QUANTITATIVE EXPRESSION ANALYSIS OF THE GENES POTENTIALLY INVOLVED IN CLAVULANIC ACID OVERPRODUCTION IN *STREPTOMYCES CLAVULIGERUS*

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

CANER AKTAŞ

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
MOLECULAR BIOLOGY AND GENETICS

SEPTEMBER 2018

Approval of the thesis:

QUANTITATIVE EXPRESSION ANALYSIS OF THE GENES POTENTIALLY INVOLVED IN CLAVULANIC ACID OVERPRODUCTION IN *STREPTOMYCES CLAVULIGERUS*

submitted by **CANER AKTAŞ** in partial fulfillment of the requirements for the degree of **Master of Science in Molecular Biology and Genetics Department, Middle East Technical University** by,

Prof. Dr. Halil Kalıpçılar
Dean, Graduate School of **Natural and Applied Sciences** _____

Prof. Dr. Orhan Adalı
Head of Department, **Biological Sciences** _____

Prof. Dr. Gülay Özcengiz
Supervisor, Biological Sciences Dept., **METU** _____

Examining Committee Members:

Prof. Dr. Mustafa Akçelik
Biology Dept., Ankara University, Ankara _____

Prof. Dr. Gülay Özcengiz
Biological Sciences Dept., METU _____

Assoc. Prof. Dr. Sezer Okay
Biology Dept., Çankırı Karatekin University, Çankırı _____

Assoc. Prof. Dr. Çağdaş Devrim Son
Biological Sciences Dept., METU _____

Assoc. Prof. Dr. Can Özen
Biochemistry Dept., METU _____

Date: **21.09.2018**

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last name : Caner Aktaş

Signature :

ABSTRACT

QUANTITATIVE EXPRESSION ANALYSIS OF THE GENES POTENTIALLY INVOLVED IN CLAVULANIC ACID OVERPRODUCTION IN *STREPTOMYCES CLAVULIGERUS*

Aktaş, Caner
MSc., Department of Molecular Biology and Genetics
Supervisor : Prof. Dr. Gülay Özcengiz

September 2018, 84 pages

Streptomyces clavuligerus is the producer of the medically important β-lactam antibiotics, including cephamycin C (CC) and the potent β-lactamase inhibitor clavulanic acid. (CA). We have already undertaken an extensive comparative proteomic analysis by using both 2-DE-MALDI-MS and GeLC-MS approaches between an industrial CA overproducer, namely DEPA and the reference strain NRRL3585. In this context, we documented several differentially expressed (over- and under-represented) proteins accounting for high CA titers in the former strain to include both common and unique ones. In this work, a total of 26 proteins were selected among the proteins that belonged to different functional categories such as secondary metabolite production, regulators of secondary metabolism, primary metabolism, stress response and several unknowns, and subjected to gene expression analysis. Thus, the aim of the present study is to reveal quantitative levels of the transcripts representing the proteins selected among over- and under-represented ones of these categories.

Keywords: *Streptomyces clavuligerus*, clavulanic acid overproduction, proteomics, quantitative gene expression

ÖZ

STREPTOMYCES CLAVULİGERUS DA KLAVULANİK ASİT AŞIRI ÜRETİMİNDE GÖREVLİ OLABİLECEK GENLERİN KANTİTATİF GEN İFADESİ ANALİZİ

Aktaş, Caner
Yüksek Lisans, Moleküler Biyoloji ve Genetik
Tez Yöneticisi : Prof. Dr. Gülay Özcengiz

Eylül 2018, 84 sayfa

Streptomyces clavuligerus, sefamisin C (CC) ve güçlü bir β-laktamaz inhibitörü olan klavulanik asit (CA) gibi tıbbi açıdan önemli β-laktam antibiyotiklerin üreticisidir. Laboratuvarımızda iki ayrı *Streptomyces clavuligerus* suşu için (endüstriyel bir suş olan, CA aşırı üreticisi suş DEPA ve referans bir suş olan NRRL3585) için 2-DE-MALDI-MS ve GeLC-MS ile oldukça detaylı karşılaştırmalı proteomik analizler gerçekleştirilmiş ve farklı ifade edilen proteinler tanımlanarak. yüksek CA titresine karşılık gelen proteinler saptanmıştı. Bu proteinler sekonder metabolit üretimi, sekonder metabolizma regülatörleri, primer metabolizma, stres yanıtı ve fonksiyonu bilinmeyenler gibi farklı fonksiyonel kategorilere aittir. Mevcut araştırmanın amacı, bu kategoriler için seçilen, seviyesi artmış ya da azalmış 26 ayrı proteine karşılık gelen genlere ait transkriptlerin nicel düzeylerini belirlemektir.

Anahtar Kelimeler: *Streptomyces clavuligerus*, klavulanik asit aşırı üretimi, proteomik, nicel gen ifadesi

To My Family

ACKNOWLEDGMENTS

I would first like to thank my thesis advisor Prof. Dr. Gülay Özcengiz. The door to Prof. Özcengiz office was always open whenever I ran into a trouble spot or had a question about my research or writing. She consistently allowed this paper to be my own work but steered me in the right the direction whenever she thought I needed it.

I must express my very profound gratitude to my parents for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis. This accomplishment would not have been possible without them. Thank you.

I sincerely thanks to my fellow lab mates Ozan Ertekin, İlayda Baydemir, Meltem Kutnu, Nazlı Hilal Erdal, Samet Aytekin and Naz Kocabay whom we shared our finest hours and supported each other whenever need arose. Special thanks to my former lab mate Ayşenur Biber for good support and insight provided during thesis writing.

I acknowledge the METU Research Fund [BAP-01-08-2017-005] which allow this study to be conducted.

TABLE OF CONTENTS

ABSTRACT v

ÖZ vi

ACKNOWLEDGMENTS viii

TABLE OF CONTENTS ix

LIST OF TABLES xi

LIST OF FIGURES xii

LIST OF ABBREVIATIONS xiii

CHAPTERS

INTRODUCTION 1

1.1 Actinobacteria and the Genus *Streptomyces* 1

1.2 *Streptomyces clavuligerus* and Secondary Metabolite Production 3

1.3 β-lactam Antibiotics and Aminoglycosides 4

1.3.1 β-lactam Antibiotics 4

1.3.2 Cephamycin C 6

1.3.3 Clavulanic Acid 7

1.3.4 Aminoglycoside Antibiotics 8

1.4 Biosynthesis of Cephamycin C, Clavulanic Acid, and Aminoglycosides 8

1.4.1 Biosynthesis of Cephamycin C 8

1.4.2. Biosynthesis of Clavulanic Acid 10

1.5 Quantitative Gene Expression Analysis 11

1.5.1 Reverse Transcription Quantitative Polymerase Chain Reaction 12

1.5.2 Alternative Methods for Quantitative Gene Expression by RT-qPCR 14

1.6 Former Proteomic Findings from Our Laboratory 15

1.7 Aim of the Present Study 16

MATERIALS AND METHODS 19

2.1. Bacterial Strains 19
2.2 Storage Media and Culture Conditions 20
2.3 Biological Activity Assays 21
2.3.1 Modified Antibiotic Susceptibility Testing (AST) 21
2.3.2 Antibiotic Concentration Estimation 21
2.3.3 Construction of the Standard Curves 22
2.4 Gene Expression Experiments 23
2.4.1 Total RNA Extraction 23
2.4.2 cDNA Synthesis with Reverse Transcription PCR (RT-PCR) 23
2.4.3 Real-Time Polymerase Quantitative Chain Reaction 24
2.4.4 Relative Quantification of RT-qPCR Data with Efficiency Correction 26
2.4.5 Statistical Analysis 27
RESULTS AND DISCUSSIONS 29
3.1 A Comparison Between S. *clavuligerus* NRRL 3585 and DEPA With Respect to the Antibiotics They Do Produce 29
3.2 Quantitative Gene Expression Analysis of *S. clavuligerus* NRRL 3585 and DEPA 33
3.2.1 Optimization of RT-qPCR experiments 33
3.2.2 Quantitative Gene Expression Profile Comparisons 36
3.3 Comparison of Differential Gene Expression with Differential Protein Abundance 39
3.3.1 Negatively Correlated Gene Expression and Protein Abundance 40
3.3.2 Positively Correlated Gene Expression and Protein Abundance 43
3.3.3 No Correlation Between Gene Expression and Protein Abundance. 48
CONCLUSION 53
REFERRENCES 55
APPENDIX A 65
APPENDIX B 67

LIST OF TABLES

Table 1.1 List of Selected Significantly Differentiating Proteins.....17

Table 2.1 Bacterial Strains and their Properties.....19

Table 2.2 RT-PCR Reaction Protocol.....24

Table 2.3 Components of RT-qPCR master mix.....25

Table 2.4 Reaction protocol for RT-qPCR.....26

Table 3.1 Average Cq and Tm values of optimized PCR condition.....35

Table 3.2 Amplification efficiencies and E values of genes.....36

Table 3.3 Negatively correlated gene expression vs protein abundance.....42

Table 3.4 Positively correlated gene expression vs protein abundance.....45

Table 3.5 Not correlated gene expression vs protein abundance.....49

LIST OF FIGURES

Figure 1.1 Schematic representation of life cycle of Streptomyces species……2

Figure 1.2 Structure of the β-lactam ring, and acylamino side ring fused to β-lactam ring…4

Figure 1.3 Peptidoglycan cell wall formation in bacteria……6

Figure 1.4 Structure of clavulanic acid and representative structures of penicillins and cephalosporins…….....7

Figure 1.5 Aspartate pathway in S. clavuligerus...…...……..…….....…….....…...9

Figure 1.6 Pathway of the formation of clavulanic acid and other clavams……..11

Figure 1.7 Diagram showing one-step RT-qPCR vs two-step RT-qPCR.....13

Figure 2.1 Calibration curve.....22

Figure 3.1 Clavulanic Acid activity assay.....30

Figure 3.2 Cephamycin C activity assay.....30

Figure 3.3 Aminoglycoside activity assay.....31

Figure 3.4 Comparison of Clavulanic Acid Production of S. clavuligerus NRRL 3585 and DEPA strains.....32

Figure 3.5 Comparison of Cephamycin C production of S. clavuligerus NRRL 3585 and DEPA strains.....33

Figure 3.6 Bar graphs showing relative gene expression levels between standard and DEPA strains according to the functional categories.....39

LIST OF ABBREVIATIONS

TSB	Tryptic Soya Broth
TSA	Tryptic Soya Agar
SA	Starch Asparagine
CA	Clavulanic Acid
CC	Cephamycin C
LB	Lysogenic Broth
PCR	Polymerase Chain Reaction
qPCR	Quantitative Polymerase Chain Reaction
RT-qPCR	Real Time Quantitative Polymerase Chain Reaction
MS	Mass Spectrometry
LC	Liquid Chromatography
Mw	Molecular Weight
kDa	Kilo Daltons
GC	Guanine - Cytosine
PBP	Penicillin Binding Protein

CHAPTER 1

INTRODUCTION

1.1 Actinobacteria and the Genus *Streptomyces*

The phylum *Actinobacteria* is one of the largest taxonomical group recognized within the Bacteria domain. The phylum includes pathogens (notably, species of *Corynebacterium, Mycobacterium, Nocardia, Propionibacterium*, and *Tropheryma*), soil inhabitants (e.g., *Micromonospora* and *Streptomyces* species), plant commensals (e.g., *Frankia* spp.), and gastrointestinal commensals (*Bifidobacterium* spp.). The name actinomycete (order: *Actinomycetales*) is used as synonym for Actinobacteria, is a combination of two ancient greek word ray (aktis or aktin) and fungi (mukès) for *Actinobacteria*, the growth of which is characterized by tip extension and branching hyphae like fungus. In fact, *Actinobacteria* are considered as a transition form between bacteria and fungus though comparison is only superficial because *Actinobacteria*'s chromosome is organized into nucleoid and they are susceptible to antibacterial agents (Barka et al., 2016). *Actinobacteria* are Gram-positive filamentous bacteria with high guanine-cytosine (G-C) content in their genomes. *Actinobacteria* comprised of mainly soil dwelling microorganisms although aquatic species are also present. The soil samples are dominated by the *Streptomyces* spp, hence the genus is the largest genus identified in phylum and many are identified as organic material recyclers (Flärdh and Buttner, 2009). Life cycle of *Streptomyces* spp begins when a typical spore encounters favorable conditions, it germinates, one or

two germ tubes emerge and grow to form hyphae. These tubes grow by tip extension and branching to form a vegetative mycelium, often finding their way deep into the surrounding substrate. In response to a shift from favorable conditions (i.e. nutrient depletion), both morphological changes and secondary metabolite production are initiated. Aerial hyphae break the surface tension, escaping the vegetative mycelium, and grow into the air. The aerial hyphae are then divided, forming prespore compartments. These parts then develop thick spore walls, accumulate grey spore pigment and show other features of bacterial spores (Flärdh & Buttner, 2009). However, these spores are far less resilient to adverse conditions than the spores produced by *Bacilius* species, but they can survive for longer time in a dessicated state, indicating that they are specialized for dispersion purposes (Elliot, Buttner, & Nodwell, 2008). When these spores disperse and find another favorable condition, the life-cycle of *Streptomyces* starts again. The full life cycle is schematically represented below (Figure 1.1).

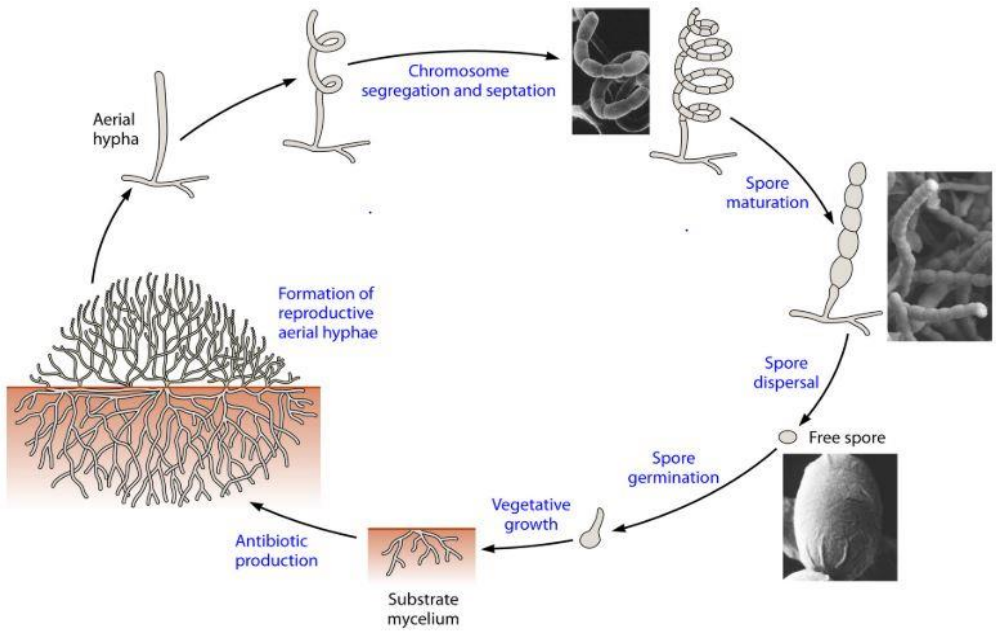


Figure 1.1. Schematic representation of life cycle of *Streptomyces* species (Barka, 2016).

The secondary metabolite synthesis and antibiotic production in *Streptomyces* coincides with aerial hyphae development (Elliot et al., 2008). As the formation of aerial hyphae is triggered by a shift to unfavorable conditions and environmental stresses that could occur in soil, where the majority of *Streptomyces* live in an immotile manner, the notorious production of those chemicals in response to such conditions is a plausible adaptation employed by this genus (Thomson et al., 2002).

1.2 *Streptomyces clavuligerus* and Secondary Metabolite Production

Microbial secondary metabolites are low-mass organic compounds synthesized by microorganisms during the idiophase (late phase) of exponential growth. They are not essential for the growth of the producer organism. Unlike primary metabolites, absence of secondary metabolites does not result in an immediate death, but serve in diverse survival functions like competitive weapons used against other bacteria, fungi, amoebae; metal transporting agents; agents of symbiosis between microbes; and as differentiation effectors (Demain et al., 2000).

Streptomyces clavuligerus was discovered and named by Higgens and Kastner (1971), clavuligerus being comprised of two ancient greek word, *clavula* (little club) and *igerus* (bearing). It was first isolated from South American soil samples and deposited in Agricultural Research Service Collection as NRRL 3585 and in the American Type Culture Collection, Rockville, Md., as ATCC 27064. *S.clavuligerus* have a linear chromosome consisting of a conserved "core" region and variable "arm" regions (Hopwood, 2006). It has a unique densely-packed 1.8-Mb linear megaplasmid pSCL4 (1,796,117 bp, 71.85% GC) which contains 25 putative secondary metabolite gene clusters, in addition to its 6.8-Mb chromosome (6,736,475 bp, 72.69% GC) which also contains 23 such clusters. The megaplasmid found in *S. clavuligerus* is by far the largest linear megaplasmid ever sequenced, and houses genes for novel bioactive metabolites (Medema et al., 2010).

S. clavuligerus has also 3 other linear plasmids pSCL1 (10,266 bp, 71.96% GC), pSCL2 (149,326 bp, 70.07% GC), pSCL3(442,792 bp, 70.77% GC) sequenced and

described by Medema et al. (2010) and (Song et al., 2010). They have also found that pSCL3 and pSCL4 has gene clusters for non-ribosomal peptide synthestases. *S. clavuligerus* is renowned for its vast array of secondary metabolite production capacity. Many of these metabolites are used in medicine, for instance, two important β-lactam antibiotics, cephamycin C and clavulanic acid (Medema et al., 2011). Cephamycin C is a β-lactam antibiotic and like cephalosporins, based on a cephem nucleus and possesses a methoxy group at the 7-alpha position (Bussari et al., 2008) . Clavulanic acid is a β-lactamase inhibitor and a weak β-lactam antibiotic (Mellado et al., 2002).

1.3 β-lactam Antibiotics and Aminoglycosides

1.3.1 β-lactam Antibiotics

β-lactam antibiotics are considered as broad-spectrum antibiotics, i.e. they can affect a number of species non-specifically. Currently, there are four major subclasses which are (i) penicillins (penams) like penicillin G, methicillin, ampicillin (ii) cephalosporins (cephems) like cephalothin; cefamandole, cefotaxime (iii) carbapenems; and (iv) monobactams. All members of β-lactam antibiotics contain beta-lactam ring, a four-membered heterolytic ring, fused with second ring, namely acylamino side ring which is variable and determining the type of the β-lactam antibiotic (Figure 1.2). Penam/penem antibiotics have five membered secondary ring, while cepham/cephem antibiotics have six membered ring (MacKenzie, 2007).

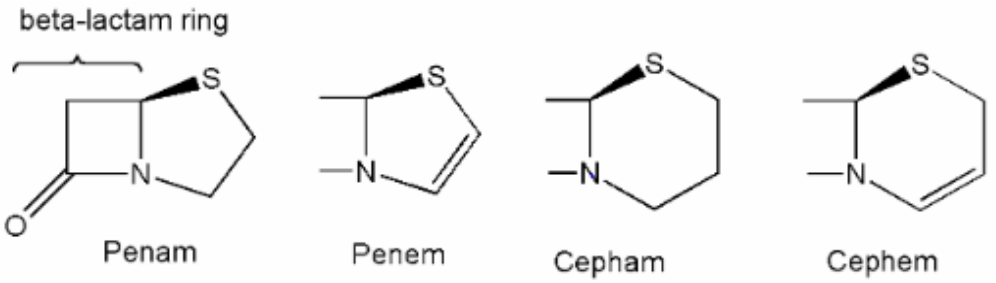


Figure 1.2. Structure of the β-lactam ring, and acylamino side ring fused to β-lactam ring (MacKenzie, 2007).

β-lactam antibiotics could only be useful on growing cultures of bacteria as their mechanism of action is the interference with bacterial cell wall synthesis during cell division.

Biosynthesis of the cell wall involves three distinct stages occurring at three distinct sites in bacterial cells (Figure 1.3). First stage is biosynthesis of pentapeptide named UDP-N-acetylmuramyl (UDP-MurNAc) and UDP-N-acetylglucosamine (UDP-GlcNAc) precursors, at the cytoplasm. Second stage is linkage of UDP-MurNAc of to the undecaprenyl pyrophosphate, a transport lipid, and the formation of lipid I. Then lipid I is converted to lipid II by the addition of UDP-GlcNAc at the plasma membrane. The modified lipid II is next transported to extracellular space and transpeptidation and transglycosilation reactions are carried out by the penicillin-binding proteins (PBPs) and transglycosylases which allow the sugar moiety carried by lipid II to be incorporated to the newly synthesized peptidoglycan chain at the immediate vicinity of plasma membrane, which then forms peptidoglycan cell wall. This is the moment at which β-lactam ring can exert its effect as it structurally resembles the D-alanyl-D-alanine moieties within the peptidoglycans and rendering PBPs by binding it, which than stops peptidoglycan chain formation. Traditionally it is believed that a leakage of cellular content (bacteriolysis) is the cause of cell death (MacKenzie, 2007; MacKenzie et al., 2007), but it was shown to be more complex since the peptidoglycan hydrolases, autolysin enzymes do also take part in the hydrolysis of the peptidoglycan chain after the PBPs are blocked (Michael A. Kohanski, 2010)(Pinho et al., 2013).

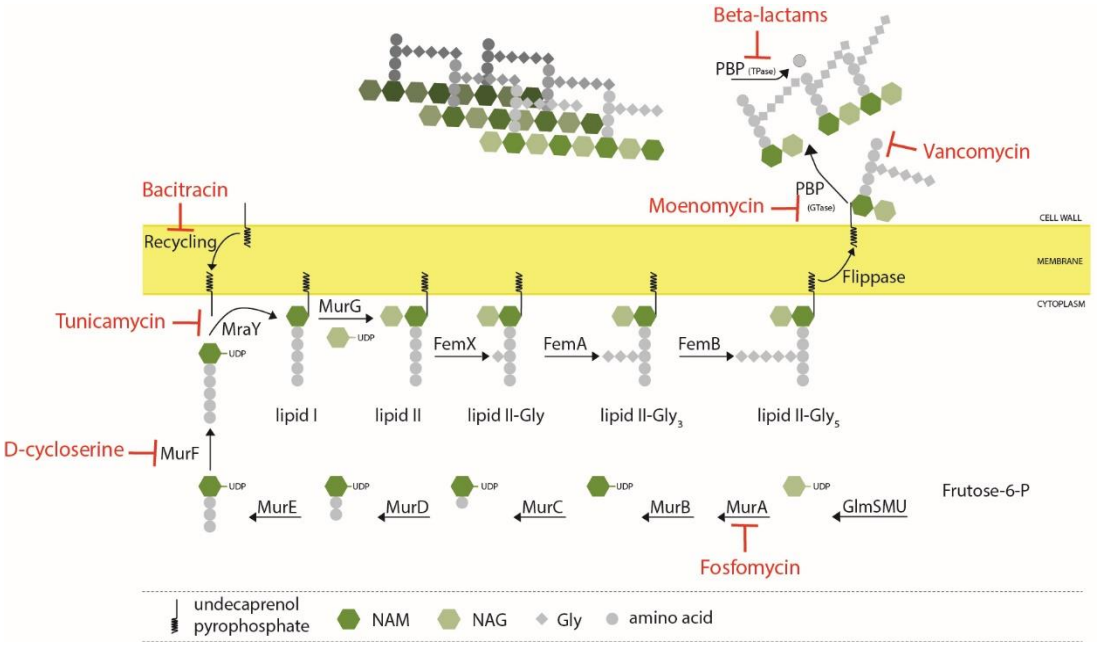


Figure 1.3. Peptidoglycan cell wall formation in bacteria.(ITQB, 2018; Pinho et al., 2013; Scheffers et al., 2005).

1.3.2 Cephamycin C

Cephamycins are composed of β-lactam ring fusion to six-carbon cephem ring. Despite their resemblance to cephalosporins, they have higher resistance to bacterial degradation caused by β-lactamase. There are three cephamycins, cephamycin A and B classified as broad-spectrum antibiotics and cephamycin C which is particularly effective against Gram-negative bacteria. Although it is considered as a β-lactam antibiotic, it is relatively ineffective against gram-positive bacteria (Daoust et al., 1973). The resistance against β-lactamases comes from the methoxy substitution at the C-7 position (Figure 1.5) (MacKenzie, 2007; MacKenzie et al., 2007). Cephamycin C is an extracellular broad spectrum β-lactam antibiotic produced by *S. clavuligerus*, *S. cattleya* and *Nocardia lactamdurans*. It is not used in clinical applications directly, instead its semisynthetic derivatives like cephalothin; cefamandole, cefotaxime are used which have greater resistance against class A β-lactamases (Bussari et al., 2008).

