS FERREIRA
SPRING 2022

A thesis
submitted in partial fulfillment
of the requirements
for a baccalaureate degree
in Biochemistry and Molecular Biology
with honors in Biochemistry and Molecular Biology

Reviewed and approved* by the following:

Timothy Meredith
Associate Professor of Biochemistry and Molecular Biology
Thesis Supervisor

Teh-hui Kao
Distinguished Professor of Biochemistry and Molecular Biology
Honors Adviser

* Electronic approvals are on file.

ABSTRACT

Streptomyces is a genus of Gram-positive bacteria commonly found in soil environments. An important attribute of *Streptomyces* is its ability to produce secondary metabolites with different bioactive properties, including antibacterial compounds. Around 50 percent of the antibiotics currently in use are derived from the secondary metabolites of streptomyces. These metabolites are encoded by large biosynthetic gene clusters within the bacteria's genome, and a single species can have more than 30 gene clusters. Whole genomic sequencing has identified silent gene clusters encoding secondary metabolites of unknown biological activity. Expression of these gene clusters and characterizing their products could serve as a source of novel antibiotics. The goal of this research is to explore a method of activation of silent gene clusters, through tying secondary metabolite synthesis to primary metabolite production and thus necessary for growth and development. Through the use of CRISPR plasmids, the *thyX* gene was excised from *S. avermitilis* creating an auxotroph. The *thyX* gene was reinserted adjacent to silent secondary metabolite gene cluster of interest, and once removed from the thymidine/thymine supplemental media, *S. avermitilis* develops suppressor mutations that allow it to now express *thyX* along with the gene cluster of interest to be expressed. In a proof of principle study, suppressor mutants were generated for the SA20 (encoding oligomycin) and SA7 (encoding avermectin) gene clusters. Suppressor mutations were then generated for each gene cluster and the production of their compound was screened with a yeast bioassay for oligomycin and thin layer chromatography for avermectin. Sequencing of suppressor mutants identified most mutants engaged in local chromosome duplication of the *thyX* gene to survive genetic pressures. Another mutant possessed a mutation at the *rpme1* gene locus. Future studies are needed to identify a mutant that produces an overabundance of secondary metabolites in order to widely apply this method in silent gene cluster activation.

TABLE OF CONTENTS

LIST OF FIGURES iii

LIST OF TABLES iv

ACKNOWLEDGEMENTS v

Chapter 1 1

Introduction 1
1.1 Streptomyces characteristics 1
1.2 Genome of Streptomyces 2
1.4 Polyketide Synthetases 3
1.5 Avermectin 4
1.6 Oligomycin 6
1.7 Antibiotic resistance 7
1.7 Scope and significance of the project 8

Chapter 2 Materials and Methods 10

Materials 10
2.1 Media 10
2.2 Strains 10
Methods 11
2.3 Genetic manipulation of *Streptomyces* 11
2.4 Generation of suppressor mutants 12
2.5 Isolating genomic DNA 13
2.6 Avermectin TLC assay 14
2.7 Oligomycin yeast screen assay 14
2.8 rtPCR 15

Chapter 3 Results 16

Evaluation of Mutagenesis protocols: Spontaneous verse dP induced 16

Chapter 4 Sequencing Data 20

Chapter 5 Avermectin TLC 23

Chapter 6 Discussion and Future Directions 27

BIBLIOGRAPHY 30

LIST OF FIGURES

Figure 1 *S. avermitilis* genome map.....2

Figure 2 Avermectin mechanism of action.....5

Figure 3 Oligomycin mechanism of action.....6

Figure 4 Experimental strategy……..……..9

Figure 5 CRISPR plasmids used.....12

Figure 6 Structure of mutagen dP……...17

Figure 7 SA20 yeast assay (1).....18

Figure 8 SA20 yeast assay (2).....19

Figure 9 SA20 Suppressor 2 genome map.....20

Figure 10 *rpmE1* mutation site…….....21

Figure 11 SA7 avermectin TLC (1)…….....24

Figure 12 SA7 avermectin TLC (2)…….....25

Figure 13 SA7 avermectin TLC (3)…….....26

LIST OF TABLES

Table 1 SA20 yeast assay data (1)…….....18

Table 2 SA20 yeast assay data (2)…….....19

Table 3 rtPCR Data table…….…….……..22

ACKNOWLEDGEMENTS

I would like to thank Dr. Meredith, my undergraduate thesis advisor, for his mentoring and support. Thank you for allowing me to a part of this research project and for investing time and energy into my development as an undergraduate researcher and future scientist. Thank you to Dr. Komazin-Meredith for helping me and supporting me in the lab with techniques and resources. I could not have accomplished much without your help and support.

Chapter 1

Introduction

1.1 Streptomyces characteristics

Streptomyces is a genus of Gram-positive bacteria belonging to the phylum Actinobacteria. They are soil saprobes and share morphological characteristics more similar to fungi than bacteria, growing as a filamentous mycelium with aerial hyphae and sporulating to reproduce[1]. They are the largest genus within the Actinobacteria phylum with over 900 described species[2]. The formation of spores allows the bacteria to survive under harsh environments, and species are found in soils across the globe. That, along with their complex genomes and ability to produce a myriad of secondary metabolites, allows them to be some of the most versatile, successful, and important soil microorganisms. As soil saprobes *Streptomyces* play a vital role in the decomposition and recycling of organic matter and thus the maintenance of a healthy ecosystem. For example, the species *S. coelicolor* produces abundance of catalytic enzymes, including 60 proteases, 13 chitanases, and 8 cellulases[3]. The breakdown of chitin provides key nutrients for the bacteria.

By far, the most commercially and scientifically important attribute of *Streptomyces* is its ability to produce a variety of secondary metabolites with different bioactive properties, including antibacterial, antifungal, antiparasitic, antitumor and immunosuppressant and antihypertensive compounds[4]. In nature, these compounds serve as a defense mechanism to help the bacteria compete against other soil microorganisms and aid in a mutualistic association with plants whereby the bacteria kills off pathogens in exchange for nutrients[3]. In some cases, *Streptomyces* itself can be a plant pathogen, such as in scab diseases of potato caused by *Streptomyces turgidiscapies*. Actinomycetes make around 2⁄3 of the known

antibiotics that are produced by microorganisms, of those 80% are made by members of the genus Streptomyces[5]. The species *S. griseu*s was the *Streptomyces* to be grown commercially for the ultra-important antibiotic and namesake streptomycin, discovered in 1942, the first antibiotic developed to treat tuberculosis one of the most common and widespread infections in the world[3].

1.2 Genome of Streptomyces

Recent advances in whole genome sequencing technology have allowed for sequencing and characterization of various *Streptomyces* species. The genes that encode the unique characteristics of the bacteria give *Streptomyces* a uniquely large genome. Most *Streptomyces* chromosomal DNA molecules are 6-10 Mb long (*E. coli* 4-5 Mb), are linear with terminal inverted repeats (telomeres), and have over 70% GC content[6], higher than nearly all other organisms, setting them apart from other well-known bacteria such as *Bacillus subtilis* and *Escherichia coli*. The pan-Streptomyces genome contains at least 34,000 genes[1]. The core region of the gene is highly conserved between species and contains most of the essential genes. The subtelomeric region of the chromosome can be highly variable in length from species to species, ranging from 49 to 132,910 bp, and contains about half of the genes associated with secondary metabolites[4]. Additionally, the telomeric regions contain most of the mobile and insertion elements within the genome.

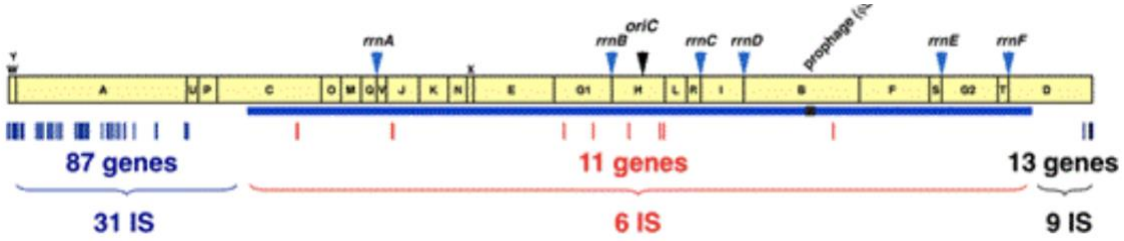

Figure 1: A map of the *S. avermitilis* linear genome (9,025,608 bases). The vertical bars are transposase genes, the thick blue bar is highlighting the central core region of the genome. The origin of replication and several rRNA operons are also annotated (from Ikeda et al., 2014).

