Spring 4-19-2018

Effect of Mutation of cvn7 conservon genes Sco6796 and Sco6798 of Streptomyces coelicolor.

Caroline Fulmer
caroline.fulmer@otterbein.edu

Effect of Mutation of *cvn*7 conservon genes Sco6796 and Sco6798 of *Streptomyces coelicolor*.

Caroline N. Fulmer
Department of Biology and Earth Science
Otterbein University
Westerville, Ohio 43081

May 2018

Submitted in partial fulfillment of the requirements for
graduation with distinction

Advisory Committee:

Jennifer Bennett, Ph.D. Distinction Advisor	_____ Advisor's Signature
Simon Lawrance, Ph.D. Second Reader (or Co-advisor if applicable)	_____ Second Reader's Signature
Jeffrey Lehman, Ph.D. Distinction Representative	_____ Distinction Rep's Signature

Abstract

Streptomyces coelicolor is a gram-positive Actinobacterium. It is a filamentous soil organism that performs complex multicellular development. Its physiological and morphological differentiation produces active secondary metabolite antibiotics that have pharmaceutical applications. Studies of *S. coelicolor* show that cyclic di-GMP proteins regulate physiological and morphological development. *S. coelicolor* has 13 copies of an operon that are collectively named "conservon" because they are highly conserved. The operons are named *cvn1-13* and genes in each operon are denoted *cvnA-D*. RNA-Seq data of a cyclic di-GMP phosphodiesterase double mutant showed differential expression between wild type and many of the operons in the double mutant. Previous studies from other laboratories provide evidence that deleting the entire *cvn7* operon may not exhibit a phenotypic change, but overexpressing one gene of the operon could. The objective of this study is to determine the function of *S. coelicolor* operon *cvn7*, specifically genes Sco6796 (*cvnC7*) and Sco6798 by analyzing overexpression mutants. Bioinformatics analyses of *cvn7* have been performed including multiple sequence alignments, BLAST and protein domain maps. Concurrently, a transposon mutant of *cvnD* in another conservon operon of *S. coelicolor* has been used to analyze the function of the conservon. The analysis of the *cvnD* transposon mutant presented a phenotype with a difference in pigment production, and a spore formation defect. Distinct mutant phenotypes of each of these genes could shed light on the function of each gene, the complete conservon, and their possible role in development of *S. coelicolor*.

1. Introduction

Streptomyces coelicolor is a gram-positive Actinobacteria. It is a filamentous soil organism that performs complex multicellular development of both physiological and morphological differentiation and produces active secondary metabolites. These metabolite products include antibiotics which can have important applications in pharmaceutical industries. Studies of *S. coelicolor* have shown that genes encoding cyclic-di-GMP proteins are under control of *bldD* which provides evidence for their potential involvement in physiological and morphological development (McCormick and Flärdh 2012) Due to the useful characteristics of *S. coelicolor*, its whole genome has been sequenced. The sequencing of its genome showed that there are 13 copies of an operon, collectively named the "conservon" because the operons are highly conserved. The 13 operons are named *cvn1-13* and genes in each operon are denoted *cvnA-D* from 5' to 3'.

Previous studies of *S. coelicolor* have shown varied effects of deleting these operons of the conservon. Takano et al., (2011) made null mutants lacking all 4 or 5 coding sequences of the operon for each of the 13 copies of *cvn* in *S. coelicolor* which resulted in wild-type phenotypes for all *cvn* except *cvn1* which presented an abnormal phenotype. *Cvn1* null mutant lacking all 5 coding sequences produced a pleiotropic phenotype only under high glucose concentrations. High glucose concentrations increase formation of aerial mycelium and production of actinorhodin and undecylprodigiosin pigment antibiotics in cvn1 mutant while these conditions decreased aerial mycelium and pigment production in the wild type. Null mutants for each coding sequence of *cvn1* showed the same phenotype as *cvn1* mutant while the other coding

sequences (*cvnB1, cvnC1* and *cvnD1*) presented wild-type phenotypes (Takano et al., 2011). Komatsu et al., (2006) made complete deletions of *cvn9* and *cvn10* presented wild-type phenotypes while individual deletions of *cvnA9, cvnD9*, and *cvnA10* exhibited abnormal phenotypes conditionally over producing the antibiotic actinorhodin (blue-colored pigment) and aerial mycelium and spore formation under high glucose conditions (Komatsu et al., 2003). These studies provide evidence that while mutating all of *cvn7* may not exhibit a phenotypic change, deletion or amplification of just *cvnC7* or Sco6798 could exhibit a mutant phenotype.

The main objective of my study is to determine the function of *S. coelicolor* operon *cvn7*, specifically genes Sco6796 and Sco6798. This operon is of interest because RNA-Seq data of *S. coelicolor* showed differential gene expression between the wild type and a double mutant strain containing transposon insertions that inactivated both the *rmdA* and *rmdB* genes. *RmdA* and *RmdB* are cyclic di-GMP phosphodiesterases; and the double mutant strain exhibited differential gene expression for *cvn7* and many of the other operons of the conservon. *Cvn2, cvn3, cvn7, cvn9* and *cvn11* all showed significant under expression while *cvn1, cvn4, cvn8, cvn10* and *cvn13* showed significant overexpression. *Cvn7* showed significant expression of - 4 to -5 Log2 fold change between the wild type and the double mutant strain that is defective for cyclic di-GMP signaling (Bennett, unpublished). Gene Sco6796 is of particular interest because it is gene C in the operon *cvn7* and according to strepdb (http://strepdb.Streptomyces.org.uk) contains a conserved domain of unknown function, or alternatively, the bacterial regulatory protein arsR family. Members of this family contain a DNA-binding helix-turn-helix motif, which implies that the protein could be involved in gene expression. Gene Sco6798 is of particular interest because it is unnamed and at the end of the *cvn7* operon. It contains a conserved GAF domain

which is present in many cGMP and c-di-GMP specific phosphodiesterases (http://strepdb.Streptomyces.org.uk). These studies provide evidence that while mutating all of *cvn7* may not exhibit a phenotypic change, deletion or overexpression of just *cvnC7* or Sco6798 could exhibit a mutant phenotype.

A concurrent analysis of a transposon mutant of Sco5292 *(cvnD5*) presents a phenotype with a difference in pigment production but appears to sporulate like wild type *S. coelicolor* at 6 days growth. Distinct mutant phenotypes of each of these genes could shed light on the function of each gene and the complete conservon in *S. coelicolor*.

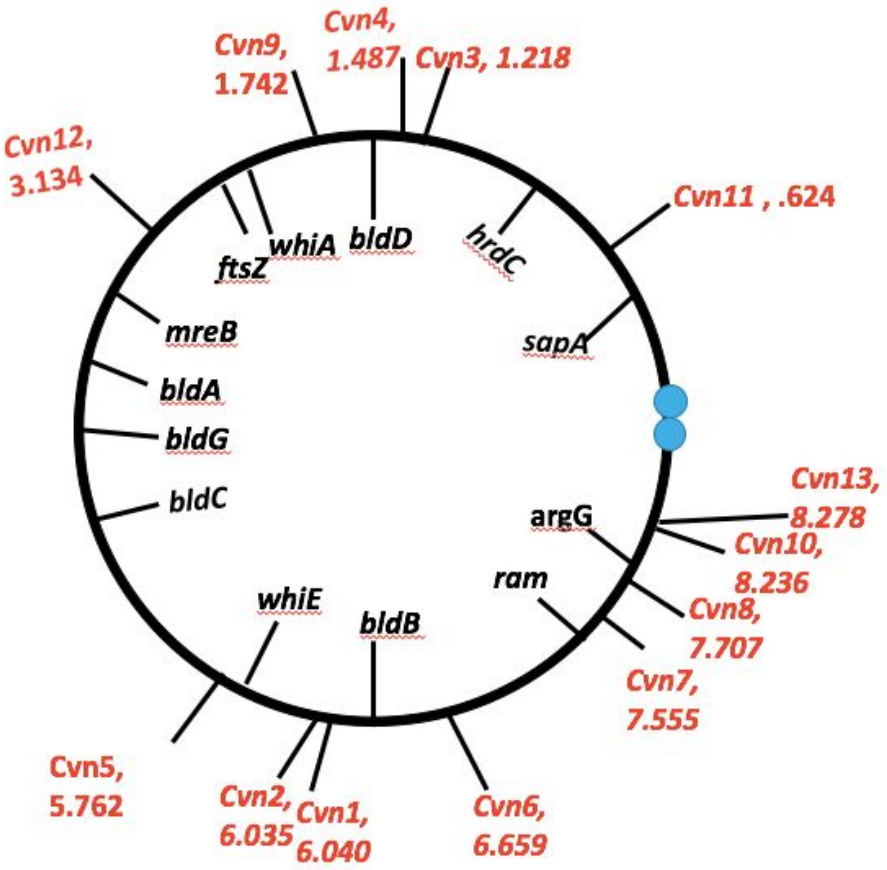


Figure 1. Chromosome map of *Streptomyces coelicolor* labeled with all 13 *cvn* operons (shown in red) and other known developmental genes for reference (shown in black).

2. Materials and Methods

2.1 BioInformatics, Protein Domains and Multiple sequence alignment.

The *Streptomyces* Annotation Server (http://strepdb.Streptomyces.org.uk) was used to look at operon 5 and 7, specifically genes C, D and the unnamed gene at the end of each operon. Each gene's conserved domains, hypothetical product, amino acid sequence and nucleotide sequence were noted. Protein sequences for *cvnD5* and each gene in *cvn7* were run in the SMART (Simple Modular Architecture Research Tool) and Pfam databases to analyze their protein domains. NCBI BLAST(basic local alignment tool search tool) was also used to analyze *cvnD5, cvnC7* and Sco6798. In addition, a multiple sequence alignment was run in Clustal Omega for *cvnD1-13* and the unnamed gene at the end of each operon.

2.2 Examination of Visual Phenotypes

Streptomyces wild type MT1110 and *Streptomyces* mutant NBS60 strains were streaked onto Soy Flour Mannitol Agar (referred to as SFM or MS Agar) plates to single colonies using the standard quadrant streak method. Sterile toothpicks were used. Plates were incubated at 30 degrees celsius. Plates were photographed on days three, four and five. Notes were taken of colony morphology for each mutant, and characteristics such as pigmentation and the presence of aerial hyphae were noted. Plates were resteaked by picking a single colony after 5 days.

2.3 Microscopy, Phase Contrast and Fluorescence

Phase-contrast and fluorescence microscopy were performed on restreaked MT1110 and NBS60 bacteria on day three of growth. Sterile tweezers were used to prepare a cover slip lift of restreaked MT1110 and NBS60 bacterial growth as shown in Bennett et al (2018). A sterile coverslip was picked up with sterile tweezers and placed on an SFM agar plate where bacterial growth was dense. The coverslip was then placed cell material facing down onto a microscope slide with a 12.5 µl drop of 50% glycerol. Phase-contrast microscopy was then performed by placing the prepared slide on the microscope stage and adding a drop of immersion oil to the center of the cover slip. Slides were viewed with 100X phase objective lens and the sample's ability to form spores, as well as spore size, shape, and overall numbers of spores were analyzed. Fluorescence microscopy was also performed according to the protocol of Bennett et al (2018). A coverslip was placed on an SFM agar plate where bacterial growth was dense and pressed lightly before being removed. The coverslip was then flooded with ice-cold methanol. It was then washed three times in Phosphate Buffered Saline (PBS). A mixture of 1% Propidium Iodide and 1% Wheat Germ Agglutinin (WGA-FITC) prepared in 50% glycerol were used to stain the DNA and newly forming cell wall, respectively. Propidium Iodide stains DNA red to detect chromosomal DNA location and Wheat Germ Agglutinin (WGA-FITC) stains newly forming cell walls green . Slides were observed under a 100X objective lens using a phase-contrast Microscope equipped with the filter set for propidium iodide (SWB) and Wheat Germ Agglutinin (WG).

