ANALYSIS OF PRENYLTRANSFERASE GENE EXPRESSION
IN *PSEUDO-NITZSCHIA MULTISERIES*

A Thesis

Presented to the faculty of the Department of Chemistry
California State University, Sacramento

Submitted in partial satisfaction of
the requirements for the degree of

MASTER OF SCIENCE

in
Chemistry
(Biochemistry)

by
Brenden Jiang

SPRING
2019
ANALYSIS OF PRENYLTRANSFERASE GENE EXPRESSION
IN PSEUDO-NITZSCHIA MULTISERIES

A Thesis

by

Brenden Jiang

Approved by:

__________________________, Committee Chair
Thomas J. Savage, Ph.D.

__________________________, Second Reader
Linda M. Roberts, Ph.D.

__________________________, Third Reader
Nicholas N. Ewing, Ph.D.

__________________________
Date

iii
Student: Brenden Jiang

I certify that this student has met the requirements for format contained in the University format manual, and that this thesis is suitable for shelving in the Library and credit is to be awarded for the thesis.

__________________________, Graduate Coordinator
Susan Crawford, Ph.D. Date

Department of Chemistry
Abstract

of

ANALYSIS OF PRENYLTRANSFERASE GENE EXPRESSION

IN PSEUDO-NITZSCHIA MULTISERIES

by

Brenden Jiang

Domoic acid is a potent neurotoxin produced by the genus Pseudo-nitzschia. It is responsible for causing amnesic shellfish poisoning, a condition capable of devastating marine wildlife and humans with severe neurological complications once ingested. Much of the biosynthetic pathway, its regulation, and the biological function of domoic acid remains to be elucidated. Recent research has found \( N \)-geranyl-L-glutamic acid as the true precursor to domoic acid. This is formed by the condensation of two metabolites, geranyl diphosphate and glutamate. The former, geranyl diphosphate, is synthesized through a condensation of isopentenyl diphosphate and dimethylallyl diphosphate. This initial step is catalyzed by a geranyl diphosphate synthase, a member of the prenyltransferase family. Here, expression of three prenyltransferase genes (\( P_{semu2} \), \( P_{semu3} \), and \( P_{semu4} \)) was analyzed through quantitative polymerase chain reaction in three strains of \( P. multiseries \) isolated from different growth phases to correlate relative abundance with domoic acid accumulation. Of the three transcripts, \( P_{semu3} \) and \( P_{semu4} \) were previously identified as potential geranyl diphosphate synthases in homology modeling studies. \( P_{semu2} \) and \( P_{semu3} \) consistently had higher relative expression over
Psemu4 in all three strains. Prenyltransferase expression also decreased as domoic acid concentration reached an apparent maximum in late-log and stationary phase. This demonstrates that prenyltransferase transcripts Psemu2 and Psemu3 are highly expressed in log phase, and possibly may play a role in domoic acid biosynthesis.

_______________________, Committee Chair
Thomas J. Savage, Ph.D.

_______________________
Date
ACKNOWLEDGEMENTS

I would first like to express my sincere gratitude to my thesis advisor, Dr. Thomas Savage for the continuous support, guidance, mentoring, and patience. I am grateful to have worked with him in exploring the wonders of *Pseudo-nitzschia* and domoic acid. Thank you for shaping me into a stronger scientist.

I would also like to thank my committee members, Dr. Linda Roberts and Dr. Nicholas Ewing for their comments, support, and advice when I went astray.

I would like to acknowledge from the chemistry department: John Disney, Christian Martinez, and Kathleen Le. Thank you for always accommodating needs of the project. Furthermore, from the biology department: Sulie Ober and Caity Fox. Thank you for always accommodating me for the gene expression analysis step.

I would also like to acknowledge Dr. Katherine McReynolds and her students: Dustin Dimas, Cory Vierra, and James Cerney for their endless support in the laboratory.

A thank you to my fellow lab mates, Adriana Badarou, Morgan Campbell, and Chris Peery for their help during critical points in experimentation.

A special thanks to Charleen Fain-Keslar, Janice Yonekura, and Mike Namba for always pushing me to achieve the best.

Thank you to my family and friends for their support. Most importantly, thank you to my wife, Kayli Dam, whose unwavering love and support for my endeavors has helped keep me moving forward in the face of the greatest challenges.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acknowledgements</td>
<td>vii</td>
</tr>
<tr>
<td>List of Tables</td>
<td>x</td>
</tr>
<tr>
<td>List of Figures</td>
<td>xi</td>
</tr>
<tr>
<td><strong>Chapter</strong></td>
<td></td>
</tr>
<tr>
<td>1. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>1.1 Overview of Diatoms</td>
<td>2</td>
</tr>
<tr>
<td>1.2 <em>Pseudo-nitzschia</em> (<em>P-n</em>)</td>
<td>3</td>
</tr>
<tr>
<td>1.2.1 Distribution</td>
<td>3</td>
</tr>
<tr>
<td>1.2.2 Life Cycle</td>
<td>5</td>
</tr>
<tr>
<td>1.2.3 Toxicity</td>
<td>8</td>
</tr>
<tr>
<td>1.3 Domoic Acid Chemistry and Biochemistry</td>
<td>11</td>
</tr>
<tr>
<td>1.3.1 Original Isolation &amp; Use</td>
<td>11</td>
</tr>
<tr>
<td>1.3.2 General Isoprenoid Biochemistry</td>
<td>11</td>
</tr>
<tr>
<td>1.3.3 Prenyltransferases</td>
<td>13</td>
</tr>
<tr>
<td>1.3.4 Domoic Acid Biosynthesis</td>
<td>14</td>
</tr>
<tr>
<td>1.4 Measuring Gene Expression with Quantitative Real Time PCR (qPCR)</td>
<td>15</td>
</tr>
<tr>
<td>1.4.1 Overview of qPCR</td>
<td>15</td>
</tr>
<tr>
<td>1.4.2 Primer Design</td>
<td>20</td>
</tr>
<tr>
<td>1.4.3 Amplification Efficiency</td>
<td>21</td>
</tr>
</tbody>
</table>
1.5 Summary .........................................................................................................23

2. MATERIALS AND METHODS .........................................................................25
   2.1 Reagents and Consumables ............................................................................25
   2.2 Primer Design .................................................................................................25
   2.3 Polymerase Chain Reaction ............................................................................26
   2.4 Quantitative Polymerase Chain Reaction .......................................................27
   2.5 Amplification Efficiency ................................................................................28
   2.6 Gene Expression Data Analysis ......................................................................28

3. RESULTS AND DISCUSSION ...........................................................................30
   3.1 Primer Design and Efficiency Validation .......................................................30
   3.2 Analysis of Prenyltransferase Gene Expression Through qPCR ....................41

4. CONCLUSION .....................................................................................................46

References ....................................................................................................................49
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Tables</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Identification of prenyltransferases in <em>P-n multiseries</em></td>
<td>26</td>
</tr>
<tr>
<td>2. RoboCycler® Gradient 96 from Stratagene Experiment Conditions</td>
<td>27</td>
</tr>
<tr>
<td>3. Bio-Rad CFX96 Touch™ Real-Time PCR Detection System Experiment Conditions</td>
<td>28</td>
</tr>
<tr>
<td>4. Primer Design Set A</td>
<td>31</td>
</tr>
<tr>
<td>5. Amplification efficiency results for Primer Set A</td>
<td>37</td>
</tr>
<tr>
<td>6. Primer Design Set B</td>
<td>38</td>
</tr>
<tr>
<td>7. Amplification efficiency results for Primer Set B</td>
<td>41</td>
</tr>
<tr>
<td>8. Strains of <em>P-n multiseries</em> for gene expression analysis</td>
<td>42</td>
</tr>
<tr>
<td>9. Reference gene sequences for Histone H3 and Dynein</td>
<td>42</td>
</tr>
<tr>
<td>10. Results of gene expression analysis</td>
<td>43</td>
</tr>
<tr>
<td>11. Relative Quantitation of <em>Psemu</em> in MLML2 and 15190C3 normalized to MLML1</td>
<td>44</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figures</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Diatoms under a microscope</td>
<td>2</td>
</tr>
<tr>
<td>2. Map of $P-n$ species found worldwide</td>
<td>4</td>
</tr>
<tr>
<td>3. <em>Pseudo-nitzschia</em></td>
<td>5</td>
</tr>
<tr>
<td>4. Structure of Frustule</td>
<td>6</td>
</tr>
<tr>
<td>5. Asexual and Sexual Reproduction in Diatoms</td>
<td>7</td>
</tr>
<tr>
<td>6. Structure of Domoic Acid</td>
<td>9</td>
</tr>
<tr>
<td>7. Domoic Acid and Kainic Acid</td>
<td>10</td>
</tr>
<tr>
<td>8. Isoprenoid synthesis pathway</td>
<td>12</td>
</tr>
<tr>
<td>9. Short-chained <em>trans</em>-prenyltransferases and their carbon products</td>
<td>14</td>
</tr>
<tr>
<td>10. Schematic for DA Synthesis</td>
<td>15</td>
</tr>
<tr>
<td>11. Basic Components for qPCR Analysis</td>
<td>16</td>
</tr>
<tr>
<td>12. SYBR Green and Hydrolysis probe fluorescence mechanism</td>
<td>17</td>
</tr>
<tr>
<td>13. Melt curve analysis</td>
<td>18</td>
</tr>
<tr>
<td>14. Example amplification plot</td>
<td>22</td>
</tr>
<tr>
<td>15. Example Plot of Cq vs DNA copy number</td>
<td>23</td>
</tr>
<tr>
<td>16. Agarose gel electrophoresis of amplicons from Primer Design Set A</td>
<td>32</td>
</tr>
<tr>
<td>17. <em>Psemu2</em> Primer Pair A Efficiency Validation</td>
<td>34</td>
</tr>
<tr>
<td>18. <em>Psemu3</em> Primer Pair A Efficiency Validation</td>
<td>35</td>
</tr>
<tr>
<td>19. <em>Psemu4</em> Primer Pair A Efficiency Validation</td>
<td>36</td>
</tr>
</tbody>
</table>
20. Agarose gel electrophoresis of amplicons from Primer Design Set B .............. 38
21. *Psemu3* Primer Pair B Efficiency Validation ................................................... 39
22. *Psemu4* Primer Pair B Efficiency Validation ................................................... 40
23. Agarose gel electrophoresis of amplicons from gene expression analysis....... 45
Chapter 1