1.3 Secondary Metabolites of streptomyces

Streptomyces secondary metabolites are structurally diverse and complex; over 20 gene clusters can be found within a single species. They are produced from common intermediates condensed into more complex structures through two distinct chemical pathways. About half to three quarters of the secondary metabolite gene clusters are associated with nonribosomal peptide synthetase (NRPS) and polyketide synthases (PKS)[2]. The other secondary metabolites include ribosomal peptides, bacteriocins, terpenoids, aminoglycosides and others[4]. How the bacteria acquire such a large collection of secondary metabolite genes, genes unnecessary for primary metabolism, is not completely known, though some theories include horizontal gene transfer and duplication of related genes. Complete sequencing and genome mining have shown that these genes constitute a large part of coding capacity, making up between 6-10% of the organism's genome[2]. Genome analysis of the model *Streptomyces* type strain *S. avermitilis* genome allowed for the identification and location of 38 secondary (PKS) pathway metabolite gene clusters[4]. The clusters revealed some previously discovered compounds and some of unknown structure and function. About half or more of the genes related to secondary metabolites were found at the end of the chromosomes in the telomeric region while the other half are scattered throughout the genome. The total length of these clusters occupies about 6.4% of the genome of *S. avermitilies.*

1.4 Polyketide Synthetases

Polyketide synthetases are multifunctional enzyme complexes that construct polyketides and are involved in the biosynthesis of a wide variety of natural products[7]. Indeed, PKS is the primary mechanism by which many important drugs are made including the antibiotic erythromycin, the cholesterol drug lovastatin, and even the cancer chemotherapy drug epothilone. Polyketide synthetases (PKS) are grouped into three types: Type I PKS, Type II PKS, and Type III PKS. Type I PKS are mainly found in bacteria

and sometimes fungi, and the primary type utilized by *Streptomyces*. Type I PKS are multifunctional proteins with several functional domains and functions as an assembly line system whereby the nascent polyketide is shuttled from module to module and each module's enzyme catalyzes a chemical reaction to elongate the polyketide[8]. A PKS consists of 3 modules each with distinct domains. Firstly, the starting or loading module AT-ACP, an acyl-Coenzyme A is transferred by an acyl carrier protein to the acyltransferase domain AC. The elongation module varies in length and can contain the following domains: ketosynthase (KS), acyltransferase (AT), dehydratase (DH), enoylreductase (ER), and ketoreductase (KR). In elongation, the nascent polyketide is transferred from one domain to the other where it is chemically modified by the distinct enzymes listed[9]. Finally, the termination module which consists of a thioesterase (TE) domain either hydrolyzes or cyclizes the polyketide to finish elongation. The modular nature of assembly line PKS allows a broad diversity of complex polyketide products. A single PKS can contain up to 30 modules, the repetition of the same modules within a PKS, such as the clustering of KS domains, is believed to be the result of gene duplication[10]. The evolution of PKS genes is also a subject of research, but sequence similarity has shown an evolutionary relationship between PKS and genes such as endiyne synthases, polyunsaturated fatty acid synthases, and glycolipid synthases.

1.5 Avermectin

Avermectin and its derivatives are secondary metabolites with anthelminthic activity isolated from *Streptomyces avermitilis*. *S. avermitlis* was first discovered from a soil sample in Japan in 1977[11]. Iveremectin, the semisynthetic derivative of avermectin, has been used since the 1980s in animal and human medicine. The discovery and subsequent utilization of avermectins as a commercial drug won the Nobel Prize for Physiology and Medicine in 2015 for its "extraordinary efficacy" in treating River blindness and Lymphatic filariasis, among other parasitic diseases affecting animals and humans. Avermectins are 16 membered macrocyclic lactones derived by pentacyclic polyketides[4]. The

biosynthetic gene cluster for avermectin is 7th in sequence order of biosynthetic gene clusters found in the *S. avermitilis* genome and is about 80 kb long. In addition to its potent antihemnthic properties, avermectins are also effective against nematodes, insects, and spiders making them a prime drug used in pest control against a variety of invertebrates[2]. Avermectins work by disrupting the electrical signals of nerve impulses in invertebrates specifically targeting their chloride gated ion channels and amplifying the effects of glutamate, which culminates in paralysis and death[12]. Mammals do not possess chlorine gated ion channels therefore, avermectin is nontoxic to humans and animals. Avermectin and other PKS gene clusters in *S. avermitilis* are not active and generating compounds under standard laboratory conditions, resulting in long processes of strain optimization to increase levels of production to sufficient quantities for study[13].

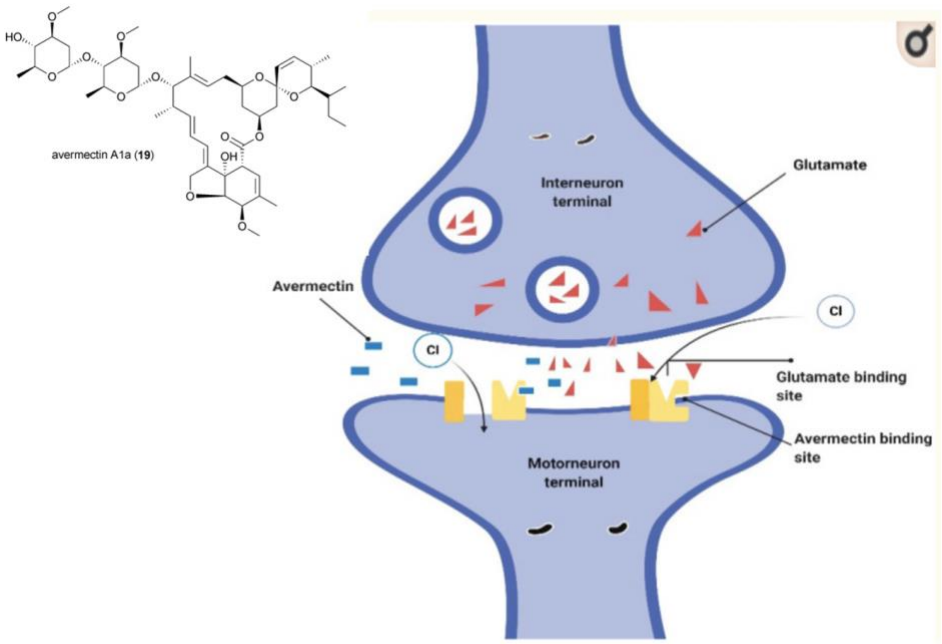

Figure 2. The chemical structure of avermectin as well as its mechanism of action (from El-Saber Batiha et al., 2020).

1.6 Oligomycin

Oligomycin is a polyketide macrocyclic lactone also produced by *S. avermitilis*. The oligomycin gene cluster is the 20th biosynthetic gene cluster in the *S. avermiltis* genome, the cluster being around 100 kb long encoding a type I PKS with 11 protein coding genes[4]. Oligomycin was originally discovered in *Streptomyces diastatochromogenes* in 1954[5]. Oligomycin is an antifungal compound; it targets eukaryotic mitochondria and inhibits F_oF_1-ATP synthase by blocking its proton channel (F_o domain), stopping mitochondrial electron transport chain. There are a total of 9 oligomycin isoforms known, all containing 26 membered macro lactone rings but with slightly different functional groups at different positions[5]. The three major isoforms are Oligomycin A, B, and C. Like avermectin, the oligomycin gene cluster is not active in most *Steptomyces* species under standard laboratory conditions.

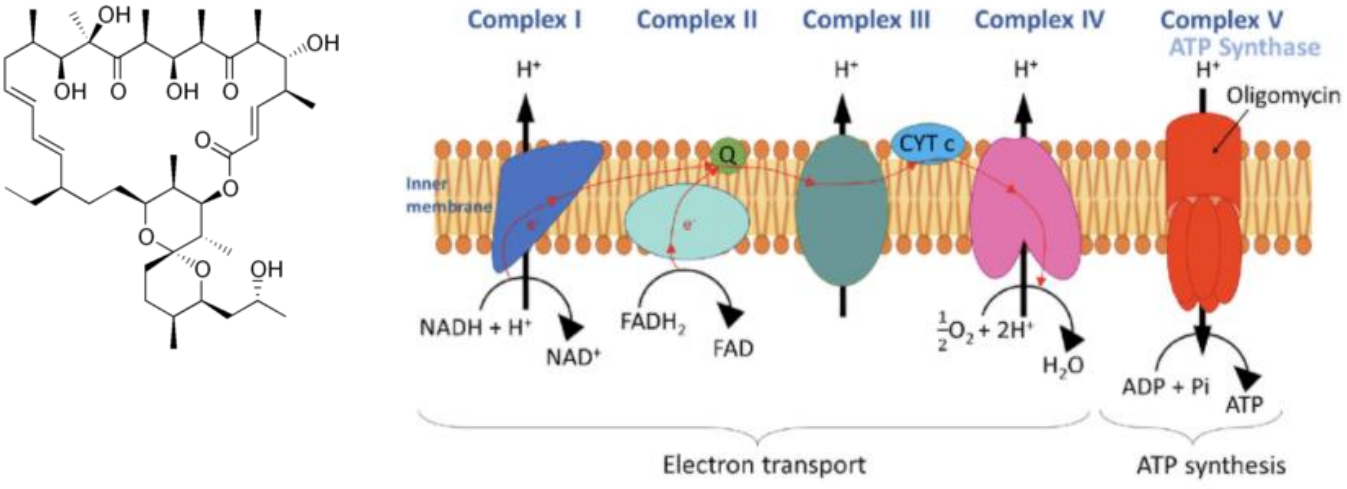

Figure 3. a. The chemical structure and b. mechanism of action of oligomycin (from Ratajczak et al., 2019).