2.4 SCO6796 and SCO6798 Amplification Primer Design

Primers were designed to clone Sco6796 and Sco6798 into the pL97 over-expression shuttle vector shown in figure 2. Four primers were designed. Primers were designed to be rich in AT content for better specificity in GC rich *Streptomyces*. Sco6796 and Sco6798 forward primers have added NdeI restriction sites for cloning into the NdeI restriction site on the pL97 plasmid. Sco6796 and Sco6798 reverse primers have HindIII restriction sites needed for cloning into the HindIII restriction site on the pL97 plasmid. In addition all primers have a 7 nucleotide end of DNA sequence next to the restriction site to ensure 50-100% cleavage efficiency. The following primer sequences were ordered;

SCO6798NdeIFwd 5'- GGAATTCCATATGTACAGACGCTGCTCACCCTCGTAC- 3',

SCO6798 Reverse primer,

SCO6798HindIIIRev 5'- CACCAAAAAGCTTCCGGTCCATCGAGCCTTTCTTAAG-3'

SCO6796 Forward Primer,

SCO6796NdeIFwd 5'-GGAATTCCATATGCGAGATGGCGATGCTCGTCAAGAG-3',

SCO6796 Reverse Primer

SCO6796HindIIIRev 5'- CACCAAAAAGCTTATGACCGTGGTGAGCAGTTCCTC-3'

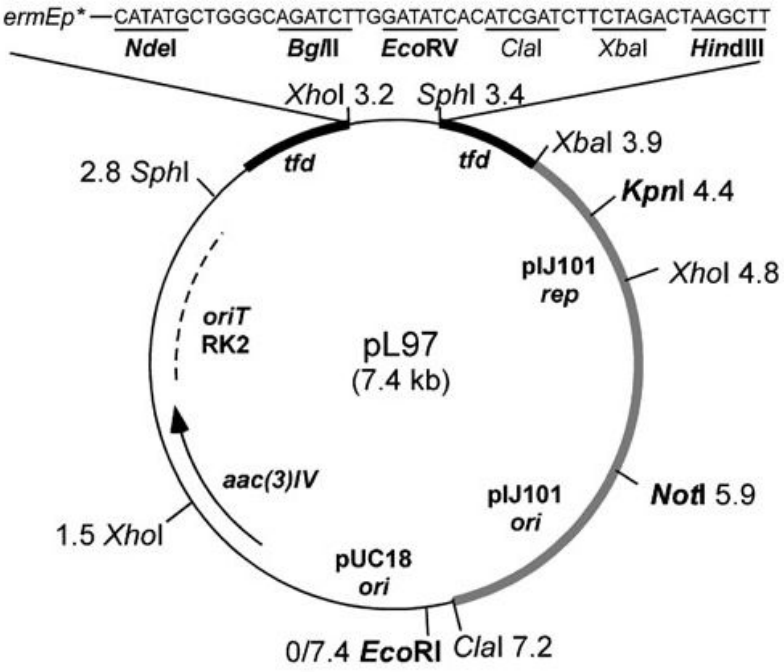


Figure 2 . pL97 over-expression shuttle vector. Genes were cloned into the multiple cloning site using the NdeI and HindIII sites to place the gene under the control of the ermEp promoter for overexpression.

2.5 PCR Amplify Gene

Ordered Primers were resuspended and then mixed with Pfx Amplification buffer, dNTP mix, $MgSo_4$, primer mix (forward and reverse primer), DNA template, platinum Pfx DNA polymerase and water to perform PCR amplification. MT1110 genomic DNA was used as the PCR template. Amplification was performed with Sco6798 NdeIfwd and Sco6798 HindIIIrev primers in one reaction and Sco6796 NdeIfwd and HindIIIrev primers in a second reaction. .

PCR products were run on a DNA gel with New England Biolabs Gel loading dye purple (6x) and stained with ethidium bromide to confirm the presence of product.

2.6 Restriction enzyme digest of PCR product and phenol-chloroform extraction

A restriction digest was run on the confirmed PCR products. The PCR products were combined with New England Biolabs 10X NEBuffer 2, sterile water, New England Biolabs HindIII and NdeI enzyme. The digest was incubated for 2 hours at 37 degrees Celsius, more of each enzyme was added and it was incubated for an additional hour. The product was then extracted using phenol-chloroform extraction followed by an ethanol precipitation.

2.7 *Plasmid Prep, restriction enzyme digest and phenol-chloroform extraction*

PL97 overnight LB Apramycin culture was prepared using Plasmid DNA Purification QIAprep Spin Miniprep Kit and a Microcentrifuge. The purified PL97 was then digested with New England Biolabs 10X NEBuffer 2, sterile water, New England Biolabs HindIII and NdeI enzyme. The digest was incubated for 2 hours at 37 degrees Celsius, more of each enzyme was added and it was incubated for an additional hour. The product was then extracted using phenol-chloroform extraction followed by an ethanol precipitation.

2.8 Clone into pL97 Vector

The PCR products for Sco6796 and Sco6798 were cloned into plasmid pL97. This was performed according to New England Biolabs Ligation Protocol with T4 DNA Ligase (M0202). Water, T4 buffer (DNA ligase), and plasmid pL97, PCR product, enzyme (T4 DNA ligase)

combined and incubated at room temperature for cohesion of the compatible sticky ends of the PCR product and vector. The reactions were then heated for inactivation of the enzyme.

2.9 Introduce into mating strain of E.coli and then into Streptomyces

The ligation reaction will be transformed into DH5α chemically competent cells according to the protocol from Invitrogen. The plasmid DNA will then be purified from DH5 alpha using the QIAprep Spin Miniprep Kit and a Microcentrifuge. The plasmid will then be digested with New England Biolabs NdeI and HindIII restriction enzymes. The digest will be run on a gel with New England Biolabs Gel loading dye purple (6x) and stained with ethidium bromide to confirm that the insert has been successfully ligated into the pL97 vector . Then the plasmid (pCF1) will be transformed into a mating strain of *E. coli* referred to as ET12567/pUZ8002. The *E.coli* mating strain contains a helper plasmid to mobilize DNA into *Streptomyces*. The bacteria will be plated on SFM plates and over layed with Apramycin (16-20 hrs after) and nalidixic acid to kill *E.coli* cells. The resulting colonies will be resteaked on SFM containing Apramycin to maintain selection of the pCF1 plasmid.

2.91 Analyze Mutant Candidates, Visual and Microscopic Phenotype.

Mutant candidates will be analyzed visually and by using phase-contrast and fluorescence microscopy. This will be performed according to the protocol stated in 2.3 Microscopy, Phase Contrast and Fluorescence of *S*co5292 wild-type and mutant.

3. Results

3.1 BioInformatics, Protein Domains and Multiple sequence alignment.

Operon 7 is shown below in Figure 3. It encodes for five proteins, genes A-D and an unnamed gene Sco6798 at the end of the opron. The operon is 4,059 nucleotides long. It is positioned from nucleotide 7,555,369 to 7,559,428 towards the end of the 8,667,507 nucleotide long genome.

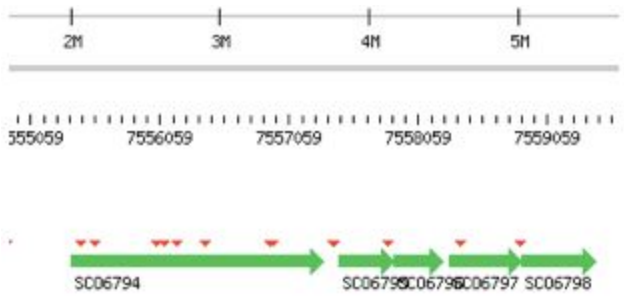


Figure 3. Operon 7 of *Streptomyces c.*

Operon 5 is shown in figure 4. Operon 5 also encodes five genes, A-D and similar to operon 7 has an unnamed gene Sco5293 at the end of the operon. The operon is 5,858 nucleotides long. It is positioned from nucleotide 5,762,200 to 5,768,058.

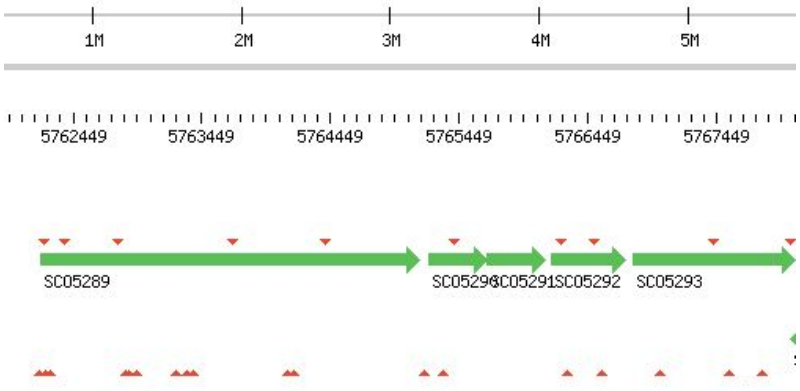


Figure 4. Operon 5 of *Streptomyces c.*

The SMART Database and Pfam analysis of *cvnA7* protein domains are shown in figure 5 below. *CvnA7* has two transmembrane helix regions, three low complexity regions and a coiled region. It also has an HATPase_c domain form position 285 to 398 with an E value of .0099. HATPase_c can be defined as a Histidine kinase-like ATPase domain and is found in ATP binding proteins.

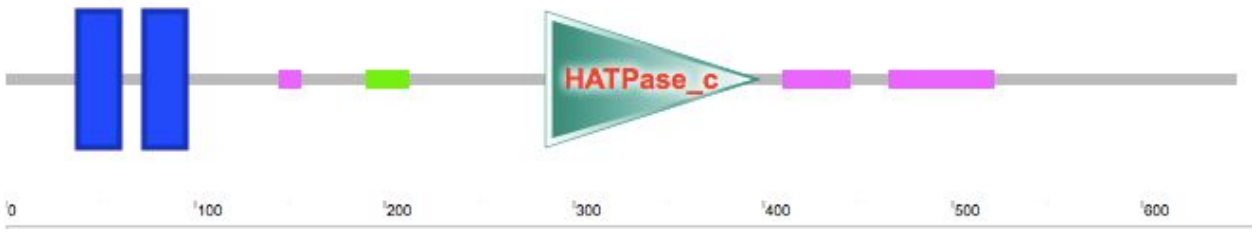


Figure 5. Protein domains of *CvnA7.*

Protein domain analysis of *CvnB7* is shown in figure 6. It shows a Roadblock/LC7 domain spanning from 14 to 104 with a significant E value of 7.816e-22. This domain in other

organisms has been thought to be involved in gliding motility but in non-motile *Streptomyces coelicolor*, the protein may be involved in structure or regulation.