INTRODUCTION

Diatoms are a class of phytoplankton found in all bodies of water worldwide. They are essential to the aquatic environment and provide a wealth of resources for marine wildlife. Within this family, one genus has the ability to synthesize harmful neurotoxins. The eventual release of this biotoxin into ocean waters consequently damages the environment and impacts wildlife health. This genus is *Pseudo-nitzschia* (*P-n*). It synthesizes the potent neurotoxin domoic acid (DA), an agent responsible for causing amnesic shellfish poisoning (ASP).\(^1\) Research has suggested the biosynthetic pathway leading up to DA production involves prenyltransferases, an enzyme family known for catalyzing carbon condensations.\(^2\)

Prenyltransferases catalyze the formation of isoprenoids. These isoprenoid metabolites have been recognized as precursors to DA.\(^2\) An increase in prenyltransferase expression could proportionally increase the amount of precursor materials available for DA synthesis. Therefore, prenyltransferase transcript abundance can then be directly correlated with DA concentration. The goal of this study was to measure expression of *P-n multiserie* genes encoding prenyltransferases at the mRNA level through Quantitative Real-Time Polymerase Chain Reaction (qPCR) to develop a better understanding of the biosynthetic pathway leading to DA production. Prenyltransferase transcript abundance was measured at different *P-n* growth phases and correlated with DA accumulation.
1.1. Overview of Diatoms

Diatoms are a diverse group of single-celled microalgae found thriving in aquatic regions worldwide. They reside in the euphotic zone where they receive enough sunlight for photosynthesis. They can also be found in upwelling currents taking advantage of nutrients ferried up from ocean depths. In the marine ecosystem, diatoms contribute to biogeochemical cycling of carbon, nitrogen and silica. One property that distinguishes diatoms from other phytoplankton is their cell wall.

Diatoms have a unique, intricate silica-based cell wall known as the frustule. This exoskeleton provides exceptional mechanical strength against predators. It also provides a high degree of structural fluidity, allowing the cell to take on various shapes (Figure 1). Diatoms can adopt two major structural shapes, centric and pennate. Centric diatoms have radial symmetry while pennate have bilateral symmetry. In addition to providing structure to diatoms, the frustule also has a role in their life cycle.

![Diatoms under a microscope](image)

**Figure 1: Diatoms under a microscope.** The frustule allows diatoms to adopt various structural shapes. Diatoms have either radial or bilateral symmetry.
All diatoms undergo vegetative cell division and sexual reproduction. A majority of their lifecycle consists of cell division, where each cycle divides the frustule into two daughter cells. Repetition of cell division eventually reduces the overall size of the diatom. Diatoms can restore their maximum size by undergoing sexual reproduction. Two diatoms form gametes and engage in reproduction to form a new zygote. The zygote develops into an auxospore and restores the frustule to its original size. When large groups of diatoms are simultaneously undergoing reproduction, it can lead to a phenomenon known as an algal bloom.

Algal blooms occur when environmental factors such as nutrients and water temperature are optimal for growth. These occurrences can cover entire coastlines for extended periods depending on the amount of nutrients in the water. Long-term algal blooms block sunlight from reaching vegetation on the seafloor. They also create dead zones in the ocean and cause harmful discoloration abnormalities in the water. Algal blooms of diatoms capable of producing biotoxins can cause devastating environmental and health impacts. One example of a toxigenic genus responsible for some of these environmental and health impacts is *Pseudo-nitzschia*.

1.2 *Pseudo-nitzschia (P-n)*

1.2.1 Distribution

*P-n* is a pennate diatom genus found in aquatic regions worldwide. It was first discovered in 1987. Since then, the number of species described has steadily increased from 15 in the late 1980s to 37 in the 2010s. At that time, 12 of the 37 *P-n* species were confirmed to be toxigenic. Recent studies have found 15 new species, totaling 52. Half
(26) of the total have been found to be toxigenic.\textsuperscript{14} \textit{P-n} has garnered widespread attention due to its toxigenicity and overwhelming presence on coastlines (\textbf{Figure 2}).\textsuperscript{15}

\textbf{Figure 2: Map of \textit{P-n} species found worldwide.} Species in bold are cultures that have demonstrated DA production.\textsuperscript{14,16}
1.2.2 Life Cycle

*P-n* is pennate diatom with bilateral symmetry. This symmetry allows cells to form chains when grouped in large colonies (**Figure 3**). Its frustule is composed of two components: epitheca and hypotheca.\textsuperscript{17} These components are also known as valves.\textsuperscript{10} The larger valve is the epitheca whereas the smaller is the hypotheca. The slits that exist in between the valves are called raphe. The raphe assists the cell with mobility and ability to transverse across surfaces. The fibulae are silica bridges that reinforce raphe, making the cell structure more cohesive and robust (**Figure 4**).\textsuperscript{4}

![Figure 3: Pseudo-nitzschia. *P-n* pennate cells chain-linked with others cells.\textsuperscript{18}](image-url)
The frustule plays a large role in diatom reproduction. In vegetative cell division, the epitheca and hypotheca separate. One daughter cell takes the epitheca half and the second takes the hypotheca. The daughter cell that takes the epitheca synthesizes a new complementary hypotheca to complete its frustule. The second daughter cell receiving the hypotheca will synthesize a new complementary half at a smaller size. The pathway following the initial epitheca results in \( P-n \) maintaining its cell size. The second pathway starting with the parental hypotheca produces cells decreasing in size (Figure 5). A cell that becomes too small loses its ability to further divide.\(^4\) This is known as the first cardinal point. This first cardinal point varies across \( P-n \) species. Arriving at the first cardinal point allows \( P-n \) cells to begin preparations for sexual reproduction and restore themselves to their original size.
Figure 5: Asexual and Sexual Reproduction in Diatoms. Diatoms undergo two pathways for reproduction as described by Russell Habermann.\textsuperscript{19}

$P$-$n$ cells can regenerate their frustule by undergoing sexual reproduction. Two cells that have reached their first cardinal point will begin to form gametes.\textsuperscript{20-21} Cells gradually align in parallel to each other to facilitate interaction. Valves begin to pair and allows the frustule to open. Fusion occurs and a zygote will begin to expand to form a new auxospore. After two to four days of development, a matured $P$-$n$ cell will exit the auxospore.\textsuperscript{4} Once the frustule has been restored to its original size, $P$-$n$ cells can restart the process of vegetative cell division.

As previously mentioned, algal blooms occur when large groups of diatoms undergo reproduction. This phenomenon also applies to $P$-$n$. The timing of $P$-$n$ blooms has historically been difficult to predict due to influences from a variety of environmental factors such as warm temperature, high salinity, or high nutrients.\textsuperscript{22} Warmer water temperatures in general increase growth by influencing germination, nutrient uptake, and
other physiological processes.\textsuperscript{23} $P$-$n$ species are also halotolerant and grow in a wide range of salinities.\textsuperscript{24-25} The optimal concentration varies between species, but studies have found positive correlation of high salinity levels with growth.\textsuperscript{4} Ocean acidification has become another factor promoting growth due to the enrichment of CO$_2$ in the atmosphere.\textsuperscript{23} Nutrient enrichment in waters also has a positive correlation with growth.\textsuperscript{26} The key macronutrient is nitrogen, a component essential for overall growth and amino acid synthesis. Understanding the factors that contribute to the growth of $P$-$n$ improves prediction of timing and intensity for harmful algal blooms.