1.7 Antibiotic resistance

Although the use of *Streptomyces* secondary metabolites has been a powerful tool for the development of antibiotics, antibiotic resistant infections are on the rise and so are the demand for new antibiotics. Worldwide, about 17 million people die annually of bacterial infections[3]. The poor administration and misuse of antibiotics has led to the rapid development of resistance, making key drugs that were once highly effective, termed "magic bullets", nearly obsolete. Antibiotic resistant infections can cause severe illness and even death. The CDC estimates that each year in the United States, at least 2 million people become infected with bacteria that are resistant to antibiotics and at least 23,000 people die as a direct result of these infections[15]. Some common human pathogens that are on the rise include *Staphylococcus aureus* a common cause of skin infections and the causative agent of MRSA, as well as multi drug resistant *Mycobacterium tuberculosis*, which is a growing problem in both developing and developed countries[16]. Resistance arises naturally via mutations or is acquired from resistance genes via horizontal gene transfer of plasmid DNA[16]. Mechanisms of resistance include drug efflux, antibiotic degradation enzymes or modification of drug targets[16]. Multi drug resistant pathogens cause more difficult to treat diseases which prolong illness and require the use of more expensive and toxic drugs for treatment. Both the over prescription and underuse of antibiotics over prolonged periods leads to the development of resistant strains. The development of completely novel antibiotics is difficult despite urgent need, drug companies have little interest in developing novel drugs designed to be used for short periods of time. Nevertheless, companies have increased efforts to develop novel antibiotics with novel drug targets and chemical structures[16]. Even still at the pace at which resistance is growing within the next ten years some infections will have no available therapies, therefore there is a clear need to find new antimicrobial compounds, including novel natural products.

1.7 Scope and significance of the project

Whole genomic sequencing has identified a variety of silent gene clusters encoding for secondary metabolites of unknown biological activity within *Streptomyces*, i.e., metabolites that are not usually produced under standard laboratory conditions. These so called "silent gene clusters" could be a potential source for novel antibiotics. Of the 38 biosynthetic secondary metabolite gene clusters discovered in the genome of *S. avermitilis* around 14 of them produce unknown and uncharacterized gene products[3]. Several ways of activation and expression of silent gene clusters have been developed and experimented including co-culture, ribosomal engineering, chromatin remodeling, insertion of a strong synthetic promotors, and insertion of a gene cluster into another organism or strain for expression[16]. While some success has been realized, complimentary methods are needed to improve the success rate for silent gene cluster activation. It is possible to clone large biosynthetic gene clusters from *Streptomyces* into secondary organisms for heterologous expression using bacterial artificial chromosomes (BACs) or transformation associated recombination cloning (TAR)[25]. Both these methods have yielded some success but are extremely laborious. Another method involves the insertion of strong synthetic promoters to activate the gene clusters of interest. Recent advances have made use of CRISPR technology to more efficiently knock in promoters; however, comprehensive insertion of multiple promoters is still limited[25]. The goal of this research is to explore another method of activation of silent gene clusters, through making secondary metabolite synthesis tied to primary metabolite production and thus necessary for growth and development. Reinsertion of the thymidine gene (*thyX*) in different locations will allow for the creation of different suppressor mutants. The deletion and reintroduction of the *thyX* gene causes stress on the organism, and once removed from supplemental media the bacteria must develop suppressor mutations to grow. Mapping suppressor mutations will provide insight into how these silent secondary metabolite gene clusters, and the different biosynthetic pathways that are encoded by them, are regulated by *Streptomyces*. This information cannot be gained by inserting multiple synthetic promoters or other

methods of inducing expression. The insertion of promoters is less reliable in the expression of silent gene clusters because each cluster has different segments read in different directions with multiple promoter sites. This strategy is also advantageous because the same producing organism is used, and the codons are optimized. This is not the case if the gene were recombined into another bacteria such as *E. coli*

In some *Streptomyces* microorganisms carrying interesting gene clusters of secondary metabolites, systematic analysis of interesting metabolic pathways cannot be applied because the productivity of the metabolites is too low[4], a problem which could be solved through this technique of inducing production. In addition to learning more about the mechanisms that regulate silent gene clusters, which are currently not well understood, this research could lead to the practical development of compounds with the potential to serve as new pharmaceuticals and a source of novel antibiotics. New antibiotics are needed because of the growing threat of antibiotic resistance.

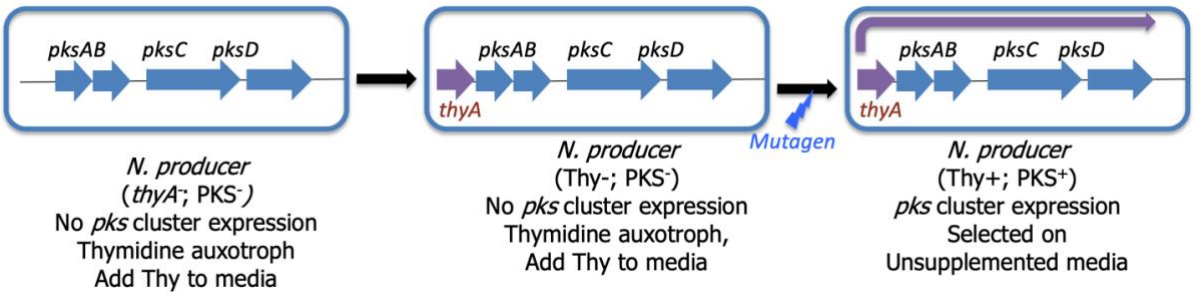

Figure 4. Experimental strategy for activation of silent gene cluster by reinsertion of the *thyX* gene.

Chapter 2

Materials and Methods

Materials

2.1 Media

Streptomyces was grown on ISP2 media containing yeast extract 4 g/L, malt extract 10 g/L, and dextrose 4 g/L. ISP5 media contained L-asparagine 1 g/L, glycerol 10 g/L, dipotassium hydrogen phosphate 1 g/L, 1mL of trace salt solution. Yeast was grown on YPG media containing 10 g/L yeast extract, 20 mL/L glycerol, and 20 g/L peptone. Secretion media consisted of yeast extract 4 g/L, malt extract 10 g/L, soluble starch 10 g/L, 2 g calcium carbonate, and agar 20 g/L. A blend of thymine and thymidine, each at a final concentration of 50 μg/mL (Thy50), was added to either liquid broth or agar plates to allow growth of thymine auxotrophs when necessary.

2.2 Strains

Streptomyces avermitilis strain ATCC 31267

Saccharomyces cerevisiae (yeast) strain W303a

SA7::ThySG ΔThyX – refers to *S. avermitilis* strains that have had their *thyX* gene reinserted into a region adjacent to the avermectin (SA7) gene cluster. This parent strain is used to generate SA7 suppressors

SA20::ThySG ΔThyX - refers to *S. avermitilis* strains that have had their *thyX* gene reinserted into a region adjacent to the oligomycin (SA20) gene cluster. This parent strain is used to generate SA20 suppressors

Methods

2.3 Genetic manipulation of *Streptomyces*

Deletion and reinsertion of thyX gene using CRISPR plasmids

Firstly, a plasmid containing a CRISPR Cas9 system was created to excise enough of the thymidine gene (*thyX*) from *S. avermitilis* to make it nonfunctional and create an auxotroph. The *thyX* gene sequence is run in the CRISPy website tool to narrow down search for CRISPR sites within the *thyX* gene, one flanking each end of the gene so that CRISPR cuts twice removing a large portion of the gene.

The CRISPR nit cassette with a sgRNA sequence protospacer homologous to the CRISPR cut sequence was created, and with a protospacer adjacent motif (PAM) site that will tell CAS9 where to bind. A repair template 10 bp from the chosen cut sites is also constructed, with 1 kb of homology upstream and downstream from the site so that the cells repair machinery can stitch the chromosome back together once CRISPR has cut the *thyX* gene out.

These two CRISPR systems along with a promotor, terminator, guide RNAs, transcriptional terminator, apramycin resistance gene, *E. coli* replication factor, and *Streptomyces* replication factor must all be fused onto a single plasmid. This is done through two rounds of cloning and two assembly steps because each CRISPR cassette must be PCRed individually. Once the plasmid is assembled, it is introduced into *Streptomyces* through *E. coli* conjugation.

In order to re-insert a functional *thyX* gene, another CRISPR nit plasmid must be constructed with the same procedure as above but with only one sgRNA antisense, because only a single cut is sufficient to introduce the gene in the gene cluster of interest. Therefore, only one round of cloning is required.

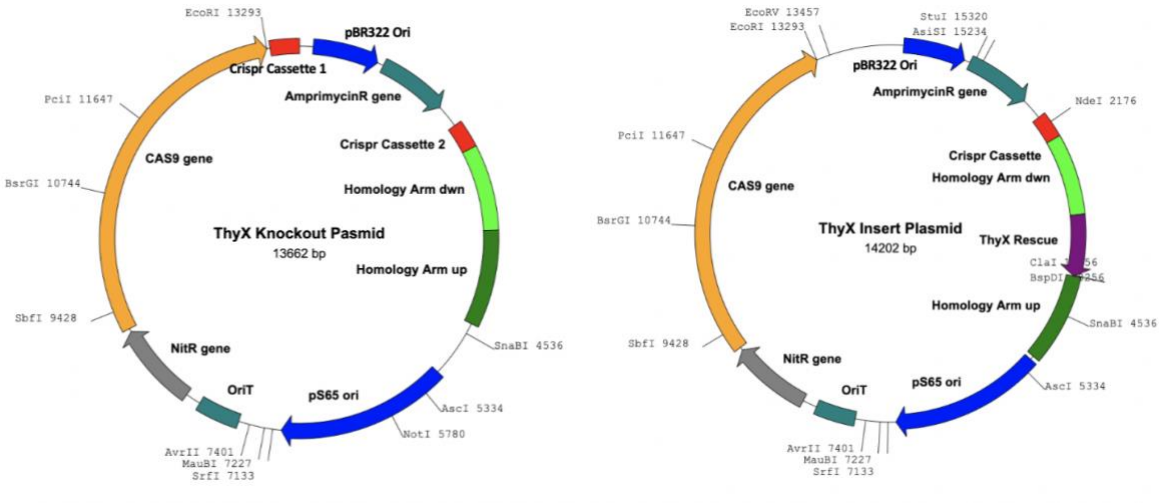

Figure 5. The CRISPR plasmids used to delete and then reinsert the *thyX* gene.