1.3.3 Clavulanic Acid

Clavulanic acid (CA), as a novel clavam metabolite, is an β-lactamase inhibitor isolated first from *S. clavuligerus* ATCC 27064 (NRRL3585). In an investigation on secondary metabolite production by a culture of this strain, a potent beta lactamase inhibition was observed, and the compound responsible for the activity was named as clavulanic acid, as derived from the name of its producer (Reading et al., 1977). It is also a broad spectrum β-lactam antibiotic, effective against Gram-positive bacteria as well as Gram-negative bacteria but shows weak antibacterial activity. CA with its 3R, 5R stereo-chemistry is the only one among the clavam metabolites in showing the β-lactamase inhibitory activity. All the others have the 3S, 5S stereochemistry and show no β-lactamase inhibition, although some have antibacterial or antifungal properties. This compound was found to be an analog of the basic penicillin structure (Figure 1.4). In CA molecule, an oxygen atom is substituted for sulphur, a characteristic of penicillins. As such, CA derivatives all contain an oxazolidine ring structure. The combined effective inhibition of β-lactamase and antibacterial activity of CA make it very important both clinically and economically (Bussari et al., 2008; Saudagar et al., 2008).

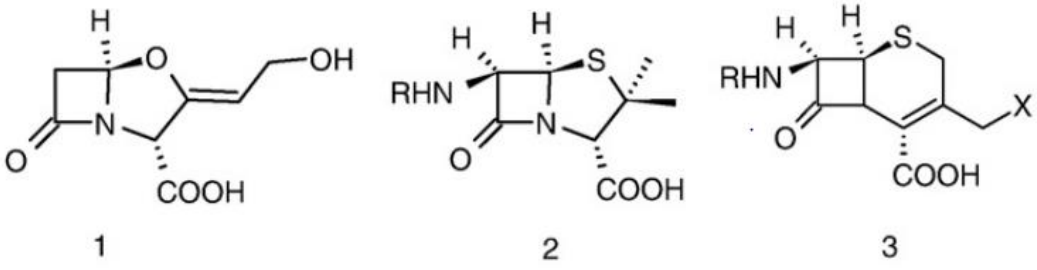


Figure 1.4. Structure of clavulanic acid (1), and representative structures of penicillins (2) and cephalosporins (3).

It acts as a pseudosubstrate, occupying the active site of β-lactamases for sufficiently long time to prevent the degradation of co-administrated β-lactam antibiotics. At higher concentrations, CA binds irreversibly (Fisher et al., 1978).

1.3.4 Aminoglycoside Antibiotics

Aminoglycosides are made up of amino-modified glycoside (sugar) molecules and used as antibacterial agents that exert their effects by inhibiting protein synthesis. They are effective against gram-negative aerobe bacteria and some species of the anaerobic species of the genus *Bacillus*, but not effective against gram-positives and anaerobic gram-negatives. They require only short contact time, and are most effective against susceptible bacterial populations that have short doubling times. Aminoglycosides are internalized by active transport. Protein synthesis inhibition is achieved through energy-dependent irreversible binding of aminoglycosides to bacterial ribosomes and distruption of protein elongation at 30S ribosomal subunit (Mingeot-Leclercq et al., 1999). Aminoglycosides also cause disruption of functional integrity of bacterial plasma membrane. The genus *Streptomyces* have reported to produce aminoglycoside antibiotic namely streptomycin (Hong et al., 2007).

1.4 Biosynthesis of Cephamycin C, Clavulanic Acid, and Aminoglycosides

1.4.1 Biosynthesis of Cephamycin C

L-α-aminoadipic acid (L-α-AAA) is one of the precursor for the cephalosporin-type molecules produced by actinomycetes (Whitney et al., 1972). It is a non-proteogenic aminoacid that does not contribute to the polypeptide chain during protein synthesis. L-valine (L-Val) and L-cysteine (L-Cys), common aminoacids, are also necessary for the for the biosynthesis of cephalosporins (Liras et al., 2006). L-α-aminoadipic acid (L-α-AAA) and L-lysine (L-Lys) are produced through aspartate pathway. The primary metabolite precursor aspartate molecule is converted into the β-aspartyl phosphate by aspatokinase enzyme. Than aspartate semialdehyde dehydrogenase enzyme then converts β-aspartyl phosphate to aspartate β-semialdehyde. At this point, aspartate pathway branches into two.

First branch is catalyzed by homoserine dehydrogenase to produce homoserine from aspartate β-semialdehyde, which is a precursor for threonine and methionine. Second branch leads to the diaminopimelate production which is a precursor for both L-lysine and cell wall (Figure 1.5). L-lysine is converted into α-aminoadipic acid

semialdehyde by the lysine-6-aminotransferase (LAT) enzyme (Coque et al., 1991) after which α-aminoadipic acid semialdehyde is spontaneously converted into piperidine-6-carboxylic acid (P6C), which is a cyclic form of same molecule. Finally, piperideine-6-carboxylate dehydrogenase (P6C-DH) enzyme converts P6C to L-α-AAA (Liras and Martin, 2006; Ozcengiz et al., 2013).

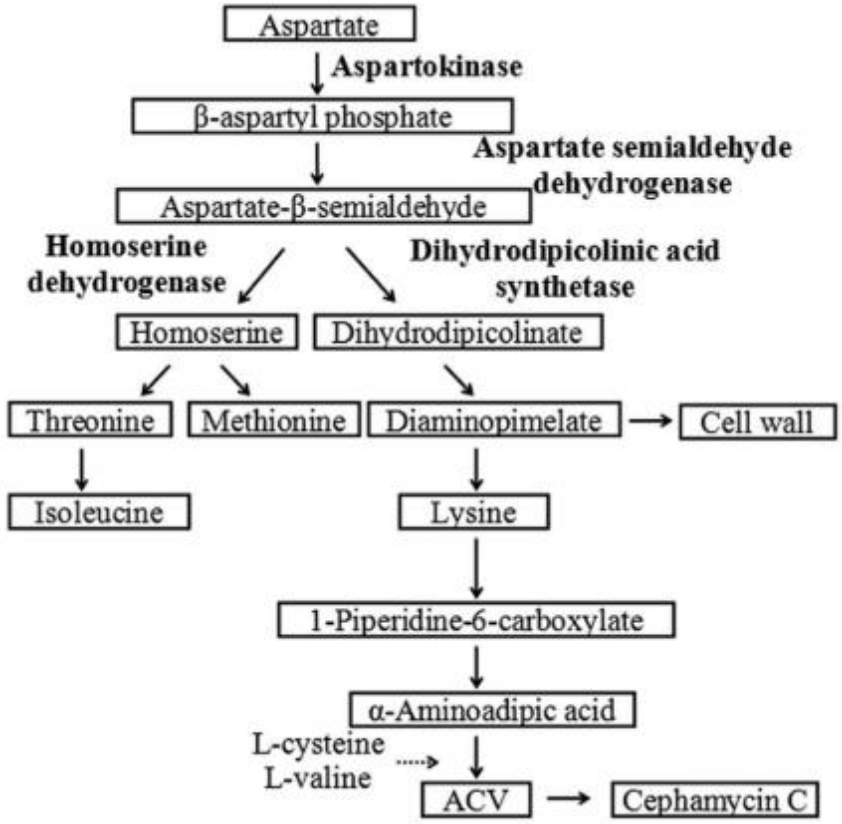


Figure 1.5. Aspartate pathway in *S.clavuligerus* (Ozcengiz et al., 2010)

L-α-AAA is the precursor of L-δ(α-aminoadipyl)-L-cysteinyl-D-valine (ACV) together with L-Val and L-Cys. The reaction is catalyzed by the ACVS, a non-ribosomal peptide synthesase (Liras and Martin, 2006).

ACV is then reduced and converted into isopenicillin N by isopenicillin N sythesase (IPNS) (Liras and Martin, 2006; Özcengiz and Demain, 2013). The two genes, *pcbAB* and *pcbC*, coding for ACVS and IPNS respectively, located next to each other in cephamycin C gene cluster which suggests a co-evolutionary process (Hamed et al., 2013). Isopenicillin N is then converted to penicillin N by isopenicillin N

epimerase (Figure 1.6), which is the precursor of deacetoxycephalosporin C (DAOC). DAOC is than converted into deacetylcephaloporin C (DAC) which is the precursor of cephalosporin C and DAC is converted into O-carbamoyl-DAC which is the precursor of cephamycin C (Liras and Martin, 2006).

1.4.2. Biosynthesis of Clavulanic Acid

S. clavuligerus synthesizes clavulanic acid (CA) from glyceraldehyde-3-phosphate (G3P), a sugar, and L-arginine (L-Arg), an amino acid (Bachmann et al., 1998; Reading and Cole, 1977). Alanylclavam, clavam-2-carboxylate, 2-formyloxym-ethylclavam, and 2-hydroxymethylclavam, are the other 5S clavams produced by *Streptomyces* spp. Genomic studies also revealed that *Streptomyces* spp. other than *S. clavuligerus* possesses necessary genes and might be able to produce CA, but only *S. clavuligerus* is able to produce both CA and other 5S clavams (Jensen, 2012). The clavulanic acid synthesis pathway starts with the production of N^2 –(2-carboxyethyl) arginine (CEA) by the carboxyethylarginine synthase. The enzyme condensates the L-Arg and G3P into CEA. The second step in CA biosynthesis is conversion of N 2-(2-carboxyethyl) arginine to deoxyguanidino proclavaminate (DGPC) which is catalyzed by β-lactam synthetase (Bls1/Bls2). This step introduces a β-lactam ring into its structure (Bachmann et al., 1998). Clavaminic acid synthase (CAS) enzyme then oxidizes DGPC to guanidino proclavaminic acid (GPC). GPC is processed by proclavaminic acid amidino hydrolase (PAH) which only catalyzes the guanidino proclavaminic acid GPC to proclavaminic acid (PCA). The CAS enzyme remarkably catalyzes the remaining two steps required for the production of clavaminic acid (Hamed et al., 2013; Jensen, 2012). The pathway then branches into two, one branch

leads to the production of clavulanic acid by the conversion of clavaminic acid to N-glycylclavaminic acid by the activity of glycylclavaminic acid synthase (GCAS) enzyme (Arulanantham et al., 2006). Conversion of N-glycylclavaminic acid to clavaldehyde is not yet fully understood, and inversion of 3S,5S stereochemistry of N-glycylclavaminic acid into 3R,5R configuration of clavaldehyde remained to be

unknown (Jensen, 2012). Clavaminic acid is also the precursor of other clavams that have 3S, 5S stereochemistry (Figure 1.6).

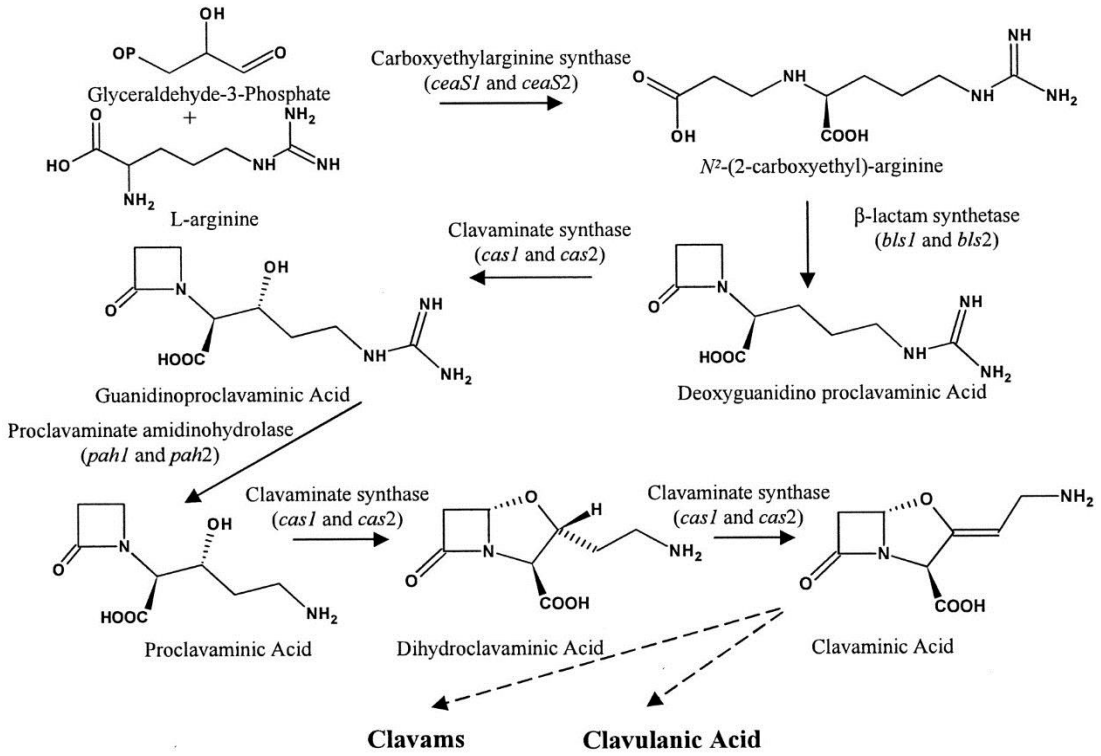


Figure 1.6. Pathway for the formation of clavulanic acid and other clavams (Tahlan et al., 2004).

1.5 Quantitative Gene Expression Analysis

Gene expression is the fundamental process by which the genotype gives rise to phenotype. The information stored in gene is transcribed and translated into functional products, i.e. RNA and protein. The transcription is a process that synthesize a mRNA molecule from a DNA template. The expression level of gene is the level of mRNA, synthesized in cell (Brueckner et al., 2009). For quantification mRNA levels, many methods have been developed over years. One of the first employed method is Northern blotting taking its name from its similarity to the first blotting technique, the Southern blot, named for the biologist Edwin Southern

(Alberts, 2008; Alwine et al., 1977). The term 'Northern blot' refers to the capillary transfer of RNA from the electrophoresis gel to the blotting membrane. Requirement of high concentration of RNA and lack of actual quantification are two major drawbacks in expression analysis. Another method, which is widely used nowadays, is RT-qPCR. This technique involves cDNA synthesis by reverse transcription and amplification of cDNA by quantitative PCR. In quantitative PCR, amplification of cDNA is tracked by fluorescence signal either generated by hybridization probe or intercalating dye, as amplification proceeds (Freeman et al., 1999; Radonic et al., 2004). With a standard curve constructed with serially diluted known concentrations of DNA, qPCR can provide an absolute measurement of the cDNA. These two methods are viable for low throughput studies. High throughput studies, like gene expression profilling where thousands of genes studied, require more sample to be processed in a short time. Transcriptomic technologies like DNA microarray is utilized to generate large data. DNA microarray or DNA chips are series of modified oligonucleotides attached to a solid surface, that are capable of hybridization with sample and emmits fluorescence afterwards. Drawbacks for this method are (i) measure the relative activity of previously identified target genes, (ii) generation of background noise. More reliable method, whole transcriptome shotgun sequencing or in short RNA sequencing (RNAseq) uses next-generation sequencing (NGS) tools to reveal the precence and quantity of the RNA in sample (Chu et al., 2012; Wang et al., 2009).

1.5.1 Reverse Transcription Quantitative Polymerase Chain Reaction

The quantitative real time polymerase chain reaction (RT-qPCR) has become method of choice for quantification of mRNA, described as gold standard and remains the most sensitive method of choice (Bustin, 2000). However, it suffers lack of standardization in choice of starting material, i.e. total RNA or mRNA as well as the efficiency of reverse transcription. The RT-qPCR assay can be performed either as a one-tube which combine reverse transcription and PCR in a single tube and buffer, using a reverse transcriptase along with a DNA polymerase known as “one step RT-

qPCR" or the "two step RT-qPCR" where the reverse transcription and PCR steps are performed in separate tubes, with different optimized buffers, reaction conditions, and priming strategies (Figure 1.7). The choices of reverse transcription enzyme and cDNA priming strategy do affect the quality of results obtained from qPCR (Bustin et al., 2004). Both one-step and two-step RT-qPCR have inherent advantages and disadvantages and should be utilized according to a specific scenario. Advantages of one-step are less experimental variations, minimal risk of contamination due to pipetting, and reproducibility of results. But since the priming is done with specific primers, a further study with the sample is impossible. Also, it is impossible to optimize each of the two-step reaction. The main advantage to two-step RT-PCR is reverse transcription is done in a separate tube, primed with random hexamers and oligo(T)s, therefore nearly all mRNA can be used for generation of cDNA. This allows the utility of the same sample in further studies. Other advantages include individual optimization of each step and high sensitivity even when the quantity of RNA sample is low. Major disadvantage of two-step RT-qPCR is time consumption (Wacker et al., 2005).

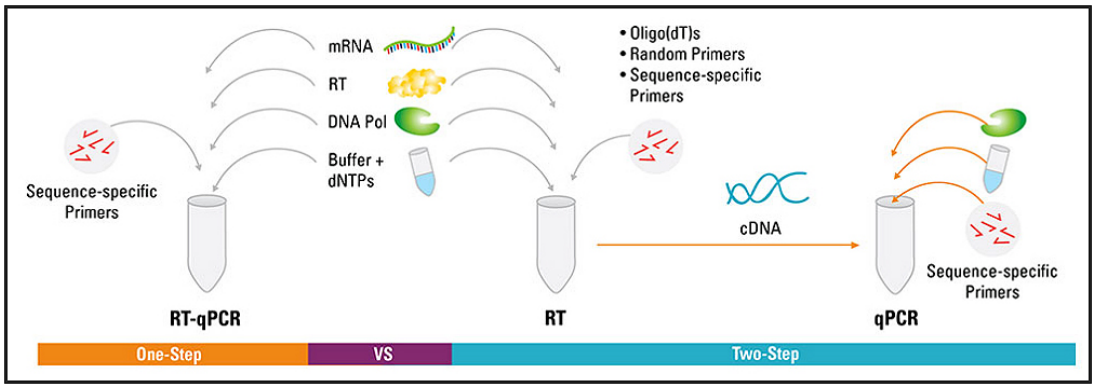


Figure 1.7. Diagram showing one-step RT-qPCR vs two-step RT-qPCR (Bustin, 2004).

1.5.2 Alternative Methods for Quantitative Gene Expression by RT-qPCR

The thermal cycler used in RT-qPCR reactions calculates two values during amplification cycle: The threshold level of detection at which a reaction reaches fluorescence intensity above background and the Quantification cycle (Cq), the amplification cycle at which sample reaches to the threshold. The Cq value used in quantitative gene expression studies (Huggett et al., 2014). The quantification cycle is represented also as Cp and Ct in the current literature, while all of them referring to the same entity. The Cq is preferred according to the MIQE guidelines published in 2009 (Bustin et al., 2009). Numerical data acquired from thermal cycler is processed to quantify expression levels of designated targets (genes) among different biological samples. Data is expressed as fold change or fold difference of gene expression levels between samples as it is more relevant than stating the copy number of each gene (Livak et al., 2001). There are three common methods to analyze the RT-qPCR gene expression experiments. Each one has different assumptions and to be used after series of experiments to show if those assumptions are validated. First one is the comparative Cq (ΔΔCq) method for relative quantification. It can be used to determine relative expression changes between experimental samples and reference sample. However, this method assumes that the amplification efficiencies of targets (genes) and internal control (gene) is similar Therefore, in order to use ΔΔCq method, a series of validation experiments should show that the amplification efficiencies of the target and the internal control gene(s) are approximately equal (Schmittgen et al., 2008).

The second method is relative standard curve method which is similar to the comparative Cq method. It can also be used to determine the relative expression differences between experimental and reference samples. It is useful when relative expression levels are questioned but the assumption in comparative Cq method is not met, in which the amplification efficiencies between target and internal control genes are not approximately equal. In order to use this method; a series of dilutions for target genes and internal control genes are created and used in RT-qPCR to calculate the amplification efficiencies of each gene. After amplification efficiencies are

calculated a new mathematical model is introduced to calculate relative expression levels between genes that have different amplification efficiencies. The amplification efficiency is represented as E, which corresponds to the total number of copies produced after each amplification reaction. E value is equal to 2 if the calculated efficiencies are 100% and after each amplification cycle DNA concentration in the reaction vessel doubles. The relative expression ratio (R) of a target gene is calculated based on E and Cq deviation between sample and control groups in comparison to an internal control gene. The method is also known as the efficiency correction method and first introduced by Michael W. Pfaffl from Technical University of Munich Institute of Physiology (Pfaffl, 2001).

The third method is Standard Curve (SC) method, which is used to calculate the absolute quantity of target gene in each sample. To calculate the quantity, the real-time PCR detection software measures the amplification of each target and the standard dilution series with known concentrations. The simple SC method, similar to the ΔΔCq method, assumes that both the standard and target amplification efficiencies are approximately equal. But this time quantification gives exact copy number for each target rather than fold difference of target copy number between two experimental group (Larionov et al., 2005). It is the least used method among all three as it requires series of dilutions for each target gene to calculate E value and a dilution of DNA standard with a known concentration that is calculated by other means (Huggett and O'Grady, 2014).

For absolute quantification, a new method called digital PCR is now increasing its popularity as the PCR kinetics are better understood and technology advances (Bustin et al., 2013).

1.6 Former Proteomic Findings from Our Laboratory

In the previous studies in our laboratory, the standard strain of *S. clavuligerus*, namely NRRL 3885 and an industrial strain of *S. clavuligerus*, namely DEPA, which was developed through successive random mutagenesis and selection program by the DEPA Pharmaceuticals Co. (İzmit, Turkey) were subjected to proteomic analyses in

order to reveal their differential protein expression profiles. The iterative cycles of random mutagenesis introduced to the industrial strain DEPA have caused random damage to the DNA through DNA strand breakage, recombination, addition, deletion or substitution of bases (Unsaldı, 2016). *S. clavuligerus* DEPA used for CA manufacturing process in Turkey produces at least 100-fold more CA relative to the wild type *S. clavuligerus* (Unsaldi et al., 2017).

Two different approaches, non-gel based LC-MS/MS and 2DE gel coupled to MALDI-TOF/MS, were exploited to acquire comprehensive understanding of global protein change profiles of these stains during the stationary phase during which a morphological change and secondary metabolite production prevail. By this study, 13 protein identified with 2DE gel with MALDI-TOF/MS and 13 protein identified with LC_MS/MS a total of 26 proteins were found to be differentially expressed between two strains. These proteins belonged to different functional categories like primary metabolism, secondary metabolism, hypothetical, stress related, DNA replication, repair, transcription etc.(Unsaldı, 2016).

1.7 Aim of the Present Study

The present study was based on above-mentioned proteomics study in that the transcript levels of the genes corresponding to the selected proteins among differentially expressed ones identified from *S. clavuligerus* DEPA were quantitatively determined. (Table 1.1). Quantitative gene expression analysis with RT-qPCR was the method employed for this purpose. We aimed at revealing the correlation between mRNA and protein levels by measuring gene expression levels exactly at the time of differential protein representation between the two samples (standard strain versus industrial strain), during the stationary phase when the secondary metabolite production and morphological differentiation are the most prominent in *Streptomyces* spp.

Table 1.1 The list of the proteins selected among over- and underrepresented proteins identified in the proteomics study performed by Ünsaldı (2016).