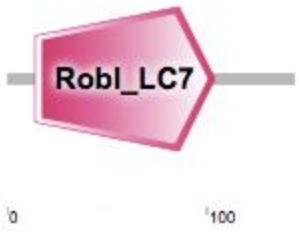


Figure 6. Protein domains of *CvnB7*.

CvnC7 protein domain analysis is shown in figure 7. *CvnC7* contains a Pfam DUF742 domain which is an uncharacterized protein domain where the function is unknown. This domain is positioned from 14 to 125 and has an E-value of 4.4e-39. In addition features not shown in the diagram include the alternative to the DUF742 domain, a HTH_ARSR domain. This domain is a helix-turn-helix, Arsenical Resistance Operon Repressor and is in transcription regulators in the arsR family. That said this domain has an E-value of 55.9 so it is not nearly as significant as the DUF742 domain.

Figure 7. *CvnC7* possible protein domains.

CvnD7 protein domain is shown in figure 8. *CvnD7* has a pfam ATP_bind_1 protein domain. The domain is positioned from 1 to 170 and has a significant E-value of 7.9e-25. This domain belongs to the GPN-loop GTPase family. GPN-loop GTPase family is a class of phosphate-binding loop NTPases that is homodimeric, regardless of the presence of nucleotides(Gras S et al.,2007). Komatsu et al., 2006, similarly found that *CvnD9* transforms from a GDP-bound form to a GTP-bound form which may in turn associate with an unidentified effector protein in order to activate downstream signaling.

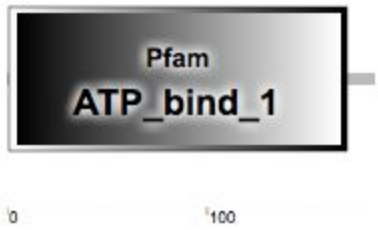


Figure 8. *CvnD7* protein domains.

Analysis of SCO6798 is shown in figure 9. Sco6798 has a GAF domain positioned 36 to 182 and has an E-value of 0.016220682678788. The GAF domain is defined as a domain present in phytochromes and cGMP-specific phosphodiesterases. GAF domains function in protein binding and it is now believed that GAF is a sensory domain.

Figure 9. Sco6798 protein domain.

Protein domain analysis of *cvnD5* is below in figure 10. *CvnD5* has a pfam ATP_bind_1 domain positioned from 19 to 188. This domain has an E-value of 4.7e-28. This domain belongs to the GPN-loop GTPase family. *CvnD7* processes this same protein domain.

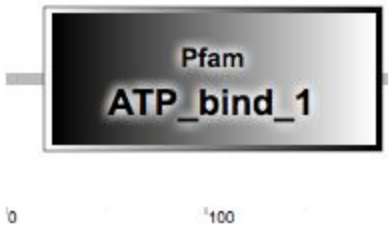


Figure 10. *CvnD5* protein domain.

The NCBI BLAST (basic local alignment tool search tool) analysis of *cvnC7* is shown in Tables 1 and 2. This matches multiple discrete regions of protein sequence similarity between *cvnC7* and sequences in an NCBI database. *CvnC7* top ten alignment matches appear in Table 1 below. *CvnC7* top ten alignment matches excluding *Streptomyces* appear in Table 2 below. *CvnC7* is highly conserved among *Streptomyces* and other organisms within the Actinobacteria class.

Genus	Query Coverage	E-value	% Identity	% Similarity
DUF742 domain-containing protein [*Streptomyces coelicolor*]	100%	8e-83	99%	100%
DUF742 domain-containing protein [*Streptomyces violaceoruber*]	98%	2e-81	99%	100%
MULTISPECIES: DUF742 domain-containing protein [*Streptomyces*]	98%	7e-81	98%	100%
DUF742 domain-containing protein [*Streptomyces sp. MH60*]	98%	1e-80	98%	99%

multi-component regulatory system-3 [*Streptomyces sp. NRRL F-4711*]	100%	1e-80	97%	98%
MULTISPECIES: DUF742 domain-containing protein [*Streptomycetaceae*]	98%	4e-80	97%	98%
DUF742 domain-containing protein [*Streptomyces sp. CCM_MD2014*]	98%	9e-80	97%	98%
hypothetical protein SAM23877_6431 [*Streptomyces ambofaciens ATCC 23877*]	100%	2e-79	95%	98%
MULTISPECIES: DUF742 domain-containing protein [*Streptomyces*]	98%	3e-79	97%	98%
MULTISPECIES: DUF742 domain-containing protein [*Streptomyces*]	98%	3e-79	97%	98%

Table 1. Top ten alignment scores for Protein-Protein BLAST (blastp) of *cvnC7*.

Genus	Query Coverage	E-value	% Identity	% Similarity
protein of unknown function DUF742 [*Actinobacteria bacterium OK006*]	98%	5e-72	88%	95%
protein of unknown function DUF742 [*Actinobacteria bacterium OV320*]	100%	7e-68	84%	92%
DUF742 domain-containing protein [*Actinobacteria bacterium OK074*]	100%	2e-62	80%	88%
DUF742 domain-containing protein [*Kitasatospora albolonga*]	96%	1e-51	72%	80%
DUF742 domain-containing protein [*Microtetraspora glauca*]	100%	5e-51	68%	80%
DUF742 domain-containing protein [*Allosalinactinospora lopnorensis*]	98%	2e-51	67%	75%

DUF742 domain-containing protein [*Streptacidiphilus albus*]	90%	2e-47	65%	78%
multi-component regulatory system-3 [*Kitasatospora aureofaciens*]	96%	1e-56	59%	70%
DUF742 domain-containing protein [*Actinoplanes sp. N902-109*]	98%	4e-45	58%	74%
protein of unknown function DUF742 [*Stackebrandtia nassauensis DSM 44728*]	99%	1e-44	61%	72%

Table 2. Top ten alignment scores for Protein-Protein blastp of *cvnC7* excluding *Streptomyces.*

The NCBI BLAST (basic local alignment tool search tool) analysis of Sco6798 is shown in Tables 3 and 4. Sco6798 top ten alignment matches appear in Table 3 below. Sco6798 top ten alignment matches excluding *Streptomyces* appear in Table 4 below. Sco6798 is highly conserved among GAF domain-containing *Streptomyces* and some histidine kinases in other organisms within the Actinobacteria class.

Genus	Query Coverage	E-value	% Identity	% Similarity
MULTISPECIES: GAF domain-containing protein [*Streptomyces*]	100%	8e-132	100%	100%
GAF domain-containing protein [*Streptomyces violaceoruber*]	100%	2e-131	99%	100%
GAF domain-containing protein [*Streptomyces sp. MH60*]	100%	2e-127	97%	98%
GAF domain-containing protein [*Streptomyces sp. CCM_MD2014*]	100%	5e-127	96%	97%
MULTISPECIES: GAF domain-containing protein [*Streptomyces*]	100%	2e-120	92%	94%

GAF domain-containing protein [*Streptomyces sp. NRRL F-4707*]	100%	1e-119	92%	93%
GAF domain-containing protein [*Streptomyces sp. NRRL F-5650*]	100%	4e-119	91%	93%
GAF domain-containing protein [*Kitasatospora aureofaciens*]	100%	9e-118	95%	96%
GAF domain-containing protein [*Streptomyces sp. M1013*]	100%	2e-115	96%	97%
GAF domain-containing protein [*Streptomyces sp. FXJ7.023*]	100%	2e-115	88%	91%

Table 3. Top ten alignment scores for Protein-Protein BLAST (blastp) of Sco6798.

Genus	Query Coverage	E-value	% Identity	% Similarity
GAF domain-containing protein [*Kitasatospora aureofaciens]*	100%	9e-118	95%	97%
GAF domain protein [*Actinobacteria bacterium OK006*]	98%	2e-80	66%	75%
GAF domain-containing protein [*Actinobacteria bacterium OV320*]	99%	4e-75	69%	76%
GAF domain-containing protein [*Streptacidiphilus albus*]	98%	4e-68	61%	67%
GAF domain-containing protein [*Actinoplanes friuliensis*]	89%	3e-66	63%	71%
histidine kinase [*Kitasatospora aureofaciens*]	95%	9e-66	63%	69%
GAF domain-containing protein [*Microtetraspora glauca*]	98%	1e-65	58%	66%
GAF domain-containing protein [*Actinoplanes sp. N902-109*]	87%	3e-63	61%	69%
histidine kinase [*Kitasatospora albolonga*]	90%	2e-61	59%	67%
GAF domain-containing protein	93%	4e-58	56%	64%

[*Glycomyces tenuis*]				

Table 4. Top ten alignment scores for Protein-Protein blastp of Sco6798 excluding *Streptomyces.*

The NCBI BLAST (basic local alignment tool search tool) analysis of *cvnD5* is shown in Tables 5 and 6. *CvnD5* top ten alignment matches appear in Table 5 below. *CvnD5* top ten alignment matches excluding *Streptomyces* appear in Table 6 below. *CvnD5* is highly conserved among ATP/GTP-binding proteins in *Streptomyces* and in other organisms within the Actinobacteria class.

Genus	Query Coverage	E-value	% Identity	% Similarity
MULTISPECIES:ATP/GTP-binding protein [*Streptomycetaceae*]	100%	2e-135	100%	100%
MULTISPECIES: ATP/GTP-binding protein [*Streptomyces*]	100%	2e-135	99%	100%
ATP-binding protein [*Streptomyces sp. CS113*]	100%	5e-135	99%	100%
ATP-binding protein [*Streptomyces sp. M1013*]	100%	6e-135	99%	100%
MULTISPECIES: ATP-binding protein [*Streptomyces*]	100%	8e-135	99%	100%
ATP-binding protein [*Streptomyces violaceus rubidus*]	100%	1e-134	99%	99%
ATP/GTP-binding protein [*Streptomyces coelicoflavus ZG0656*]	100%	1e-134	99%	100%
MULTISPECIES: ATP-binding protein [*Streptomyces*]	100%	2e-134	98%	100%

ATP-binding protein [*Streptomyces sp. CCM_MD2014*]	100%	2e-134	99%	99%
ATP-binding protein [*Streptomyces diastaticus*]	100%	2e-134	98%	99%

Table 5. Top ten alignment scores for Protein-Protein BLAST (blastp) of *cvnD5*.

Genus	Query Coverage	E-value	% Identity	% Similarity
MULTISPECIES: ATP-binding protein [*Actinobacteria*]	98%	1e-126	94%	96%
ATP-binding protein [*Kitasatospora albolonga*]	99%	2e-112	82%	92%
ATP/GTP-binding protein [*Microtetraspora glauca*]	99%	3e-109	82%	90%
ATP-binding protein [*Actinobacteria bacterium OV450*]	99%	1e-107	78%	90%
protein of unknown function ATP binding protein [*Actinobacteria bacterium OK074*]	98%	8e-107	82%	89%
ATP-binding protein [*Actinobacteria bacterium OK074*]	98%	1e-106	82%	89%
ATP-binding protein [*Streptacidiphilus oryzae*]	91%	1e-90	77%	85%
ATP/GTP-binding protein [*Streptosporangium nondiastaticum*]	99%	2e-90	73%	84%
ATP-binding protein [*Kitasatospora aureofaciens*]	93%	1e-86	72%	82%
ATP-binding protein [*Kitasatospora cheerisanensis*]	93%	1e-86	72%	82%

Table 6. Top ten alignment scores for Protein-Protein blastp of *cvnD5* excluding *Streptomyces*.