2.4 Generation of suppressor mutants

Spontaneous suppression

The parent strain SA20::ThySG ΔThyX was struck onto ISP2 plates containing thymidine and grown at 30ºC for three days. Then 6-8 colonies were ground up and put in 100 ml of liquid ISP2 media with thy50. These cultures were grown for 24 hours at 30ºC. The culture was then spun down at max speed and washed 2x with ISP2 without thy50. It was then resuspended in 25 ml PBS and used to inoculate 8 flasks 250 mL flasks by adding 500 µl of washed cells into 100 ml of ISP2 media (6 flasks without thy, one control with thy50, one control without *Streptomyces*).

Chemically induced suppression

To isolate suppressors using the chemical mutagen dP [6-(β-D-2-Deoxyribofuranosyl)-3,4-dihydro-8H-pyrimido-[4,5-c][1,2]oxazin-7-one (catalog number PY7270, Berry and Associates Biosearch Technologies) , 6-8 colonies of SA7::ThySG ΔThyX from the ISP2Thy50 plate were grounded up and grown in 20 ml of ISP2 Thy50 for 24-48 hours. They were washed 2 times with PBS and resuspended in few mLs of PBS. 100 µl aliquots were added to 1 ml of TBS Thy50 with either 0, 1, 10, or 100 µg/ml of dp mutagen. This was done in duplicates. Cultures were grown overnight at 30° C, washed 2 times with PBS and plated on large ISP2 plates lacking Thy50 supplement to isolate suppressors. Duplicate samples were plated on ISP2 Spec50Thy50 plates as control for mutagen efficiency by analyzing the rate of spectinomycin resistance. Candidate suppressor colonies able to grow without Thy50 were regrown in liquid ISP2 without Thy50, and stored as glycerol stocks in ISP2 with 20% glycerol final concentration at -80 ° C.

2.5 Isolating genomic DNA

300 µL of cell biomass was placed into a centrifuge tube with 300 µL of P1 buffer (10 mM TRIS-HCl, 10 mM EDTA, and 100 µg/mL RNAse A), and 600 µL of silica beads. The sample was bead beated at 3,800 rpm for 20-30 s. 100 µL of the supernatant was treated with 10 µL of lysozyme (4 ng per µL) and incubated for 30 minutes at 37° C to complete cell disruption. The GES method was then adapted for genomic DNA isolation[23]. 500 µL of 5x GES was added, mixed, then 250 µL of chilled ammonium acetate (7.5M) was added and the sample was vortexed then chilled on ice for 10 minutes. The supernatant was removed and separated into two tubes of equal volume, cold isopropanol was added at a 1:2 ratio, supernatant to isopropanol. Tubes were inverted then centrifuged at max speed for 20 s.

Isopropanol was decatenated. The DNA pellets were rinsed twice with 70% ethanol 900 µL, re-pelleted by spinning max speed 5 minutes. The ethanol was dried off and the DNA resuspended in 50 µL of elution buffer. DNA concentration was determined via nanodrop. DNA was sent for next gen Illumina sequencing at the MiGS Microbial Genome Sequencing Center (Pittsburgh, PA).

2.6 Avermectin TLC assay

SA7 suppressors strains were grown in 24 well plates from stock with 2 mL ISP2 media and sterile glass beads, shaking at 28° C. 1 mL of secretion media was pipetted into wells of 98 well plate to make agar plugs using a multichannel pipette. 10 µL of liquid culture from 24 well plates was pipetted onto the agar plugs and the plate was left sealed with parafilm and suppressors left to grow for 10 days at 28° C. Agar plugs were extracted using chloroform and left to shake for 4 hours at 30° C. Afterwards, chloroform was left to dry overnight in the chemical fume hood and avermectin was resuspended in methanol for TLC spotting. 2 µL of Avermectin extract were spotted onto TLC normal phase silica plates and placed into TLC chamber containing mobile phase (70 mL hexanes, mL isopropyl, 2.5 mL methanol). Avermectin was visualized under UV light.

2.7 Oligomycin yeast screen assay

SA20 suppressor strains are spotted on ISP5 agar media plate: first colonies were picked from stock plates, ground in 300 µL of ISP5 media and then 5 µL was spotted. Wild type *S. avermitilis* was also spotted as well as 2 different concentrations of oligomycin (3 µL of 1 mg/mL and of 0.1 mg/mL). The plate was left to incubate for 3 days at 30° C. The *Saccharomyces cerevisiae* yeast strain W303a was grown in YPG at 30° C over night until OD of 0.6. A soft agar YPG overlay (0.7% agar) was warmed at

50° C, then slightly cooled and yeast was added at 1:100 dilution. A soft overlay containing *Saccharomyces cerevisiae* YPG was poured on top of the ISP5 plate containing the SA20 suppressors. The plate was incubated at 30°C for 24 hours and then analyzed for zones of clearing.

2.8 rtPCR

PowerUp SYBR Green Master Mix (Applied Biosystems) was used for the real time PCR, with gene-qPCR specific primer sets (GXK2358 SGThyX-FOR, GXK2359 SGThyX-REV, GXK2360 SAV_2444-FOR, GXK2361 SAV_2444-REV) and measured using an Applied Biosystems 7300 Real-Time PCR machine by running standard cycling mode with annealing temperature of 55° C. All data were measured in triplicate and relative expression levels were calculated using the $2\text{-}\Delta\Delta^{CT}$ method.

Chapter 3

Results

Evaluation of Mutagenesis protocols: Spontaneous verse dP induced

The chemical compound dP is a deoxycytidine base analog, whose incorporation into DNA is known to induce both GC-->AT and AT-->GC transitions in model organisms such as *Escherichia coli* (Nucleic Acids Res. 1997 Apr 15; 25(8): 1548–1552). As such, its broad mutation spectrum makes it a potentially ideal mutagen in many organisms, including GC rich organisms such as Streptomyces sp. To be active, however, it must first be activated by tri phosphorylation so that DNA polymerase can utilize it as a substrate. While thymidine kinase from *E. coli* can activate and then utilize dP, it was unknown whether *S. avermitilis* would be able to (mis)incorporate the base analog. We tested three concentrations of dP (1, 10, or 100 µg/ml), along with no dP control, to measure the spontaneous mutation rate. Mutagenesis efficiency was evaluated by screening for resistance to the antibiotic spectinomycin, which targets the ribosome[24]. Inclusion of dP raised the mutation rate from 10 to 1000-fold, with cell death rate increasing from less than 0.1% to up to 1% with the highest concentration of dP used. This indicated dP is an efficient mutagen in Streptomyces. For isolation of *thyX* auxotrophy suppressors, both 10 and 100 µg/ml dP treatments were used to isolate suppressors.

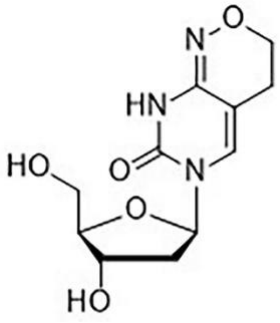

Figure 6. Chemical structure of dP. [6-(β-D-2-Deoxyribofuranosyl)-3,4-dihydro-8H-pyrimido-[4,5-c][1,2]oxazin-7-one

Screening of SA20 suppressors.

In order to visualize secretion of oligomycin by different SA20 suppressors and test their biological activity, a yeast layer two-layer agar overlay assay was used. *S. cerevisiae* is a eukaryotic organism that possesses mitochondria and performs oxidative phosphorylation. The production of oligomycin by *S. avermitlis* kills off the yeast by stopping electron transport, and zones of clearing are visible around the spotted bacteria where oligomycin is secreted. The larger the zone of clearing the more oligomycin produced, and assumedly the more active the gene cluster for oligomycin. The ISP5 media contains glycerol as a carbon source which cannot be fermented by the yeast, therefore they must perform oxidative phosphorylation in order to produce energy. The size of the zone of clearing was measured and used to identify suppressors that were producing a large quantity of oligomycin compared to the wt control strain. Two separate experiments were conducted to screen SA20 suppressors.

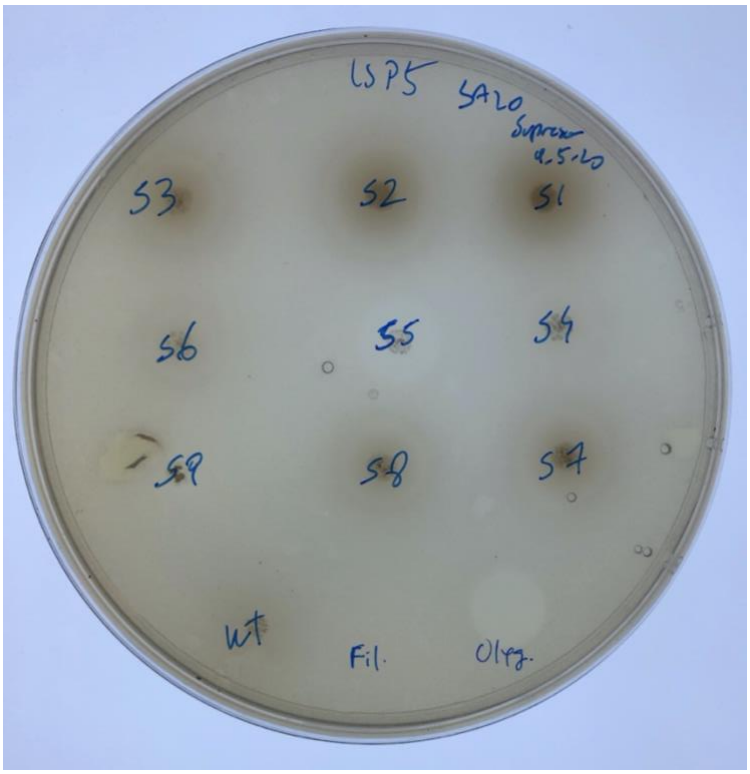

Figure 7. Batch 1: SA20 Suppressors spotted on ISP5 agar plate to test oligomycin bioactivity and secretion. 9 suppressors are spotted along with a wt strain and oligomycin.