Fuction	Uniprot Accession Number	Protein Name	Method	Relative Pro. Ab.*	Related Gene
DNA related	B5GPP5 (a)	Anti-sigma factor	2-D gel	7.30	SCLAV_2541
DNA related	B5GPP5 (b)	Anti-sigma factor	2-D gel	3.30	SCLAV_2541
DNA related	E2QA55	LexA Repressor	LC-MS/MS	2.78	lexA
DNA related	E2PZ58	TetR-family transcriptional regulator	LC-MS/MS	-2.33	SCLAV_5246
hyp/trans/stress	E2Q5C6	Protein translocase subunit SecA	2-D gel	24.10	secA
hyp/trans/stress	E2PWW5(a)	Putative M28-family peptidase	2-D gel	-2.70	SCLAV_3034
hyp/trans/stress	E2PWQ8	DUF2587 domain-containing protein	2-D gel	-2.80	SCLAV_2986
hyp/trans/stress	E2PWW5(b)	Putative M28-family peptidase	2-D gel	-6.10	SCLAV_3034
hyp/trans/stress	E2Q3P1	Trigger factor	2-D gel	-37.00	Tig
Other Function	B5H160	Siderophore-interacting protein	2-D gel	4.70	SCLAV_0843
Other Function	B5H0U2	DNA-binding protein	LC-MS/MS.	2.78	SCLAV_0130
Other Function	D5SJT7	Bacterial luciferase domain-containing protein	LC-MS/MS	-3.72	SCLAV_p0693
Other Function	B5H4A8	Rhs element Vgr protein	2-D gel	3.80	SCLAV_0043
Other Function	B5GZC6	Amidohydrolase	LC-MS/MS.	-2.02	SCLAV_5667
Primary Metabolism	D5SLL7	Methionine synthase II	LC-MS/MS	-3.72	SCLAV_p1324

Fuction	Uniprot Accession Number	Protein Name	Method	Relative Pro. Ab.*	Related Gene
Primary Metabolism	B5GZV9	E1-alpha branched-chain alpha keto acid dehydrogenase	LC-MS/MS	-2.02	bkdA1
Primary Metabolism	E2PWQ0	Branched-chain alpha keto acid dehydrogenase E1 beta subunit	LC-MS/MS	-2.58	bkdB1
Primary Metabolism	B5GL42	UDP-glucose 4-epimerase (carbonhydrate met)	2-D gel	-3.80	SCLAV_4282
Primary Metabolism	B5H0X6	Enoyl-[acyl-carrier-protein] reductase [NADH]	2-D gel	-4.50	fab1
Primary Metabolism	B5GSL5 (b)	4-hydroxyphenylpyruvate dioxygenase	2-D gel	-4.90	SCLAV_2046
Primary Metabolism	B5GSL5 (a)	4-hydroxyphenylpyruvate dioxygenase	2-D gel	-35.70	SCLAV_2046
Secondary Metabolism	E2PUT1	Polyprenyl synthetase	LC-MS/MS	2.28	SCLAV_4756
Secondary Metabolism	E2PZ87	Thioredoxin reductase	LC-MS/MS	3.90	SCLAV_5275
Secondary Metabolism	B5GLB4	7-alpha-cephem-methoxylase P8 chain	LC-MS/MS	2.91	cmcJ
Secondary Metabolism	E2Q9C7	Gamma-butyrolactone biosynthesis protein	LC-MS/MS	-3.36	avaA2
Secondary Metabolism	E2Q9A9	Polyketide synthase	LC-MS/MS	-4.81	SCLAV_0453

*Relative protein abundance in folds

CHAPTER 2

MATERIALS AND METHODS

2.1. Bacterial Strains

S. clavuligerus strains and other bacterial species that were used in this study are given in Table 2.1.

Table 2.1. Bacterial strains and their properties

Strains	**Description**	**Source or Reference**
S. clavuligerus NRRL 3585	Wild type, cephamycin, clavulanic acid and clavam producer	Prof. J. Piret, Northeastern University, USA
S. clavuligerus DEPA	Industrial; clavulanic acid overproducer	DEPA Pharmaceuticals Co., İzmit, Turkey
Escherichia coli ESS 2235	β-lactam supersensitive strain; indicator organism for β-lactam activity in bioassays	Prof. J. Piret, Northeastern University, USA
Staphylococcus epidermidis ATCC 14990	Susceptible to aminoglycosides; indicator organism for kanamycin activity in bioassays	Prof. M. Akçelik,, Ankara University, Turkey
Klebsiella pneumoniae ATTC 29665	Resistant to Penicillin G. indicator organism for clavulanic acid activity in bioassays	Prof. P. Liras, INBIOTEC, Leon, Spain

DEPA has been acquired through the induction of iterative random mutations and selection at the laboratories of the DEPA Pharmaceuticals Company. *E. coli* ESS 2235 and *K. pneumoniae* ATTC 29665 and *S. epidermidis* ATTC 14990 are commonly used in antimicrobial activity testing.

2.2 Storage Media and Culture Conditions

E. coli strains were grown in Lysogeny Broth (LB) or on agar plates (LA) in a rotary shaker (200 rpm) at 37^{o} C (APPENDIX A). *S. epidermidis* and *K. pneumoniae* strains were grown in Tryptic Soy Broth (TSB) (APPENDIX A) in a rotary shaker (200 rpm) at 37^{o} C. All three strains were kept at 4^{o} C for 3 weeks for short term storage either in agar plates or as a suspension culture. For long term storage, 2 ml of suspension culture is centrifuged at 11000 g for 10 minutes, supernatants discarded and the pellets were resuspended in 1 ml of 50% glycerol solution (v/v) in a 2 ml microcentrifuge tube. *S. clavuligerus* mycellium stocks were stored in Ultra Low Temperature (ULT) freezer at -80^{0} C. Mycelium stocks were prepared by taking 1 ml of suspension culture into 2 ml of microcentrifuge tube and mixing with 1 ml of 50% glycerol solution (v/v) without any centrifugation. Seed cultures were prepared by adding 200-400 µl of frozen mycelium stock into 50 ml of TSB in 250 ml total volume of baffled flasks and incubated at 28^{o} C at 220 rpm for 24-72 hours. Bacterial growth was measured by optical density (OD) measurement of liquid culture. Measurement method was adopted from Malmberg et al. (1993). 6:1:1 volume ratio of dH_2O : HCl (2,5M) : Seed culture. Samples were mixed in a glass tube and the mixture is subjected to ultrasonication (Ultrasonic Processor, Cole Palmer) for 3 x 30 seconds at 50% amplitude. Then the OD measured at 600nm (OD_{600}). Fermentation cultures were prepared when the OD_{600} value of seed culture reached 0.7 - 0.8, a 30 ml of the sample was precipiated by centrifugation at 4000 g for 10 min at 4^{o} C. Supernatants were discarded and the pellets were washed by fresh Starch-Asparagine (SA) medium (APPENDIX A) and resuspended in 100 ml of SA medium in 500 ml baffled flask for fermentation process for 96 hours.

2.3 Biological Activity Assays

Both wild-type NRRL 3585 and industrial overproducer DEPA were incubated for 96 hours at 28^{o} C at 220 rpm in SA medium. Supernatants are collected at 0, 24, 48, 72, and 96 hour points during batch growth..

2.3.1 Modified Antibiotic Susceptibility Testing (AST)

The supernatants of the samples kept at -80 ^{o}C were thawed on the ice on the day of bioassay. Sterile glass-made petri plates, 17 cm in diameter, were used for the bioassay experiments. A 100 ml of Tryptic Soy Agar (TSA) (%2 agar) per petri plate was cooled to 45 – 47 ^{o}C before the addition of indicator organisms. Indicator organisms were grown in 50 ml TSB up to an OD_{600} value of 1.0 and 3.3 ml of these cultures was added per 100 ml of TSA. For cephalosporin and aminoglycosidase activity measurements, *E. coli* ESS 3235 and *S. epidermidis* ATCC 14990 were the bacterial strains used as the indicator organisms, respectively. For calvulanic acid activity measurements, *K. pneumoniae* was used and penicillin G of 100 µl (10 µl/mg final concentration) was added per 100 ml of TSA to which *K. penumoniae* added. TSA and culture mixtures were left for solidification and 5 mm diameter holes were made on petri dishes. These holes were filled with 60 µl of supernatants taken from *S.clavuligerus* samples. Plates were kept at refrigerator at 4 ^{o}C for 2 hours and incubated at 30 ^{o}C for 12 – 15 hours. After incubation, the diameter of inhibition zones are measured with a ruller (in mm).

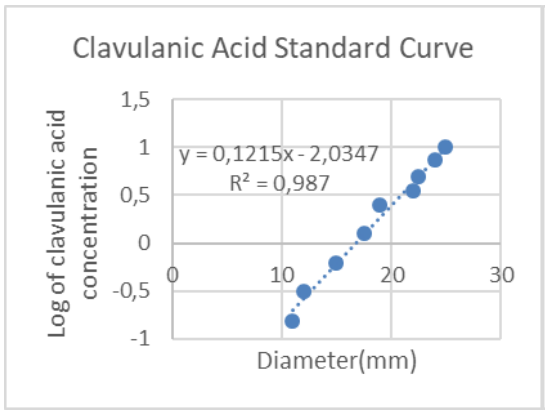
2.3.2 Antibiotic Concentration Estimation

Cephamycin C, clavulanic acid, and aminoglycoside concentrations in the supernatants were estimated using the respective commercial antibiotics as the standards and calculated according to linear regression equations obtained from standard curves. 100 mg/ml of potassium clavulanate, cephalosporin C, and kanamycin stock solutions were used for serial dilutions to generate the standard curves Potassium clavulanate was dissolved and diluted with 1 M MOPS (pH=6.8), cephalosporin C was diluted with 1 M NaOH, and kanamycin was diluted with

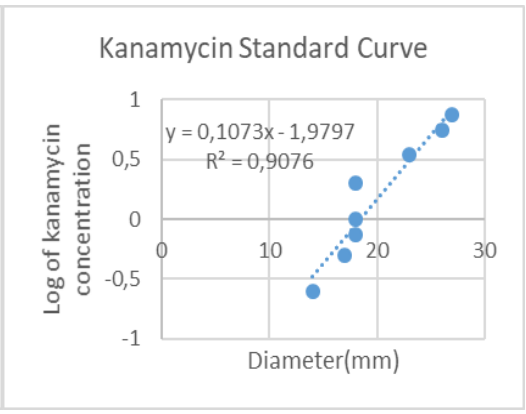
distilled water. 60 µl of the respective dilutions were poured into the petri dishes containing the indicator strains, as mentioned above.

2.3.3 Construction of the Standard Curves

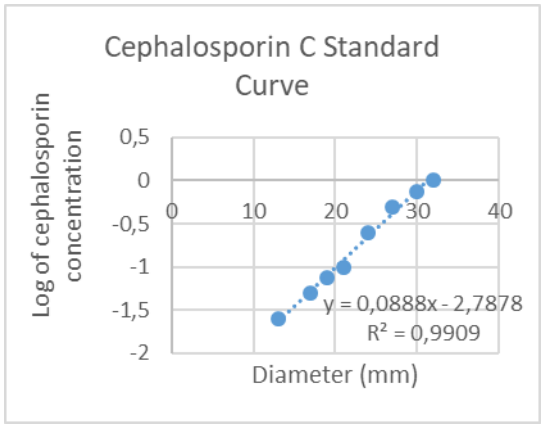
Serial dilutions of the commercial antibiotics to known concentrations were used for standart curve construction. Calibration curves were as follows.



A



B



C

Figure 2.1. Calibration curve construction. For each at least three independent runs were performed on 17cm glass plates containing Tyriptic Soy Agar. Data shown here represent the average of these triplicates. (A) clavulanic acid standard curve, (B) kanamycin standard curve, and, (C) cephalosporin C standard curve

2.4 Gene Expression Experiments

For gene expression analysis relying on quantification of mRNA transcripts at a given time, absolute and relative quantification methods are available, as mentioned in section 1. In this study, the relative quantification method was used. The method starts by obtaining high quality total RNA samples and followed by, RT-PCR measurement and statistical analysis.

2.4.1 Total RNA Extraction

Total RNA of the *S. clavuligerus* samples were extracted from pellets collected at 48 hours of fermentation in SA medium. Sample pellets were incubated in 37^{o} C for 10 minutes in TE buffer with lysozyme (APPENDİX A) for homogenization. For Total RNA extraction, Macherey-Nagel NucleoSpin® RNA extraction kit was used and the suggested protocol were optimized due to low yield and quality of RNA samples. These modifications for the recommended procedure were as follows: (i) 450 µl of buffer RA1 is used instead of suggested 200 µl, (ii) to adjust RNA binding condition 450 µl 70% Ethanol (diluted with DEPC threated water) is used instead of 200 µl. Quality and quantity of RNA samples are measured by UV spectrometer (BioDrop, UK). For this measurement, 1 µl of RNA sample is used. An A_{260}/A_{280} ratio in a range of 1.8-2.0 ± 0.1 was preferred. A ratio of$1.8 < A_{260}/A_{280} < 2.2$ gave RNA samples that are free from contaminants like salt, phenol, carbohydrates, peptides, or aromatic compounds, thereof referring to a pure RNA. Integrity of RNA samples were detected by running the samples on native agarose gels. Invisible bands, and a smear indicated a high chance of degradation of total RNA which is unreliable to start gene expression analysis. Two sharp bands which represents 23S and 16S rRNA is preferential for downstream analysis. RNA samples were stored in -80^{o} C ULT Freezer.

2.4.2 cDNA Synthesis with Reverse Transcription PCR (RT-PCR)

For cDNA synthesis from total RNA samples, Bio-Rad iScript^{TM} cDNA Synthesis kit was used. This kit includes three tubes, which contain all the reagents required for successful reverse transcription.

First tube, 5X iScript™ Reaction Mix, contains salts required for reverse transcriptase, also blend of oligo(dT) and random hexamer primers. Second tube, iScript™ reverse transcriptase contains a modified Moloney murine leukemia virus (MMLV) reverse transcriptase, optimized for reliable cDNA synthesis over a wide dynamic range of input RNA. The enzyme is provided in a form mixed with an RNase inhibitor. Third tube contains nuclease-free water. Reaction protocol for successful cDNA synthesis is shown below (Table 2.2). Concentrations of synthesized cDNA were measured by UV Spectrometer (BioDrop, UK) and stored at a 4 °C refrigerator.

Table 2.2. RT-PCR Reaction Components and Protocol

Component	Volume per Reaction (µl)
5X iScript™ Reaction Mix	4
iScript™ reverse transcriptase	1
Nuclease-free Water	Variable (total volume 20 µl)
RNA template	Variable (100 ng to 1 µg RNA)
Reaction Protocol	
Priming	5 minutes at 25 °C
Reverse transcription	20 minutes at 46 °C
RT inactivation	1 minutes at 95 °C
Cooling	Hold at 4 °C

2.4.3 Real-Time Polymerase Quantitative Chain Reaction

For Real-Time Quantitative Polymerase Chain Reaction (RT-qPCR), not to be confused by Reverse-Transcription PCR, cDNA synthesized from total RNA samples of *S. clavuligerus* were used as template for RT-qPCR. For preparation of RT-qPCR Master Mix, SsoAdvanced™ Universal SYBR® Green Supermix (BioRad, USA) were used. SsoAdvanced Universal SYBR® Green supermix is 2X concentrated and ready-to-use reaction master mix optimized for dye-based real-time PCR on any real-time PCR instrument (ROX-independent and ROX-dependent).

It contains antibody-mediated hot-start Sso7d fusion polymerase, dNTPs, MgCl2, SYBR® Green I dye, enhancers, stabilizers, and a blend of passive reference dyes (including ROX and fluorescein). Sso7d fusion protein is a patented technology of BioRad. Sso7d is a 7 kDa protein obtained from *Sulfolobus solfataricus*, binds to dsDNA without any sequence specificity. Sso7d is covalently linked to a DNA polymerase and is able to stabilize the polymerase-template complex without compromising the structural integrity, thermal stability, or catalytic activity of the enzyme. The Sso7d fusion polymerase has several advantages such as high inhibitor tolerance, increased processivity, and efficient amplification of GC rich regions or secondary structure over naked DNA polymerase. (BioRad, 2013) For optimization of RT-qPCR six different concentrations of primers, 25, 100, 150, 200, 250 ve 350 nM were tried. Full list of components for RT-qPCR master mix is given below (Table 2.3).

Table 2.3. Components of RT-qPCR master mix

Ingredients	Amount
SsoAdvanced Universal SYBR Green Supermix	5 µL
Forward primer	150 nM
Reverse primer	150 nM
cDNA template	Up to100 ng
dH_2O (PCR grade)	Up to 10 µL

Reaction mixtures were prepared in Low profile, White, 0.2 ml 8-tube PCR strips (BioRad, USA). Each strip was containing 3 biological replicates of cDNA samples taken from NRRL3585 and DEPA strains of *S. clavuligerus*, and 1 tube of no template control (NTC), having dH_2O instead of cDNA template in order to detect background noise and 1 tube of RNA template control (RTC), having total RNA sample instead of cDNA in order to assure that RNA samples are not contaminated by the genomic DNA of *S. clavuligerus* and to detect primer dimer formation.

Melting Curve Analysis is also a crucial step for detecting any double stranded DNA like primer dimers contaminating DNA and non-specific amplification of template in SYBR Green I detection chemistry, therefore included in RT-qPCR protocol. Two independent experiments of RT-qPCR have done for each target gene and each strip. For PCR detection, CFX Connect^{TM} Real-Time PCR Detection System (BioRad, USA) were used. Reaction protocol for the RT-qPCR is given below (Table 2.4).

Table 2.4. Reaction protocol for RT-qPCR

1.	Denaturation-Holding stage		95^{o} C	30 seconds	1 cycle
2.	Amplification-Cycling stage	denaturation	95^{o} C	10 seconds	
3.		annealing	58 - 60^{o} C	60 seconds	
4.		extention	---	---	
5.		Return to 2			40 cycle
6.	Melting curve stage		65^{o} C – 95^{o} C	0,5^{o} C, increment 0.05 sec for each increment	1 cycle
7.	Cooling stage		40^{o} C	30 seconds	1 cycle

2.4.4 Relative Quantification of RT-qPCR Data with Efficiency Correction

A relative quantification is based on the relative expression of a target gene versus a reference gene. To investigate the changes in gene expression, the relative expression ratio is adequate for the most purposes. The mathematical model presented below was used to determine the relative quantification of a target gene in comparison to a reference gene.

$$R = (E_{target})^{\Delta Cq\ target} / (E_{reference})^{\Delta Cq\ reference}$$

The Relative Expression Ratio (R) of a target gene is calculated as based on E and the Cq deviation of an unknown sample versus a control and expressed in comparison to a reference gene. Amplification Efficiency of each gene is calculated from the slope of a standard curve. In order to obtain slope of the standard curve, 10-fold (2-5 fold for genes with low expression) serial dilution of the cDNA template is made and the respective Cq values for each dilution is acquired. A RT-qPCR standard curve is graphically represented as a semi-log regression line plot of Cq value vs. log of input cDNA concentration. Bio-Rad CFX manager Software 3.1 (BioRad, USA) was used for graph acquisition and calculating the amplification efficiency.

A reference gene with stable expression levels is essential as an internal control to normalize and monitor the variations in expression of the interested gene(s). The relative expression of the target gene is normalized by at least one non-regulated reference gene expression, which is mostly a housekeeping gene. In this study, reference gene was *hrdB* which codes for the vegetative principle sigma factor σ^{70}, which is the major sigma factor during exponential growth. It is commonly used as an internal control gene for *S. clavuligerus* gene expression studies.

2.4.5 Statistical Analysis

For statistical analysis, Cq values collected from NRRL3585 and DEPA strains were used to calculate the R values. The data were evaluated by individual t-test for comparison of each gene expression at T_{48}, after 48 hours of fermentation, between control and sample groups and one-way ANOVA was used to repeat the analysis in GraphPad Prizm 6.0 Software (GraphPad, CA). The level of significance was stated as *($p<0.05$), ** ($p<0.01$) and *** ($p<0.001$) in the graphs. For the correlation of relative gene expression and relative protein abundance, non-parametric Spearman's rank-order correlation test was used (SPSS, USA).

CHAPTER 3

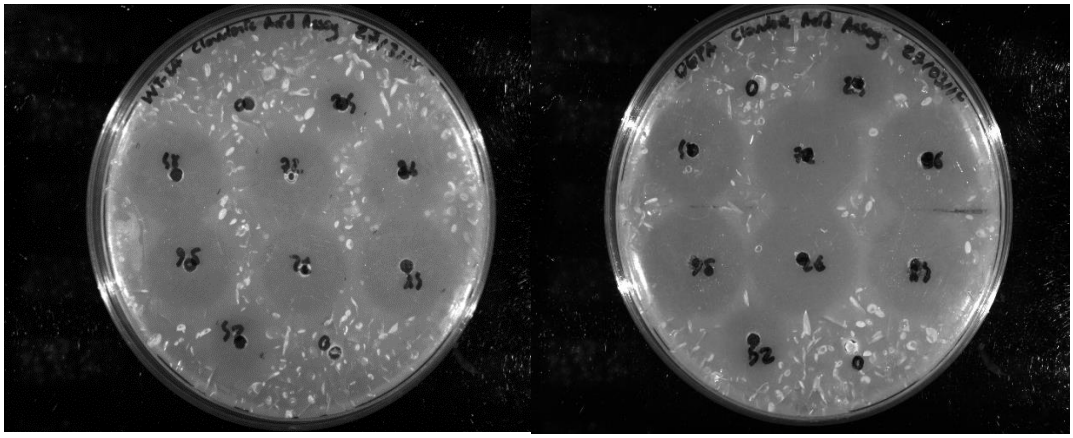
RESULTS AND DISCUSSIONS

3.1 A Comparison Between S. *clavuligerus* NRRL 3585 and DEPA With Respect to the Antibiotics They Do Produce

According to the former proteomics study, the proteins required for CA biosynthesis were more abundant in DEPA strain than the standart strain while the proteins required for the CC biosynthesis were in low abundance in this strain. The protein, namely putative aminoglycoside 2-N-acetyltransferase Aac2 belonging to aminoglycoside 2-N-acetyltransferases of AAC(2') family that confers aminoglycoside resistance were much less abundant (22.7 fold) in DEPA strain as well though *S. clavuligerus* is known not to harbor any aminoglycoside biosynthetic operons (Unsaldı, 2016).

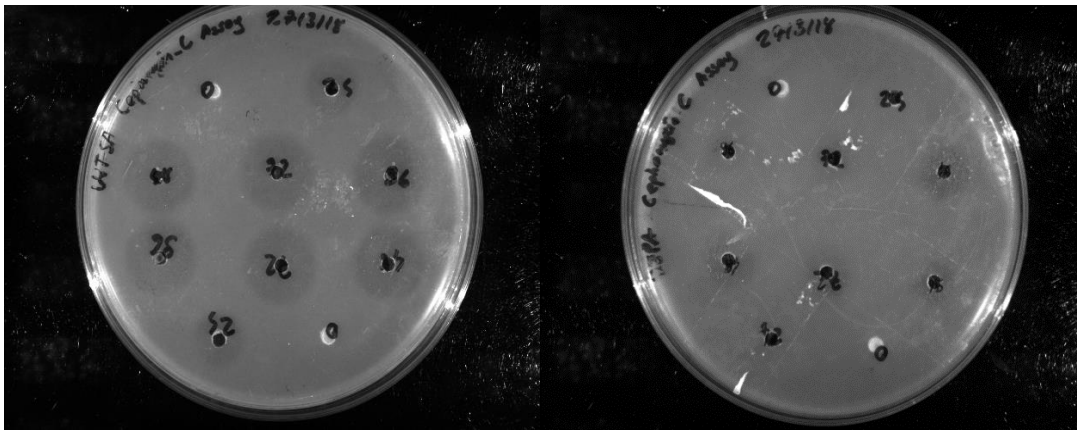
The antibiotic resistance genes that prevent the suicide of the microbes are well known to reside in the respective biosynthetic operons. (Mak et al., 2014)**.** Thus, this finding brought about the question as to if aminoglycoside biosynthesis is suppressed in favor of highly elevated CA production in DEPA strain. The antibiotic activity bioassays were therefore conducted for both strains in order to reveal the relation between low and high abundance of specific proteins in DEPA and over- and under-production of relevant antibiotics. NRRL3585 and DEPA strains were grown in SA medium, the samples of cultures were collected at 24 h interval and the supernatants were used to determineantibiotic production levels by suitable bioassays described in Section 2.3.1.

The results are shown in Figures 3.1, 3.2 and 3.3, for CA, CC and aminoglycoside antibiotics, respectively.



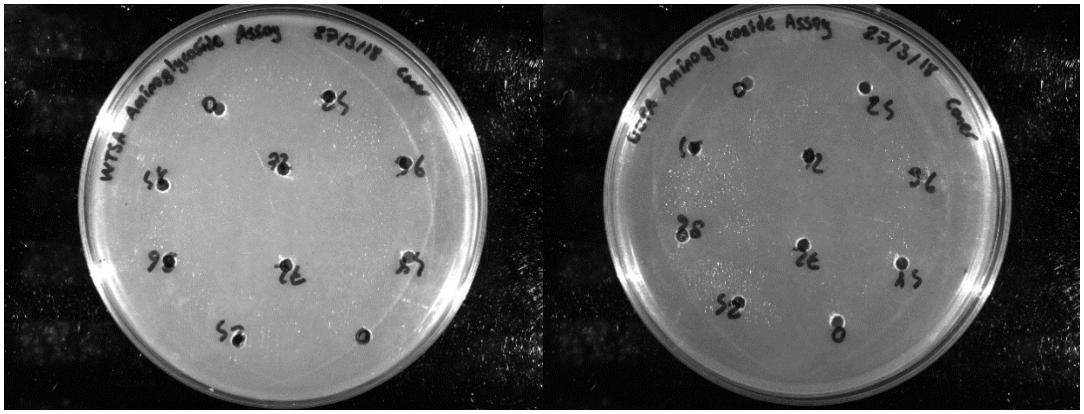
(A) (B)

Figure 3.1. Clavulanic acid bioassay. Growth inhibition zones as a function of time in tryptic soy agar plate when the culture supernanatants of NRRL3585 (A) and (B) DEPA were used against the test organism *K. pneumoniae* ATTC 29665.