Clustal Omega multiple sequence alignment program used seeded guide trees and HMM profile-profile techniques to generate alignments between genes *cvnD1-13*. The multiple sequence alignment appears below in figure 11. The columns with an asterisk are positions where a single fully conserved residue exists between all 13 genes of *cvnD*. The Columns with a colon present are positions with conservation between amino acid groups of similar properties and columns with a period are amino acid positions with conservation between amino acid groups of weakly similar properties.This multiple sequence alignment was used to create a percent identity shown below in table 7. The multiple sequence alignment showed that genes *cvnD1-13* have percent identity spanning from 43.41 to 70.45. *CvnD7* and *cvnD5* have a percent identity of 58. 96. The phylogeny created from this analysis appears below in figure 12. It is indicates common evolutionary origins between *cvnD1-13* genes.

```
CLUSTAL O(1.2.4) multiple sequence alignment


SCO2882      ----------------------VVAGGFGVGKTTLIGSVSEVRPLRMEEPITQASAGVD      37
SCO7466      --------------------VKILVVGHFAVGKTTLIGSISEIEPLSTEETMTQAAESVD      40
SCO6940      ----------------------MIAGGFGAGKTTAVGAVSEITPLSTEEYLTEASADVD      37
SCO6797      ----------------------VAGGFGVGKTTFVGAVSEIAPLSTEELLTTVSAATD      36
SCO7419      ----------------------VVVGGFGVGKTTMVRSVSEIRPLNTEETMTQAGEAVD      37
SCO0585      ---------------------VGGFGVGKTTLVRSVSEIRPLNTEEVMTQAGVGVD      35
SCO5292      ------VQPWQTDPTRAPVATKIVVAGGFGVGKTTLVTAVSEITPLQTEALMTEASEETD      54
SCO5541      ----------------------VVAGGFGVGKTTFVGAVSEINPLRTEAVMTSASAGID      37
SCO5537      ----------------------VVAGGFGVGKTTFVGAVSEINPLRTEAVMTSASAGID      37
SCO1157      MRSERSERSDDTDGDTTALALKILVAGGFGVGKTTLVGAVSEIRPLRTEELLSEAGRLVD      60
SCO1399      MIFG---RSQRGKPPVEPVTLKILVAGGFGVGKTTLVGAVSEIRPLRTEELLTEAGRPVD      57
SCO6066      --------------------VKILIAGGFGVGKTTMVGSVSEIAPLRTEEPLTAAGLDVD      40
SCO1627      ---------------------VGGFGVGKTTMVGSVSEIRPLTTEETMTQAGIGVD      35
                                  .* *..**** : ::**: **  *  :: ..   *

SCO2882      DLRGTPHKTTTTVAMDFGRIHMAGG--RLALYLFGLPGQSRFQPLWEDLAEGALGCLVLA      95
SCO7466      DLKGVQGKTTTTVAMDFGRLTISDR---VVLYLFGTPGQQRFVQMWEDMARGALGALVLV      97
SCO6940      SLAGLASKRTTTVAFDFGRLSLPDAPMPLELFLFGTPGQDRFVDLWYDLSRGAVGAVVLV      97
SCO6797      NLDGIENKLETTVAMDFGRLTLDPE---HVLYLFGTPGQERFWFMWDELCEGALGAVIIA      93
SCO7419      DISEVRGKSATTVAFDFGRISLDAR---NVLYLFGAPGQERFWFLWDRLFSGTLGAVVLV      94
SCO0585      ETGGVTSKTTTTVAFDFGRISLNDR---MVLYLFGAPGQERFWFLWDRLFSGTLGAVVLV      92
SCO5292      DLTATPGKLTTTVAMDFGRITLDDD---LVLYLFGTPGQQRFWFMWDDLVRGAIGAVVLA      111
SCO5541      DLTHAGDKTTTTVAMDFGRITLDQD---LILYLFGTPGQDRFWFMWDDLVRGAIGAVVLC      94
SCO5537      DLTHTGDKTTTTVAMDFGRITLDQD---LILYLFGTPGQDRFWFMWDDLVRGAIGAIVLV      94
SCO1157      DTDGIDRKVTTTVAMDFGRITIRSG---LSLYLFGTPGQDRFWFLWDELAQGALGAVVLA      117


SCO1399      DVSGVEGKHTTTVAMDFGRITLRED---LVLYLFGTPGQERFWFMWDELSEGALGAVVLA      114
SCO6066      DLAGIEEKRATTVALDFGRITIGRD---VVLYLFGTPGQQRFWFMWNDLAIGALGAVVLV      97
SCO1627      DNYGSASKTATTVAMDFGRISITDE---LVLYLFGTPGQERFWFLWNGLFEGALGAVVLV      92
             .      *  ****:****: :         *:*** ***.**  :*  :  *::*.:::

SCO2882      DTRDLDASHDALGLLDGAGIPYAVAINTFPGSP-SYPEAELREALALEPATPLTYCDARD      154
SCO7466      DPERLADSFPVIDLIEHYGLDYAIAVNHFDGAP-LRDQRALREALDLLDDTPVVTCDARD      156
SCO6940      DTRRLESSFTPVSFFEDIGLPFVVAVNRFDGAH-RYHPEQVRAALELPAAVPVVTCDARE      156
SCO6797      DTRRLEESFAAVDFFEERGLGFVVAINEFDGSY-RYDPEEVRAAMDLHPEIPIVRCDARI      152
SCO7419      DTRRIDDSWYAIDRLEHHGTPFIVACNDFGGP--EHSPEAVREALDLDPRVPLVSCDARS      152
SCO0585      DTRRMDDSWYAIDRLEHHRTPFVVAVNRFDDDTARHSLDEIRLALALPDHVPLIDCDARV      152
SCO5292      DTRRLKDCWPALDYFESCGLPYVVAVNHFDGSE-LFEAEDVREALTIPAHIPVMIMDARR      170
SCO5541      DTRRLADCFPAVDYFENSGLPFVVALNGFDGQQ-PYSPDEVREALQIGPDTPIITTDARH      153
SCO5537      DTRRLADCFPAVDYFENSGLPFVIALNGFDGQQ-PYNPDEVREALQIGPDTPIITTDARH      153
SCO1157      DTRRLADCFPAVDYFEHRHIPFVVAVNCFTGAR-RHDEYSVLRALDLDPGTPVVLCDARD      176
SCO1399      DTRRLEDCFAAIDYFERRSIPFVVGVNCFEGSA-RYPGETVRQALDLDEDVPLVMCDARD      173
SCO6066      DVRRPESSFAAIDFFERRGIPFVVAVNGFHGRH-PYPAAEIREALALPEHVQVLLCDARE      156
SCO1627      DTRRLEVSFDVMGRLEERGVPFVVAVNTFPDAP-RYPVAELRTALDLPEEIPILDCDVRR      151
             * .    .   :. ::     : :. * * .         :  *: :     :   *.*

SCO2882      RTSSLHALITLTEYLSEHRSAALLESRR--------      182
SCO7466      EKSSAAALTTLVRYLLDRTR----------------      176
SCO6940      PNHVAGVLLALVDHAVNCAATTNRAGHTPSTTLQDA      192
SCO6797      SSSGVQTLLTLVRHLIAHTPAPTSGSGAGVRP----      184
SCO7419      RQSSKQVLITLVEHLQAHYAGPTARPRQEFA-----      183
```

```
SCO0585     RTSGKHVLVTLVDHLYQLALAQESAS----------     178
SCO5292     RISAIETLLSLVGHALDETPE--------------     191
SCO5541     RADAKSALITLVEHALMARLR--------------     174
SCO5537     RADAKSALITLVEHALMARLR--------------     174
SCO1157     RDSGKEVLIRLVEYAGRMHTARLLDSVR--------     204
SCO1399     RESVKEVLIGVVQHAMTQAADRRRAVPT--------     201
SCO6066     RTSSRDVLIALIDQLIDAAAGAAAS-----------     181
SCO1627     RASSRDVLMTLMRLLHSLAMKGALT-----------     176
                  .*   :
```

Figure 11. Clustal Omega Multiple sequence alignment of *cvnD1-13*.

Percent Identity Matrix - created by Clustal2.1

1: SCO2882	100.00	53.18	43.41	47.78	50.56	50.85	49.43	52.87	52.87	54.14	49.72	52.25	55.68
2: SCO7466	53.18	100.00	49.71	50.00	52.91	50.88	54.55	51.45	52.60	53.41	51.14	53.41	53.80
3: SCO6940	43.41	49.71	100.00	53.80	47.54	52.54	55.17	55.17	54.60	52.49	56.91	57.30	52.84
4: SCO6797	47.78	50.00	53.80	100.00	53.30	50.85	58.96	58.38	57.23	57.22	59.44	54.80	54.55
5: SCO7419	50.56	52.91	47.54	53.30	100.00	70.45	53.76	54.91	53.76	54.44	58.33	55.37	58.29
6: SCO0585	50.85	50.88	52.54	50.85	70.45	100.00	52.33	54.07	52.91	55.37	57.06	55.68	61.93
7: SCO5292	49.43	54.55	55.17	58.96	53.76	52.33	100.00	67.24	66.67	55.50	58.95	54.24	50.00
8: SCO5541	52.87	51.45	55.17	58.38	54.91	54.07	67.24	100.00	97.13	60.34	59.77	54.60	51.74
9: SCO5537	52.87	52.60	54.60	57.23	53.76	52.91	66.67	97.13	100.00	59.20	58.62	54.02	51.16
10: SCO1157	54.14	53.41	52.49	57.22	54.44	55.37	55.50	60.34	59.20	100.00	65.67	56.91	54.55
11: SCO1399	49.72	51.14	56.91	59.44	58.33	57.06	58.95	59.77	58.62	65.67	100.00	61.88	56.82
12: SCO6066	52.25	53.41	57.30	54.80	55.37	55.68	54.24	54.60	54.02	56.91	61.88	100.00	61.93
13: SCO1627	55.68	53.80	52.84	54.55	58.29	61.93	50.00	51.74	51.16	54.55	56.82	61.93	100.00

Table 7. Clustal Omega Multiple sequence alignment of *cvnD1-13* percent identity matrix.

Figure 12. Phylogenetic tree from Clustal Omega Multiple sequence alignment of cvnD1-13.

Clustal Omega multiple sequence alignment between unnamed genes at the end of each operon appears below in figure 13. The columns with an asterisk are positions where a single fully conserved residue exists between unnamed gene at the end of each operon. The Columns with a colon present are positions with conservation between amino acid groups of similar properties and columns with a period are amino acid positions with conservation between amino acid groups of weakly similar properties. The multiple sequence alignment of unnamed genes at the end of each operon show significantly fewer similarities than the multiple sequence alignment of *cvnD1-13* genes. This multiple sequence alignment was used to create a percent identity shown below in table 8. The multiple sequence alignment showed that unnamed genes at the end of each operon have percent identity spanning from 12.17 to 50.86. The unnamed genes at the end of each operon 7 and 5 have a percent identity of only 26.35. The phylogeny created from this analysis appears below in figure 14. It indicates common evolutionary origins between unnamed genes at the end of each operon.