Suppressor Number	Wild type control	S1	S2	S3	S4	S5	S6	S7	S8	S9
Zone of clearing (mm)	0	10	22	17	14	15	13	18	17	0

Table 1. Measured zones of clearing batch 1

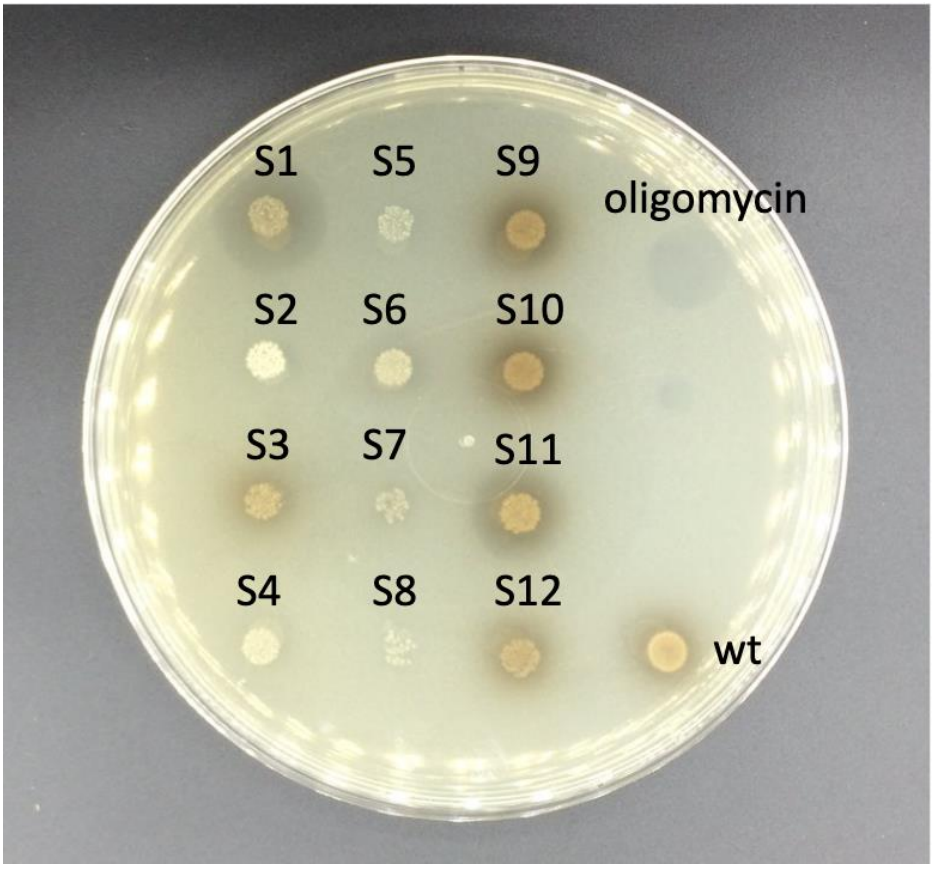

Figure 8. Batch 2:12 different SA20 suppressors along with a wild type *S. avermitilis* strain. Note the suppressors vary in their ability to grow properly, secrete oligomycin and secrete pigment.

Suppressor Number	Wild type control	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12
Zone of clearing (mm)	0	20	8	11	10	0	16	0	0	15	15	18	13

Table 2. Measured zones of clearing batch 2

Chapter 4

Sequencing Data

Chromosomal DNA was isolated from the representative bioactive suppressor strains and prepped for whole genome Illumina sequencing. Genetic analysis of suppressor strains was conducted to identify interesting mutations that lead to upregulation of biosynthetic gene cluster. Five SA20 suppressors were chosen to sequence based on oligomycin bioactivity screening. The sequencing data was analyzed using Geneious software and assembled using the reference *S. avermitilis* genome[7].

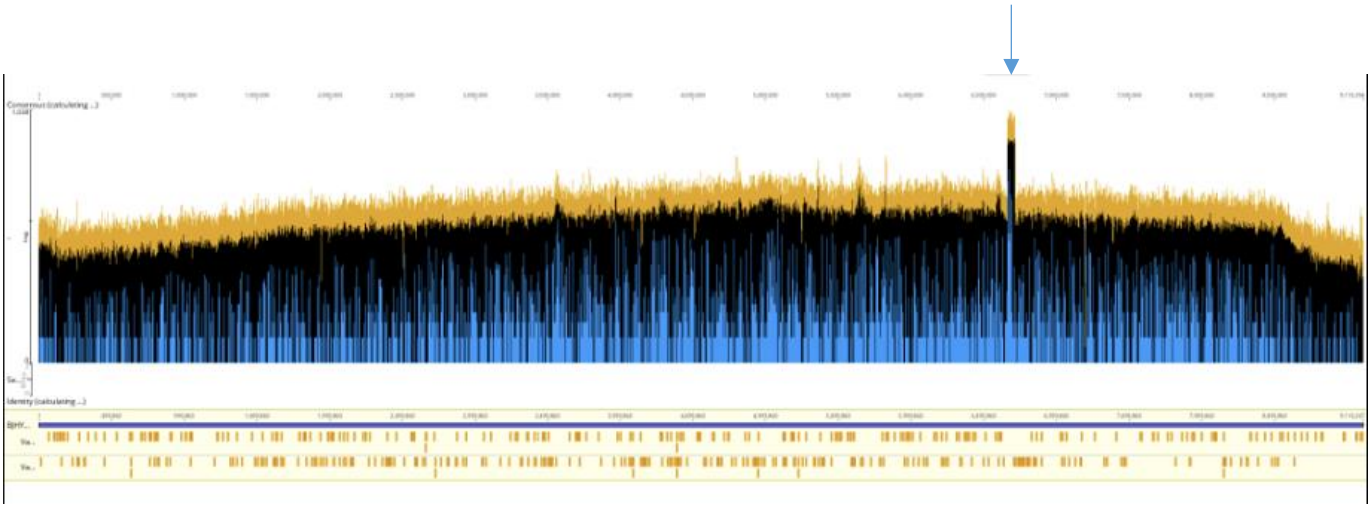

Figure 9

Genome map of suppressor 2. X-axis is genome position (linear genome with telomeres), and Y-axis is sequencing coverage. Blue arrow indicates the site of *thyX* gene insertion within the oligomycin biosynthetic gene cluster.

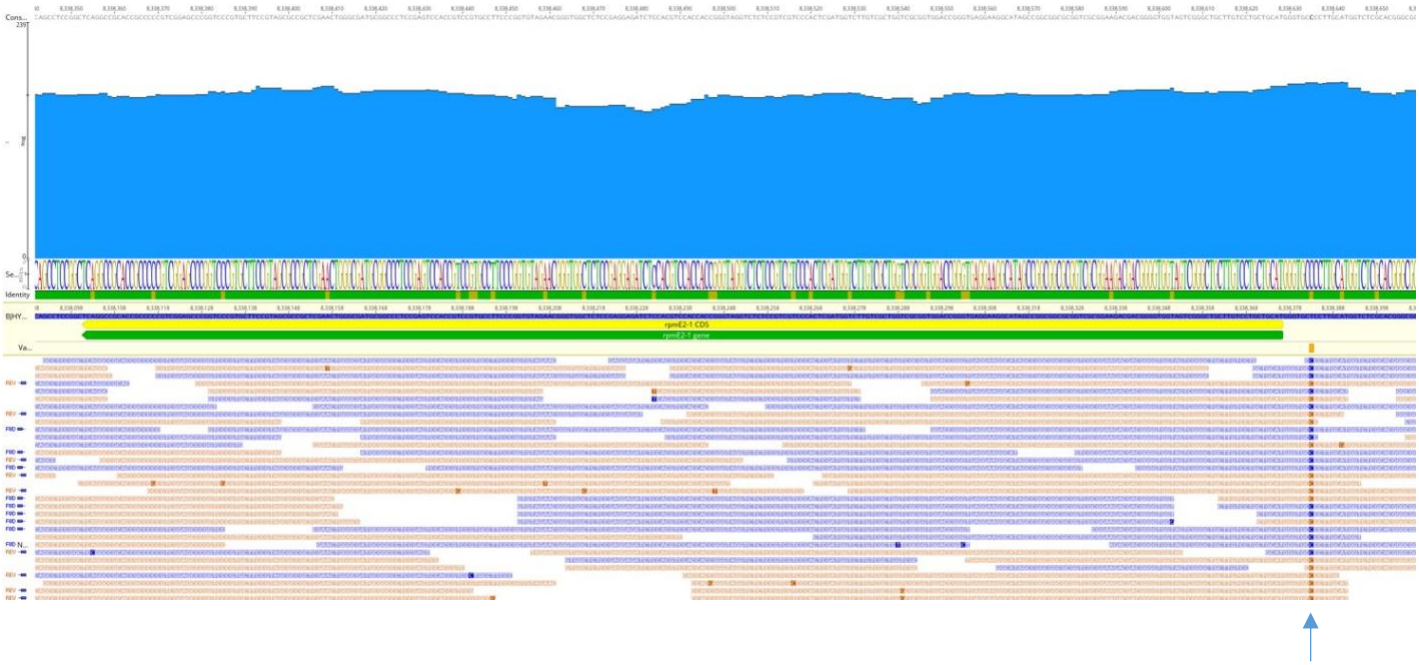

Figure 10. A snapshot of the suppressor 2 genome map at the sight of the *rpme1* gene locus. Arrow indicates the site of a single nucleotide polymorphism.