\textbf{1.2.3 Toxicity}

The $P$-$n$ genus has toxigenic species that can synthesize DA, the agent responsible for causing ASP (\textit{Figure 6}).\textsuperscript{27} The first incidence of toxicity was reported in 1987. Several people immediately fell ill after consuming blue mussels from Prince Edward Island, Canada.\textsuperscript{24} Over 150 individuals were hospitalized within 48 hours. Patients reported gastrointestinal pain, nervous system abnormalities, and varying degrees of memory loss.\textsuperscript{28} Some had difficulty recalling events after dining, while others could not remember their experiences from years before. All patients experienced vomiting, nausea, diarrhea, and abdominal cramps within 24 hours.\textsuperscript{29} More severe symptoms developed as time elapsed. These patients experienced permanent memory loss, seizures, coma, and confusion within 48 hours.\textsuperscript{27} There are no cures for the neurological effects caused by ASP. Once DA has been consumed, the individual must cycle it out through their system. Without a treatment for ASP, DA mitigation became a focus for preventing ASP.
Figure 6: Structure of Domoic Acid.

Domoic Acid

DA was traced to an algal bloom of *P-n multiseries* in waters where mussels were harvested. Shellfish such as mussels filter-feed from seawater containing plankton including *P-n*. Once *P-n* is ingested, DA accumulates within their digestive tracts without visible harm. Monitoring programs were quickly developed around the world to routinely test shellfish tissue for DA and issue fishery closures when DA levels were high. By promptly closing fisheries prior to harvest and preventing contaminated shellfish from entering the market, this has protected the general public from encountering DA.

Regulatory programs have been an integral oversight entity for protecting human health from ASP. However, DA intoxication continues to affect marine wildlife such as seabirds, sea otters, sea lions, and whales. Marine animals are most vulnerable when algal blooms occur. Contaminated shellfish are left widely accessible for marine animals to consume. As a result, wildlife intoxication has increased throughout the last decade. California sea lions are often intoxicated by feeding on anchovies and krill.
mammals such as Minke whales, dwarf sperm whales, and North Atlantic right whales were also found to have developed ASP-like symptoms. Ultimately, intoxicated animals succumb to the effects of ASP.

DA’s ability to cause neurological complications originates from its structural similarities to kainic acid and glutamate (Figure 7). Kainic acid has been known to bind and activate kainate receptors since the early 1950s. It was first used as an anthelmintic, but concern was immediately raised when it induced prolonged excitatory responses in rats. Kainic acid was soon identified as a lethal neurotoxin inducing and maintaining neuron depolarization leading to eventual cellular death. DA mimics kainic acid’s binding to kainic receptors, but at a 100-fold greater binding affinity. This causes depolarization to become repetitive and intensifies activation of signaling pathways dependent on calcium.

![Figure 7: Domoic Acid and Kainic Acid](image)