(A) (B)

Figure 3.2. Cephamycin C bioassay. Growth inhibition zones as a function of time in tryptic soy agar plate when the culture supernanatants of NRRL3585 (A) and (B) DEPA were used against the test organism *E.coli* ESS 2235



(A) (B)

Figure 3.3. Aminoglycoside bioassay. Growth inhibition zones as a function of time in tryptic soy agar plate when the culture supernanatants of NRRL3585 (A) and (B) DEPA were used against the test organism *S.epidermidis* ATCC 14990. As seen in Figure 3.3, neither the standard strain, nor the industrial one produced any aminoglycoside antibiotic. This finding further proved that aminoglycoside biosynthetic genes are really missing from *S. clavuligerus*.

Antibiotic levels in culture fluids were next estimated by consulting to the calibration curves established (Section 2.3.3). After the inhibition zone diameters measured in mm were converted to antibiotic concentrations, time versus antibiotic production curves were obtained for both strains.

As expected CA production in DEPA strain was incomparably higher than NRRL 3585. At 72nd h of cultivation when it peaks in this strain, it was over 10 times higher. (Figure 3.4). DEPA strain was reported to produce CA in at least 100 times increased levels, as detected with HPLC analysis (Unsaldi et al., 2017). HPLC method is analytic method for the quantification of the potency of antibiotics, but unable to determine the bioactivity. Plate assay is used to determine both bioactivity and the concentration of antibiotics, but it is not an analytic method to give absolute quantification. The determination of concentration is based on estimation according to the standart curves generated in plate assay as well, and in agar plates diffusion problems arise in restricted area so that the measurement confidence is lost especially

when inhibition zone diameters gets larger, any errors in measuring milimeters resulting in very high deviations.

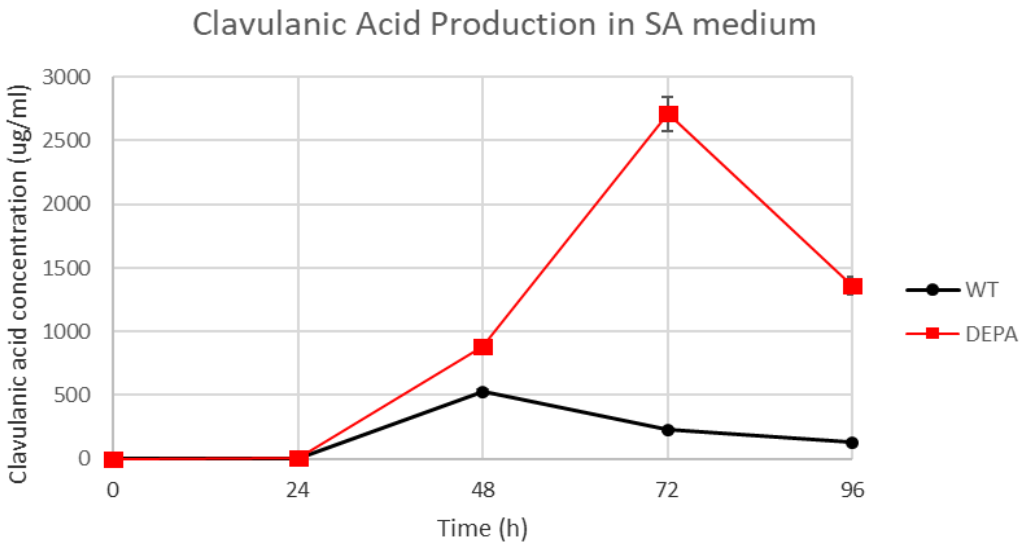


Figure 3.4. CA production by *S.clavuligerus* NRRL 3585 and DEPA strains when incubated in SA medium for 96 hours. The samples were collected at 24 hour intervals.

While CA production is much higher, the CC production of DEPA strain was much lower in DEPA strain than in NRRL 3585. It was at least 6 times lower after 48 hours, and more than 20 times lower at 72 hours. Although the production in both strains decreased in 96 h cultures, NRRL 3585 was still producing much more CC than the DEPA strain. These findings were quite in accordance with those proteomics which showed that the respective biosynthetic enzymes as well as those providing precursors for CC biosynthesis are profoundly underrepresented in DEPA strain.

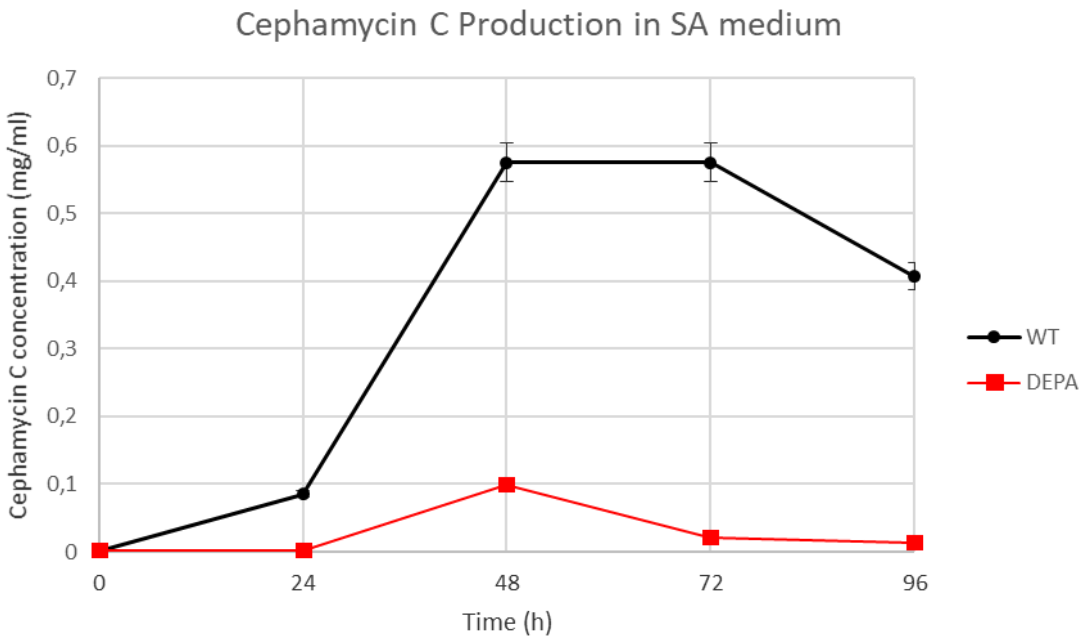


Figure 3.5. CC production by *S. clavuligerus* NRRL 3585 and DEPA strains when incubated in SA medium for 96 hours. The samples were collected at 24 hour intervals.

3.2 Quantitative Gene Expression Analysis of *S. clavuligerus* NRRL 3585 and DEPA

3.2.1 Optimization of RT-qPCR experiments

The optimization of the RT-qPCR is crucial for acquiring accurate and reproducible results from the experiments and eliminating user bias. In this study, two step RT-qPCR method is used in order to increase sensitivity and allow for cDNA storage for further use. Total RNA template is used for cDNA synthesis with reverse transcription in a separate reaction as the first step and cDNA is used as template for qPCR in the second step. The primer concentration, cycling conditions and bivalent

ion concentration like $MgCl_2$ are the parameters for the optimization in second step. In this study, the commercially available SYBR Green supermix is used, and it has defined amount of bivalent ion and buffer as well as DNA polymerase. Only interchangeable parameter was primer concentration. 25, 100, 150, 200, 250 ve 350 nM of primers were tried. For all of the genes including the internal control, 150 nM primer concentration was found the optimum for a 10 µL reaction mix. In addition, an annealing temperature of 58 ^oC was found to be most appropriate and, 40 cycles of amplification were chosen to generate the most reliable results. The amplification curves and melting peaks as well as standard curves for each gene are given in APPENDİX B. Average Cq and Tm values are listed below (Table 3.1). Amplification efficiencies of the genes are also listed as E values and tabulated below (Table 3.2).

Table 3.1. Average Cq and Tm values for optimized PCR conditions

	Cq			**Tm ºC**		
	Target	NTC	RTC	Target	NTC	RTC
hrdB	18.52	-	-	85.50	-	-
lexA	21.40	-	-	91.00	-	-
SCLAV_0043	19.38	-	-	89.00	-	-
SCLAV_5275	21.16	-	-	90.00	-	-
cmcJ	19.91	-	-	90.50	-	-
SCLAV_4282	23.88	-	-	85.50	-	-
SCLAV_2046	23.62	-	36.07	89.00	-	-
ectD	24.02	-	-	90.50	-	-
tig	24.49	-	-	89.00	-	-
fab1	24.75	-	-	90.00	-	-
SCLAV_1945	26.65	-	-	91.00	-	-
bkdA1	25.74	-	-	90.50	-	-
SCLAV_0130	23.04	-	-	85.00	-	-
SCLAV_0453	21.52	39.58	39.23	88.50	-	-
SCLAV_0843	25.12	-	-	90.00	-	-
SCLAV_2541	28.43	-	39.47	89.00	-	-
SCLAV_2986	22.99	-	-	88.00	-	-
SCLAV_3034	25.45	-	-	90.50	-	-
SCLAV_4756	21.48	-	-	91.00	-	-
SCLAV_5256	29.75	-	-	91.50	-	-
SCLAV_5667	23.95	-	-	90.00	-	-
avaA2	21.45	-	-	89.00	-	-
bkdB1	27.67	-	-	92.00	-	-
SCLAV_p0693	39.15	-	-	87.50	-	-
SCLAV_p1324	34.71	-	-	90.50	-	-
secA	22.07	-	-	89.00	-	-

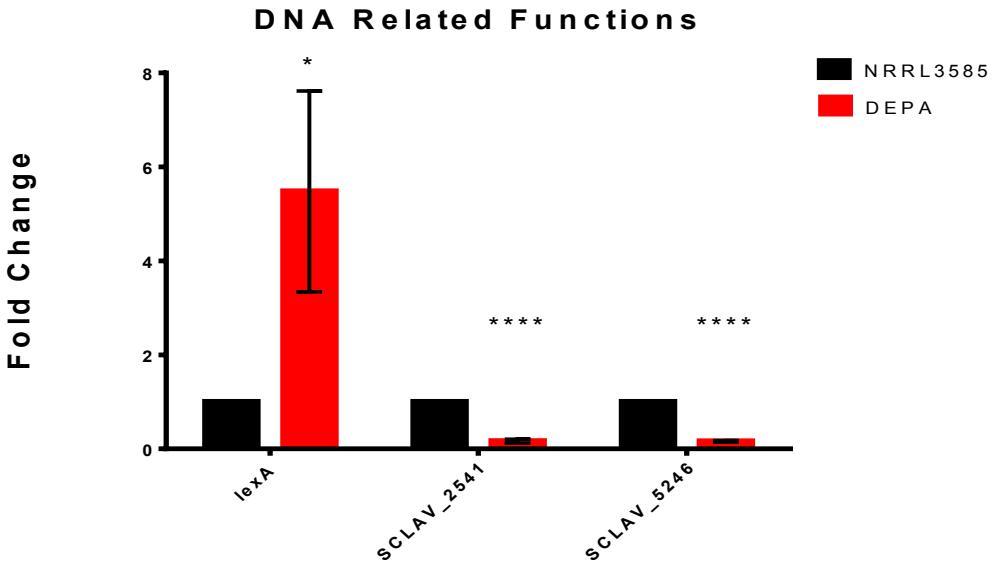
Table 3.2 Amplification efficiencies and E values for the genes studied

Gene	Efficiency (%)	E value	Gene	Efficiency (%)	E value
hrdB	86.4	1.728	***SCLAV_0453***	94.4	1.888
lexA	110.9	2.218	***SCLAV_0843***	99.5	1.99
SCLAV_0043	116.3	2.104	***SCLAV_2541***	104.2	2.084
SCLAV_5275	105.2	2.326	***SCLAV_2986***	90.3	1.806
cmcJ	104.9	2.098	***SCLAV_3034***	95.9	1.918
SCLAV_4282	82.4	1.648	***SCLAV_4756***	99.9	1.998
SCLAV_2046	95.7	1.914	***SCLAV_5256***	84.8	1.696
ectD	107.9	2.158	***SCLAV_5667***	95.5	1.91
tig	98.2	1.964	***avaA2***	92.3	1.846
fab1	105.8	2.102	***bkdB1***	94.3	1.886
SCLAV_1945	91.1	2.004	***SCLAV_p0693***	94.7	1.894
bkdA1	100.2	1.822	***SCLAV_p1324***	91.3	1.826
SCLAV_0130	105,3	2.106	***secA***	97.5	1.95

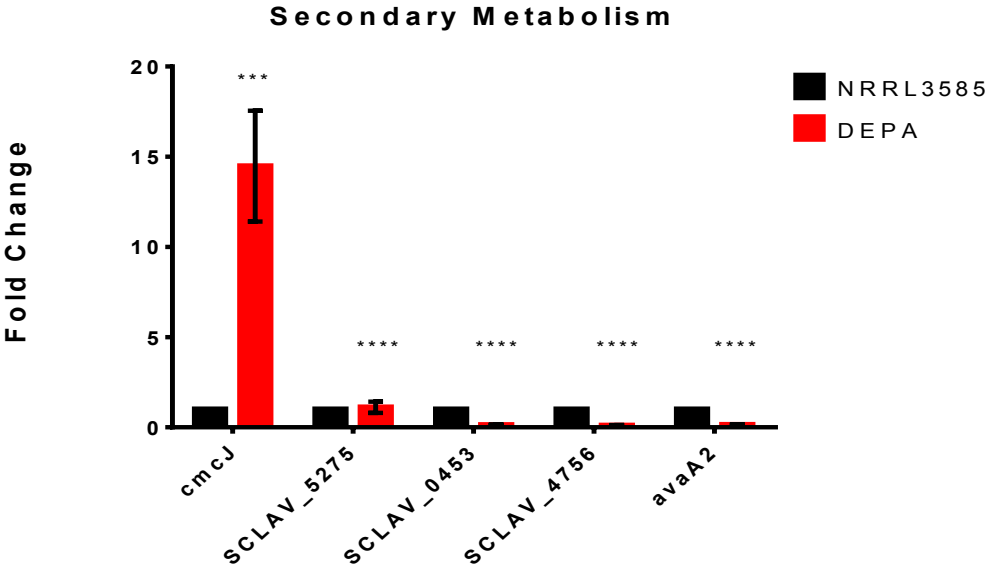
The Cq value bigger than 35.00 in NTC and RTC for all genes show that RT-qPCR experiments done without gDNA contamination and Tm bigger than 85 ^{0}C is the indication of no primer dimer formation for the reactions which are required for the optimization of reactions and validation of the RT-qPCR experiments (Bustin et al., 2009).

3.2.2 Quantitative Gene Expression Profile Comparisons

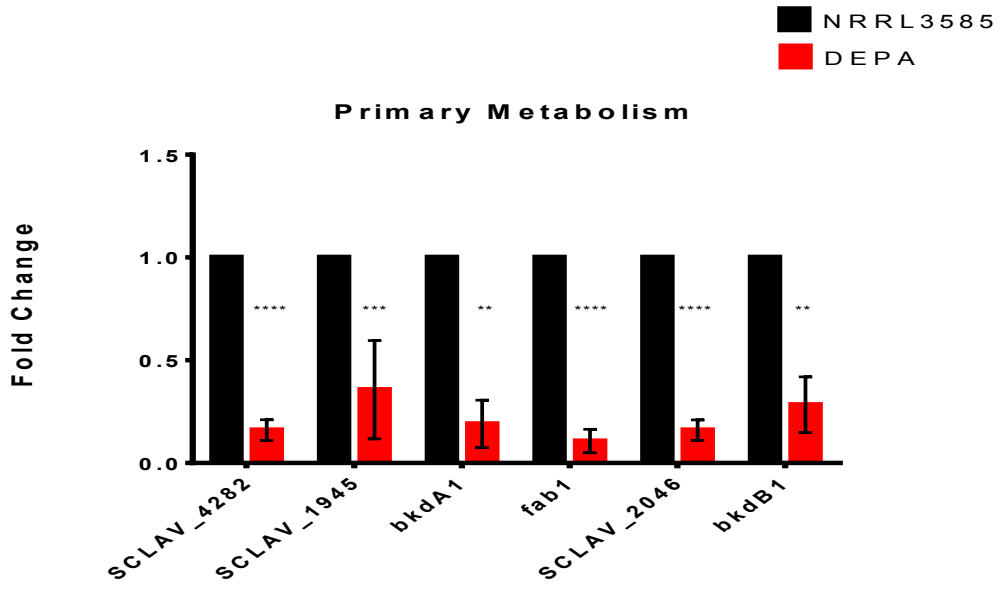
The relative gene expression differences are acquired with one-way ANOVA and individual t-tests the results were in bar graphs as the fold change with respect to S. clavuligerus NRRL 3585 strain for each gene. The results are tabulated in (Figure 3.6) The black bar (1-fold) represent the base expression level in *S. clavuligerus* NRRL 3585 strain for each gene. Two of the analyzed genes (*SCLAV_p0693*, and *SCLAV_p1324*) are missing in graph because of their outlier values. The studied genes were grouped into relevant functional categories. The fold changes in relative gene expression are also tabulated in APPENDIX B to include values missing from the histograms.


DNA Related Functions
Fold Change
8
6
4
2
0
*
****
****
NRRL3585
DEPA
lexA
SCLAV_2541
SCLAV_5246

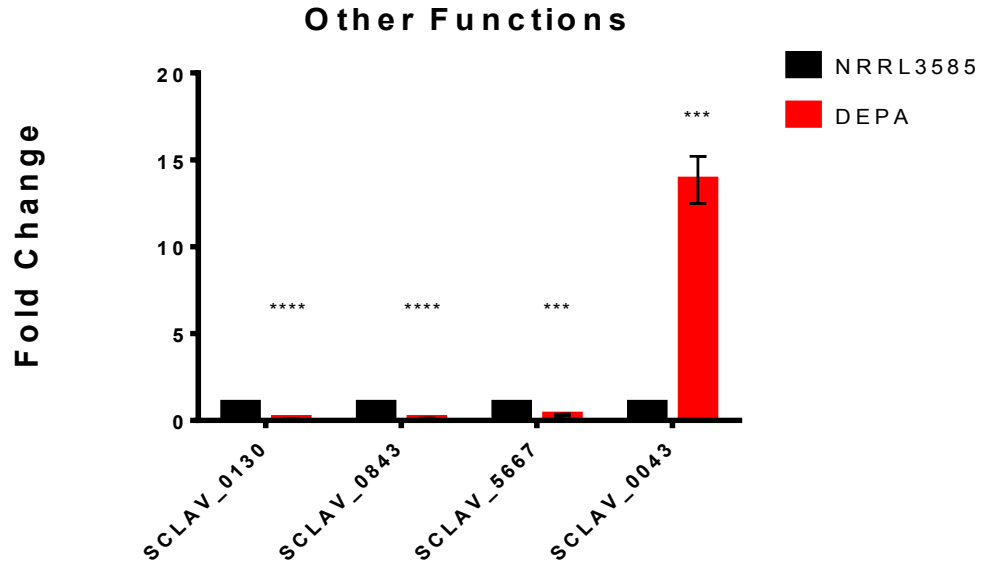
1


Secondary Metabolism
Fold Change
20
15
10
5
0
***
****
****
****
****
NRRL3585
DEPA
cmcJ
SCLAV_5275
SCLAV_0453
SCLAV_4756
avaA2

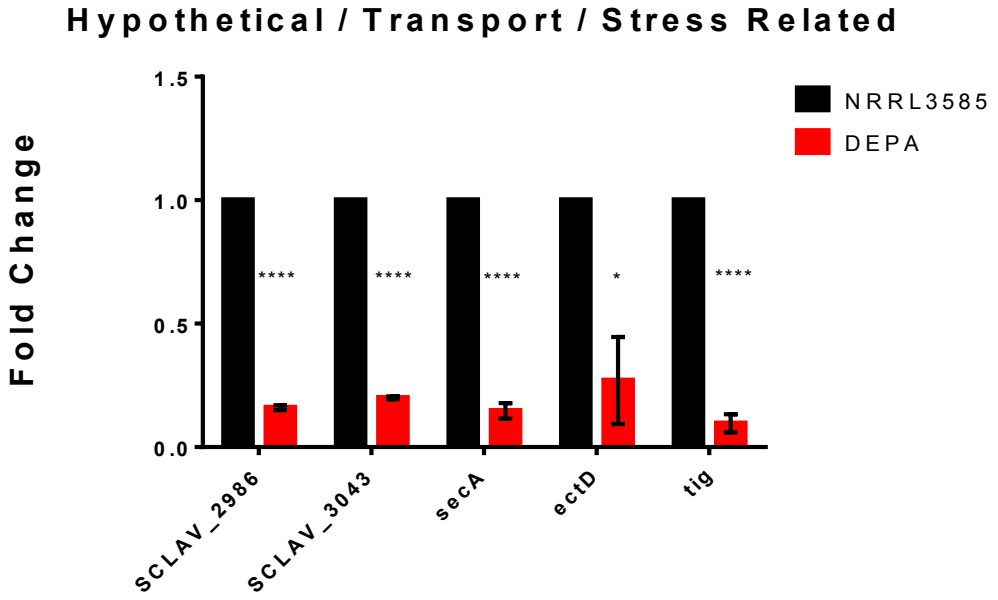
2


NRRL3585
DEPA
Primary Metabolism
Fold Change
1.5
1.0
0.5
0.0
****
***
**
****
****
**
SCLAV_4282
SCLAV_1945
bkdA1
fab1
SCLAV_2046
bkdB1

3


Other Functions
NRRL3585
DEPA
Fold Change
20
15
10
5
0
****
****
***
***
SCLAV_0130
SCLAV_0843
SCLAV_5667
SCLAV_0043

4



5

Figure 3.6 Bar graphs showing relative gene expression levels between the standard and DEPA strains according to the functional categories, (1) DNA repair, recombination, transcription; (2) Secondary metabolism; (3) Primary metabolism; (4) Other functions; and (5) Hypothetical, transport, stress related functions.

3.3 Comparison of Differential Gene Expression with Differential Protein Abundance

Results acquired in gene expression analysis was used to compare mRNA levels protein abundance. Few studies on mRNA–protein correlations in bacterial cells have been published. It has been shown that protein and mRNA abundances are not following a normal distribution (Futcher et al., 1999; Ghaemmaghami et al., 2003). Therefore, the Spearman rank coefficient (r_s) is suitable to describe the correlation between mRNA and protein levels (Maier et al., 2009). The Spearman's rho value of Spearman rank coefficient (rs) has been evaluated which can vary from −1 (perfect negative correlation) through 0 (no correlation) to +1 (perfect positive correlation).

Our correlation results consistent with previous similar study in the literature (Koussounadis et al., 2015; Vogel et al., 2011).

Systematic analyses on such organisms incorporating transcriptome or proteome profiling can be quite revealing in functional analysis and elucidate the common regulators. Although consistency between mRNA and proteins is often inherently assumed in many studies, examples of divergent trends are also frequently observed. In one of such systemic analysis done with *S. coelicolor*, the model organism for *Streptomyces* genus, which coupled shotgun proteomics and DNA microarray-derived transcriptome data revealed that the prominent patterns in both protein and mRNA domains are surprisingly well correlated. Despite this overall correlation, by employing a systematic concordance analysis, it is estimated that over 30% of the analyzed genes exhibited significantly divergent patterns, of which nearly one-third displayed even opposing trends (Jayapal et al., 2008). The differences between mRNA and protein synthesis/degradation mechanisms are prominent in microbes while reaffirming the plausibility of such mechanisms acting in a concerted fashion at a protein complex or sub-pathway level.

Genes were selected from LC-MS/MS and 2DE gel electrophoresis done on DEPA and standard strains with relative protein abundances show more than 2.00 folds of change (Unsaldı, 2016). Selected genes that are listed in introduction section before. Gene expression experiments are conducted RT-qPCR with 6 replicate each (3 biologicals and 2 technical replicates). According to results, gene expression levels classified in three main groups, (i) the negatively correlated gene expression and protein abundance, (ii) the positively correlated gene expression and protein abundance, and (iii) no correlation between gene expression and protein abundance.