rtPCR to amplify the sight of the *thyX* insertion was performed in order to confirm the duplication seen in suppressor 2 genomic sequencing data (Figure 7), and also to screen other suppressors for local duplication of the *thyX* gene insert. Gene duplication is a common method for auxotrophy suppression in bacteria. Additionally, *Streptomyces* often duplicate genes naturally even when not under genetic or selection pressures.

	thyX gene				SAV_2444 gene
Suppressor Number	Average count	delta Ct	deltadeltaCt	avg fold change	Average ct
1	16.14	-3.28	-2.78	6.89	19.43
6	17.23	-0.56	-0.06	1.04	17.80
10	17.94	-0.45	0.05	0.96	18.39
11	17.74	-0.12	0.38	0.76	17.86
19	18.40	-0.11	0.38	0.76	18.51
21	18.35	-0.48	0.02	0.98	18.83
23	17.67	-0.73	-0.23	1.17	18.41
5	15.84	-3.28	-2.78	6.87	19.12
SA20::ThyXSGΔThyX (parent strain)	17.63	-0.50	0	1	18.13

Table 3. rtPCR data. Chromosomal DNA was also used as template for rtPCR of the *thyX* gene. Note SAV_244 is used as a control gene for the rtPCR data. Suppressors with an average fold change in gene copy significantly greater than 1 are highlighted.

Chapter 5

Avermectin TLC

In order to test the production of avermectin by SA7 generated suppressors, Thy suppressors were grown in 96-well plate agar plugs containing optimal growth media and then avermectin was extracted using chloroform and spotted on a silica plate for TLC analysis. The avermectin molecule is UV active and therefore can be detected by TLC. The relative intensity of the bands gives an indication of the amount of avermectin being produced by each suppressor. The presence of several bands in each lane (compared to the single band in the control lane) can be attributed to the production of several isoforms of avermectin by the bacteria. The control avermectin is only one form of the compound (Figure 2).

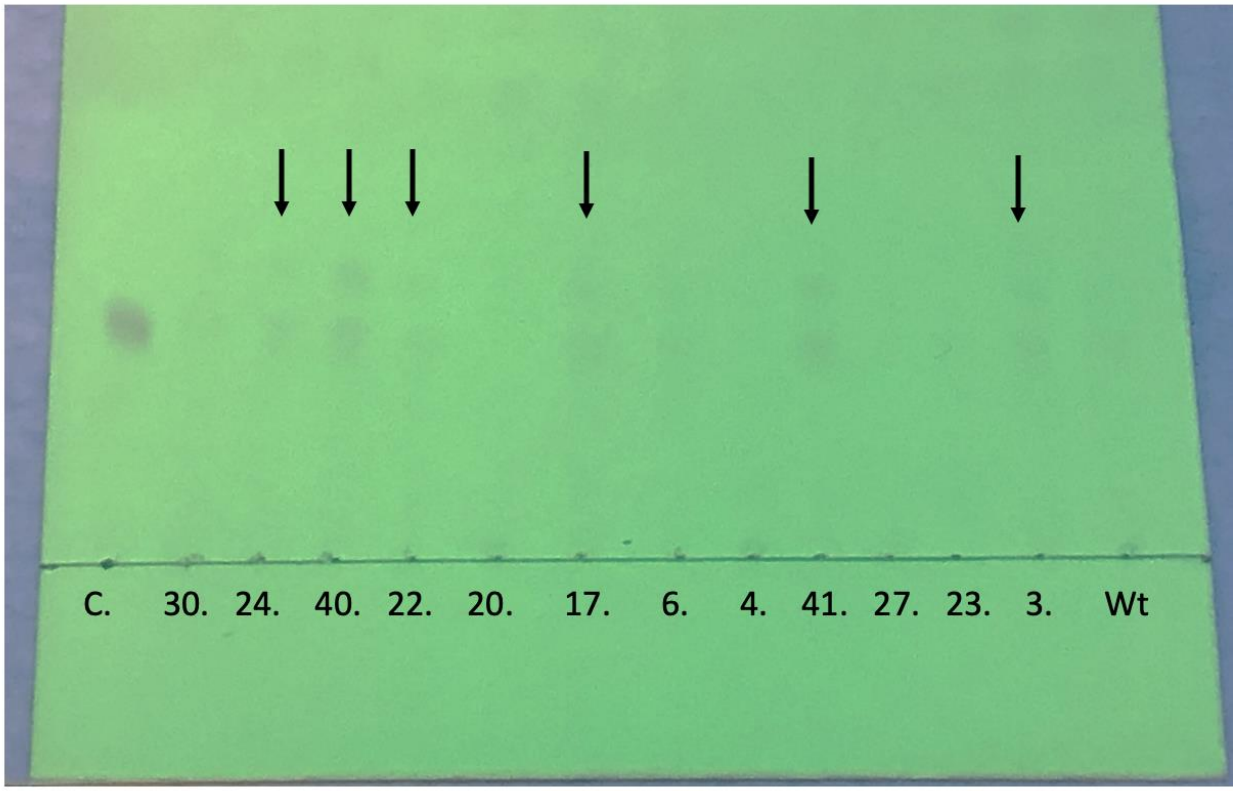

Figure 11. Avermectin extracts from 12 suppressor SA7 strains spotted on a TLC silica plate , along with wild type extract and one 2 µL of mg/mL avermectin. Suppressors with visible avermectin production are highlighted with the black arrows.

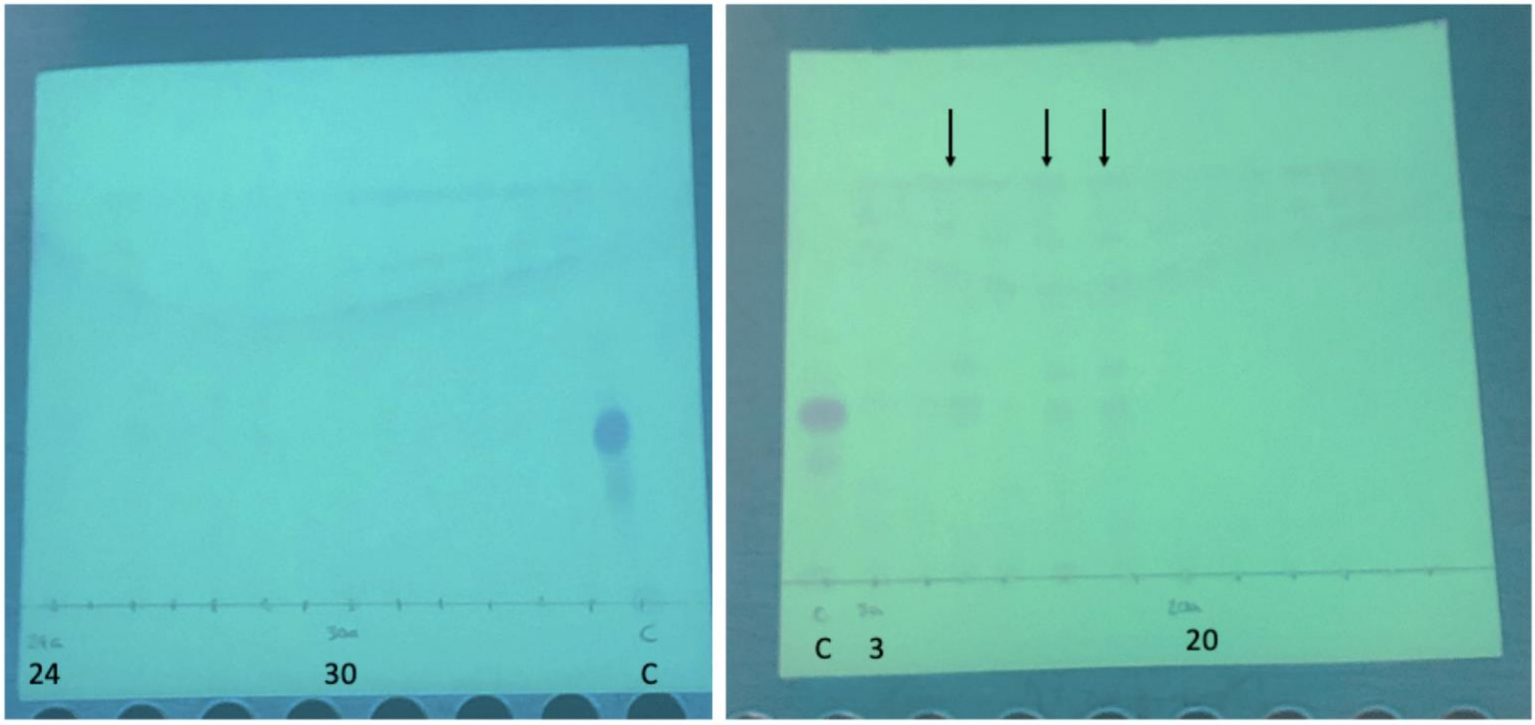

Figure 12. TLC extracts run in duplicates of six, SA7 suppressors were tested. High levels of avermectin are indicated by the black arrows.