Stimulated kainate receptors release calcium ions into the neuron where calcium-dependent enzyme pathways can be activated. Various signaling pathways become continuously activated, prolonging excitation. Excess of calcium concentration also promotes the release of glutamate. Neurons begin to swell as a result which disrupts

```
cellular function. This eventually causes neurons to enter apoptosis and causes the neurological complications observed in ASP.

1.3 Domoic Acid Chemistry and Biochemistry

1.3.1 Original Isolation & Use

DA was first isolated from Chondria armata, a seaweed found off the southern coast of Japan. In 1958, it was used to treat gastrointestinal discomfort caused by ascaris and pinworm. Oral administration of 20 mg successfully eradicated the parasite without causing memory loss. This dosage was substantially less than what the restaurant patrons ingested in Canada. As a result, the established maximum safe concentration for DA is 0.5 mg/kg of body weight or 20 μg DA/g of shellfish tissue.

1.3.2 General Isoprenoid Biochemistry

Domoic acid is a member of a general class of metabolites known as isoprenoids. Isoprenoids, also known as terpenoids, are the most abundant and structurally diverse class of natural carbon compounds. Approximately 55,000 compounds have been identified in bacteria, archaea, and eukaryotes, all fulfilling metabolic and physiological roles. For example, plants use isoprenoid metabolites for a variety of functions including respiration, photosynthesis, and growth and reproduction. Isoprenoids also exist as prenyl-lipids used as metabolites for the membrane structure found in archaebacteria, defense molecules in plants, and cholesterol in humans. The diversity of biochemical functions that isoprenoids are involved in signifies their fundamental importance in all species.
Isoprenoids are derived from isopentenyl diphosphate (IPP) formed through the mevalonate pathway or the 1-deoxy-d-xylulose 5-phosphate/2-C-methyl-D-erythritol 4-phosphate pathway (Figure 8). Plants and algae can utilize both pathways to synthesize IPP. Isomerases convert IPP to dimethylallyl diphosphate (DMAPP) in preparation for isoprenoid biosynthesis. Subsequent 1-4 condensation of DMAPP and one or more IPP molecules produces prenyl diphosphate molecules such as geranyl diphosphate (GPP, C\textsubscript{10}), farnesyl diphosphate (FPP, C\textsubscript{15}), and geranylgeranyl diphosphate (GGPP, C\textsubscript{20}). These prenyl diphosphates are intermediates for a variety of complex isoprenoids. The enzymes that catalyze the condensation of IPP and DMAPP to form prenyl diphosphates are known as prenyltransferases.

Figure 8: Isoprenoid synthesis pathway. Mevalonate Pathway (Left) and 1-Deoxy-d-Xylulose 5-Phosphate/2-C-Methyl-D-Erythritol 4-Phosphate Pathway (Right).
1.3.3 Prenyltransferases

Prenyltransferases are widely distributed in nature. There are three main classes of prenyltransferases. Isoprenyl diphosphate synthases are responsible for generating linear polymers at defined chain lengths. Protein prenyltransferases catalyze isoprenyl diphosphate transfers between proteins. General prenyltransferases cyclize reactions with IPP. Prenyltransferases are also categorized by stereoisomers trans- and cis-. Trans-prenyltransferases synthesize products up to C50 while Cis- are known to generate even longer chains.

Trans-prenyltransferases are further categorized into those that produce short-, medium-, or long-chain products. Short-chain trans-prenyltransferases are limited by the maximum carbon length product they can produce (Figure 9). Geranyl diphosphate synthase (GPPS) synthesizes the simplest of polyisoprenoid compounds with C10 carbon chains. Farnesyl diphosphate synthase (FPPS) adds another C5 group on geranyl diphosphate (GPP) to form farnesyl diphosphate (FPP), which is a precursor to metabolites such as cholesterol and other steroids. Geranylgeranyl diphosphate synthase (GGPPS) adds another C5 group to FPP to form geranylgeranyl diphosphate (GGPP), which is a precursor to carotenoids, retinoid, and chlorophylls. The largest short-chained trans-prenyltransferase is farnesylgeranyl diphosphate synthase. It can synthesize carbon chains of C25 which are used in membrane lipids.
1.3.4 Domoic Acid Biosynthesis

Several efforts by researchers have described portions of biosynthetic pathway leading to DA production. It has been determined that N-geranyl-L-glutamic acid (L-NGG) is the precursor to DA.\textsuperscript{51} L-NGG is formed by the condensation of GPP and glutamate.\textsuperscript{2, 52} GPP is synthesized by a condensation of IPP and DMAPP catalyzed by a prenyltransferase (Figure 10). GPP then condenses with l-glutamate to form l-NGG.\textsuperscript{53} L-NGG is further modified by oxygenases to form DA.\textsuperscript{54} The genome of \textit{P-n} encodes four trans-prenyltransferases, but product specificity (GPP, FPP or GPP) of each prenyltransferase is not apparent based strictly on sequence analysis. Thus, the challenge becomes identifying which of the four \textit{trans}-prenyltransferases catalyzes the formation of GPP and thus is involved in the biosynthesis of DA.
1.4 Measuring Gene Expression with Quantitative Real Time PCR (qPCR)

1.4.1 Overview of qPCR

Quantitative PCR (qPCR) is a standard technique that provides a wealth of information from applications of basic research through medicinal diagnostics. It has been accepted as the standard for gene expression analysis. It is based on traditional PCR, but expands on the technique by automating quantitation of amplified products after each cycle. Measuring amplification after each cycle provides real-time data and allows the user to monitor progress. Reactions are prepared using the same basic components: buffer, deoxynucleotides, DNA template, primers, and DNA polymerase (Figure 11). In addition to the conventional PCR components, a fluorescent dye or marker is required for qPCR instruments to measure amplification. Amplicons, the products of amplification or replication events, are measured by detecting fluorescence intensities emitted from a DNA binding dye or custom primer probe after each cycle.
Figure 11: Basic Components for qPCR Analysis. mRNA must be converted to cDNA through reverse transcription for qPCR analysis. The components to perform a reverse transcription reaction are shown on the left (mRNA, reverse transcriptase, DNA polymerase, buffers, and primers). Once cDNA is synthesized, the qPCR reaction uses similar components except reverse transcriptase.\(^{58}\)

Binding dyes and primer probes are types of markers used to quantify amplification. An example of a non-specific binding dye is SYBR Green. It is common, inexpensive, easy to use, and capable of generic detection of dsDNA.\(^ {59}\) It binds to all dsDNA and emits fluorescence. For more specificity, custom primers probes are designed to bind the target sequence downstream from the starting point carrying a quenched-fluorescence segment.\(^ {60}\) Fluorescence is detected when DNA polymerase overwrites the primer and releases the quenching portion. Either of these fluorescence systems can be used to obtain quantitation data on polynucleotide abundance (Figure 12).
Figure 12: SYBR Green and Hydrolysis probe fluorescence mechanism. The green circles represent molecules that emit fluorescence. SYBR Green dye (left) is unbound during the annealing phase. Once dsDNA is formed, the dye binds and emits fluorescence. Hydrolysis probes (right) are custom, fluorophore-quencher labeled primers designed to bind to the target sequence. Once the probe is hydrolyzed, the fluorophore is released and fluorescence is emitted.\textsuperscript{61}

While relatively inexpensive and simple compared to custom primer probes, non-specific binding dyes have disadvantages. Non-specific dyes bind to all amplified dsDNA
including any amplified products from sample contamination and primer dimerization. The instrument cannot discriminate between fluorescence from contaminant amplification and primer dimers and fluorescence from target transcript amplification. However, contamination and dimerization can be evaluated by performing a melt curve analysis following amplification. The thermal cycler gradually denatures all dsDNA and measures fluorescence emitted after each increment of temperature increase. As dsDNA denatures, the binding dye is released. Emitted fluorescence decreases as a result and the change in fluorescence can be plotted against temperature (Figure 13). A single peak should result from denaturing one amplicon. A curve with multiple peaks indicates that more than one amplicon was denatured. The melt curve analysis is a critical post-amplification step for qPCR experiments using binding dyes to evaluate contamination or formation of multiple amplicons.

![Figure 13: Melt curve analysis](image)

**Figure 13: Melt curve analysis.** A melt curve analysis provides two graphs: fluorescence against temperature (left) and change in fluorescence against temperature (right). As temperature increases, fluorescence decreases once amplicons denature and release binding dye. Contamination and dimerization can be evaluated by the number of peaks resulting from fluorescence changes. Sample 1 and 3 had one peak each, indicating only one amplicon denatured. Sample 2 had two peaks, indicating that two amplicons denatured.
There are several parameters that must be considered when setting up the instrument for qPCR analysis. First, a standard experiment protocol similar to conventional PCR must be developed. These parameters are dependent on the reagents used in each sample, temperatures for polymerase activity and primer annealing, and time allocated for each step. Second, the fluorescent marker such as SYBR Green must be defined in the protocol for the instrument to detect fluorescence. The instrument can also be set to measure a passive reference dye to normalize fluorescence between samples after amplification. A common passive reference dye used is carboxy-X-rhodamine (ROX). Before a run begins, the instrument determines baseline fluorescence by measuring background noise. It then uses the baseline value to calculate a threshold value where fluorescence from amplification has reached a statistically significant point above the baseline.