3.3.1 Negatively Correlated Gene Expression and Protein Abundance

In the first group such divergent correlation pattern has been observed (Table 3.3). It is known that mRNA and protein profiles can be diverge in time (Fournier et al., 2010; Vogel et al., 2011), thus it is possible that positive correlations might have

existed when different time scales are used, considering the fact that after 48 hours the organisms undergoes major physiological changes, antibiotic production rapidly declines, gene expression levels very possibly rapidly decline after this point and protein degradation occurs afterwards. Negative correlation between protein abundance and gene expression levels could be the result of translationally regulated protein expression where protein abundance possibly exerting a negative feedback control and downregulate mRNA synthesis.

It should also be taken into account that those genes (Table 3.1) are not directly related to the biosynthetic pathways but stress response and other regulatory mechanisms of organism. Regulatory proteins and sRNAs can act as translational modulators. For instance, in *E. coli*, genes encoding R-proteins (rpl, rps) are spread in several operons. A regulatory R-protein represses translation of some cystrons by binding to its own mRNA at a region contiguous with the Shine-Dalgarno sequence (Golding et al., 2005; Maier et al., 2009). The target RNA sites are similar to their corresponding binding sites in the rRNA but with a lower affinity. Only when all rRNA is assembled into ribosomes, R-proteins bind to their mRNA so as to stop translation. This case is also express in the decrease the gene expression in the correspondent increase in the protein abundance.

Table 3.3. Negatively correlated gene expression versus protein abundance levels and Spearman's correlation table

Protein Functions	Uniprot Accession Number	Protein Name	Method	Relative Pro. Ab.	Related Gene	Gene Expression
DNA related	B5GPP5 (a)	Anti-sigma factor	2-D gel	7.30	SCLAV_2541	-6.20
DNA related	B5GPP5 (b)	Anti-sigma factor	2-D gel	3.30	SCLAV_2541	-6.20
Other Function	B5H160	Siderophore-interacting protein	2-D gel	4.70	SCLAV_0843	-8.26
Other Function	B5H0U2	DNA-binding protein	LC-MS/MS.	2.78	SCLAV_0130	-8.52
Secondary Metabolism	E2PUT1	Polyprenyl synthetase	LC-MS/MS	2.28	SCLAV_4756	-9.23

*Relative protein abundance in folds

Spearman's Correlation

			Gene expression	Protein abundance
Spearman's rho	Gene expression	Correlation Coefficient	1.000	-.821*
		Sig. (1-tailed)	.	.044
		N	5	5
	Protein abundance	Correlation Coefficient	-.821*	1.000
		Sig. (1-tailed)	.044	.
		N	5	5

*. Correlation is significant at the 0.05 level (1-tailed).

Both isoforms of anti-sigma factor protein (represented as a, and b), siderophore-interacting protein, DNA-binding protein, and polyprenyl synthetase protein are found to be up-regulated in DEPA strain by Unsaldı (2016) with 2-D gel and LC-MS/MS by more than 2 folds (Table 3.3). Relative gene expression experiments by RT-qPCR show 6 to 9 times decrease in mRNA levels. By the spearman's rank correlation test strong negative correlation (rho = -0.821) is found between these proteins and expression of relevant genes.

Names and UniProtKB accession numbers of first group are follows: Anti-sigma factor (B5GPP5), siderophore-interacting proteins (B5H160), DNA binding protein (B5H0U2), and polyprenyl synthetase (E2PUT1). Anti-sigma factor is regulating gene expression at the level of transcription by antagonizing with the sigma transcription factors. (Hughes et al., 1998). Siderophores are iron chelators of microorganisms which do also have roles in other metals chelation and oxidative stress (Johnstone and Nolan, 2015). Siderophore-interacting proteins are responsible to reduce the iron-siderophore complexes to release iron (Li et al., 2015). Therefore, siderophore-interacting proteins have role in regulation of functions carried out by siderophore itself. Since fermentation media are designed to cause stress to induce secondary metabolite production the siderophore-interacting protein could have some roles in antibiotic overproduction. DNA binding protein (B5H0U2) belongs to lambda repressor-like, DNA-binding domain superfamily is a transcriptional regulator according to UniProtKB database. Polyprenyl synthetase is an enzyme responsible for synthesis of isoprenoids which could be among the secondary metabolite precursors produced by the *S. clavuligerus* though there is no elaborate study of this enzyme to the date.

3.3.2 Positively Correlated Gene Expression and Protein Abundance

In the second group, there was positive correlations between differential gene expression and relative protein abundance (Table 3.4). According to the Spearman's correlation, there is a positive correlation (r_s = 0.578) between gene expression and protein abundance levels. Many of the proteins in this group has a catalytic activity and involved directly in biosynthetic pathways as opposed to first group proteins which mostly have regulatory functions.

RT-qPCR experiments revealed that the extent of alterations in mRNA levels are lower than those in protein levels. There are possible explanations for these results. First, the rapid changes in transcript levels have to be buffered post-transcriptionally to maintain homeostatic steady-state protein levels (Liu et al., 2016). Second, the cells have to sense external signals through surface receptors and distinguish noisy

signals such as rapid environmental fluctuations from signals that should actually result in intracellular state changes, such as adaptation to stress or differentiation into a new cell type. On the other hand, cells have to confer robustness to internal noise caused by genomic variation or due to stochastic initiation of transcription (Chalancon et al., 2012; Gandhi et al., 2011). The strategies for responding to external and internal signals could be different, while external signals can be buffered before causing transcriptional changes (Chalancon et al., 2012; Hornung et al., 2008). Variation at the DNA level can directly affect transcript levels, when transcription factor binding sites were altered.

Table 3.4. Positively correlated gene expression versus protein abundance levels and spearman's correlation table

Protein Functions	Uniprot Accession Number	Protein Name	Method	Relative Pro. Ab.*	Related Gene	Gene Expression
Other Function	B5GZC6	Amidohydrolase	LC-MS/MS.	-2.02	SCLAV_5667	-3.24
Primary Metabolism	B5GZV9	E1-alpha branched-chain alpha keto acid dehydrogenase	LC-MS/MS	-2.02	bkdA1	-5.18
DNA related	E2PZ58	TetR-family transcriptional regulator	LC-MS/MS	-2.33	SCLAV_5246	-6.15
Primary Metabolism	E2PWQ0	Branched-chain alpha keto acid dehydrogenase E1 beta subunit	LC-MS/MS	-2.58	bkdB1	-3.49
hyp/trans/stress	E2PWW5(a)	Putative M28-family peptidase	2-D gel	-2.7	SCLAV_3034	-5.01
hyp/trans/stress	E2PWW5(b)	Putative M28-family peptidase	2-D gel	-6.1	SCLAV_3034	-5.01
Secondary Metabolism	E2PWQ8	DUF2587 domain-containing protein	2-D gel	-2.8	SCLAV_2986	-6.25
Primary Metabolism	E2Q9C7	Gamma-butyrolactone biosynthesis protein	LC-MS/MS	-3.36	avaA2	-6.88
Primary Metabolism	B5GL42	UDP-glucose 4-epimerase	2-D gel (AK39)	-3.8	SCLAV_4282	-6.35
Secondary Metabolism	B5H0X6	Enoyl-[acyl-carrier-protein] reductase [NADH]	2-D gel	-4.5	fab1	-9.42
Primary Metabolism	E2Q9A9	Polyketide synthase	LC-MS/MS	-4.81	SCLAV_0453	-7.25
hyp/trans/stress	B5GSL5 (a)	4-hydroxyphenylpyruvate dioxygenase	2-D gel	-35.7	SCLAV_2046	-6.24
Primary Metabolism	B5GSL5 (b)	4-hydroxyphenylpyruvate dioxygenase	2-D gel	-4.9	SCLAV_2046	-6.24
hyp/trans/stress	E2Q3P1	Trigger factor	2-D gel	-37	tig	-10.6
DNA related	E2QA55	lexA Repressor	LC-MS/MS	2.78	lexA	5.48
Secondary Metabolism	E2PZ87	Thioredoxin reductase	LC-MS/MS	3.9	SCLAV_5275	1.11

*Relative protein abundance in folds

Spearman's Correlation

			Gene Expression	Protein Abundance
Spearman's rho	Gene Expression	Correlation Coefficient	1.000	.578*
		Sig. (1-tailed)	.	.015
		N	14	14
	Protein Abundance	Correlation Coefficient	.578*	1.000
		Sig. (1-tailed)	.015	.
		N	14	14

*. Correlation is significant at the 0.05 level (1-tailed).

Names and UniProtKB accession numbers for the second group of proteins are as follows; One categorized as "other function", Amidohydrolase (B5GZC6), or amidase, is a type of hydrolase that acts on ester and amide bonds at the carbon or phosphorus centers of the substrate (Seibert, 2005). Amidohydrolases have diverse members with diverse substrates, but their main function is to hydrolyze the amide or ester groups located at carbon or phosphorus centers of the substrate. Since the protein identified as amidohydrolase has not been characterized specifically, requires further studies are needed to understand its function in *S. clavuligerus* and the reason for its downregulation in DEPA strain. Another protein categorized as "DNA related function" was TetR-family transcriptional regulator (E2PZ58), which can act as a repressor and activator. *Streptomyces* species possesses more than 100 TetR-family protein coding genes. these TetR-family proteins participate in for the initiation of morphogenesis and antibiotic production (Guo et al., 2013; Hillerich et al., 2008). Three proteins were categorized as hypothetical protein; putative M28-family peptidase (E2PWW5 with 2 isoforms; a and b), 4-hydroxyphenylpyruvate dioxygenase (B5GSL5 with 2 isoforms; a and b), and Trigger factor (E2Q3P1) is a chaperone preventing newly-synthesized proteins from misfolding or aggregating (Wong et al., 2004). It also act as a cell division activator (Langlois, 2003). Three other proteins in this group were related with secondary metabolite production and represented down-regulated ones; gamma-butyrolactone biosynthesis protein (E2Q9C7) which has been shown to serve as quorum-sensing signaling molecules for

activating antibiotic production in *Streptomyces* species (Du et al., 2011), enoyl-[acyl-carrier-protein] reductase (B5H0X6), which play essential role in fatty acid biosynthesis and aromatic polykedite antibiotics (Tang et al., 2006), and polyketide synthase (E2Q9A9), multifunctional enzyme for the stepwise generation of polyketides which constitute a group of antibiotics, including erythromycin (Jenke-Kodama et al., 2006). Down-regulation of polyketide synthesis was thought to favor CA overproduction in DEPA strain through the elimination of the biosynthesis of polyketide antibiotic(s). The rest of the proteins included those that are functioning in primary metabolism: E1-alpha branched-chain alpha keto acid dehydrogenase (B5GZV9), and its branched-chain alpha keto acid dehydrogenase E1 beta subunit (E2PWQ0), the activity of which provide fatty acid precursors for polyketide antibiotic biosynthesis in *Streptomyces* species (Denoya et al., 1995), UDP-glucose 4-epimerase (B5GL42), which is responsible for the reversible conversion of galactose to glucose 1-phosphate in *S. clavuligerus* (Thoden et al., 1996). Although *S. clavuligerus* does not utilize glucose and galactose as sole carbon sources (Pérez-Redondo et al., 2010) the genes for transport and utilization of glucose and galactose are intact in *S. clavuligerus*. Moreover, there exists 100% sequence coverage and 76% identity between tunicamycin biosynthesis protein TunF of *S. chartreusis* (Wyszynski et al., 2012), and *S. clavuligerus* UDP-GlcNAc 4-epimerase (Ünsaldı, 2016). In genome projects of *S. clavuligerus* (Medema et al. 2010), no tunicamycin-related antibiotic biosynthetic cluster could be found. Still it is reported a tunicamycin-like complex present in this organism (Kenig et al., 1979).

Enoyl-[acyl-carrier-protein] reductase (B5H0X6), which play essential role in fatty acid biosynthesis and aromatic polykedites i.e antibiotics(Tang et al., 2006), and 4-hydroxyphenylpyruvate dioxygenase (B5GSL5 with two isoforms a and b) catalyzes the second step in the pathway of tyrosine catabolism (Johnson-Winters et al., 2003). Thioredoxin reductase is a disulfide reductase that transfer electrons from pyridine nucleotides via a flavin carrier to disulfide-containing substrates., This enzyme helps the function of isopenicillin-N-synthase, a key enzyme in the biosynthesis of β-lactam compounds for the maintenance of the redox state of their cysteine amino acid residues its tripeptide substrate 8-(L-a-aminoadipyl)-L-cysteinyl-

D-valine (ACV) (Cohen et al., 1993) thus, an elevation in its level in a CA-overproducing strain is meaningful. LexA repressor negatively regulates transcription of the genes to in SOS global regulatory network which functions in DNA repair. RecA and LexA are two proteins governing the response to DNA damage in bacteria. In *E. coli*, RecA stimulates cleavage of the LexA repressor, yjereby inducing more than 40 genes that comprise the SOS global regulatory network (Butala et al., 2009). Its upregulation in DEPA strain is quite expected since it would prevent error-prone DNA repair during successive rounds of mutagenesis.

3.3.3 No Correlation Between Gene Expression and Protein Abundance.

Third group of proteins displayed "no correlation between gene expression and protein abundance level" (Table 3.5). There are presumably at least three reasons for the poor correlations generally reported in the literature between the level of mRNA and the corresponding protein. First, there are many complicated and varied post-transcriptional mechanisms involved while mRNA molecules are converted into proteins. These mechanisms are not yet sufficiently well defined to compute protein concentrations from mRNA levels. Second reason for a general lack of correlation between mRNA and protein abundance is that the proteins have much longer half life than mRNA molecules in prokaryotic organisms. Protein turnover significantly depends on a number of different conditions (Glickman et al., 2002). and there is significant heterogeneity even among proteins with similar functions (Pratt et al., 2002). Third, there has been reported a significant amount of error and noise in both protein and mRNA experiments that limit our ability to obtain a clear picture (Greenbaum et al., 2003).

Table 3.5 Not correlated gene expression versus protein abundance and Spearman's correlation table

Protein Functions	Uniprot Accession Number	protein name	Method	Relative Pro. Ab.*	Related Gene	Gene Expression
Other Function	B5H4A8	Rhs element Vgr protein	2-D gel	3.8	SCLAV_0043	13.84
Secondary Metabolism	B5GLB4	7-alpha-cephem-methoxylase P8 chain	LC-MS/MS	2.91	cmcJ	14.48
Other Function	D5SJT7	Bacterial luciferase domain-containing protein	LC-MS/MS	-3.72	SCLAV_p0693	1924.56
Primary Metabolism	D5SLL7	Methionine synthase II	LC-MS/MS	-3.72	SCLAV_p1324	165.12

*Relative protein abundance in folds

Spearman's Correlation

			Gene Expression	Protein Abundance
Spearman's rho	Gene Expression	Correlation Coefficient	1.000	.400
		Sig. (1-tailed)	.	.300
		N	4	4
	Protein Abundance	Correlation Coefficient	.400	1.000
		Sig. (1-tailed)	.300	.
		N	4	4

Names and UniProtKB accession numbers of third group of proteins were as follows: Rhs element Vgr protein (B5H4A8); Rhs elements are accessory repetitious sequences which are the major source for chromosomal rearrangements in laboratory cultures of bacteria. Vgr (Val-Gly dipeptide repetition) is located upstream of the core regions and found only in RhsE and RhsG elements of *E. coli* and it is suggested that *vgr* genes might be involved in horizontal gene transfer or more precisely in transfer of a foreign DNA into the cell given that *vgr* genes described so far are usually located close to the genes that might have been horizontally transferred from other species (Wilderman et al., 2001). Since Vgr protein is related with the phage

tail proteins [as suggested by STRING data (Ünsaldı, 2016)], from which Type 6 secretion systems are known to have evolved, the presence of Rhs element Vgr protein in *S. clavuligerus* DEPA strain might provide clue for the existence of a contact-dependent type 6 secretion (T6SS)-like system in this organism (Ünsaldi et al., 2017).

Positive regulator (*CcaR*) is a transcriptional regulator protein that induces both CC and CA biosynthetic clusters in *S. clavuligerus* by binding the promoter regions of certain genes (Liras et al., 2008). 7-alpha-cephem-methoxylase P8 chain (B5GLB4), is a methoxylase enzyme that catalyzes one of the two last steps in CC production (Oster et al., 2006). Its upregulation should be a manifestation of CcaR upregulation, as extensively discussed in Ünsaldı et al., 2017. Otherwise, a significant decrease in CC production was quite prevalent from underrepresented proteins of CC biosynthetic pathway (Ünsaldı, 2016) and significantly decreased yields of CC in DEPA strain, as shown in the present study.

There are two types of methionine sythase, responsible for methionine synthesis by tranferring a methyl group to homocysteine in the presence of magnesium and phosphate ions (Whitfield et al., 1970), in *E. coli.* Cobalamin-dependent enzyme encoded by *metH* and the cobalamin-independent enzyme encoded by *metE* also both known as methionine synthase II (Banerjee et al., 1990). Since *E. coli* does not have cobalamin pathway, *MetH* is utilized only when exogenous cobalamin is present, moreover *E. coli* represses *MetE* in a such condition (Ferla et al., 2014). In *S. clavuligerus* DEPA, pSCL4-encoded methionine synthase II (D5SLL7; SCLAV_p1324) which is also known as 5-methyltetrahydropteroyltriglutamate-homocysteine S-methyltransferase as it transfers a methyl group from 5-methyltetrahydropteroyltriglutamate to homocysteine was found to be down regulated by 3.7 and 4.2 fold, in gel based and nongel-based (LC-MS MS) proteomics, respectively (Ünsaldı et al, 2017, Ünsaldı, 2016). Another enzyme acting in the same transsulfuration pathway is 5-methyltetrahydropteroyltriglutamate--homocysteine methyltransferase (*MetE*), 2.2 times underrepresented in DEPA strain.

Underrepresentation of both methionine synthase II (D5SLL7; SCLAV_p1324) and *MetE* is indicative of a general lowering the metabolic flux homoserine to methionine in aspartate pathway, thereby increasing the intracellular level of aspartate that would be directed towards L-arginine (CA precursor) biosynthesis. Another pSCL4-encoded protein, namely bacterial luciferase domain-containing protein (D5SJT7), is a member of luciferase-like, F_{420}-dependent oxidoreductase family. In several *Streptomyces* spp., such oxidoreductases have some roles in secondary metabolite production, tetracycline production in particular (Bown et al., 2016).

CHAPTER 4

CONCLUSION

The available body of literature shows that the correlation of mRNA and protein levels in complex samples is far from perfect, pointing towards complex and diverse regulatory mechanisms responsible for the observed differences in the quantitative relation between transcriptome and translatome. Several reasons can be accounting for the apparent poor correlation: (1) post-transcriptional parameters, (2) post-translational parameters, and (3) noise and experimental error. In order to be fully able to understand the relationship between mRNA and protein abundances, the dynamic processes involved in protein synthesis and degradation have to be better understood, i.e. is the protein level changing because of a change in the rate of protein synthesis, or mRNA, or protein turnover? Since it is known that each contributes at a varying extent for each bacterial taxon, there exists no single answer(s) to apply to whole members of the Kingdom of Bacteria to explain their relative roles, thus no formulations are yet available to fully predict or explain the relationships between mRNA and protein abundance levels in prokaryotic organisms.

Finally, given the pace of technological advancements in protein quantification, mRNA expression analysis and noise reduction, more comprehensive correlation formulations will soon be realized even on a species basis. This will allow for more robust analyses of the relationship between mRNA expression and protein abundance values.

Summarizing our findings in view of the current knowledge on cellular-state transitions in morphologically prokaryotes, two conclusions can be drawn: (1) over- and underrepresented proteins may show distinct behavior with respect to their transcript levels, which underlines regulatory and kinetic differences accounted by immediate posttranscriptional controls (attenuation, riboswitch and sRNAs) as well as protein turnover; (2) after transition into aerial hyphae development stage during which secondary metabolite production accompanies morphological differentiation, respective mRNA levels are still indicative of variation in protein levels for a majority of functionally critical proteins.

REFERRENCES

Alberts. (2008). Molecular Biology of the Cell (5 ed.). New York: Garland Science, Taylor & Francis Group.

Alwine, Kemp, & Stark. (1977). Method for detection of specific RNAs in agarose gels by transfer to diazobenzyloxymethyl-paper and hybridization with DNA probes. Proc Natl Acad Sci U S A, 74(12), 5350-5354.

Arulanantham, Kershaw, Hewitson, Hughes, Thirkettle, & Schofield. (2006). ORF17 from the clavulanic acid biosynthesis gene cluster catalyzes the ATP-dependent formation of N-glycyl-clavaminic acid. J Biol Chem, 281(1), 279-287.

Bachmann, Li, & Townsend. (1998). beta-Lactam synthetase: a new biosynthetic enzyme. Proc Natl Acad Sci U S A, 95(16), 9082-9086.

Banerjee, & Matthews. (1990). Cobalamin-dependent methionine synthase. FASEB J, 4(5), 1450-1459.

Bown, Altowairish, Fyans, & Bignell. (2016). Production of the *Streptomyces scabies* coronafacoyl phytotoxins involves a novel biosynthetic pathway with an F420 -dependent oxidoreductase and a short-chain dehydrogenase/reductase. Mol Microbiol, 101(1), 122-135.

Brueckner, Armache, Cheung, Damsma, Kettenberger, Lehmann, . . . Cramer. (2009). Structure-function studies of the RNA polymerase II elongation complex. Acta Crystallogr D Biol Crystallogr, 65(Pt 2), 112-120.

Bussari, Saudagar, Shaligram, Survase, & Singhal. (2008). Production of cephamycin C by Streptomyces clavuligerus NT4 using solid-state fermentation. J Ind Microbiol Biotechnol, 35(1), 49-58.

Bustin. (2000). Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. J Mol Endocrinol, 25(2), 169-193.

Bustin. (2004). A-Z of Quantitative PCR. La Jolla, California: International University Line,.

Bustin, Benes, Garson, Hellemans, Huggett, Kubista, . . . Wittwer. (2009). The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. Clin Chem, 55(4), 611-622.

Bustin, Benes, Garson, Hellemans, Huggett, Kubista, . . . Vandesompele. (2013). The need for transparency and good practices in the qPCR literature. Nat Methods, 10(11), 1063-1067.

Bustin, & Nolan. (2004). Pitfalls of quantitative real-time reverse-transcription polymerase chain reaction. J Biomol Tech, 15(3), 155-166.

Butala, Zgur-Bertok, & Busby. (2009). The bacterial LexA transcriptional repressor. Cell Mol Life Sci, 66(1), 82-93.

Chalancon, Ravarani, Balaji, Martinez-Arias, Aravind, Jothi, & Babu. (2012). Interplay between gene expression noise and regulatory network architecture. Trends Genet, 28(5), 221-232.

Chu, & Corey. (2012). RNA sequencing: platform selection, experimental design, and data interpretation. Nucleic Acid Ther, 22(4), 271-274.

Cohen, Yanko, Mislovati, Argaman, Schreiber, Av-Gay, & Aharonowitz. (1993). Thioredoxin-thioredoxin reductase system of *Streptomyces clavuligerus*: sequences, expression, and organization of the genes. J Bacteriol, 175(16), 5159-5167.

Coque, Martin, Calzada, & Liras. (1991). The cephamycin biosynthetic genes pcbAB, encoding a large multidomain peptide synthetase, and pcbC of Nocardia lactamdurans are clustered together in an organization different from the same genes in Acremonium chrysogenum and Penicillium chrysogenum. Mol Microbiol, 5(5), 1125-1133.

Daoust, Onishi, Wallick, Hendlin, & Stapley. (1973). Cephamycins, a new family of beta-lactam antibiotics: antibacterial activity and resistance to beta-lactamase degradation. Antimicrob Agents Chemother, 3(2), 254-261.

Demain, & Fang. (2000). The natural functions of secondary metabolites. Adv Biochem Eng Biotechnol, 69, 1-39.

Denoya, Fedechko, Hafner, McArthur, Morgenstern, Skinner, . . . Wernau. (1995). A second branched-chain alpha-keto acid dehydrogenase gene cluster (bkdFGH) from Streptomyces avermitilis: its relationship to avermectin biosynthesis and the construction of a bkdF mutant suitable for the production of novel antiparasitic avermectins. J Bacteriol, 177(12), 3504-3511.

Du, Shen, Yu, Bai, & Li. (2011). Gamma-butyrolactone regulatory system of Streptomyces chattanoogensis links nutrient utilization, metabolism, and development. Appl Environ Microbiol, 77(23), 8415-8426.

Ferla, & Patrick. (2014). Bacterial methionine biosynthesis. Microbiology, 160(Pt 8), 1571-1584.

Fournier, Paulson, Pavelka, Mosley, Gaudenz, Bradford, . . . Washburn. (2010). Delayed correlation of mRNA and protein expression in rapamycin-treated cells and a role for Ggc1 in cellular sensitivity to rapamycin. Mol Cell Proteomics, 9(2), 271-284.