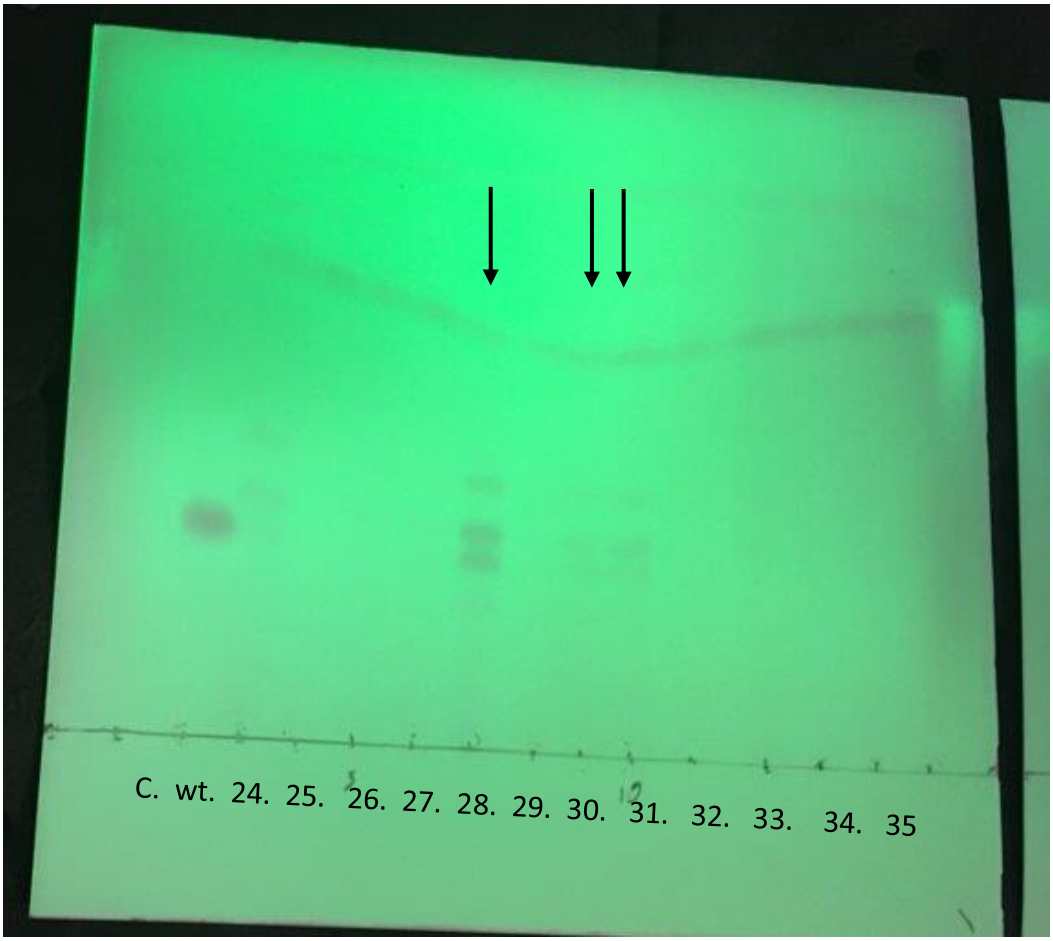

Figure 13.

TLC of 11 different SA7 suppressors, along with the wildtype strain and 2 µL of avermectin in the control lane.

Chapter 6

Discussion and Future Directions

The sequencing data of the SA20 suppressor mutant 2 shows a spike in read numbers at the site of the *thyX* gene insertion (Figure 7). This is confirmed by the rtPCR data and appears to be the results of local chromosome duplication. Suppressor 2 duplicated the *thyX* gene multiple times at that insertion site. This is likely due to overcoming selection and genetic pressures. Based on the enrichment of read counts corresponding to this region of the chromosome, it is estimated that from 12-18 copies of an ~18 kb sequence containing the *thyX* gene were present. Multiple copies of the *thyX* gene allows for greater expression of the gene if it is being weakly expressed due to its change in location within the organism's chromosome. Of the suppressor analyzed, only one other had duplicated *thyX* genes confirmed through the rtPCR data suppressor (Table 2). This is an interesting but not completely unexpected results, as *Streptomyces* have very plastic and adaptable genomes. It has been shown that the genomes of various *Streptomyces* species can undergo rapid expansion and contraction. Comparative genetic analysis of 5 different *Streptomyces* species found that specific genes are expanded in a lineage specific manner[19]. These duplicated genes can account for 4%-11% of the genome based on the species studied, with *S. avermitilis* having the smallest percent of its genome comprised of LSEs at only 4%.

LSE occurs in two distinct ways within an organism's chromosome, either through contiguous expansion or through the duplication of gene blocks. Block duplicated genes are most often found at the terminal ends of the linear chromosome of *Streptomyces*. Contiguous expansion LSE occur when a single gene is duplicated multiple times consecutively in the genome. The duplication of the *thyX* gene more closely resembles contiguous expansion. LSEs are characteristic of the *Streptomyces* genome and

contribute to species specificity and evolution within the genus. These LSE in the bacteria are the product of gene duplication and horizontal gene transfer.

Duplication of the *ThX* gene was corroborated through real time PCR (rtPCR). rtPCR allows for the quantification of a starting DNA template through a fluorescent detection of the product. rtPCR allows the determination of the initial absolute copy number of a specific gene compared to a standard or control. rtPCR showed that suppressor 1 has an average x6 fold more starting DNA of the *thyX* gene than the SA20::ThyXSGΔThyX parent strain it was generated from (Table 2). Performing rtPCR also revealed another suppressor, S1 also had duplications of the *thyX* gene insert, with around x6 fold difference in copy number compared to the parent as well. All the other suppressors had about the same average fold change compared to the parent strain with integers around 1. rtPCR is a useful and fast technique to eliminate suppressors that have duplicated *thyX* genes from downstream whole sequencing analysis. The goal of generating these suppressors is to identify one with a specific mutation that upregulates a biosynthetic gene cluster, the mutants with duplicated *thyX* genes, albeit interesting, are not helpful in accomplishing the said goal. rtPCR analysis of SA7 suppressors found no suppressors that had duplicated the *thyX* gene insert (data not shown).

One non-merodipliod SA20 suppressor sequence was found to have a mutation at the *rpmE1* gene within its chromosome (Figure 8). The *rpmE1* gene encodes the 50 S ribosomal protein L31[20], and the mutation is a single nucleotide polymorphism. *Streptomyces* are unique as they have 3 copies of this gene scattered across the genome, while other families of bacteria have a single copy. This mutation is of note because it could lead to upregulation of secondary metabolite gene cluster products, which has been observed in other *Streptomyces* species with ribosomal gene mutations. Researchers discovered a novel insertion mutation, named GI92, in the gene encoding the S12 ribosomal protein in *Streptomyces coelicolor* had a two-fold effect on the organism's physiology[21]. Firstly, the bacteria displayed a 20-fold increase in paromycin resistance. Secondly, and more importantly, when combined with another S12 mutation called K88E, the two mutations greatly enhanced the production of the polyketide secondary

metabolite antibiotics actinorhodin and undecylprodigiosin. Through *in vitro* translation experiments, they discovered that the double mutation greatly enhanced translation efficiency by six-fold and decreased errors significantly. In further experiments, suppressor mutants with ribosomal mutations might hold the key for activation and over production of silent secondary metabolites in *S. avermitilis*.

The avermectin TLC data clearly shows the production of avermectin by some suppressor strains generated more than others. SA7 suppressors 24, 40, 22, 17, 41, and 3 were producing the most compound of the 12 strains tested, and produced more than the wild-type strain by comparison (Figure 9). This method is an easy way to screen suppressors but has some limitations. The TLC is not very quantitative and it was a struggle to obtain reproducibility. Suppressors that showed notable amounts of production of avermectin in one TLC, when screened again in another subsequent TLC, had little to no visual production of avermectin. Suppressor 24 showed high levels of avermectin production in the results shown in Figure 9 but very little to no production in those shown in Figure 10. Conversely, suppressor 30 showed high production in Figure 10 but no production in the TLC in Figure 11. In the future, more TLC assays must be done to narrow down the search for best 4-5 suppressors to continue for downstream analysis. It is difficult to account for the variation because suppressors are grown in identical conditions between batches; further testing is needed to improve accuracy. Once the suppressors with the most avermectin production are identified, they will be chosen to sequence and identify novel mutations that generate more compound. Before whole genome sequencing, the *thyX* gene cluster will first be amplified using rtPCR to determine copy number and screen out SA7 suppressor with duplicated *thyX*, just as was done with SA20 suppressors. Additionally, HPLC must be performed on avermectin extracts as a more quantitative method of analyzing the products generated by the suppressors, and at the same time confirming the chemical identity of the extracts. In the future, a bioassay will be developed to see how antiparasitic and effective the suppressor secreted avermectin is; this assay must include a worm target such as *C. elegans* which are susceptible to avermectin.