The amplification cycle after which fluorescence exceeds the threshold value is defined as the quantitation cycle (Cq). Cq is a function of the amount of initial template DNA in the sample. A low Cq value indicates that statistically significant amplification was reached early in the run because the initial template concentration was high. A high Cq value indicates low initial template concentration and additional cycles were needed to produce significant fluorescence. Thus, the Cq value can be used to quantify nucleic acids within samples.

The quantitation of nucleic acids with Cq values can be performed by absolute quantitation or relative quantitation. Absolute quantification relies on a standard curve to determine nucleic acid concentration. A standard curve must be generated by measuring fluorescence from amplifying serial dilutions of known template concentrations. Cq
values are plotted against concentration to form a graphically linear relationship. Cq values from test samples can be evaluated against the standard curve to determine concentration. The reliability of this method depends on how similar amplification efficiencies are between the primers used for test sample and primers on the serial dilutions of known template concentrations. An alternative method to absolute quantification is relative quantification.

Relative quantification relies on using reference genes to compare relative expression levels. Reference genes are sequences that have a consistent rate of expression in a cell. In a single qPCR experiment, reference gene sequences are amplified alongside the gene of interest in a separate well. Cq values of gene of interest can then be compared to the Cq from the reference gene. Due to the high sensitivity of qPCR, measuring amplification of two or more reference genes are recommended to account for sporadic amplification. Mathematical models are also required to quantify the expression of the target gene to the reference gene. Normalization of multiple genes is a step required to provide an accurate comparison. The relative quantification method allows for calculated expression levels of a single gene to be compared against multiple qPCR experiments.

1.4.2 Primer Design

Primer design is a critical step for experimental specificity and data acquisition. The objective is to define a nucleotide sequence that binds specifically to the target template while avoiding risk of primer dimerization and mispriming. Primers need to be 18-24 nucleotides in length with melting temperatures between 55-65°C. GC content should be
between 40%-60% of the primer to ensure stability and binding. A primer sequence must not contain a sequence of repeats to avoid self-priming. Primer pairs must have melting temperatures within 5°C of each other and be evaluated for unwanted priming on non-targeted templates that may be included in DNA mixture samples. These qualities can be assessed using various software provided on the World Wide Web (e.g. Primer3, Primer-BLAST). Once primers are designed and obtained, they must be evaluated in qPCR to validate amplification efficiency.

1.4.3 Amplification Efficiency

Analysis of amplification efficiency validates the specificity of primers pairs and ensures amplification is consistent per cycle. There are two methods to evaluate amplification efficiency. The first method consists of running three replicates of ten-fold serial dilutions of the target DNA sequence. Clones of the target sequence are the best template models for dilutions. The basic PCR components (buffer, deoxynucleotides, primers, and DNA polymerase) must be unchanged throughout all samples, leaving the only variation to be concentration of DNA template. Ideally, the template should double after each amplification cycle. The result for complete doubling will be a plot where the ΔCq between each diluted sample is equivalent (Figure 14).
Figure 14: Example amplification plot. The ideal Cq distance between serial dilutions is illustrated with green arrows.\(^{59}\)

Amplification efficiency can be calculated once Cq values are obtained from serial dilutions and plotted against log dilutions.\(^ {62}\) Graphs should ideally have data points in a linear shape. A trend line must be used to measure the best fit line (Figure 15). The R\(^2\) value should be at least 0.99. The slope of the trend line is used to calculate amplification efficiency using the equation below. A slope of -3.32 is 100% amplification efficiency.\(^ {62}\) Efficiencies between 90-105% are accepted in qPCR studies, but a value closer to 100% is desired.

\[
\text{amplification efficiency} = \left[10^{\left(-\frac{1}{\text{slope}}\right)} - 1\right] \times 100
\]
The second method to evaluate amplification efficiency is by running sample sets through a temperature gradient. The goal of this method is to experimentally identify the optimal annealing temperature where amplification efficiency of 100% is obtained without amplifying contaminates or primer dimers. The standard qPCR components of buffer, deoxynucleotides, DNA template, primers, and DNA polymerase must be consistent through all replicates. Amplification should be measured in temperatures ranging ± 5°C from the calculated primer pair melting temperature ($T_m$). Melt curve analysis and agarose gel electrophoresis are required to evaluate the resulting amplicons.

1.5 Summary

In summary, this study investigates the biosynthetic pathway leading to the production of DA in *P-n multiseries* by investigating prenyltransferases responsible for the formation of GPP through gene expression analysis.
Seven prenyltransferases transcripts had been identified within the genome of *P-n multiseries* through homology studies. A preliminary set of primers were designed for each transcript. However, only three of the seven were evaluated (*Psemu2*, *Psemu3*, and *Psemu4*). They were selected based on positive results from previous research conducted in the group and availability of cloned expression vectors for primer validation studies. Additionally, homology modeling completed by Peery suggested that *Psemu3* and *Psemu4* exhibited catalytic pocket sizes consistent to geranyl diphosphate synthase. This study measured prenyltransferase gene expression in three strains of *P-n multiseries* isolated at different growth phases and correlated transcript abundance to DA accumulation.
Chapter 2

MATERIALS AND METHODS

2.1 Reagents and Consumables

LB Broth powder was purchased from Fisher Scientific. Agarose powder (Electrophoresis grade) was purchased from Alfa Aesar. GelRed® Nucleic Acid Gel Stains was purchased from Biotium. The 0.2 mL 8-Tube PCR Strips without Caps and 0.2 mL Flat PCR Tube 8-Cap Strips were purchased from Bio-Rad. QIAprep Spin Miniprep Kit was purchased from QIAGEN. SYBR™ Select Master Mix, custom DNA Oligomer Primers, and 0.2 mL RNase-Free PCR Tubes were purchased from ThermoFisher Scientific. Taq PCR Kit was purchased from New England Biolabs. Total RNA and cDNA from three P-n multiseries strains were provided by Dr. G. Jason Smith of Moss Landing Marine Laboratories. All water used for gene expression analysis was autoclaved before use.

2.2 Primer Design

Seven prenyltransferase transcript sequences identified in P-n multiseries were retrieved from the Joint Genome Institute (Table 1).\textsuperscript{71} Oligonucleotide sequences spanning exon-exon junctions were designed. Forward and reverse primer sequences were optimized using Primer3.\textsuperscript{72-73} The formation of primer secondary structures was assessed using mFold software.\textsuperscript{74} Re-designed primers were evaluated with NetPrimer in place of mFold.\textsuperscript{75} Primer specificity to target sequences was assessed using the National Center for Biotechnology Information (NCBI) Basic Local Alignment Search Tool (BLAST).\textsuperscript{76}
Table 1: Identification of prenyltransferases in *P-n* multiserie.

<table>
<thead>
<tr>
<th>Alias</th>
<th>Transcript Id</th>
<th>Protein Id</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Psemu1</em></td>
<td>265069</td>
<td>264726</td>
</tr>
<tr>
<td><em>Psemu2</em></td>
<td>62818</td>
<td>62475</td>
</tr>
<tr>
<td><em>Psemu3</em></td>
<td>244176</td>
<td>243833</td>
</tr>
<tr>
<td><em>Psemu4</em></td>
<td>299194</td>
<td>298851</td>
</tr>
<tr>
<td><em>Psemu5</em></td>
<td>204169</td>
<td>203826</td>
</tr>
<tr>
<td><em>Psemu6</em></td>
<td>150943</td>
<td>150600</td>
</tr>
<tr>
<td><em>Psemu7</em></td>
<td>29520</td>
<td>29177</td>
</tr>
</tbody>
</table>

2.3 Polymerase Chain Reaction

PCR was performed on the RoboCycler® Gradient 96 from Stratagene. Reaction samples were prepared in sterile 0.2 mL RNase-Free PCR Tubes. Standard Taq Reaction Buffer, dNTPs, and Taq DNA Polymerase were used from the New England Biolabs Taq PCR Kit. Sample volumes totaled to 50 µL. Components were 5 µL of 10X Standard Taq Reaction Buffer (100 mM Tris-HCl, 500 mM KCl, 1.5 mM MgCl₂), 1 µL of 10 mM dNTPs (200 µM final concentration), 2 µL of 0.5 µM Forward and Reverse Primer mix (0.05 µM final concentration), 1 µL of Miniprep isolated template DNA, 0.35 µL of 5,000 units/mL Taq polymerase (1.75 units/50 µL PCR), and sterile water added to a final volume of 50 µL. Tubes were centrifuged at 500 RPM for 5 minutes prior to amplification. PCR analysis was performed according to the parameters described in Table 2.
Table 2: RoboCycler® Gradient 96 from Stratagene Experiment Conditions.

<table>
<thead>
<tr>
<th></th>
<th>Cycles</th>
<th>Temperature</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial denaturation</td>
<td>1</td>
<td>95.0 °C</td>
<td>30 sec</td>
</tr>
<tr>
<td>Denaturation</td>
<td></td>
<td>95.0 °C</td>
<td>30 sec</td>
</tr>
<tr>
<td>Annealing</td>
<td>30</td>
<td>55.0 °C</td>
<td>30 sec</td>
</tr>
<tr>
<td>Extension</td>
<td></td>
<td>68.0 °C</td>
<td>1 min</td>
</tr>
<tr>
<td>Final Extension</td>
<td>1</td>