Freeman, Walker, & Vrana. (1999). Quantitative RT-PCR: pitfalls and potential. Biotechniques, 26(1), 112-122, 124-115.

Futcher, Latter, Monardo, McLaughlin, & Garrels. (1999). A sampling of the yeast proteome. Mol Cell Biol, 19(11), 7357-7368.

Gandhi, Zenklusen, Lionnet, & Singer. (2011). Transcription of functionally related constitutive genes is not coordinated. Nat Struct Mol Biol, 18(1), 27-34.

Ghaemmaghami, Huh, Bower, Howson, Belle, Dephoure, . . . Weissman. (2003). Global analysis of protein expression in yeast. Nature, 425(6959), 737-741.

Glickman, & Ciechanover. (2002). The ubiquitin-proteasome proteolytic pathway: destruction for the sake of construction. Physiol Rev, 82(2), 373-428.

Golding, Paulsson, Zawilski, & Cox. (2005). Real-time kinetics of gene activity in individual bacteria. Cell, 123(6), 1025-1036.

Greenbaum, Colangelo, Williams, & Gerstein. (2003). Comparing protein abundance and mRNA expression levels on a genomic scale. Genome Biol, 4(9), 117.

Guo, Zhang, Luo, He, Chen, Wen, & Li. (2013). A novel TetR family transcriptional regulator, SAV576, negatively controls avermectin biosynthesis in Streptomyces avermitilis. PLoS One, 8(8), e71330.

Hamed, Gomez-Castellanos, Henry, Ducho, McDonough, & Schofield. (2013). The enzymes of beta-lactam biosynthesis. Nat Prod Rep, 30(1), 21-107.

Hillerich, & Westpheling. (2008). A new TetR family transcriptional regulator required for morphogenesis in Streptomyces coelicolor. J Bacteriol, 190(1), 61-67.

Hong, Phornphisutthimas, Tilley, Baumberg, & McDowall. (2007). Streptomycin production by Streptomyces griseus can be modulated by a mechanism not associated with change in the adpA component of the A-factor cascade. Biotechnol Lett, 29(1), 57-64.

Hopwood. (2006). Soil to genomics: the Streptomyces chromosome. Annu Rev Genet, 40, 1-23.

Hornung, & Barkai. (2008). Noise propagation and signaling sensitivity in biological networks: a role for positive feedback. PLoS Comput Biol, 4(1), e8.

Huggett, & O'Grady. (2014). Molecular Diagnostics: Current Research and Applications (Huggett and O'Grady Eds.). LGC, Teddington, UK and Norwich Medical School, University of East Anglia, Norwich, UK: Caister Academic Press.

Hughes, & Mathee. (1998). The anti-sigma factors. Annu Rev Microbiol, 52, 231-286.

ITQB. (2018). Bacterial Cell Biology from http://www.itqb.unl.pt/labs/bacterial-cell-biology

Jayapal, Philp, Kok, Yap, Sherman, Griffin, & Hu. (2008). Uncovering genes with divergent mRNA-protein dynamics in Streptomyces coelicolor. PLoS One, 3(5), e2097.

Jenke-Kodama, Borner, & Dittmann. (2006). Natural biocombinatorics in the polyketide synthase genes of the actinobacterium Streptomyces avermitilis. PLoS Comput Biol, 2(10), e132.

Jensen. (2012). Biosynthesis of clavam metabolites. J Ind Microbiol Biotechnol, 39(10), 1407-1419.

Johnson-Winters, Purpero, Kavana, Nelson, & Moran. (2003). (4-Hydroxyphenyl)pyruvate dioxygenase from Streptomyces avermitilis: the basis for ordered substrate addition. Biochemistry, 42(7), 2072-2080.

Kenig, & Reading. (1979). Holomycin and an antibiotic (MM 19290) related to tunicamycin, metabolites of Streptomyces clavuligerus. J Antibiot (Tokyo), 32(6), 549-554.

Koussounadis, Langdon, Um, Harrison, & Smith. (2015). Relationship between differentially expressed mRNA and mRNA-protein correlations in a xenograft model system. Scientific Reports, 5, 10775.

Langlois. (2003). Identification of Streptomyces coelicolor Proteins That Are Differentially Expressed in the Presence of Plant Material. Applied and Environmental Microbiology, 69(4), 1884-1889.

Larionov, Krause, & Miller. (2005). A standard curve based method for relative real time PCR data processing. BMC Bioinformatics, 6, 62.

Li, Chen, & Bruner. (2015). Structure and Mechanism of the Siderophore-Interacting Protein from the Fuscachelin Gene Cluster of Thermobifida fusca. Biochemistry, 54(25), 3989-4000.

Liras, Gomez-Escribano, & Santamarta. (2008). Regulatory mechanisms controlling antibiotic production in Streptomyces clavuligerus. J Ind Microbiol Biotechnol, 35(7), 667-676.

Liras, & Martin. (2006). Gene clusters for beta-lactam antibiotics and control of their expression: why have clusters evolved, and from where did they originate? Int Microbiol, 9(1), 9-19.

Liu, Beyer, & Aebersold. (2016). On the Dependency of Cellular Protein Levels on mRNA Abundance. Cell, 165(3), 535-550.

Livak, & Schmittgen. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods, 25(4), 402-408.

MacKenzie. (2007). Studies on the Biosynthetic Pathways of Clavulanic Acid and Cephamycin C in Streptomyces clavuligerus. (PHD), Acta Universitatis, Uppsala, Sweeden.

MacKenzie, Kershaw, Hernandez, Robinson, Schofield, & Andersson. (2007). Clavulanic acid dehydrogenase: structural and biochemical analysis of the final step in the biosynthesis of the beta-lactamase inhibitor clavulanic acid. Biochemistry, 46(6), 1523-1533.

Maier, Guell, & Serrano. (2009). Correlation of mRNA and protein in complex biological samples. FEBS Lett, 583(24), 3966-3973.

Mak, Xu, & Nodwell. (2014). The expression of antibiotic resistance genes in antibiotic-producing bacteria. Mol Microbiol, 93(3), 391-402.

Medema, Alam, Heijne, van den Berg, Muller, Trefzer, Takano. (2011). Genome-wide gene expression changes in an industrial clavulanic acid overproduction strain of Streptomyces clavuligerus. Microb Biotechnol, 4(2), 300-305.

Medema, Trefzer, Kovalchuk, van den Berg, Muller, Heijne, Takano. (2010). The sequence of a 1.8-mb bacterial linear plasmid reveals a rich evolutionary reservoir of secondary metabolic pathways. Genome Biol Evol, 2, 212-224.

Mellado, Lorenzana, Rodriguez-Saiz, Diez, Liras, & Barredo. (2002). The clavulanic acid biosynthetic cluster of Streptomyces clavuligerus: genetic organization of the region upstream of the car gene. Microbiology, 148(Pt 5), 1427-1438.

Mingeot-Leclercq, Glupczynski, & Tulkens. (1999). Aminoglycosides: activity and resistance. Antimicrob Agents Chemother, 43(4), 727-737.

Oster, Lester, Terwisscha van Scheltinga, Svenda, van Lun, Genereux, & Andersson. (2006). Insights into cephamycin biosynthesis: the crystal structure of CmcI from Streptomyces clavuligerus. J Mol Biol, 358(2), 546-558.

Özcengiz, & Demain. (2013). Recent advances in the biosynthesis of penicillins, cephalosporins and clavams and its regulation. Biotechnol Adv, 31(2), 287-311.

Ozcengiz, Okay, Unsaldi, Taskin, Liras, & Piret. (2010). Homologous expression of aspartokinase (ask) gene in Streptomyces clavuligerus and its hom-deleted mutant: effects on cephamycin C production. Bioeng Bugs, 1(3), 191-197.

Pfaffl. (2001). A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res, 29(9), e45.

Pinho, Kjos, & Veening. (2013). How to get (a)round: mechanisms controlling growth and division of coccoid bacteria. Nat Rev Microbiol, 11(9), 601-614.

Pratt, Petty, Riba-Garcia, Robertson, Gaskell, Oliver, & Beynon. (2002). Dynamics of protein turnover, a missing dimension in proteomics. Mol Cell Proteomics, 1(8), 579-591.

Radonic, Thulke, Mackay, Landt, Siegert, & Nitsche. (2004). Guideline to reference gene selection for quantitative real-time PCR. Biochem Biophys Res Commun, 313(4), 856-862.

Reading, & Cole. (1977). Clavulanic acid: a beta-lactamase-inhiting beta-lactam from Streptomyces clavuligerus. Antimicrob Agents Chemother, 11(5), 852-857.

Saudagar, Survase, & Singhal. (2008). Clavulanic acid: a review. Biotechnol Adv, 26(4), 335-351.

Scheffers, & Pinho. (2005). Bacterial cell wall synthesis: new insights from localization studies. Microbiol Mol Biol Rev, 69(4), 585-607.

Schmittgen, & Livak. (2008). Analyzing real-time PCR data by the comparative C(T) method. Nat Protoc, 3(6), 1101-1108.

Seibert. (2005). Structural and catalytic diversity within the amidohydrolase superfamily. Biochemistry, 44(17), 6383-6391.

Song, Jeong, Yu, Fischbach, Park, Kim, . . . Kim. (2010). Draft genome sequence of Streptomyces clavuligerus NRRL 3585, a producer of diverse secondary metabolites. J Bacteriol, 192(23), 6317-6318.

Tahlan, Park, Wong, Beatty, & Jensen. (2004). Two sets of paralogous genes encode the enzymes involved in the early stages of clavulanic acid and clavam metabolite biosynthesis in Streptomyces clavuligerus. Antimicrob Agents Chemother, 48(3), 930-939.

Tang, Lee, Tang, Kim, Mathews, & Khosla. (2006). Structural and functional studies on SCO1815: a beta-ketoacyl-acyl carrier protein reductase from Streptomyces coelicolor A3(2). Biochemistry, 45(47), 14085-14093.

Thoden, Frey, & Holden. (1996). Molecular structure of the NADH/UDP-glucose abortive complex of UDP-galactose 4-epimerase from Escherichia coli: implications for the catalytic mechanism. Biochemistry, 35(16), 5137-5144.

Unsaldı. (2016). Proteomic Analysis of Two Cephamycin C Overproducer and an Industrial Clavulanic Acid Overproducer Strains of Streptomyces clavuligerus in Comparison with the Standart Strain NRRL 3585. (Phd Phd), Middle East Technical University, Ankara / Turkey.

Unsaldi, Kurt-Kizildogan, Voigt, Becher, & Ozcengiz. (2017). Proteome-wide alterations in an industrial clavulanic acid producing strain of Streptomyces clavuligerus. Synth Syst Biotechnol, 2(1), 39-48.

Vogel, Silva, & Marcotte. (2011). Protein expression regulation under oxidative stress. Mol Cell Proteomics, 10(12), M111 009217.

Wacker, & Godard. (2005). Analysis of one-step and two-step real-time RT-PCR using SuperScript III. J Biomol Tech, 16(3), 266-271.

Wang, Gerstein, & Snyder. (2009). RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet, 10(1), 57-63.

Whitfield, Steers, & Weissbach. (1970). Purification and properties of 5-methyltetrahydropteroyltriglutamate-homocysteine transmethylase. J Biol Chem, 245(2), 390-401.

Whitney, Brannon, Mabe, & Wicker. (1972). Incorporation of labeled precursors into A16886B, a novel -lactam antibiotic produced by Streptomyces clavuligerus. Antimicrob Agents Chemother, 1(3), 247-251.

Wilderman, Vasil, Johnson, & Vasil. (2001). Genetic and biochemical analyses of a eukaryotic-like phospholipase D of Pseudomonas aeruginosa suggest horizontal acquisition and a role for persistence in a chronic pulmonary infection model. Mol Microbiol, 39(2), 291-303.

Wong, & Houry. (2004). Chaperone networks in bacteria: analysis of protein homeostasis in minimal cells. J Struct Biol, 146(1-2), 79-89.

Wyszynski, Lee, Yabe, Wang, Gomez-Escribano, Bibb, . . . Davis. (2012). Biosynthesis of the tunicamycin antibiotics proceeds via unique exo-glycal intermediates. Nat Chem, 4(7), 539-546.

APPENDIX A

A.1. 1X TE Buffer

10 mM Tris, bring to pH 8.0 with HCl

1 mM EDTA

A.2. 1X TE Buffer with lysozyme

10 mM Tris, bring to pH 8.0 with HCl

1 mM EDTA

2 mg/ml Lysozyme (lyophilized)

A.3. Lysogeny broth (LB) (1L)

Dissolve following in 1L dH_2O

10 g tryptone

5 g yeast extract

10 g NaCl

autoclave at 121 ^{0}C for 15 min

A.4. Starch-Asparagine Medium (1L)

600 ml of dH_2O boiled, following added;

10 g Starch

21 g MOPS

4.4 g K_2HPO_4

Solution cooled to room temperature,

Completed to 800 ml with dH_2O

Autoclave at 121 ^{0}C for 15 min.

Add following components;

200 ml L-Asparagine (10g/L)

2 ml $MgSO_4.7H_2O$ (0.6 g/L)

1 ml Trace element solution*

A.5. *Trace element solution (1L)

1 g $FeSO_4.7H_2O$

1 g $MnCl_2.4H_2O$

1 g $ZnSO_4.7H_2O$

1.3 g $CaCl_2.3H_2O$

A.6. TSB (1L)

Dissolve following in 1L dH_2O

21 g TSB

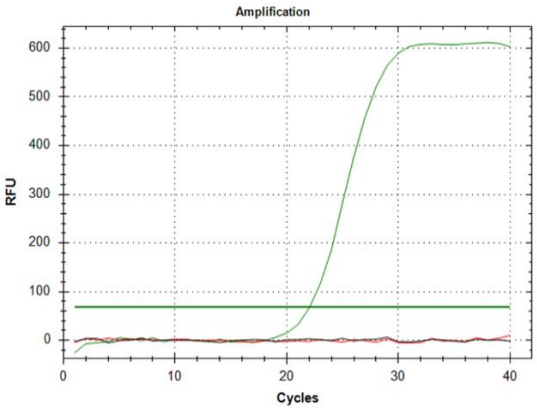
autoclave at 121 ^{0}C for 15 min

APPENDIX B

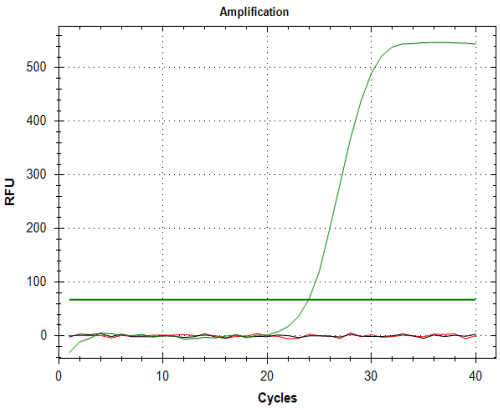
B.1. Relative gene expression in fold changes

Gene	Expression	Gene	Expression
SCLAV_5275	1.11	**SCLAV_0453**	-7.25
SCLAV_0043	13.84	**SCLAV_2046**	-6.24
cmcJ	14.48	**SCLAV_3034**	-5.01
lexA	5.48	**SCLAV_2046**	-6.24
SCLAV_5667	-3.24	**tig**	-10.60
bkdA1	-5.18	**SCLAV_2541**	-6.20
SCLAV_5246	-6.15	**SCLAV_2541**	-6.20
bkdB1	-3.49	**SCLAV_0843**	-8.26
SCLAV_3034	-5.01	**SCLAV_0130**	-8.52
SCLAV_2986	-6.25	**SCLAV_4756**	-9.23
avaA2	-6.88	**SCLAV_p0693**	1924.56
SCLAV_4282	-6.35	**SCLAV_p1324**	165.12
fab1	-9.42	**secA**	-6.79

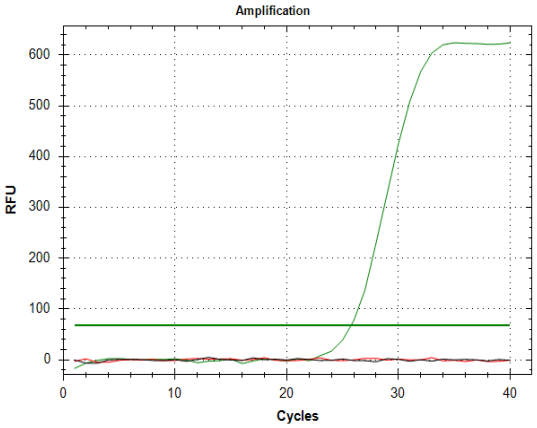
B.2. Amplification curves for *hrdB*(1), *SCLAV_0043*(2), *SCLAV_5275*(3), *cmcJ*(4), and *SCLAV_4282*(5). Target(green), NTC(red), RTC(black)



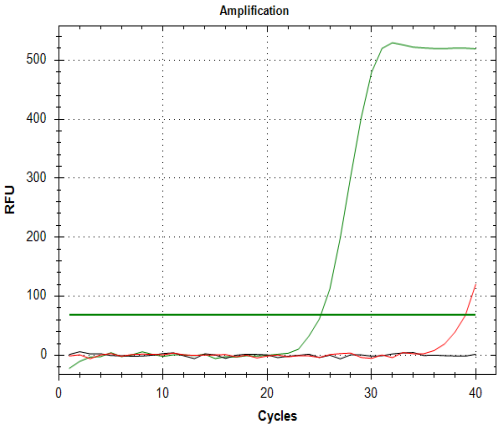
1



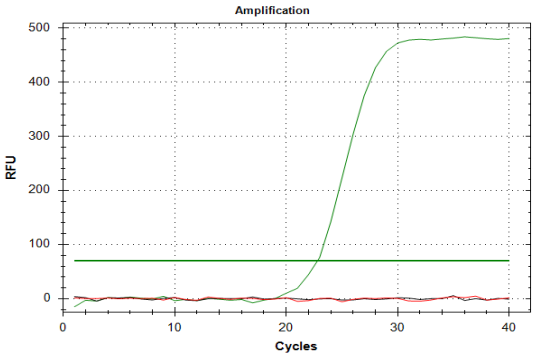
2



3

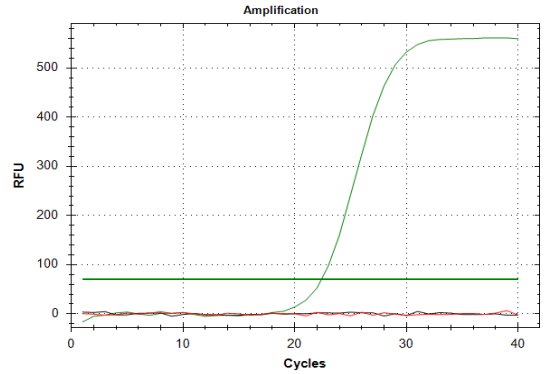


4

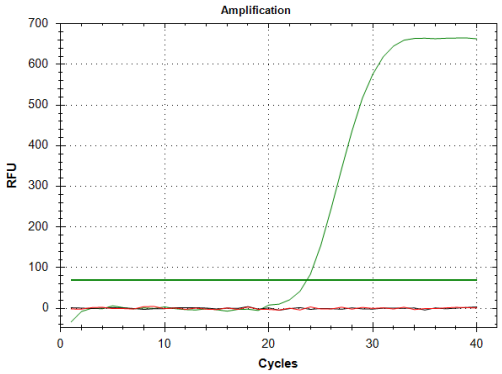


5

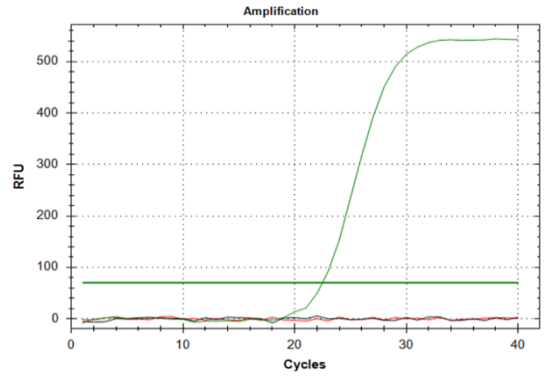
B.3. Amplification curves for *SCLAV_2046*(1), *ectD*(2), *tig*(3), *fab1*(4), and *SCLAV_1945*(5). Target(green), NTC(red), RTC(black)



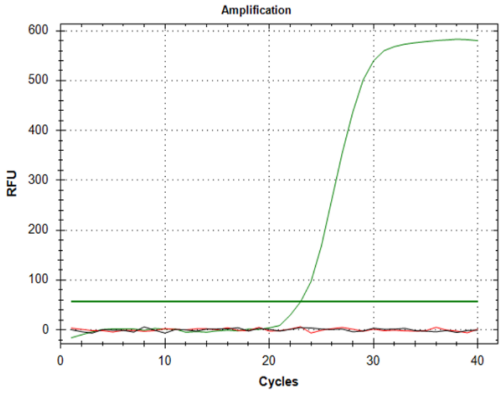
1



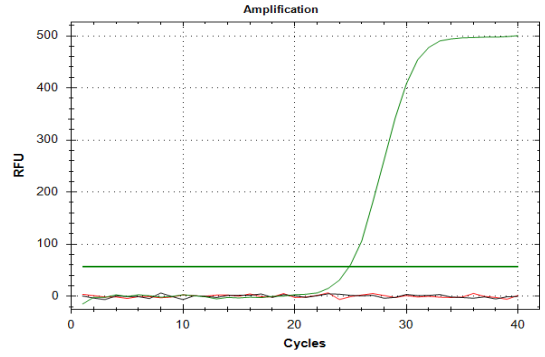
2



3

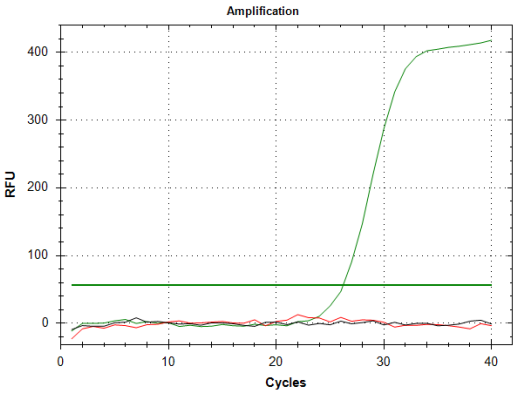


4

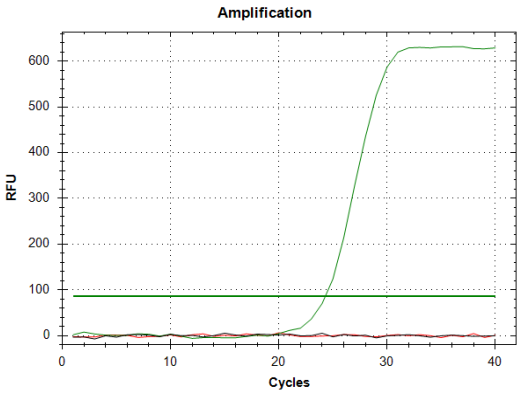


5

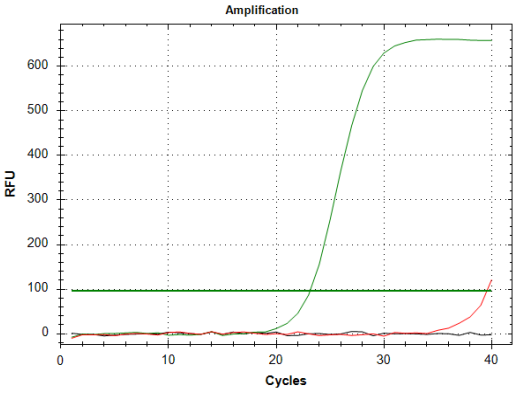
B.4. Amplification curves for *bkdA1*(1), *SCLAV_0130*(2), *SCLAV_0453*(3), *SCLAV_0843*(4), and *SCLAV_2541*(5). Target(green), NTC(red), RTC(black)