CLUSTAL O(1.2.4) multiple sequence alignment

```
SCO2883      ----------------MTLPSTETAPTGEPGRIALYAPEFAADPHAAYRSMRRTHGPLVP      44
SCO1626      ---VTPESHSPTGTGEPLLEPPPGCPAHGLGPGGLHRLHEAEDLEELYEKLREQHGPVAP      57
SCO7418      --------------------VSTDAHDVP-GAVPLGGPRFQTDPALLYRQMRREHGAVTP      39
SCO0584      VTDIDPSPHP---------VAAPGCPVHP-DAVPLAGLEYQQTPSELYRGLRAEHGAVAP      50
SCO5293      ------------------------------------------------MRKI-----L       5
SCO6798      ------------------------------------------------------------      0
SCO6939      ------------------------------------------------------------      0

SCO2883      VDLAPGVPATLVIGYYQ----ARRIL-------------NDPLRF------PAD--PRAW      79
SCO1626      ALLHDDVPMWVVLGHAE----NLHMV-------------STPSQF------CRD--SRIW      92
SCO7418      VVLDGDVPAWLVLGYRE----LHQVT-------------GDPVLF------SRD--SDLW      74
SCO0584      VLLDGGIPAWLVLGYPE----VSYVT-------------SHDELF------ARD--SRRW      85
SCO5293      VVGAGQSGLQLALGLQSHGYEVTLMSNRTADEIRTGRVMSTQCMFHTALQHERDLQLNFW      65
SCO6798      ------------------------------------------------------------      0
SCO6939      ------------------------MSQ-------------------------------       3

SCO2883      EKL-----------------------------------IPATCPVRPMMEW-RPNALRSG      103
SCO1626      TPL-----------------------------NEGMVKPDHPLMPHIAW-QPICSHAE      120
SCO7418      NQW---------------------------------ENIPDDWPLLPMIGRRQPSILYTV      101
SCO0584      NQW---------------------------------GSIPPDWPLLPYVGH-QPSVLFTE      111
SCO5293      ESQAPKIEGLGVSVAAPGSFDPGPSQRAIDWLGHLDGFAQSVDQRVKMAGWMETFA-QRG      124
SCO6798      ------------------------------------------------------------      0
SCO6939      ------------------------------------------------------------      3

SCO2883      GAEHTRYRSANTHAIDQVDQHGLRALVEQVASDAIEGFRTAGSADLLTQYSFPIAFRVLS      163
SCO1626      GDEHKRLRGAVTSAMSDLDYRELRRHIKRYTQRVVNRFCEEGRADLVSQFAEHLPMGVMC      180
SCO7418      GERHRERAAMISDALEAVDPHLLRGHAERFADELVDRLCAKGEADLVGDYAMLLPVRVLA      161
SCO0584      GEEHRRRAGVITQALAGIDQFELARDCRHLADRLIAAFAGSGRAELMSGYAHPLPMRAAV      171
SCO5293      GQ-----LVIHGAAVGDLDYFSRAYDL---------VLVSAGKGELVQM-----------      159
SCO6798      ------------------------------------------------------------      0
SCO6939      ------------------------------------------------------------      3

SCO2883      ALLGCPDEIGQR--IADGMAKIFD---TTNADQGNLILAQAVSDLVTLRRTHPGDDI---      215
SCO1626      HLLGMPEEYNDR--LVEAARDTLK--GTETAIASHAYVMEALGRLTATRRADPADDI---      233
SCO7418      RLYGFPDEQGPA--LVTALNDMID--GRERALAGQTHLGTSMARLLADRKAAPADDV---      214
SCO0584      RMCGMPHTGVETRQLVEDLRISLDAAEGDDPVAAYTRVGERIHQLVRHKRERPGPDV---      228
SCO5293      --FGRDAERSPYS------------EPQRALA-----VAYVHG-MGPRPEHPETEAVRC      198
SCO6798      ------------------------------------------MSYDPPRPA------      9
SCO6939      ------HPSDPYR------------TPQTAPPARSLPRRDMRD-TTRGPLTPP------      37
                                                                 *
```

```
SCO2883      -----TSRLALHPVRLTDEEMSHQLVTLYGAGIEPMTNLISNTILKILTDEEFSADLHAG      270
SCO1626      -----AGRLVTHPAGLTDDEVREHLRVVLLAAYEATVNLIGNVMRVVLTDPGFRAQLSGG      288
SCO7418      -----ASRMLADESGFTEEEVAQDLMVMMAAGHQPTADWIGNSLRLMLTDDRFAASLFGG      269
SCO0584      -----TSRMLTHPAGLTDEEIVQDLISVIAAAQQPTANWIGNTLRLLLTDERFALNVSGG      283
SCO5293      NLVPGVGEMFIMPT----FTTSGRADILFWEGIPGGPLDAFKDV------KDPAEHL---      245
SCO6798      ------GRLLLTP---EDREAPTRVRRLRRLGLAERADPALDA---------FAAHL---      48
SCO6939      ------GTVLSSKPPQQSLELAQRYELLNRLGVPPMANEDFDD---------LARDM---      79
                   . :                  :   .        .            :

SCO2883      LSTVRDALDAVLYTDPPMANYCISYPPYPVDVEGVLLPADQPVVISMAAANNDPALTEGV      330
SCO1626      QMTVPQAVEQSLWDEPPFSTVFAYFAKQETELGGQRIRAGDGLLLGIAPGNVDPRIRPDL      348
SCO7418      RNSVAEAMNEVLWEDTPTQNVAGRWAARDTQLGGRRIRAGDLVLLGLQGANSDPQVRTDG      329
SCO0584      RLSVGEALNEVLWLDTPTQNFIGRWAVNDTQLGGRHIRAGDCLVLGFAAANTDPQLWPEA      343
SCO5293      -SLTLELMEKFLPWEYARATK-----VELT-------DAGGT-LAGR----YAPTVRNPV      287
SCO6798      -AALAEA-------PYA---M-----VNFL-------VEGGQFFAGL----RVPEV-PPV      80
SCO6939      -ATRARF-------QYG---F-----VNLF-------LEEQTF-VGL----HQPP-----      106
                                                          .      *

SCO2883      PAGQHGGNRAHLAWSTGPHTCP--------------------ARSHAYLIAETAVTHLL      369
SCO1626      D-ASMMGNRAHLAFGGGPHECP--------------------GQDIGRAIADAGIDALL      386
SCO7418      S-ALTGGNNAHFSFGHGEHRCP--------------------FPAQEVAEVIARTGIEVVL      369
SCO0584      H--VGAENSAHLSFSNGEHRCP--------------------YPAPLLADVVARTAVETLL      382
SCO5293      GRLPGGGAVLGVA---------DVVVAN-----------DPITGQGSNS--ASKCAASYL      325
SCO6798      TRGDGTSPEFGRVQPRDHGFCPHVVVRRKALVLEDVRDYPRFAGNPVVD--A-FGIRSYL      137
SCO6939      --ADSGYVFVGRTMSRDHGWCPEVMARKKALPLHDVHASPRFSGNAVVD--A-VGIRSYF      161
                                                       .      *      :

SCO2883      DALPETDLARPAAELVWRPGPFHRALESLPVTFPAAQSAAH------------------      410
SCO1626      MRLPDLQLDCDEDDLRWRSSIASRHLVELPVRFEPRAQQDIMQQPSHAPTPERH------      440
SCO7418      DRLPDIDLAVPAGSLTRRPSPWLRGLTELPVRFTPTTALGGTSA----------------      413
SCO0584      ERLPDLVLALEPGELTWRPSIWMRGLTALPALFTPVVA---------------------      420
SCO5293      AS--IIEHGDQEFDETWMRQTFDRYWETAQHV-------TKWTNTMLAPPPEHILNLFGA      376
SCO6798      GA-PLIDSTGTVLGTVSAADVQPRTWGTAGLA--------TIK-TTAADLVRRI------      181
SCO6939      GA-PLIHDSGTVLGTVCVIDPEKRPLAEARRL-------RDIVINASAQVMDHI------      207
                                  *

SCO2883      ----------------------------------------------------------      410
SCO1626      --------APWHVGLPKPERRAQPPLPAQPPQPVSVTAAEPQQAPGAGQPRPRGAWQRFL      492
SCO7418      ----------------------------------------------------------      413
SCO0584      ----------------------------------------------------------      420
SCO5293      ATQLQPIADRFANGFNDPSDFENFF--YDPAKT-EGYLAEVSGAAGA-------------      420
SCO6798      -------ERSAEDGLPL-------------------------------------------      191
SCO6939      -------TRAPVR-----------------------------------------------      213


SCO2883      ------      410
SCO1626      LWWRGY      498
SCO7418      ------      413
SCO0584      ------      420
SCO5293      ------      420
SCO6798      ------      191
SCO6939      ------      213
```

Figure 13. Clustal Omega Multiple sequence alignment of unnamed gene at the end of each operon.

Percent Identity Matrix - created by Clustal2.1

1: SCO2883	100.00	30.07	33.17	31.92	17.39	16.89	14.71
2: SCO1626	30.07	100.00	37.56	38.69	19.15	18.45	12.17
3: SCO7418	33.17	37.56	100.00	50.86	18.66	21.19	17.61
4: SCO0584	31.92	38.69	50.86	100.00	16.55	17.69	14.62
5: SCO5293	17.39	19.15	18.66	16.55	100.00	26.35	19.78
6: SCO6798	16.89	18.45	21.19	17.69	26.35	100.00	40.56
7: SCO6939	14.71	12.17	17.61	14.62	19.78	40.56	100.00

Table 8. Clustal Omega Multiple sequence alignment of unnamed gene at the end of each cvn operon percent identity matrix.

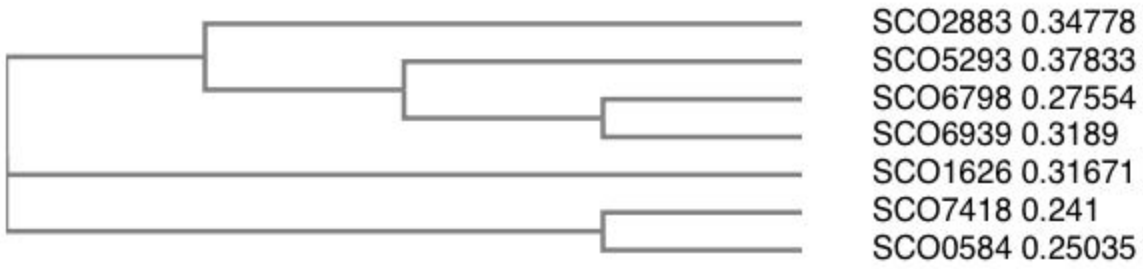


Figure 14. Phylogenetic tree Clustal Omega Multiple sequence alignment of gene at the end of each cvn operon.

3.3 Physical Phenotype of Wild type MT1110 and Sco5492 NBS60 Transposon Insertion Mutant

Sco5292 was found when scanning transposon insertion plates for visible phenotypic mutations. The mutation was then sequenced and found to be Sco5292. Sco5292 is *cvn5D*, a

gene apart of *Streptomyces* operon 5. According to Strepdb (http://strepdb.Streptomyces.org.uk) *cvn5D* is a possible ATP/GTP-binding protein. *Streptomyces* wild type MT1110 and *Streptomyces* mutant NBS60, containing a transposon disruption of gene Sco5292, were streaked onto SFM agar plates. Photographs taken on day 3 of growth are shown in figure 15. Image A has wild type *Streptomyces* growing on the left side of the plate and Sco5292 mutant growing on the right side of the plate. Image B has a plate streaked with wild type *Streptomyces* and image C has a plate streaked with just the Sco5292 mutant *Streptomyces*. The image below these are the back of the corresponding photos above them. NBS60 appears to present a difference in pigment production. While there does appear to be a difference in pigmentation there appears to be no visible difference in colony morphology or growth rate. This trend continued for days 4 and 5 of growth, shown in figures 16 and 17.