<td>68.0 °C</td>
<td>5 min</td>
</tr>
</tbody>
</table>

Amplicon products were analyzed with agarose gel electrophoresis. Nucleic acids were separated in 1% agarose with 1X TAE buffer. Visualization with UV light revealed amplicons stained with GelRed®.

2.4 Quantitative Polymerase Chain Reaction

qPCR was performed on the Bio-Rad CFX96 Touch™ Real-Time PCR Detection System. Reaction samples were prepared in triplicate with Bio-Rad 8-Tube Strips White and Bio-Rad Flat PCR Tube 8-Cap Strips. Reaction volumes totaled to 20 µL. Components were 10 µL of SYBR Green Master Mix, 8 µL of 0.5 µM of the forward and reverse primer (0.2 µM final concentration), 1 µL of DNA, and 1 µL sterile water. Ten-fold serial dilutions beginning from 1 ng/µL of template DNA was used for amplification efficiency analysis. A ten-fold dilution was prepared from each stock cDNA sample isolated from Moss Landing Marine Laboratories for gene expression analysis. Tube strips were centrifuged at 500 RPM for 5 minutes prior to amplification. qPCR analysis was performed according to the parameters described in Table 3.
<table>
<thead>
<tr>
<th>Experiment Conditions</th>
<th>Cycles</th>
<th>Temperature</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Polymerase activation</td>
<td>1</td>
<td>50.0 °C</td>
<td>2 min</td>
</tr>
<tr>
<td>Initial denaturation</td>
<td>1</td>
<td>95.0 °C</td>
<td>2 min</td>
</tr>
<tr>
<td>Denaturation</td>
<td>40</td>
<td>95.0 °C</td>
<td>15 sec</td>
</tr>
<tr>
<td>Anneal, extend, and read fluorescence</td>
<td></td>
<td>60.0 °C</td>
<td>1 min</td>
</tr>
<tr>
<td>Denaturation</td>
<td>1</td>
<td>95.0 °C</td>
<td>10 sec</td>
</tr>
<tr>
<td>Melting curve</td>
<td>1</td>
<td>65.0 °C to 95.0 °C: Increment 0.5°C</td>
<td>5 sec</td>
</tr>
</tbody>
</table>

Amplicon products were analyzed with agarose gel electrophoresis. Nucleic acids were separated in 1% agarose with 1X TAE buffer. Visualization with UV light revealed amplicons stained with GelRed®.

### 2.5 Amplification Efficiency

Cq values resulting from ten-fold serial dilutions were used to calculate primer amplification efficiency. Average Cq values were calculated for each triplicate and plotted against log dilution. A trend line was generated and the slope was used to calculate primer amplification efficiency.

\[
amplification\ efficiency = \left[10\left(\frac{-1}{slope}\right) - 1\right] \times 100
\]

### 2.6 Gene Expression Data Analysis

Relative quantification gene expression ratios were calculated with the \(2^{-\Delta\Delta CT}\) Method.\(^\text{77}\) Ct is known as the cycle threshold and is interchangeable with Cq. The following steps and equations were used to calculate a normalized expression ratio for each target cDNA.
\[ \Delta C_q \text{ (test sample)} = C_q \text{ (target sequence, test sample)} - C_q \text{ (reference gene, test sample)} \]

\[ \Delta C_q \text{ (calibrator sample)} = C_q \text{ (target sequence, test sample)} - C_q \text{ (reference gene, test sample)} \]

\[ \Delta \Delta C_T = \Delta C_q \text{ (test sample)} - \Delta C_q \text{ (calibrator sample)} \]

\[ 2^{-\Delta \Delta C_T} = \text{Normalized expression ratio} \]

The triplicate Cq data for target sequences (\textit{Psemu2, Psemu3, and Psemu4}) and reference genes (Histone H3 and Dynein) were separately averaged. The mean Cq of each target sequence was normalized to both reference genes. Expression ratios were compared between strains of \textit{P-n multiseries}. 
Chapter 3

RESULTS AND DISCUSSION

The goal of this study is to develop an understanding of the biosynthetic pathway leading to DA production in \( P-n \) by correlating prenyltransferase gene expression with DA accumulation.

3.1 Primer Design and Efficiency Validation

To begin qPCR analysis of transcript levels, primer pairs were designed and validated for \( Psemu2 \), \( Psemu3 \), and \( Psemu4 \). As previously noted, they were selected based on previous research conducted in the group and availability of cloned expression vectors for primer validation. Homology modeling also suggested that \( Psemu3 \) and \( Psemu4 \) exhibited characteristics consistent with geranyl diphosphate synthase.\(^{70}\)

Oligonucleotide sequences are ideally designed to span exon-exon junctions to increase annealing selectivity to mRNA over genomic DNA. This was possible for transcripts \( Psemu2 \) and \( Psemu4 \), but difficult for \( Psemu3 \). The exon ends on the \( Psemu3 \) transcript exhibited several G and C repeats. Presence of these repeats increases the chance for primer pairs to complementary bind. Thus, primers not spanning exon-exon junctions were designed to minimize the potential for dimerization.

Primer pairs were developed using Primer3, mFold, and BLAST. Primer3 was used for the initial primer design. It generated potential primer sequences while considering parameters such as preset \( T_m \) and resulting amplicon length. Primer3 was also used to evaluate primer penalty values, estimating potential for self-complementarity and hairpin formation. mFold was used to further evaluate potential secondary structure formation.
The last step was to perform a homology search using BLAST to evaluate potential for the primers to bind other sequences in *P-n multiseries*’ genome. When primer pairs failed to pass each step, new sequences were designed and reevaluated.

All *Psemu* primers pairs were designed to allow amplification under one thermocycling protocol. By designing for a $T_m$ of 60 °C, all test samples containing the primer pairs will undergo the annealing step at the same temperature. Similarly, product amplicons were designed to be in the range of 100-200 bp. Primers identified in Table 4 were assessed for amplification activity with PCR.

**Table 4: Primer Design Set A.**

<table>
<thead>
<tr>
<th>cDNA</th>
<th>Primer</th>
<th>Sequence</th>
<th>GC (%)</th>
<th>$T_m$ (°C)</th>
<th>Amplicon (bp)</th>
<th>Exon-Exon Spanning</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Psemu</em>2</td>
<td>A</td>
<td>F: CAAAGGCCGGAAGTGATGC</td>
<td>57.9</td>
<td>59.5</td>
<td>137</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R: TTTGTGGACGCTTCCTC</td>
<td>55.0</td>
<td>60.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Psemu</em>3</td>
<td>A</td>
<td>F: ATTTGCTGTGCGAGGGAAGA</td>
<td>50.0</td>
<td>60.0</td>
<td>153</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R: GGAAGAAATTCAGGCGTGCG</td>
<td>55.0</td>
<td>60.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Psemu</em>4</td>
<td>A</td>
<td>F: CGCGGATTGATGGTGCTAGA</td>
<td>55.0</td>
<td>60.0</td>
<td>141</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R: TGGTGGCGGACGATAATCAGTGC</td>
<td>55.0</td>
<td>60.0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The primer pairs were evaluated for activity using conventional PCR and amplicon size was confirmed with agarose gel electrophoresis. Expression vectors cloned by the lab group served as DNA templates for the amplification. A single amplified fragment around 100-200 bp was expected if primers annealed successfully. Redesign was required if resulting bands appear fragmented or if primer dimers formed around 10-100 bp. The results from this analysis gauged the success of generating amplicons with temperature conditions set at 60.0°C.
PCR amplification of \textit{Psemu2}, \textit{Psemu3}, and \textit{Psemu4} was successful. Agarose gel electrophoresis analysis of amplified fragments are shown in \textbf{Figure 16}. A molecular weight marker was used to confirm amplicon size between 100-200 bp. Each \textit{Psemu} sample revealed only one fragment per lane. This indicated that there were no other fragments amplified.

\textbf{Figure 16: Agarose gel electrophoresis of amplicons from Primer Design Set A.} Molecular weight markers are adjacent to each test sample. Expected amplicon size is between 100-200 bp.

The next step was to assess amplification efficiency. This consisted of using Primer Set A of \textit{Psemu2}, \textit{Psemu3}, and \textit{Psemu4} and amplifying ten-fold serial dilutions of known DNA concentrations. Experimental Cq values were obtained and amplification efficiency
was determined by measuring the difference between each sample. This step also
evaluated consistency of amplification as DNA concentration decreased per sample. To
account for the possibly of contamination or pipetting error, samples were prepared in
triplicate.

All amplification efficiency analyses were performed using the Bio-Rad CFX96
Thermocycler. A melt curve analysis to evaluate for contamination or primer dimers
followed after amplification. Amplicon size and purity was then evaluated using agarose
gel electrophoresis. Amplification and efficiency plots of Psemu2, Psemu3, and Psemu4
are shown in Figures 17-19.
Figure 17: *Psemu2* Primer Pair A Efficiency Validation. Amplification of template DNA serial dilutions run in triplicates presented by Relative Fluorescence Units (RFU) at each cycle. The horizontal line at 4 RFU is the calculated quantitation baseline. (Upper Graph). The mean Cq for each sample was plotted against log dilution to derive slope for amplification efficiency calculation (Lower Graph).
Figure 18: *Psemu3* Primer Pair A Efficiency Validation. Amplification of template DNA serial dilutions run in triplicates presented by Relative Fluorescence Units (RFU) at each cycle. The horizontal line at 5 RFU is the calculated quantitation baseline. (Upper Graph). The mean Cq for each sample was plotted against log dilution to derive slope for amplification efficiency calculation (Lower Graph).
Figure 19: *Psemu4* Primer Pair A Efficiency Validation. Amplification of template DNA serial dilutions run in triplicates presented by Relative Fluorescence Units (RFU) at each cycle. The horizontal line at 5 RFU is the calculated quantitation baseline. (Upper Graph). The mean Cq for each sample was plotted against log dilution to derive slope for amplification efficiency calculation (Lower Graph).