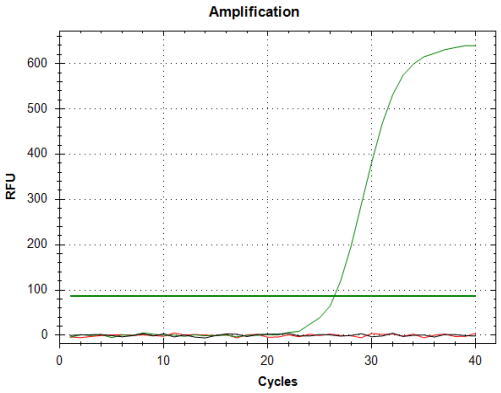
1



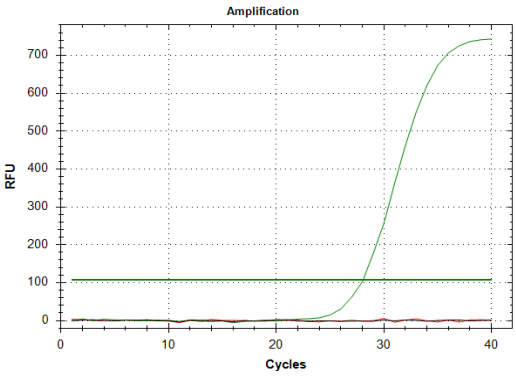
2



3

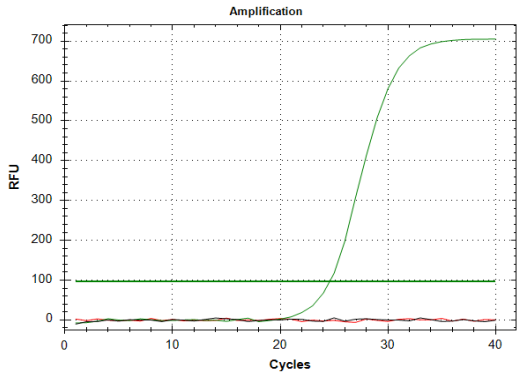


4

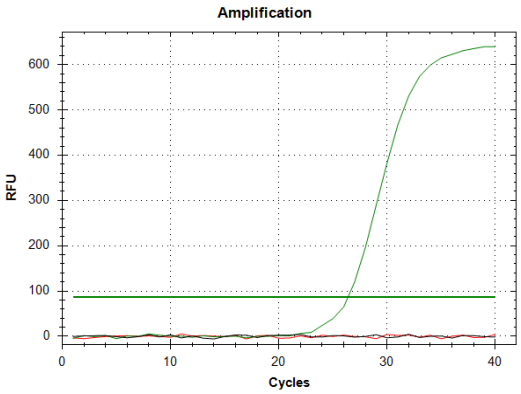


5

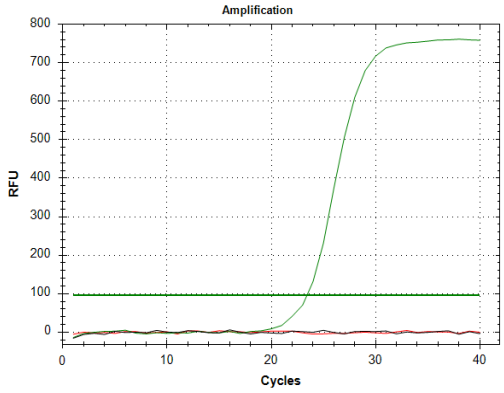
B.5. Amplification curves for *SCLAV_2986*(1), *SCLAV_3034*(2), *SCLAV_4756*(3), *SCLAV_5246*(4), and *SCLAV_5667*(5). Target(green), NTC(red), RTC(black)



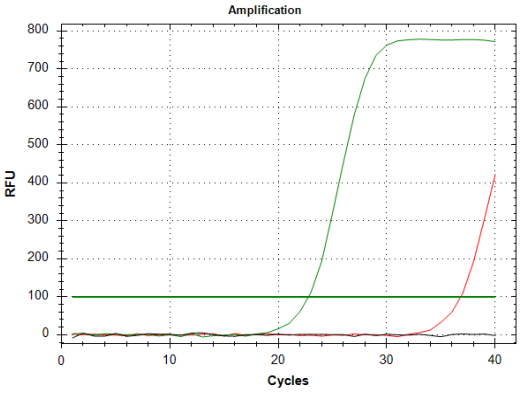
1



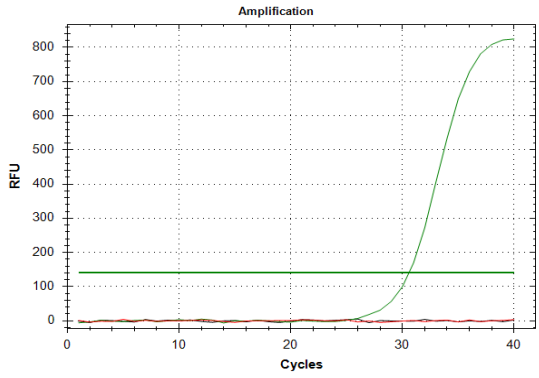
2



3

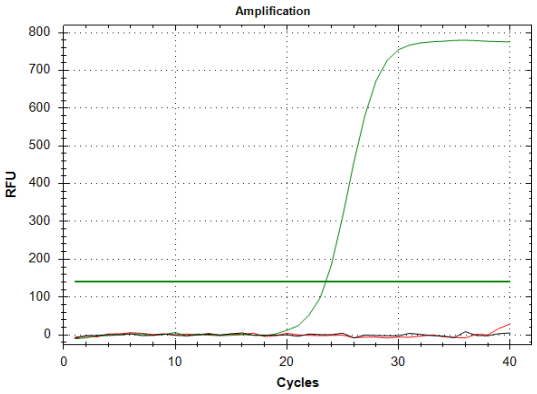


4

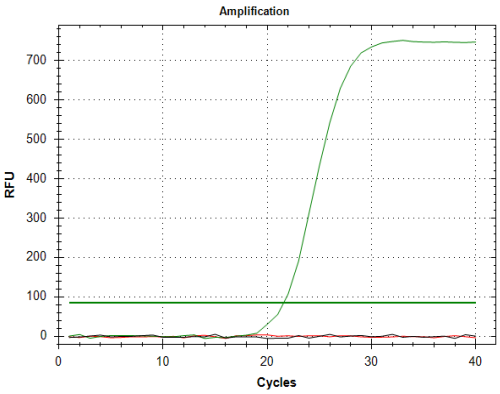


5

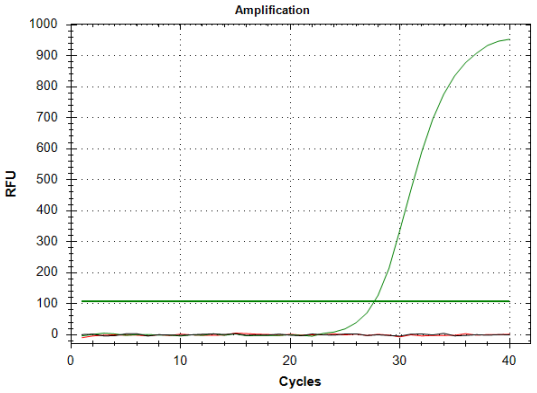
B.6. Amplification curves of *avaA2*(1), *bkdB1*(2), *SCLAV_p0693*(3), *SCLAV_p1324*(4), *secA*(5), and *lexA*(6) Target(green), NTC(red), RTC(black)



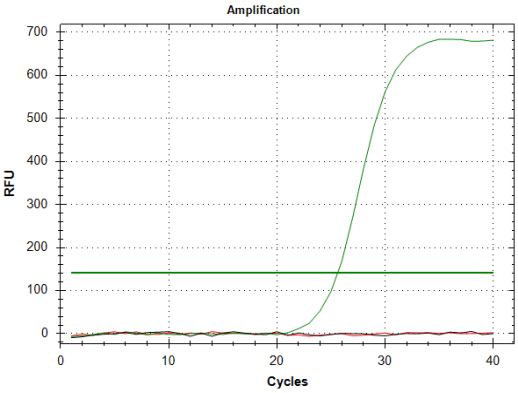
1



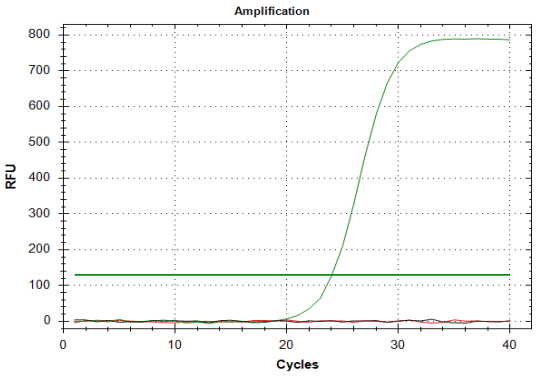
2



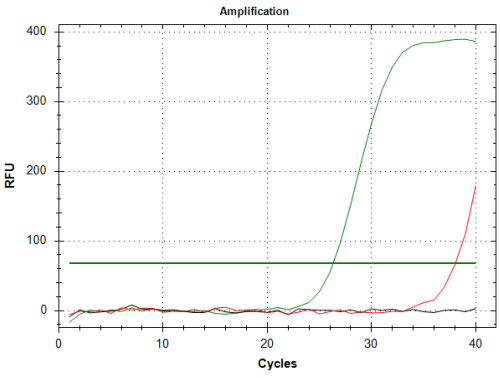
3



4

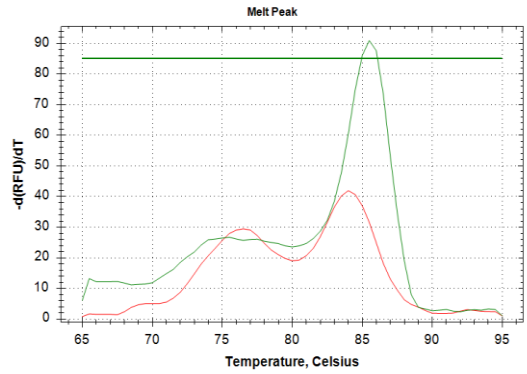


5

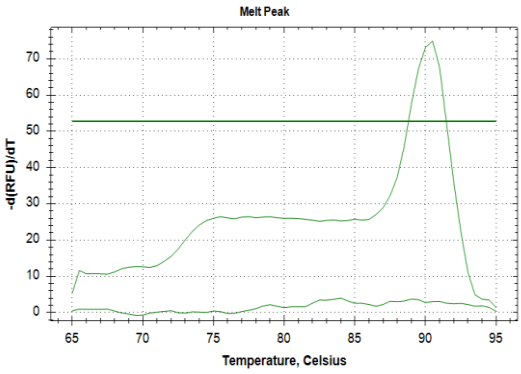


6

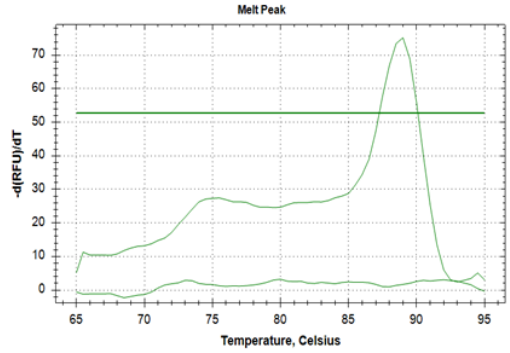
B.7. Melting peaks of *hrdB*(1), *SCLAV_0043*(2), *SCLAV_5275*(3), *cmcJ*(4), *SCLAV_4282*(5). Target(green), NTC(red)



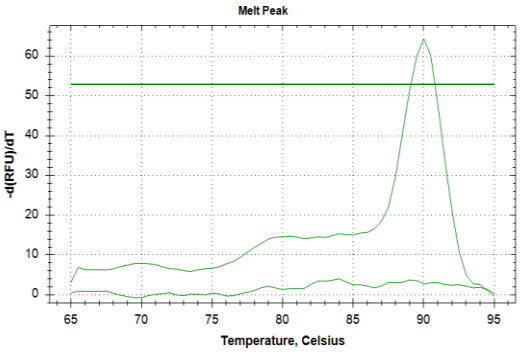
1



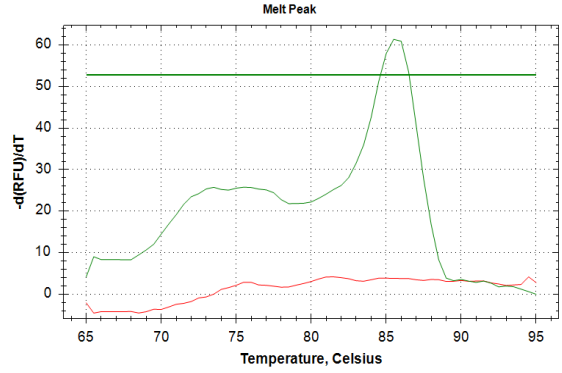
2



3

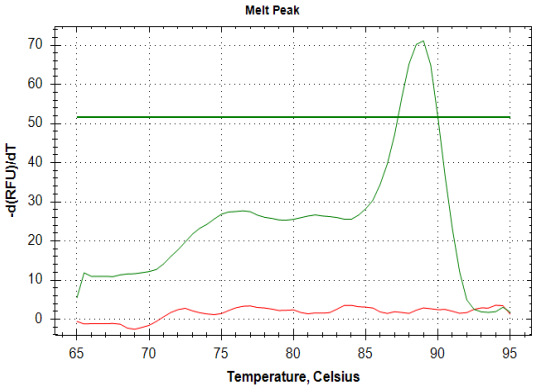


4

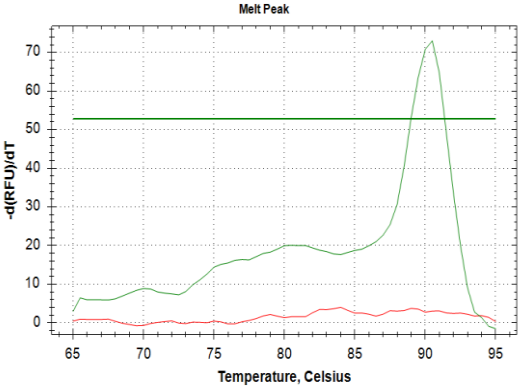


5

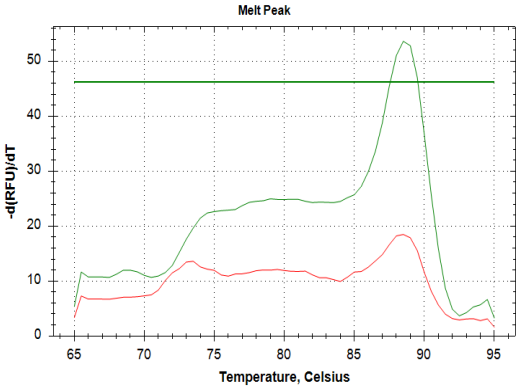
B.8. Melting peaks of *SCLAV_2046*(1), *ectD*(2), *tig*(3), *fab1*(4), *SCLAV_1945*(5). Target(green), NTC(red)



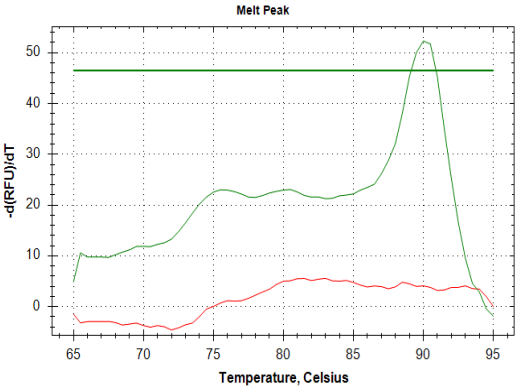
1



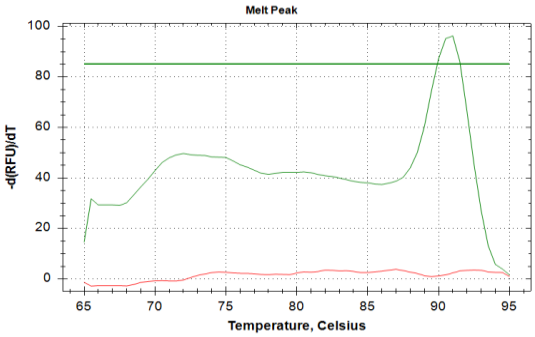
2



3

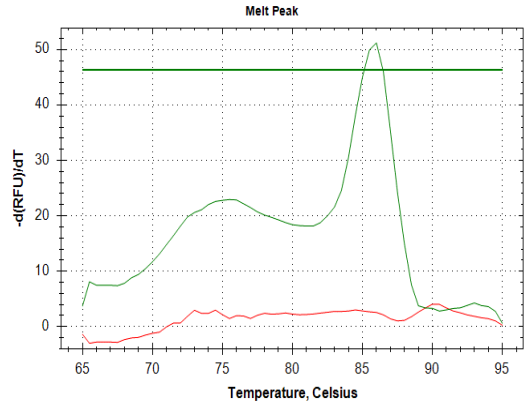


4

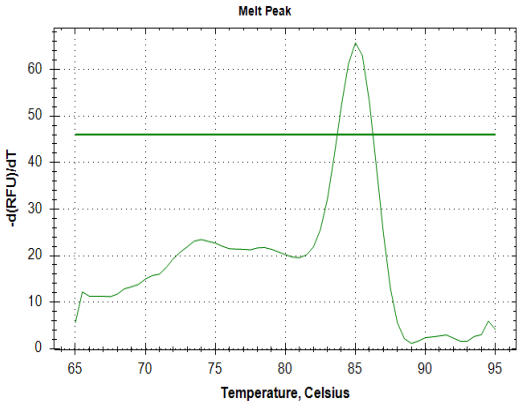


5

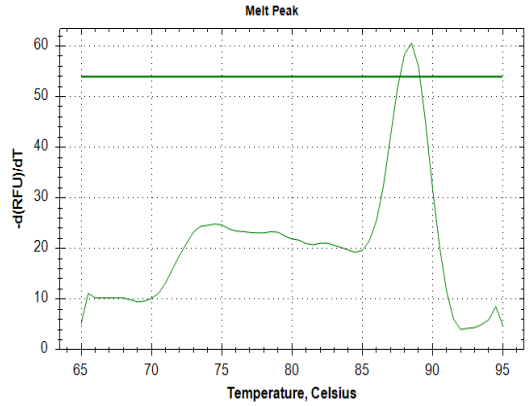
B.9. Melting peaks of *bkdA1*(1), *SCLAV_0130*(2), *SCLAV_0453*(3), *SCLAV_0843*(4), and *SCLAV_2541*(5). Target(green), NTC(red)



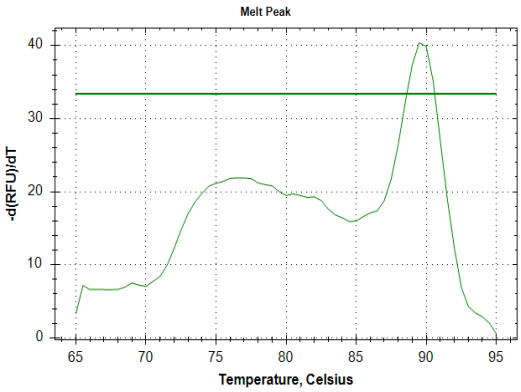
1



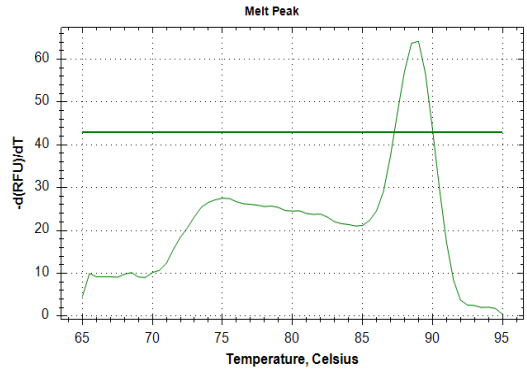
2



3

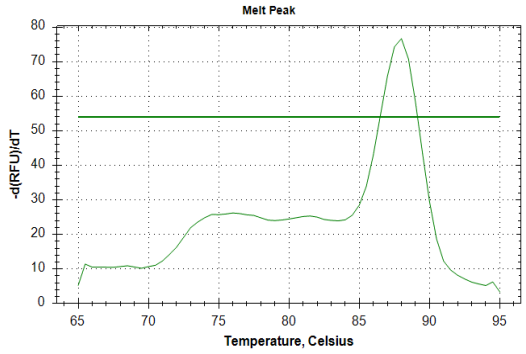


4

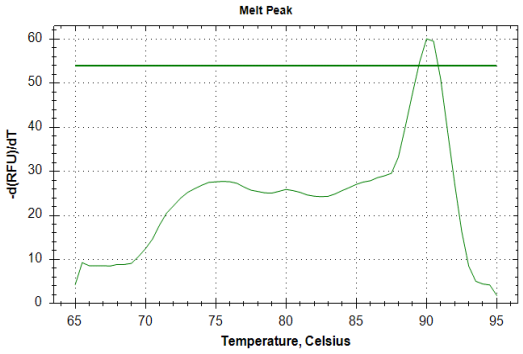


5

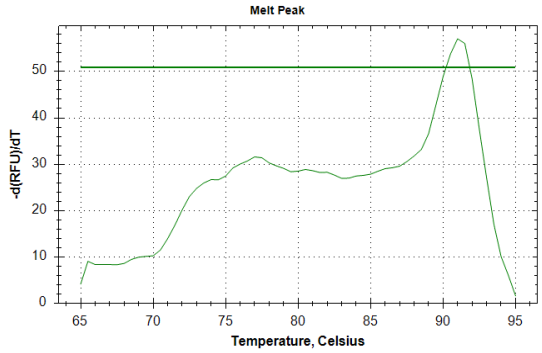
B.10. Melting peaks of *SCLAV_2986*(1), *SCLAV_3034*(2), *SCLAV_4756*(3), *SCLAV_5246*(4), and *SCLAV_5667*(5). Target(green), NTC(red)



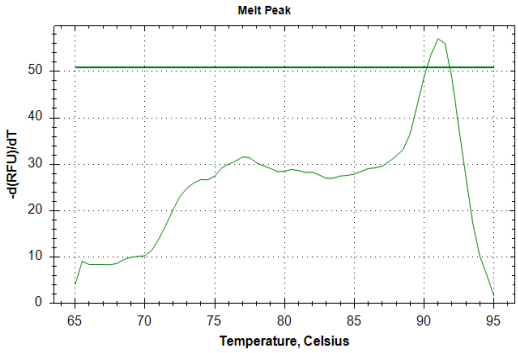
1



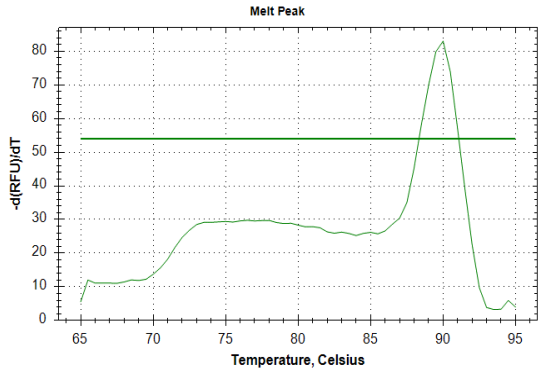
2



3

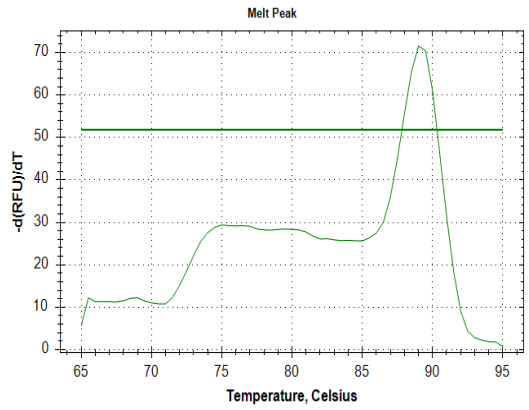


4

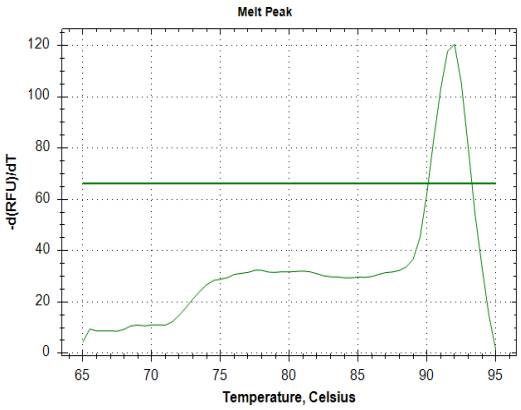


5

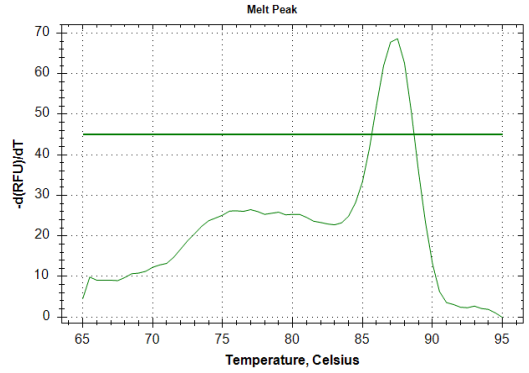
B.11. Melting peaks of *avaA2*(1), *bkdB1*(2), *SCLAV_p0693*(3), *SCLAV_p1324*(4), *secA*(5), and *lexA*(6) Target(green), NTC(red)