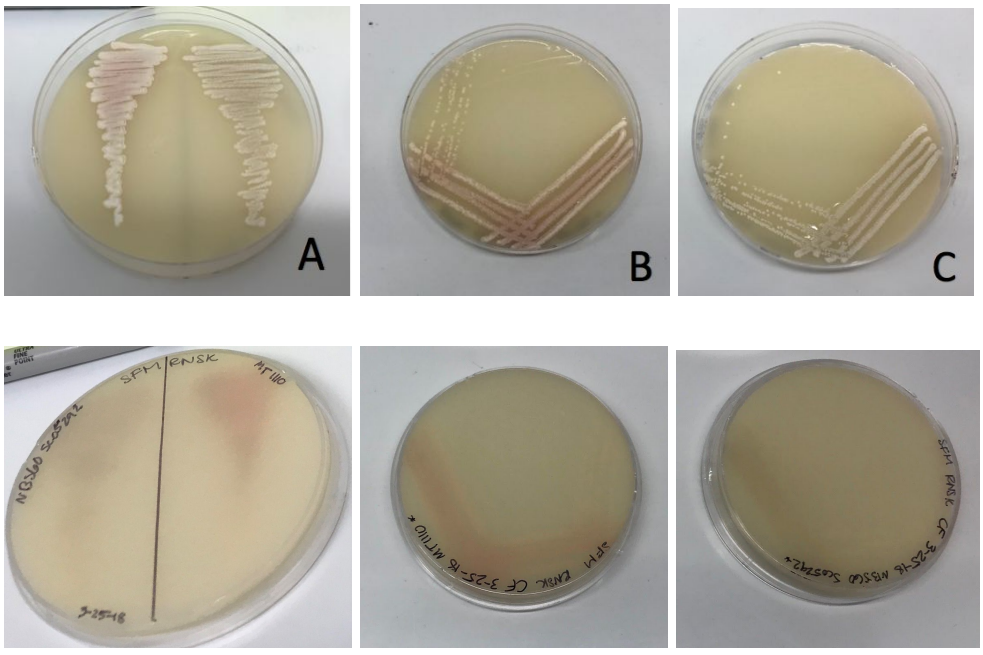


Figure 15. Day 3 of wild type and Sco5292 mutant growth.

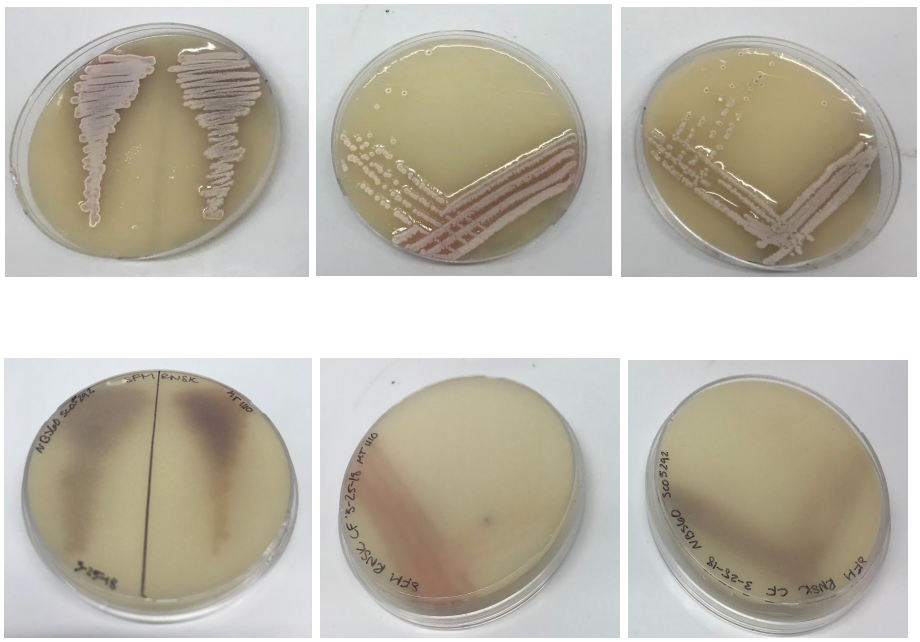


Figure 16. Day 4 of wild type and Sco5292 mutant growth.

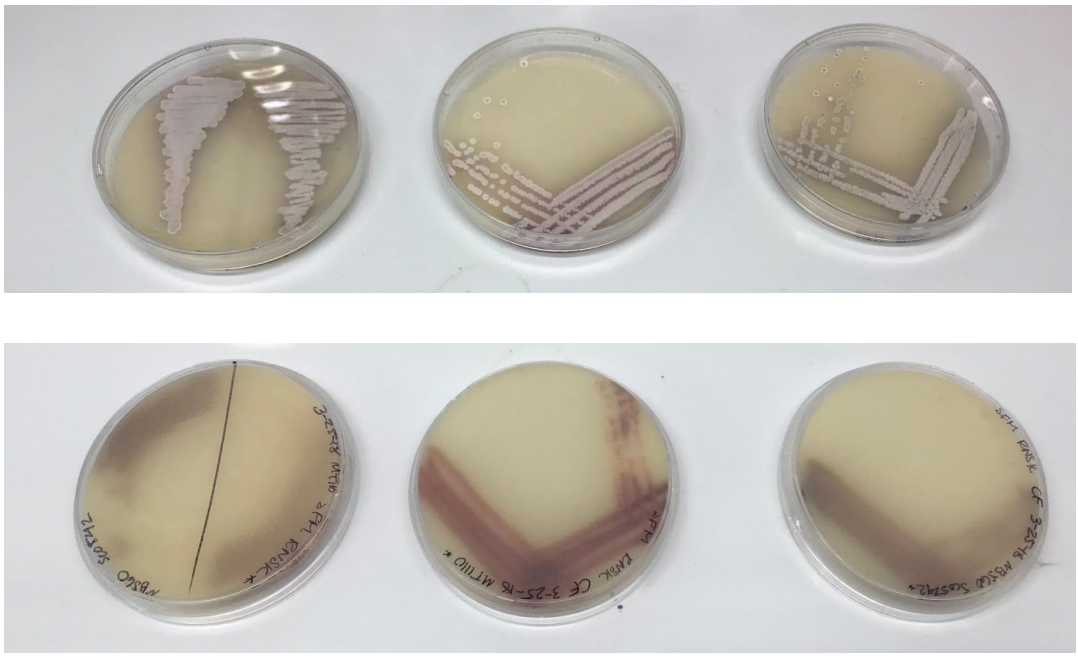


Figure 17. Day 5 of wild type and Sco5292 mutant growth.

3.3 Microscopy, Phase Contrast and Fluorescence

Phase-contrast and fluorescence microscopy were performed on MT1110 and Sco5292 mutant NBS60 bacteria on day three of growth. Phase-contrast microscopy in figures 18 and 20 showed that wild type *Streptomyces* has relatively uniform coloration, spore size and shape. Sco5292 appears to have more frequent misshapen and irregularly-sized spores. The Sco5292 mutant also appears to have more frequently lysed cells. On Day 3 of growth these samples of wild type and Sco5292 mutants were also analyzed with 1% Propidium Iodide stain using the wide green fluorescence filter and 1% Wheat Germ Agglutinin (WGA-FITC) stain using the SWB fluorescence filter. The results of this appear below in figures 19 and 20. Propidium Iodide stain makes the DNA appear red and Wheat Germ Agglutinin (WGA-FITC) shows newly developing cell walls. These photos in figures 19 and 20 show that the Sco5292 mutant has lower intensity staining of propidium iodide within many spore compartments as compared to wild type. The visibly lower intensity propidium iodide staining in the mutant indicated that there was less DNA in the cell or that the DNA in the cell was degrading.

Wild Type Sco5292

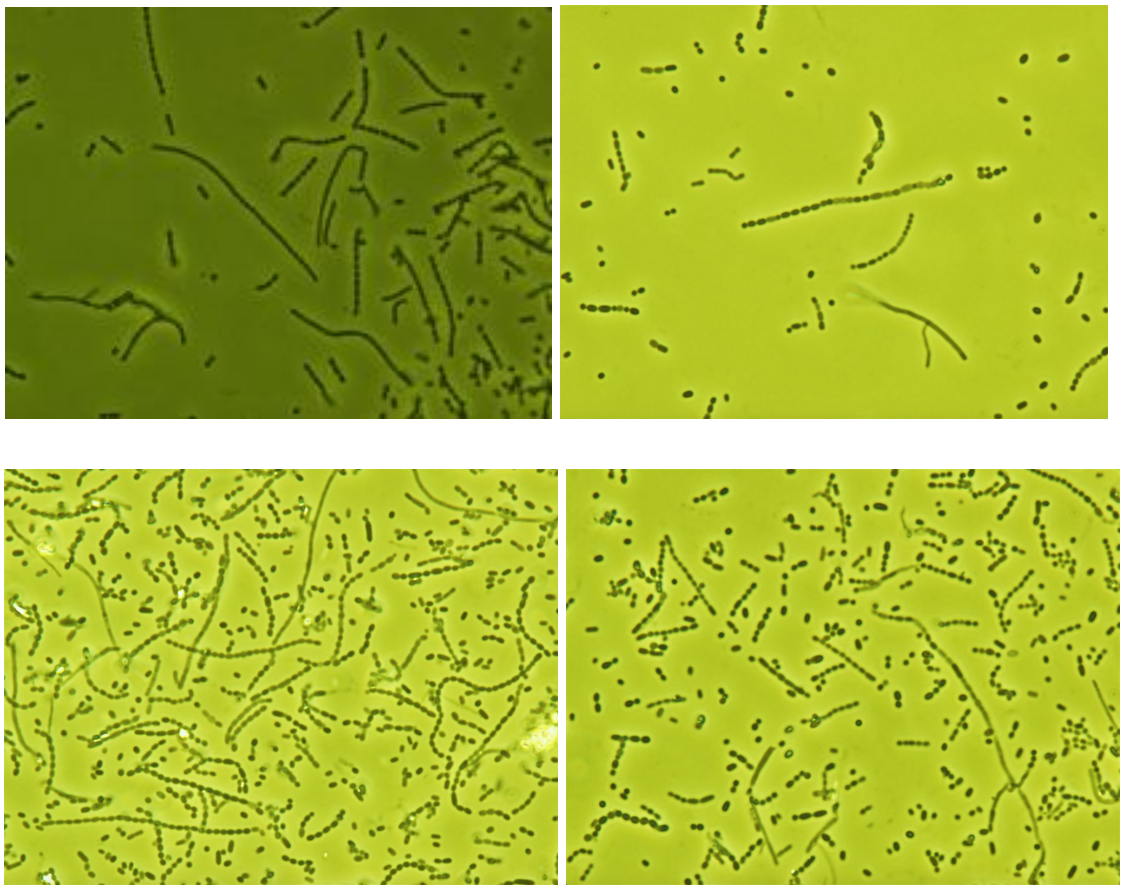


Figure 18. The left two quadrants are phase contrast photos of wild type *Streptomyces* and the right two quadrants are Sco5292 mutants.