Amplification efficiency analysis was completed for all three primer pairs. Cq values from each triplicate were averaged and used to calculate primer efficiency (Table 5). Primers targeting the *Psemu2* sequence had an efficiency of 97.6%, indicating near ideal amplification between each cycle. Primer pairs for *Psemu3* and *Psemu4* resulted in efficiencies ≥ 125%. In addition, fluorescence was measured in the NTC samples, potentially indicating the presence of contamination or dimerization. These results were undesirable for gene expression analysis and thus primers were redesigned for *Psemu3* and *Psemu4*.

Table 5: Amplification efficiency results for Primer Set A.

<table>
<thead>
<tr>
<th>cDNA</th>
<th>Primer</th>
<th>Sequence</th>
<th>Efficiency (%)</th>
<th>R²</th>
</tr>
</thead>
</table>
| *Psemu2* | A | F: CAAAGGCGGAGTGATGC  
R: TTTGTGGACGCTTCTCCT | 97.6 | 0.998 |
| *Psemu3* | A | F: ATTTGCTGTGCGAGGGAAGA  
R: GGAAGAAATTACGCGGTGCG | 125.6 | 0.997 |
| *Psemu4* | A | F: CGCGGATTGATGCTGTAAGA  
R: TGGTGCGAGCATATCATG | 128.0 | 0.993 |

Primers for *Psemu3* and *Psemu4* were redesigned following the same constraints as for Primer Set A. Sequences were designed and checked with Primer3, NetPrimer, and BLAST (Table 6). Amplicon sizes were targeted between 100-200 bp. Primers were not designed to span the exon-exon junction. Amplification activity and efficiency was evaluated under one experiment. Amplicon sizes were verified with agarose gel electrophoresis shown in Figure 20. Amplification efficiency plots and results of *Psemu3* and *Psemu4* are shown in Figures 21-22 and in Table 7, respectively.
Table 6: Primer Design Set B.

<table>
<thead>
<tr>
<th>cDNA</th>
<th>Primer</th>
<th>Sequence</th>
<th>GC (%)</th>
<th>Tm (°C)</th>
<th>Amplicon (bp)</th>
<th>Exon-Exon Spanning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Psemu3</td>
<td>B</td>
<td>F: TCAGGCCGTGCGAAACATTTG</td>
<td>50.0</td>
<td>60.0</td>
<td>165</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R: CGTGCTCGTCCAATCTTAG</td>
<td>60.0</td>
<td>60.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Psemu4</td>
<td>B</td>
<td>F: CGAGATGGGCGTCTCTTTCC</td>
<td>60.0</td>
<td>60.0</td>
<td>177</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R: CCGTAGTGTGCCTCCAGAAA</td>
<td>55.0</td>
<td>59.8</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 20: Agarose gel electrophoresis of amplicons from Primer Design Set B. Molecular weight marker are adjacent to each test sample. Expected amplicon size is between 100-200 bp.

Agarose gel electrophoresis revealed a single amplicon for Psemu3 and Psemu4. Since there was only one fragment found for each target cDNA, this indicated no other products were amplified.
Figure 21: *Psemu3 Primer Pair B Efficiency Validation*. Amplification of template DNA serial dilutions run in triplicates presented by Relative Fluorescence Units (RFU) at each cycle. The horizontal line at 3.5 RFU is the calculated quantitation baseline. (Upper Graph). The mean Cq for each sample was plotted against log dilution to derive slope for amplification efficiency calculation (Lower Graph).
Figure 22: *Pseud4 Primer Pair B Efficiency Validation.* Amplification of template DNA serial dilutions run in triplicates presented by Relative Fluorescence Units (RFU) at each cycle. The horizontal line at 3.5 RFU is the calculated quantitation baseline. (Upper Graph). The mean Cq for each sample was plotted against log dilution to derive slope for amplification efficiency calculation (Lower Graph).
Table 7: Amplification efficiency results for Primer Set B.

<table>
<thead>
<tr>
<th>cDNA</th>
<th>Primer</th>
<th>Sequence</th>
<th>Efficiency (%)</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Psemu3</td>
<td>B</td>
<td>F: TCAGGCGTGCGAAACATTTG</td>
<td>107.7</td>
<td>0.991</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R: CGTGCTCCGTCCTACCTAG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Psemu4</td>
<td>B</td>
<td>F: CGAGATGGGGGTCTACTTCC</td>
<td>102.2</td>
<td>0.999</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R: CCGTAGTGCTCGTCCAGAAA</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Primer B for Psemu3 resulted in an efficiency of 107.7% whereas Psemu4 had 102.2%. The measured fluorescence from the NTC sample remained near baseline, indicating there were no contaminates or dimerization. Amplification efficiency for Psemu3 is slightly out of the ideal range (90%-105%) for qPCR analysis. However, the results from agarose gel electrophoresis confirmed that there were no additional amplicons present.

3.2 Analysis of Prenyltransferase Gene Expression Through qPCR

Complementary DNA from three strains of P-n multiseries were provided by Dr. Jason Smith of Moss Landing Marine Laboratories for gene expression analysis. Each strain had been isolated from Monterey Bay, CA and cells from axenic cultures were collected during different growth phases. RNA was extracted from each strain, measured, and stored. cDNA was synthesized from 1 µg of collected total RNA. DA levels were also measured. Of the three strains, the highest DA concentration measured was 1.098 ng/mL. To put in context, the North American west coast P-n bloom of 2015 measured particulate DA exceeding 31.7 ng/mL.\textsuperscript{14} The strains studied here had much lower DA concentration in comparison. The characteristics of each strain including concentrations of RNA and DA were provided by Dr. Smith and presented in Table 8.
Table 8: Strains of *P-n multiseries* for gene expression analysis.

<table>
<thead>
<tr>
<th>Name</th>
<th>Growth Phase</th>
<th>µg RNA/mL</th>
<th>ng DA/mL</th>
<th>ng DA/µg RNA</th>
</tr>
</thead>
<tbody>
<tr>
<td>MLML1</td>
<td>Log</td>
<td>69.14</td>
<td>0.490</td>
<td>0.007</td>
</tr>
<tr>
<td>MLML2</td>
<td>Late Log</td>
<td>27.28</td>
<td>1.098</td>
<td>0.040</td>
</tr>
<tr>
<td>15190C3</td>
<td>Stationary</td>
<td>26.73</td>
<td>1.068</td>
<td>0.040</td>
</tr>
</tbody>
</table>

Histone H3 and Dynein were reference genes used to normalize expression. A previous study had found both as suitable normalization genes for cells transitioning from late log to stationary phase. Both primer pairs for the reference genes were not exon-exon spanning. Amplification efficiencies were 101.4% for Histone H3 and 98% for Dynein. The amplicon length for Histone H3 was 151 bp. Dynein had smaller amplicon length of 84 bp. The reference gene primer sequences are identified in Table 9.

Table 9: Reference gene sequences for Histone H3 and Dynein.

<table>
<thead>
<tr>
<th>cDNA</th>
<th>Sequence</th>
<th>GC (%)</th>
<th>Tm (°C)</th>
<th>Amplicon (bp)</th>
<th>Exon-Exon Spanning</th>
<th>Efficiency (%)</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Histone H3</td>
<td>F: GAAGCCTACCTGGTGGGTCTC</td>
<td>61.9</td>
<td>59.3</td>
<td>151</td>
<td>No</td>
<td>101.4</td>
<td>0.991</td>
</tr>
<tr>
<td></td>
<td>R: CGTCCGATCACCTTCCGTC</td>
<td>61.9</td>
<td>59.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dynein</td>
<td>F: CGAAGCCAGTAGGGTATCAAGG</td>
<td>52.2</td>
<td>57.1</td>
<td>84</td>
<td>No</td>
<td>98.0</td>
<td>0.991</td>
</tr>
<tr>
<td></td>
<td>R: CGAATCAGGTGTGTCTGGAGTCG</td>
<td>52.2</td>
<td>57.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Primers pairs for *Psemu2 A, Psemu3 B, Psemu4 B*, Histone H3 and Dynein were then used to measure transcript abundance in all three strains of *P-n multiseries* with qPCR. Each sample amplification was performed in triplicate. ΔCq values were calculated from the target gene to the reference gene Histone H3 (Table 10).
Table 10: Results of gene expression analysis. Replicates and Cq values for each sample are shown below. Average ΔCq are calculated with standard deviation.

<table>
<thead>
<tr>
<th>MLML1</th>
<th>Psemu2</th>
<th>Psemu3</th>
<th>Psemu4</th>
<th>Histone H3</th>
<th>Dynein</th>
</tr>
</thead>
<tbody>
<tr>
<td>Replicate 1</td>
<td>26.57</td>
<td>25.58</td>
<td>29.22</td>
<td>23.04</td>
<td>25.39</td>
</tr>
<tr>
<td>Replicate 2</td>
<td>26.9</td>
<td>25.54</td>
<td>29.92</td>
<td>23.43</td>
<td>25.60</td>
</tr>
<tr>
<td>Replicate 3</td>
<td>26.78</td>
<td>26.23</td>
<td>30.16</td>
<td>23.18</td>
<td>25.81</td>
</tr>
<tr>
<td>Average Cq</td>
<td>26.75 ± 0.17</td>
<td>25.78 ± 0.39</td>
<td>29.77 ± 0.49</td>
<td>23.22 ± 0.20</td>
<td>25.60 ± 0.21</td>
</tr>
<tr>
<td>ΔCq = Avg Cq - Histone H3 Cq</td>
<td>3.53 ± 0.07</td>
<td>2.57 ± 0.47</td>
<td>6.55 ± 0.40</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>MLML2</th>
<th>Psemu2</th>
<th>Psemu3</th>
<th>Psemu4</th>
<th>Histone H3</th>
<th>Dynein</th>
</tr>
</thead>
<tbody>
<tr>
<td>Replicate 1</td>
<td>35.85</td>
<td>34.77</td>
<td>38.44</td>
<td>32.99</td>
<td>35.19</td>
</tr>
<tr>
<td>Replicate 2</td>
<td>36.80</td>
<td>34.87</td>
<td>37.77</td>
<td>33.63</td>
<td>34.20</td>
</tr>
<tr>
<td>Replicate 3</td>
<td>36.78</td>
<td>34.57</td>
<td>36.45</td>
<td>32.75</td>
<td>35.05</td>
</tr>
<tr>
<td>Average Cq</td>
<td>36.48 ± 0.54</td>
<td>34.74 ± 0.15</td>
<td>37.55 ± 1.01</td>
<td>33.12 ± 0.45</td>
<td>34.81 ± 0.54</td>
</tr>
<tr>
<td>ΔCq = Avg Cq - Histone H3 Cq</td>
<td>3.35 ± 0.61</td>
<td>1.61 ± 0.32</td>
<td>4.43 ± 0.91</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>15190C3</th>
<th>Psemu2</th>
<th>Psemu3</th>
<th>Psemu4</th>
<th>Histone H3</th>
<th>Dynein</th>
</tr>
</thead>
<tbody>
<tr>
<td>Replicate 1</td>
<td>28.16</td>
<td>27.44</td>
<td>33.23</td>
<td>23.14</td>
<td>26.94</td>
</tr>
<tr>
<td>Replicate 2</td>
<td>28.20</td>
<td>27.39</td>
<td>32.04</td>
<td>22.56</td>
<td>26.99</td>
</tr>
<tr>
<td>Replicate 3</td>
<td>27.95</td>
<td>27.40</td>
<td>32.12</td>
<td>22.91</td>
<td>26.79</td>
</tr>
<tr>
<td>Average Cq</td>
<td>28.10 ± 0.13</td>
<td>27.41 ± 0.03</td>
<td>32.46 ± 0.67</td>
<td>22.87 ± 0.29</td>
<td>26.91 ± 0.10</td>
</tr>
<tr>
<td>ΔCq = Avg Cq - Histone H3 Cq</td>
<td>5.23 ± 0.35</td>
<td>4.54 ± 0.27</td>
<td>9.93 ± 0.65</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Across all three strains, Psemu3 consistently had the lowest Cq value. This indicates that there was higher template concentration of Psemu3 relative to Psemu2 and Psemu4. Psemu4 had the highest Cq value, indicating that it had the lowest template concentration in each strain. These results suggest that Psemu3 is the most expressed prenyltransferase in P-n. However, one key experimental parameter to consider that amplification efficiency of Psemu3 B was 108% whereas that of Psemu2 was 98%. The