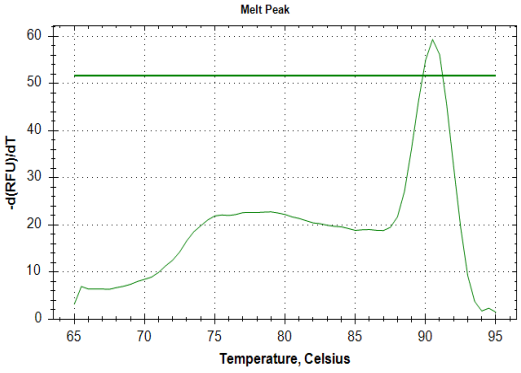
1



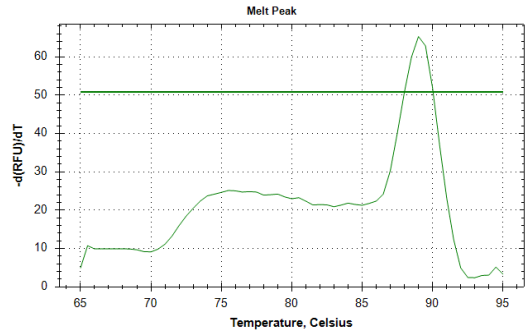
2



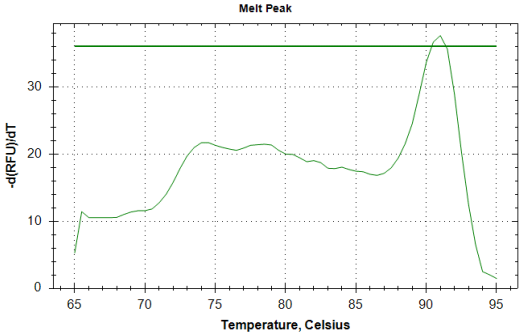
3



4

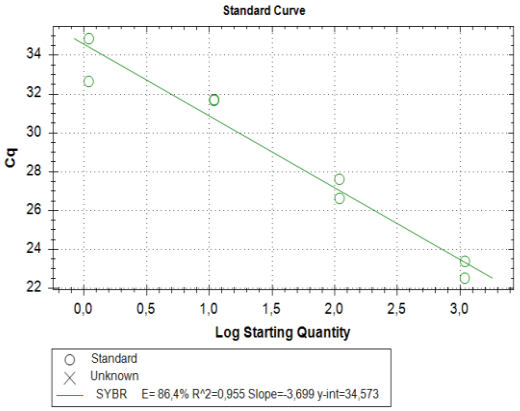


5

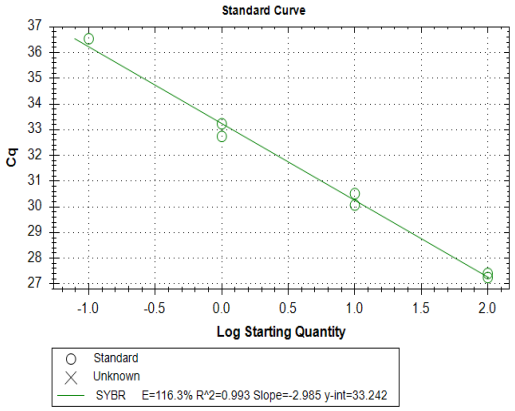


6

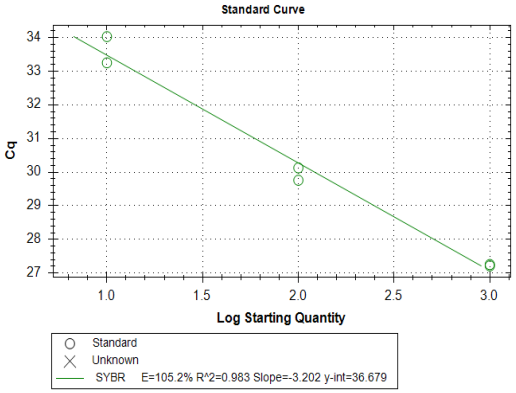
B.12. Amplification efficiency standard curves for *hrdB*(1), *SCLAV_0043*(2), *SCLAV_5275*(3), *cmcJ*(4), and *SCLAV_4282*(5)



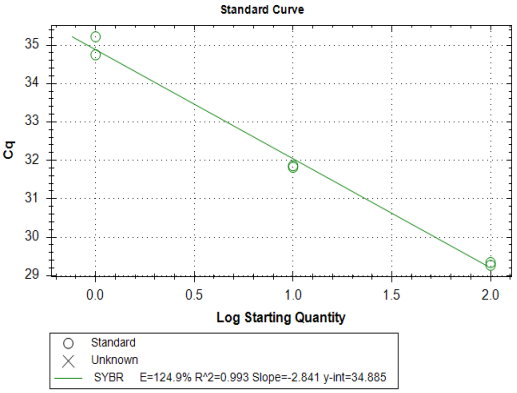
1



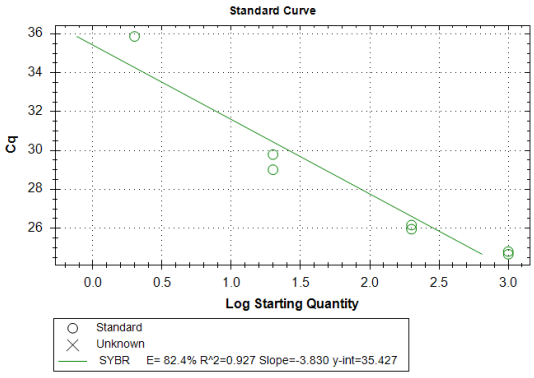
2



3

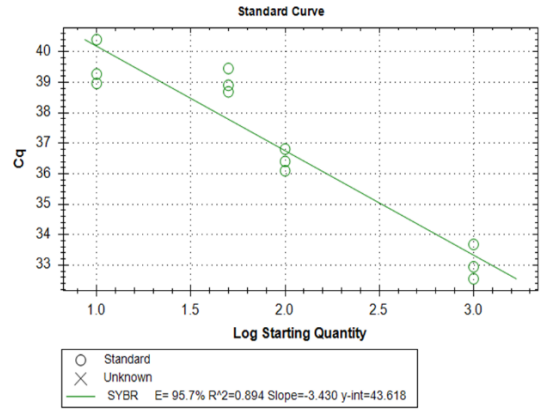


4

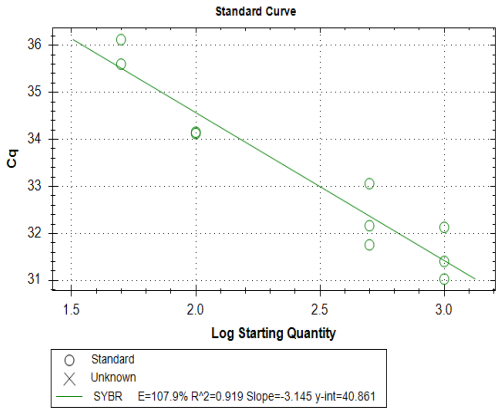


5

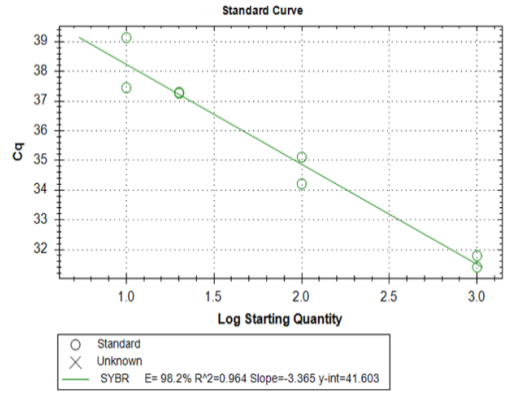
B.13. Standard curves for *SCLAV_2046*(1), *ectD*(2), *tig*(3), *fab1*(4), and *SCLAV_1945*(5)



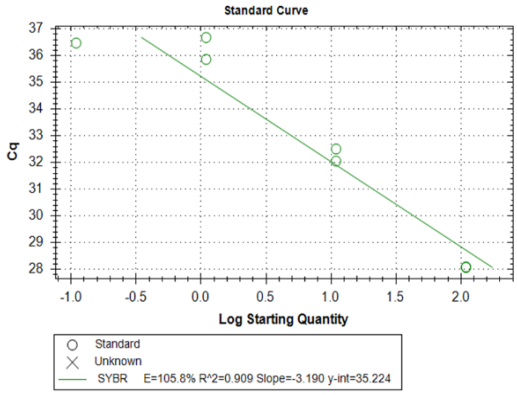
1



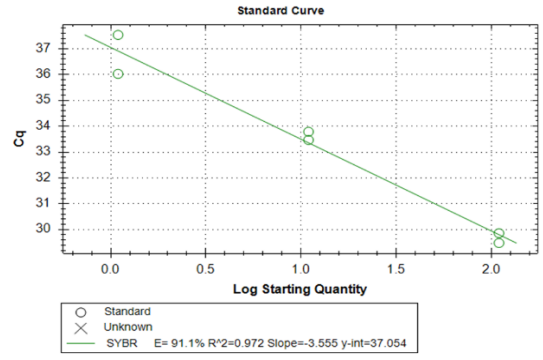
2



3

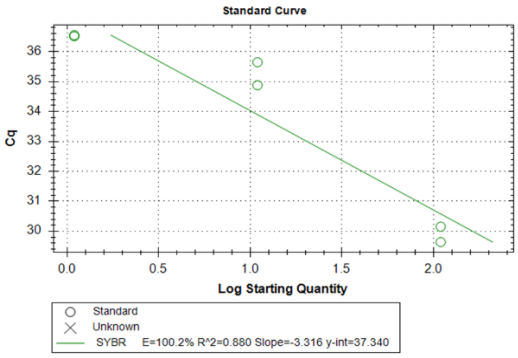


4

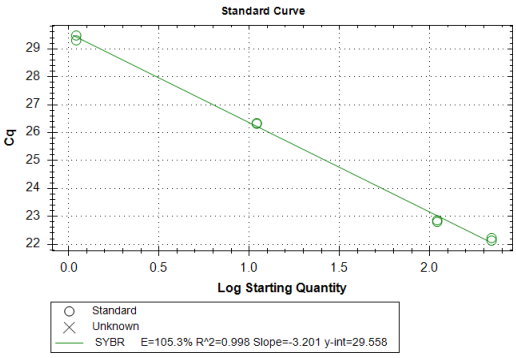


5

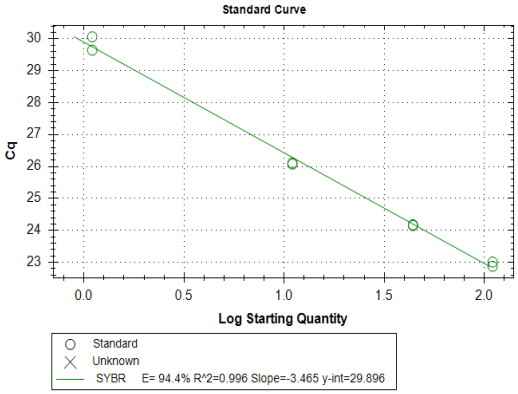
B.14. Standard curves for *bkdA1*(1), *SCLAV_0130*(2), *SCLAV_0453*(3), *SCLAV_0843*(4), and *SCLAV_2541*(5)



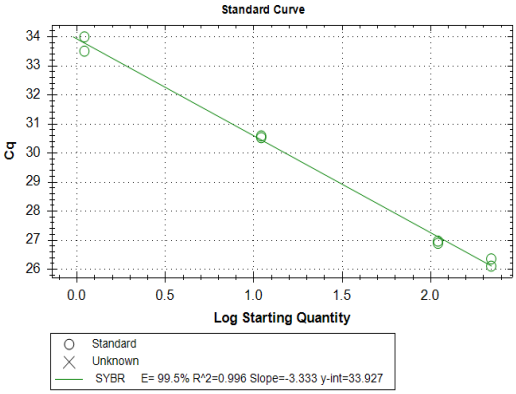
1



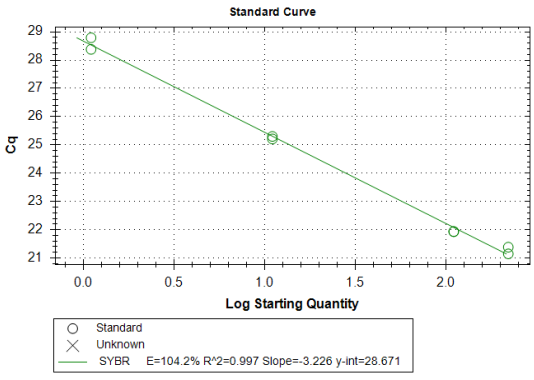
2



3

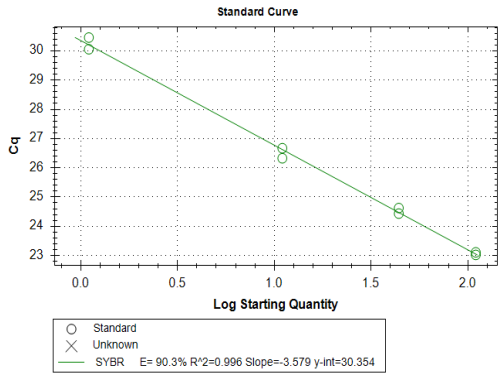


4

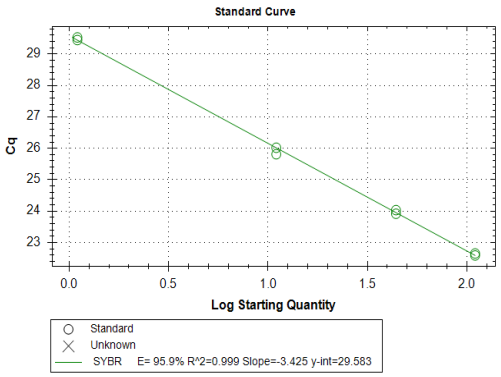


5

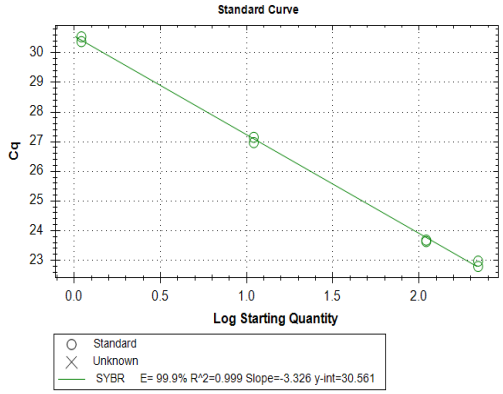
B.15. Standard curves for *SCLAV_2986*(1), *SCLAV_3034*(2), *SCLAV_4756*(3), *SCLAV_5246*(4), and *SCLAV_5667*(5)



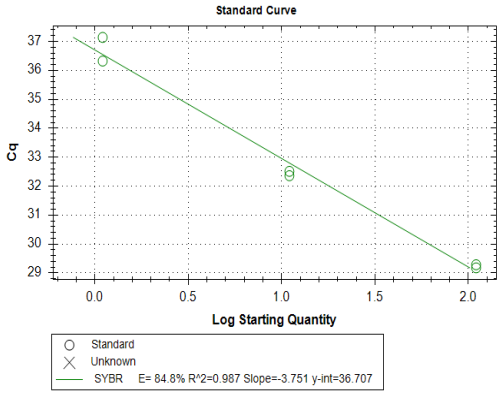
1



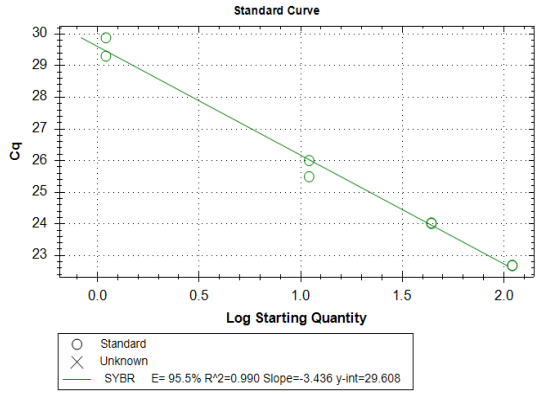
2



3

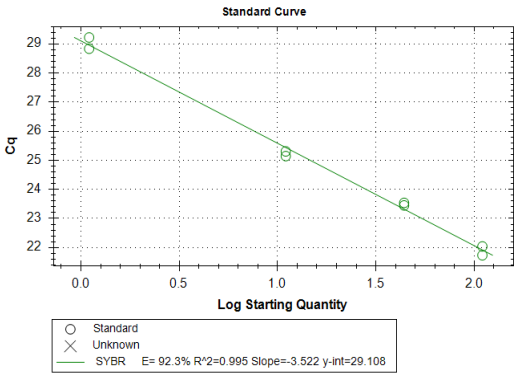


4

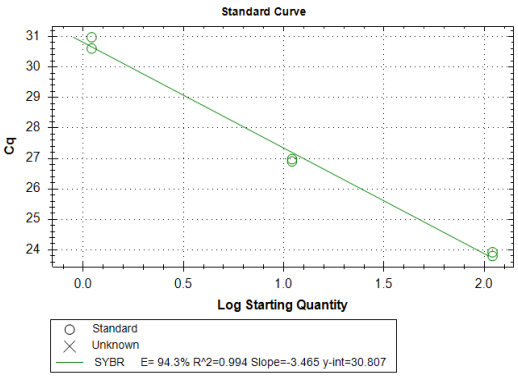


5

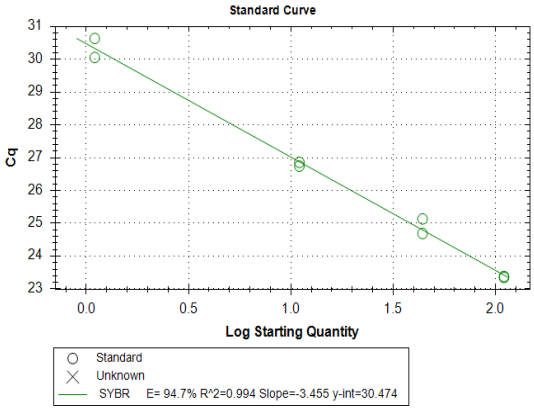
B.16. Standard curves for *avaA2*(1), *bkdB1*(2), *SCLAV_p0693*(3), *SCLAV_p1324*(4), *secA*(5), and *lexA*(6)



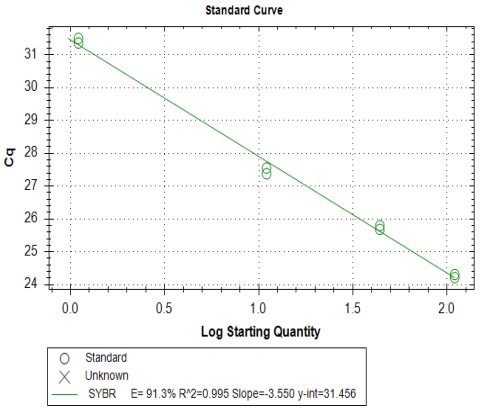
1



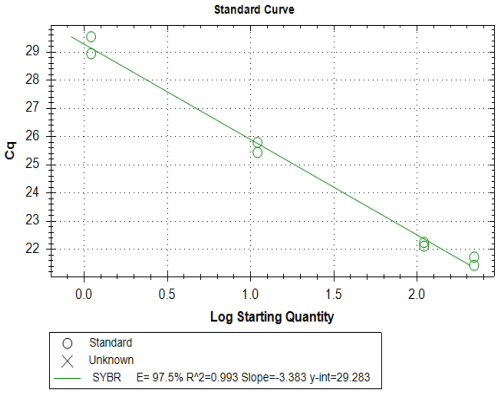
2



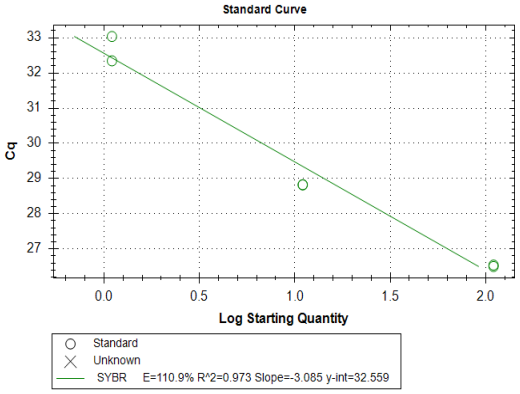
3



4



5



6

B.17. List of primers used in RT-qPCR experiments

Oligo name	Sequence 5' to 3'	Bp	Tm °C
hrdB-F	CGCGGCATGCTCTTCCT	17	58
hrdB-R	AGGTGGCGTACGTGGAGAAC	20	61
lexA-F	GATCGAAGCCGCCATCTG	18	58
lexA-R	GCCGTTCTCCCGTTTGAA	18	56
cmcj-R	GAGACCGAAGAGCTGCTGAC	20	61
cmcj-F	AGATGCGGCGACTGATGG	18	58
SCLAV_5275-F	GCAGGCACGGCAGGAAA	17	58
SCLAV_5275-R	GACCAGCGCGGCATTGA	17	58
SCLAV_0043-F	GCCCAAGATCTACCGCAATA	20	57
SCLAV_0043-R	CGTCGTCCTGGTTCTTCTG	19	59
SCLAV_4282-F	CCTACGGTGAGTCGAAGTATCT	22	60
SCLAV_4282-R	CGTTGAAGTAGCGGAGGTTG	20	59
SCLAV_2046-F	CTTCGTCTTCTCCTCGTCCT	20	59
SCLAV_2046-R	CATGCGTTCGCACAGATACT	20	57
tig-F	CTTCGTCTTCTCCTCGTCCT	20	59
tig-R	CATGCGTTCGCACAGATACT	20	57
bkdA1-F	CTCCACGCCCATCTTGT	17	55
bkdA1-R	GTGCTCGACGGGTAGAC	17	58
fab1-F	GCATGTCTCCGCGTACT	17	55
fab1-R	GCCCATCCAGTCGTACTT	18	56
SCLAV_0130-F	CGTGTACATCGAGCATCTGAC	21	60
SCLAV_0130-R	CACCCGGTCGAGCATATTG	19	59
SCLAV_P1324-F	GATATCCGCGGCTGTCTG	18	58
SCLAV_P1324-R	GGTGTTGTTGAGGTCGATGA	20	57
avaA2-F	GAGCAACTGGGCATCTTCA	19	57
avaA2-R	ACCTCCATTGTGCCACTTC	19	57
bkdB1-F	GCTCATCAGCCATGTCTCC	19	59
bkdB1-R	GGTCGCATGTAGTACGC	17	55
SCLAV_0205-F	GTACGTCCTTCAAGCAGATCAA	22	58
SCLAV_0205-R	GCCCGCAGACTGTTCAT	17	55
SCLAV_5246-F	CGGCGTGAGGAACTTCT	17	55
SCLAV_5246-R	CACCAGCTTCTCCTTGGT	18	56

Oligo name	Sequence 5' to 3'	Bp	Tm °C
SCLAV_p0693-F	CGGAGTCAGCCATGAACAG	19	59
SCLAV_p0693-R	TCCAGGGCGTCGAGATAG	18	58
SCLAV_5667-F	GGGACACCACCACGTTC	17	58
SCLAV_5667-R	AAGCCCATCTTCAGTTCGAC	20	57
SCLAV_0453-F	ATCATGGGCAACGAGTGG	18	56
SCLAV_0453-R	TTGGCCGTCATGCAGTAG	18	56
SCLAV_4756-F	TCCAGATCGTCCGGTACAAG	20	59
SCLAV_4756-R	GGTGAACGTGTCGAGCAG	18	58
SCLAV_4693-F	CCATCGCCTACTACCAGCA	19	59
SCLAV_4693-R	TACTTCCAGATCGCCTCGTC	20	59
SCLAV_2541-F	ATGGGTCTCGCCGTCAT	17	55
SCLAV_2541-R	GTGGGCCAGGTCATCCT	17	58
SCLAV_0843-F	GGGAGCACCTGTCGATCT	18	58
SCLAV_0843-R	GTTCCACTCGCGCTTCAC	18	58
secA-F	AGTCCCGCTTCTACCTCTC	19	59
secA-R	CGCGTCACCATCTTGTTCT	19	57
SCLAV_2986-F	GGTGGCTCGAAGGTCTTT	18	56
SCLAV_2986-R	GCGTATCTGCTCCAACT	17	55
SCLAV_3034-F	CGCGGAAGTGGTGAAGAC	18	58
SCLAV_3034-R	CAGGCTGATGTTGGTGATGT	20	57