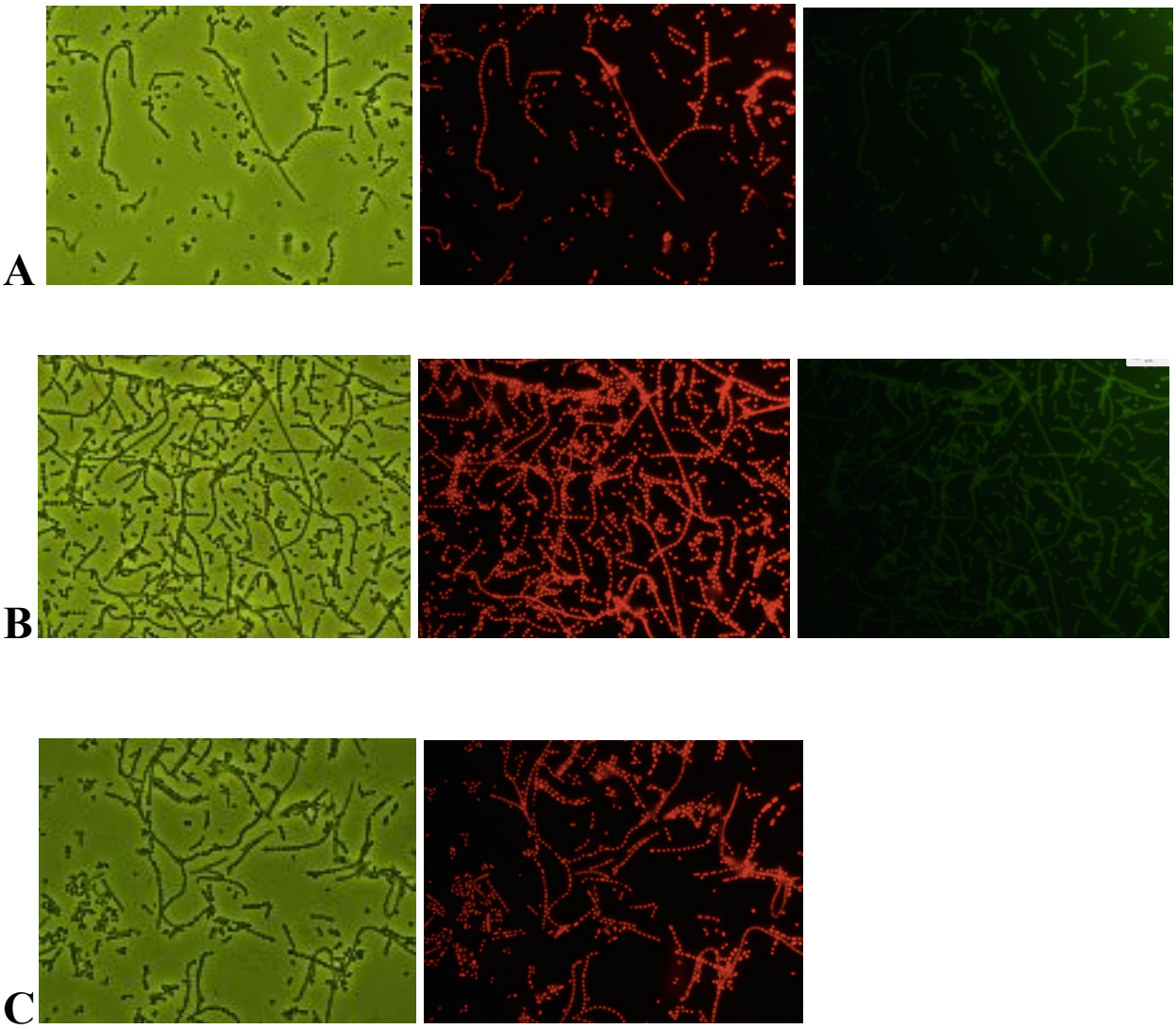
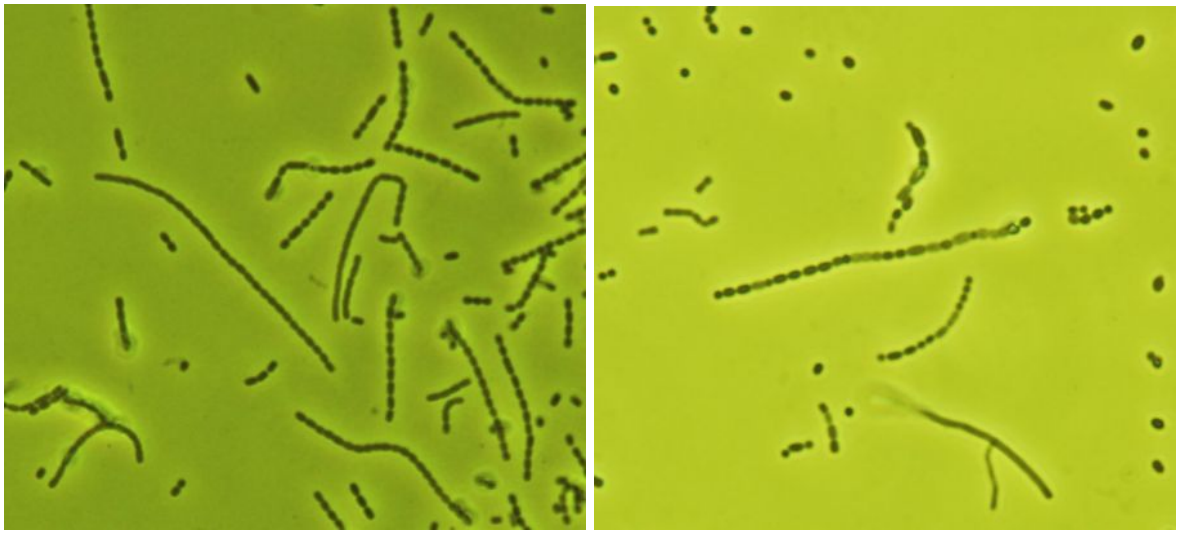
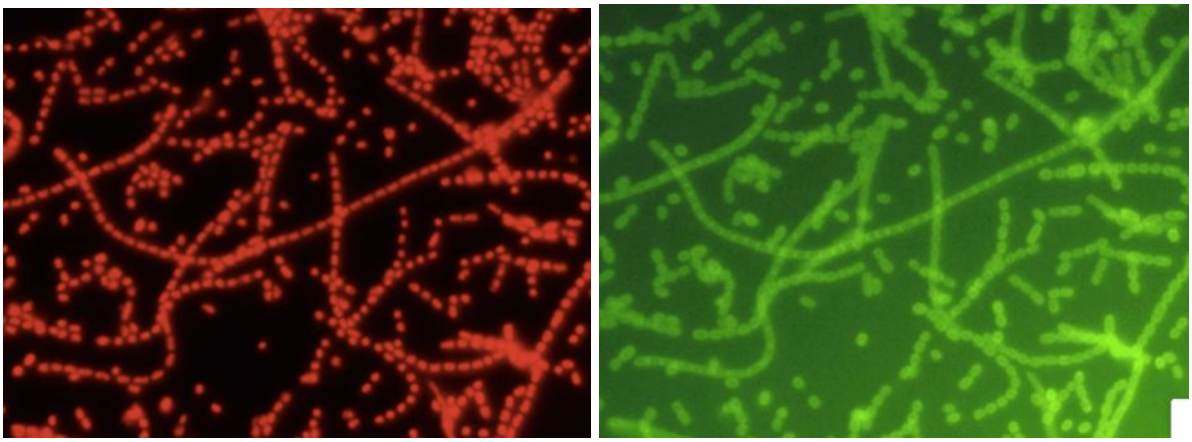


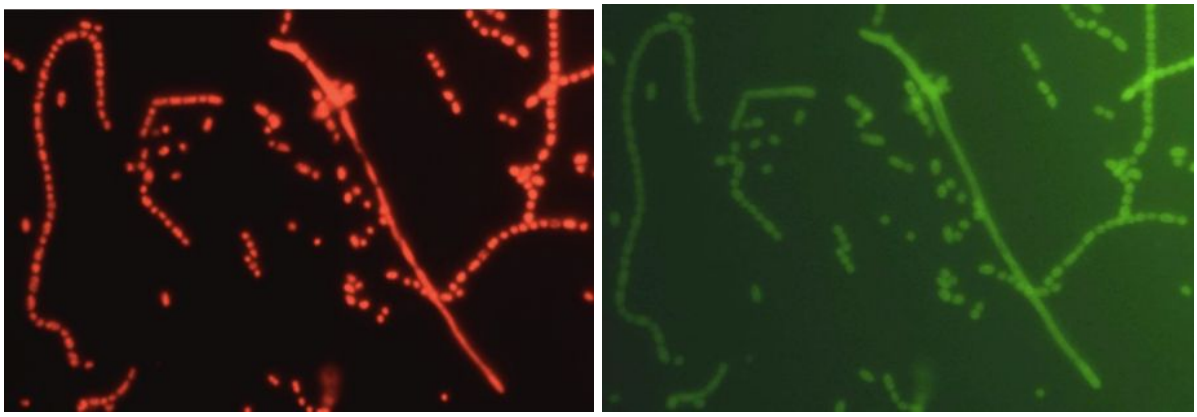
Figure 19. Phase contrast, 1% Propidium Iodide stain under GW column filter and Wheat Germ Agglutinin(WGA-FITC) stain under SWB column filter. Rows A & C contain mutant Sco5292 *Streptomyces*. Row B contains wild type *Streptomyces*.



A



B



C

Figure 20. Row A; close up phase contrast with wild type *Streptomyces* on the left and *cvnD* mutant on the right. Row B; Wild type 1% Propidium Iodide stain under GW column filter and Wheat Germ Agglutinin(WGA-FITC) stain under SWB column filter. Row C; *CvnD5* Mutant 1% Propidium Iodide stain under GW column filter and Wheat Germ Agglutinin(WGA-FITC) stain under SWB column filter.

3 .5 PCR confirmation of Amplified Genes

CvnC7 and Sco6798 were PCR amplified and the product was then confirmed by gel electrophoresis. The gel product is pictured below in figure 21. Bands in lanes 5 and 7 are of the predicted size 706 bp and bands in lanes 9 and 11 are the predicted size of 565 bp. Extra bands in lanes 5 and 7 are underneath the bands of expected size and most likely represent nonspecific binding of one or more primers to form an undesired product.