BIBLIOGRAPHY

1. Chater K. F. (2016). Recent advances in understanding *Streptomyces*. *F1000Research*, *5*, 2795. https://doi.org/10.12688/f1000research.9534.1
2. Nett, M., Ikeda, H., & Moore, B. S. (2009). Genomic basis for natural product biosynthetic diversity in the actinomycetes. *Natural product reports*, *26*(11), 1362–1384. https://doi.org/10.1039/b817069j
3. Procópio, R. E., Silva, I. R., Martins, M. K., Azevedo, J. L., & Araújo, J. M. (2012). Antibiotics produced by Streptomyces. *The Brazilian journal of infectious diseases : an official publication of the Brazilian Society of Infectious Diseases*, *16*(5), 466–471. https://doi.org/10.1016/j.bjid.2012.08.014
4. Ikeda, H., Kazuo, S. Y., & Omura, S. (2014). Genome mining of the Streptomyces avermitilis genome and development of genome-minimized hosts for heterologous expression of biosynthetic gene clusters. *Journal of industrial microbiology & biotechnology*, *41*(2), 233–250. https://doi.org/10.1007/s10295-013-1327-x
5. Masaya, N., Takashi, I., Shinchi, A., Youichi, H., Hiroshi, M., Rika, S., Kuniaki, T. (1995). Synthetic studies on oligomycins. Synthesis of Oligomycin B spiroketal and polypropionate portions. *Bulletin of the chemical society of Japan*, 68:3, 967-989
6. Kieser, T., Bibb, M. J., Buttner, M. J., Chater, K. F.,, Hopwood, D. A. (2000). *Practical streptomyces genetics*. John Innes Foundation. ISBN: 0708406238
7. Ikeda, H., Ishikawa, J., Hanamoto, A., Shinose, M., Kikuchi, H., Shiba, T., Sakaki, Y., Hattori, M., & Omura, S. (2003). Complete genome sequence and comparative analysis of the industrial microorganism Streptomyces avermitilis. *Nature biotechnology*, *21*(5), 526–531. https://doi.org/10.1038/nbt820
8. Das, A., & Khosla, C. (2009). Biosynthesis of aromatic polyketides in bacteria. *Accounts of chemical research*, *42*(5), 631–639. https://doi.org/10.1021/ar8002249
9. Dutta S, Whicher JR, Hansen DA, et al. Structure of a modular polyketide synthase. *Nature*. 2014;510(7506):512-517. doi:10.1038/nature13423
10. Wang, J., Zhang, R., Chen, X. *et al.* Biosynthesis of aromatic polyketides in microorganisms using type II polyketide synthases. *Microb Cell Fact* **19,** 110 (2020). https://doi.org/10.1186/s12934-020-01367-4
11. Nivina, A., Yuet, K. P., Hsu, J., & Khosla, C. (2019). Evolution and Diversity of Assembly-Line Polyketide Synthases. *Chemical reviews*, *119*(24), 12524–12547. https://doi.org/10.1021/acs.chemrev.9b00525
12. Olano, C., Méndez, C., & Salas, J. A. (2009). Antitumor compounds from actinomycetes: from gene clusters to new derivatives by combinatorial biosynthesis. *Natural product reports*, *26*(5), 628–660. https://doi.org/10.1039/b822528a
13. Yu, G., Jia, X., Wen, J., Lu, W., Wang, G., Caiyin, Q., & Chen, Y. (2011). Strain improvement of Streptomyces roseosporus for daptomycin production by rational screening of He-Ne laser and NTG induced mutants and kinetic modeling. *Applied biochemistry and biotechnology*, *163*(6), 729–743. https://doi.org/10.1007/s12010-010-9078-x
14. El-Saber Batiha, G., Alqahtani, A., Ilesanmi, O. B., Saati, A. A., El-Mleeh, A., Hetta, H. F., & Magdy Beshbishy, A. (2020). Avermectin Derivatives, Pharmacokinetics, Therapeutic and Toxic Dosages, Mechanism of Action, and Their Biological Effects. *Pharmaceuticals (Basel, Switzerland)*, *13*(8), 196. https://doi.org/10.3390/ph13080196
15. Ratajczak, K., Lukasiak, A., Grel, H., Dworakowska, B., Jakiela, S., & Stobiecka, M. (2019). Monitoring of dynamic ATP level changes by oligomycin-modulated ATP synthase inhibition in

SW480 cancer cells using fluorescent "On-Off" switching DNA aptamer. *Analytical and bioanalytical chemistry*, *411*(26), 6899–6911. https://doi.org/10.1007/s00216-019-02061-0

16. CDC antibiotic resistance, December 2020
17. Zaman, S. B., Hussain, M. A., Nye, R., Mehta, V., Mamun, K. T., & Hossain, N. (2017). A Review on Antibiotic Resistance: Alarm Bells are Ringing. *Cureus*, *9*(6), e1403. https://doi.org/10.7759/cureus.1403
18. Peterson, E., & Kaur, P. (2018). Antibiotic Resistance Mechanisms in Bacteria: Relationships Between Resistance Determinants of Antibiotic Producers, Environmental Bacteria, and Clinical Pathogens. *Frontiers in microbiology*, *9*, 2928. https://doi.org/10.3389/fmicb.2018.02928
19. Okada, B. K., & Seyedsayamdost, M. R. (2017). Antibiotic dialogues: induction of silent biosynthetic gene clusters by exogenous small molecules. *FEMS microbiology reviews*, *41*(1), 19–33. https://doi.org/10.1093/femsre/fuw035
20. Zhou, Z., Gu, J., Li, Y. Q., & Wang, Y. (2012). Genome plasticity and systems evolution in Streptomyces. *BMC bioinformatics*, *13 Suppl 10*(Suppl 10), S8. https://doi.org/10.1186/1471-2105-13-S10-S8
21. Owen, G. A., Pascoe, B., Kallifidas, D., & Paget, M. S. (2007). Zinc-responsive regulation of alternative ribosomal protein genes in Streptomyces coelicolor involves zur and sigmaR. Journal of bacteriology, 189(11), 4078–4086. https://doi.org/10.1128/JB.01901-06
22. Wang, G., Inaoka, T., Okamoto, S., & Ochi, K. (2009). A novel insertion mutation in Streptomyces coelicolor ribosomal S12 protein results in paromomycin resistance and antibiotic overproduction. *Antimicrobial agents and chemotherapy*, *53*(3), 1019–1026. https://doi.org/10.1128/AAC.00388-08
23. Picher, D.G., N.A. Saunders, and R.J. Owen. (1989). Rapid Extraction of Bacterial Genomic DNA with Guanidium Thiocyanate. Lett Appl Microbiol. 8:151-156
24. Demirci, H., Murphy, F., 4th, Murphy, E., Gregory, S. T., Dahlberg, A. E., & Jogl, G. (2013). A structural basis for streptomycin-induced misreading of the genetic code. *Nature communications*, *4*, 1355. https://doi.org/10.1038/ncomms2346
25. Liu, Z., Zhao, Y., Huang, C., & Luo, Y. (2021). Recent Advances in Silent Gene Cluster Activation in *Streptomyces*. *Frontiers in bioengineering and biotechnology*, *9*, 632230. https://doi.org/10.3389/fbioe.2021.632230

João G. Ramos Ferreira

ACADEMIC VITA

EDUCATION

The Pennsylvania State University, *University Park, PA* — Class of 2022
The Eberly College of Science | BS Biochemistry & Molecular Biology
Schreyer Honors College | Millennium Scholars Program
Western High School, *Davie, FL*

EXPERIENCE

Undergraduate Researcher — **University Park, PA**
The Pennsylvania State University, Dr. Timothy Meredith — *January 2019 – Present*
- Help design a plasmid using CRISPR Nit system to knock out thymidine gene using ApE software
- Constructed a plasmid to transform Streptomyces using Takara infusion kit
- Culture *Streptomyces* auxotroph and isolated DNA for genome sequencing

Summer Research Intern – GWU Summer Program for Advancing Cancer Research — **Washington, DC**
The George Washington University, Dr. Conrad Russell Cruz — *May 2019 – August 2019*
- Isolated Natural Killer cells from peripheral blood mononuclear cells using immunomagnetic negative selection beads
- Facilitated transduction of lymphocytes using retroviral vector
- Experienced in cancer cell line and immune cell culture technique
- Used flow cytometry to quantify levels of cell death using Raji cells and Natural killer cells grown in coculture with CD47 monoclonal antibody

RELEVANT COURSES

Fall 2019
- BMB 251H: Molecular / Cellular Biology I
- PPEM 424: Virus Ecology
- CHEM 210: Organic Chemistry I

Spring 2020
- MICRB 410: Immunology
- BMB 252: Molecular / Cellular Biology II
- CHEM 212: Organic Chemistry II

LEADERSHIP & INVOLVEMENT

Science LionPride — **University Park, PA**
2018- present
- Act as a representative for the Eberly College of Science, give tours to prospective students

Undergraduate tutor — **University Park, PA**
2019 - Present
- Tutor three undergraduate students in introductory microbiology

Volunteer, Cleveland Clinic Florida, — **Weston, FL**
2016-2017
- Filed and organized histology slides and worked at the information desk greeting and directing visitors/patients

ADDITIONAL SKILLS

Laboratory skills
- Cell culture (cancer cell line, and immune cell)
- Immunofluorescent staining and flow cytometry analysis
- Agarose gel electrophoresis
- PCR

Bilingual
- Fluent in Portuguese, proficient in Spanish

João G. Ramos Ferreira

ACADEMIC HONORS

Millennium Scholars Program Scholarship *2018 – present*
- Schreyer Family Honors Scholarship *2019 – present*
- Hispanic Scholarship Fund Scholar *2019 – present*
- Dean's List *2018. 2019. 2020. 2021*