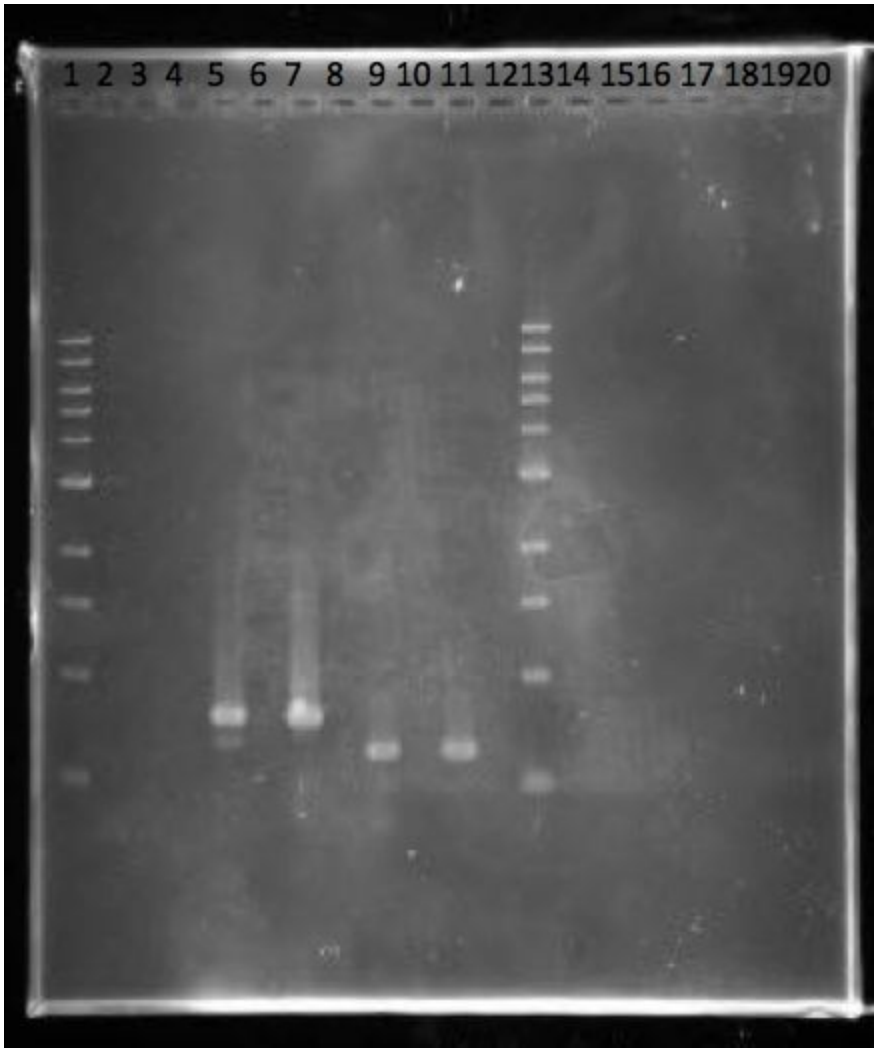


Figure 21. *CvnC7* and Sco6798 PCR product confirmation. Lanes 1 and 13 are DNA ladders, lanes 5 and 7 are Sco6798 PCR products and lanes 9 and 11 are Sco6796 PCR Products.

Discussion

The features and specific function of the 13 copies of *cvn* throughout the *Streptomyces* genome is widely unknown. Previous knockout studies have shown varied effect while deletion of a whole operon leads to a wild type phenotype for every *cvn* except *cvn1*. When just one gene in *cvn1*, *cvn9* and *cvn10* was deleted it resulted in a mutant phenotype for *cvnA1, cvnA9, cvnD9*, and *cvnA10* (Komatsu et al, 2006). This inconsistency could be due to an ability of other *cvn*

homologs to compensate for cvn operon deletions or the regulatory pathway affected by the gene deletion does not appear to be important under normal culture conditions.

In RNA-Seq, a cyclic di-GMP double mutant showed significant under expression of the *cvn7* operon compared to the wild-type. This indicates that *cvn7* is a likely target of cyclic di-GMP and it is most likely involved in one or more facets of *Streptomyces* development. Bioinformatic results reported from this study show that the genes within each *cvn* operon are diverse, implying that they each have a distinct role in *Streptomyces*. *CvnC7* has a large uncharacterized protein domain with unknown function. Blast data shows that *cvnC7*'s protein sequence is highly conserved providing evidence that the gene most likely attributes to specific characteristics of the class, potentially for their development. Analysis of the *cvnC7* amplified mutant phenotype will shed light on *cvnC7*'s function in *Streptomyces* as well as the function of this uncharacterized protein sequence and the importance of the conservon as a whole. Overexpression as opposed to deletion of this gene could reduce the possible effect of other homologs compensating for mutation.

Sco6798 is one of multiple unnamed genes at the end of different operons in the conservon. This gene has a GAF protein domain believed to have a sensory function, possibly in cGMP or other cyclic nucleotide signaling. Blast data shows this protein sequence is highly conserved in GAF domain-containing *Streptomyces* as well as in histidine kinases in other organisms within the Actinobacteria class.

This gene is not named as apart of the homologous operons and therefore may not have a homologous gene to compensate for its mutation. The effect of deletion or amplification of this gene or a similar gene at the end of the operon in the conservon has not yet been studied.

Analyzing its overexpression will shed light on its role within the operon, cyclic di-GMP signaling pathway, and *Streptomyces* development.

Concurrent analysis of the Sco5292 (*cvnD5*) transposon mutant sheds light on the function of this gene and the other homologous *cvnD* genes in the other 12 operons. It also supports transposon insertion as another viable option for analyzing other genes in the conservon. RNA-Seq analysis of the cyclic di-GMP phosphodiesterase double mutant inactivated for *rmdA* and *rmdB* did not show differential expression of *cvn5,* but microarray data from a *bldD* mutant in comparison to wild type showed differential expression for *cvnD5*, specifically. *bldD* is believed to be a key regulator of developmental genes. While *cvnD5* may or may not be regulated by cyclic di-GMP, it does play a role in development according to this study and it has not been previously analyzed. Bioinformatics data on *cvnD5* shows that it has an ATP-binding protein domain similar to *cvnD7* and other homologous genes in the conservon. This ATP-binding protein is believed to be a part of the GTPase family and therefore could play a role in ATP/GTP binding, an important factor in differentiation and development. This study observed a defective visual phenotype of the Sco5292 transposon insertion mutant grown on SFM agar plates. While there appeared to be no delay or reduction in colony grow of the mutant there was a distinct pigmentation difference. *Streptomyces* produces four antibiotics, two of which are pigmented. One produces a red pigment and the other a blue pigment. The mutant appears to lack the pink pigment. Phase-contrast and fluorescence microscopy of wild type compared to that of the Sco5292 mutant on day three of growth presented noticeably different phenotypes. Wild type *Streptomyces* had relatively uniform coloration, spore size and shape. Sco5292 appeared to have more frequent misshaped and irregularly sized spores. The Sco5292

mutant also appeared to have more lysed spore compartments. The visibly abnormal phenotype of the Sco5292 transposon insertion mutant provides evidence that *cvnD5* may have a vital role in regulation of the antibiotic producing red pigmentation and spore formation. This provides further support that genes in the conservon are important in *Streptomyces* development. Previous studies of aerial mycelium formation of *cvn9* and *cvn10* on media supplement with 2% glucose presented defective phenotypes also. Future research on *cvn5* and *cvn7* conservon mutant development in different growth conditions should be explored.

Acknowledgements

Sincere thanks to Dr. Jennifer Bennett, the Bennett lab and the Otterbein Biology and Earth Science department.

References

Bennett J, Kandell G, Kirk S, McCormick J. 2018. Visual and Microscopic Evaluation of Streptomyces Developmental Mutants. The Journal of Visualized Experiments.

Bentley, S.D., Chater, K.F., Cerdeno-Tarraga, 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.

Bourne, H.R., Sanders, D.A., and McCormick, F. (1991) The GTPase superfamily: conserved structure and molecular mechanism. Nature 349: 117–127.

Bulyha, I., Hot, E, Huntley, S., Sogaard-Andersen, L., (2011) GTPases in bacterial cell polarity and signaling, Current Opinion in Microbiology, Volume 14, Issue 6, 2011, Pages 726-733, ISSN 1369-5274, https://doi.org/10.1016/j.mib.2011.09.001. (http://www.sciencedirect.com/science/article/pii/S136952741100138X)

Bystrykh, L.V., Fernandez-Moreno, M.A., Herrema, J.K., Malpartida, F., Hopwood, D.A., Dijkhuizen, L., 1996. Production of actinorhodin-related "blue pigments" by *Streptomyces coelicolor* A3(2). J. Bacteriol. 178, 2238–2244.

Chater, K.F. (1993) Genetics of differentiation in Streptomyces. Annu Rev Microbiol 47: 685–713.

Biolabs, N. E. (n.d.). Cleavage Close to the End of DNA Fragments. Retrieved from https://www.neb.com/tools-and-resources/usage-guidelines/cleavage-close-to-the-end-of-dna-fragments

Sun N, Wang Z-B, Wu H-P, Mao X-M and Li Y-Q. (2012) Construction of over-expression shuttle vectors in Streptomyces. Annals of Microbiology Dec 2012, Volume 62, Issue 4, pp 1541-1546 10.1007/s13213-011-0408-1

Elliot, M., Damji, F., Passantino, R., Chater, K., and Leskiw, B. (1998) The bldD gene of *Streptomyces coelicolor* A3(2): a regulatory gene involved in morphogenesis and antibiotic production. J Bacteriol 180: 1549–1555.

Gordon, N.D., Ottaviano, G.L., Connell, S.E., Tobkin, G.V., Son, C.H., Shterental, S., and Gehring, A.M. 2008. Secreted-protein response to sU activity in *Streptomyces coelicolor*. J. Bacteriol. 190, 894–904.

Gras S, Chaumont, V., Fernandez, B., Carpentier, F., Charrier-Savournin, F., Schmitt, S., Pineau, C., Flament, D., Hecker, A., Forterre,P., Armengaud, J., and Housset, D. "Structural Insights into a New Homodimeric Self-Activated GTPase Family." EMBO Reports, vol. 8, no. 6, 2007, pp. 569–575., doi:10.1038/sj.embor.7400958.

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.

Komatsu, M., 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.

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ and Higgins DG ClustalW and ClustalX version 2. (2007) *Bioinformatics 2007 23(21): 2947-2948.* *doi:10.1093/bioinformatics/btm404*

Mamoru Komatsu,† Hideaki Takano,□Tomonari Hiratsuka, Yuji Ishigaki,□Kyoko Shimada, Teruhiko Beppu and Kenji Ueda, 2006. Proteins encoded by the conservon of *Streptomyces coelicolor* A3(2) comprise a membrane-associated heterocomplex that resembles eukaryotic G protein-coupled regulatory system. *Life Science Research Center, College of Bioresource Sciences, Nihon University, 1866 Kameino, Fujisawa 252-8510, Japan.*

McCormick JR, Flärdh K. Signals and regulators that govern Streptomyces development. FEMS Microbiol Rev. 2012 Jan;36(1):206-31. doi: 10.1111/j.1574-6976.2011.00317.x. Epub 2011 Dec 2. Review. PubMed PMID: 22092088; PubMed Central PMCID: PMC3285474.

Takano, H., Hosono, K., Beppu, T., and Ueda, K. (2003) Involvement of sigmaH and related sigma factors in glucose- dependent initiation of morphological and physiological development of *Streptomyces griseus*. Gene 320: 127– 135.

Takano H, Hashimoto K, Yamamoto Y, Beppu T, Ueda K. 2011. Pleiotropic effect of a null mutation in the cvn1 conservon of *Streptomyces coelicolor* A3(2). Gene. 477(1–2): Gust, B., Challis, G.L., Fowler, K., Kieser, T., 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. USA 100, 1541–1546.

Welsh, K.M., Trach, K.A., Folger, C., and Hoch, J.A. (1994) Biochemical characterization of the essential GTP-binding protein Obg of *Bacillus subtilis*. J Bacteriol 176: 7161–7168.