10% amplification efficiency difference between the two primer pairs may have resulted in $Psemu_3$ having higher apparent expression. Optimizing $Psemu_3$ primers may lead to $Psemu_2$ being observed as the most highly expressed prenyltransferase of the three sequences studied.

Fold difference of expression was calculated with the $2^{-\Delta\Delta CT}$ method by normalizing MLML2 and 15190C3 to MLML1 (Table 11). The reference gene Cq used in the calculation was from Histone H3. This was performed knowing that MLML1 had the lowest concentration DA and was isolated in log phase. MLML2 (late log) and 15190C3 (stationary) could be compared to develop a general understanding of how DA concentration could correlate to growth across different strains. Expression of all three transcripts decreased as growth approached stationary phase. DA concentration also plateaued at ~1 ng/mL after reaching late log phase. This suggests gene expression for prenyltransferases declines as cells grow, but DA levels are sustained for the three strains studied.

Table 11: Relative Quantitation of $Psemu$ in MLML2 and 15190C3 normalized to MLML1.

<table>
<thead>
<tr>
<th></th>
<th>Relative Quantitation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$Psemi_2$</td>
</tr>
<tr>
<td>MLML2</td>
<td></td>
</tr>
<tr>
<td>MLML2 ΔCq – MLML1 ΔCq</td>
<td>0.18 ± 0.61</td>
</tr>
<tr>
<td>Fold difference (2^{-\Delta CT})</td>
<td>0.58-1.35</td>
</tr>
<tr>
<td>15190C3</td>
<td></td>
</tr>
<tr>
<td>15190C3 ΔCq – MLML1 ΔCq</td>
<td>1.7 ± 0.36</td>
</tr>
<tr>
<td>Fold difference (2^{-\Delta CT})</td>
<td>0.24-0.39</td>
</tr>
</tbody>
</table>
Amplicon size was verified with agarose gel electrophoresis (Figure 23). Only one fragment was amplified. There is one exception with the molecular weight marker mixing into the \textit{Psemu2} lane of 15190C3.

\textbf{Figure 23: Agarose gel electrophoresis of amplicons from gene expression analysis.} Molecular weight marker adjacent to test samples. Starting from the left, MLML1, MLML2, and 15190C3. Wells from left to right are molecular weight marker, \textit{Psemu2}, \textit{Psemu3}, and \textit{Psemu4}. Expected amplicon size is between 100-200 bp.
Chapter 4

CONCLUSION

DA is a potent neurotoxin responsible for causing ASP in marine wildlife and humans. Although there has not been another recent incident with humans and ASP, harmful algal blooms remain a threat for coastlines and shellfish fisheries. Much of the biosynthetic pathway for domoic acid in *P-n multiseris* remains to be elucidated. However, recent research has brought the field closer to understanding the steps. It has been demonstrated that short-chained prenyltransferases are likely candidates responsible for catalyzing the formation of GPP, eventually leading to the synthesis of *N*-geranyl-L-glutamic acid.\(^2\,^{51}\) The goal of this study was to analyze gene expression of prenyltransferases in *P-n multiseris* and identify correlations between transcript abundance with DA accumulation. This research also continues investigating the findings from a previous homology modeling study that suggested *Psemu3* and *Psemu4* may have catalytic pocket sizes consistent with geranyl diphosphate synthase.\(^70\) Of the three transcripts studied (*Psemu2*, *Psemu3*, and *Psemu4*), notable gene expression patterns were discovered across all three *P-n multiseris* strains.

The first observation was the consistent relative transcript abundance among each *Psemu* in all three samples analyzed. *Psemu3* had the most expression followed by *Psemu2* and *Psemu4*. This was shown by *Psemu3* consistently having lower Cq values in each strain relative to the other two. It is important to note that Cq values for *Psemu2* were very similar to *Psemu3*. The difference in apparent expression may be a result of differences in amplification efficiency rather than a true difference in expression levels.
Nevertheless, the gene expression results allow us to conclude that *Psemu2* and *Psemu3* transcripts are expressed more than *Psemu4*.

The second observation comes from analysis of expression after normalizing Cq values to the MLML1, the least toxic strain based on measured DA concentration. Expression ratios of *Psemu2*, *Psemu3*, and *Psemu4* from MLML2 and 15190C3 were normalized to MLML1 using the $2^{-\Delta\Delta Ct}$ method. It was found that prenyltransferase gene expression decreased as *P-n multiseries* grew into stationary phase. One limitation is that this comparison was performed on three different strains isolated from different growth phases. Therefore, these results conclude that relative prenyltransferase expression decreases as *P-n multiseries* grow into stationary phase for this specific set of strains under the conditions where these samples were collected.

The final observation is the correlation between RNA concentration and DA levels for these strains. Again, a limitation of this observation is that the strains are all different. MLML1 had the highest concentration of RNA, but the lowest amount of DA. Given that this strain was isolated in its log phase, DA production may not have accelerated yet. Strains MLML2 and 15190C3 had equal amounts of RNA and DA concentrations. This suggests RNA and DA production plateaus as growth exits log phase. These measurements allow us to conclude that DA production begins at log phase and stabilizes before entering the stationary phase for this specific set of strains.

The results from this study show that *Psemu2* and *Psemu3* are highly expressed in log phase when DA biosynthesis appears to begin. The gene expression analysis did not indicate whether *Psemu3* and *Psemu4* encodes geranyl diphosphate synthase. It also did
not indicate that prenyltransferases are involved at regulatory steps for biosynthesis since expression declined when DA accumulation increased as growth approached stationary phase. In light of these observations, prenyltransferases may still play a role in DA biosynthesis. It is possible that instead of synthesizing transcripts on-demand for DA biosynthesis, prenyltransferases are readily available in the cell to catalyze the formation of GPP.

The research conducted here was similar to Boissonneault’s gene expression microarray study on *P-n multiseries*. In their effort, 5,265 transcripts were studied with microarray analysis and six were confirmed to have been up-regulated during silicon-limited conditions when DA accumulates. However, only one of the prenyltransferase transcripts (Psemu7, which is most homologous to a long-chain prenyltransferase) was represented on the microarray. Our study was the first attempt to specifically study gene expression of prenyltransferases in *P-n multiseries* and correlate it to DA accumulation. Future work includes the continuation of gene expression studies for *Psemu2, Psemu3* and *Psemu4* on individual strains of *P-n multiseries* isolated at different growth phases. By studying individual strains, prenyltransferase expression can be normalized under the same strain and correlated to DA accumulation.
References


8. What are diatoms? [https://diatoms.org/what-are-diatoms](https://diatoms.org/what-are-diatoms).


18. NOAA awards $88,000 in grant funding to respond to West Coast harmful algal bloom outbreak. 2015.

19. Moen, S. What Good is a Diatom.
   
   [http://www.seagrant.umn.edu/newsletter/2013/01/what_good_is_a_diatom.html](http://www.seagrant.umn.edu/newsletter/2013/01/what_good_is_a_diatom.html).


23. Wells, M. L.; Trainer, V. L.; Smayda, T. J.; Karlson, B. S. O.; Trick, C. G.;
Kudela, R. M.; Ishikawa, A.; Bernard, S.; Wulff, A.; Anderson, D. M.; Cochlan,
W. P., Harmful algal blooms and climate change: Learning from the past and
25. Thessen, A. E.; Dortch, Q.; Parsons, M. L.; Morrison, W., Effect of Salinity on
2005*, **41** (1), 21-29.
Heil, C. A.; Kudela, R.; Parsons, M. L.; Rensel, J. E.; Townsend, D. W.; Trainer,
V. L.; Vargo, G. A., Harmful algal blooms and eutrophication: Examining
linkages from selected coastal regions of the United States. *Harmful Algae* **2008**, 
8 (1), 39-53.
27. Muller, C.; Glamuzina, B.; Pozniak, I.; Weber, K.; Cialla, D.; Popp, J.; Cinta
Pinzaru, S., Amnesic shellfish poisoning biotoxin detection in seawater using pure
outbreak of toxic encephalopathy caused by eating mussels contaminated with
de Boer, J., Marine toxins: chemistry, toxicity, occurrence and detection, with

30. Wesolek, N.; Parent-Massin, D.; Roudot, A.-C., Evaluating the performance of
sampling plans for phycotoxins in shellfish: Improvement of an existing method.
Harmful Algae 2014, 34, 1-6.

31. Bejarano, A.; Van Dolah, F. M.; Gulland, F. M.; Schwacke, L., Exposure
assessment of the biotoxin domoic acid in California sea lions: Application of a

32. Fire, S. E.; Wang, Z.; Berman, M.; Langlois, G. W.; Morton, S. L.; Sekula-Wood,
E.; Benitez-Nelson, C. R., Trophic transfer of the harmful algal toxin domoic acid
as a cause of death in a minke whale (Balaenoptera acutorostrata) stranding in

33. Fire, S. E.; Wang, Z.; Leighfield, T. A.; Morton, S. L.; McFee, W. E.; McLellan,
W. A.; Litaker, R. W.; Tester, P. A.; Hohn, A. A.; Lovewell, G., Domoic acid
exposure in pygmy and dwarf sperm whales (Kogia spp.) from southeastern and

34. Leandro, L. F.; Rolland, R. M.; Roth, P. B.; Lundholm, N.; Wang, Z.; Doucette,
G. J., Exposure of the North Atlantic right whale Eubalaena glacialis to the

35. Costa, L. G.; Giordano, G.; Faustman, E. M., Domoic acid as a developmental


40. Berman, F. W.; Murray, T. F., Domoic acid neurotoxicity in cultured cerebellar granule neurons is mediated predominantly by NMDA receptors that are activated as a consequence of excitatory amino acid release. *J. Neurochem.* 1997, 69 (2), 693-703.


45. Tholl, D., Biosynthesis and Biological Functions of Terpenoids in Plants. 


58. Scientific, T. Basic Principles of RT-qPCR. 
   


63. Microfluidic PCR & qPCR_qPCR Melting curve analysis.

   [https://www.elveflow.com/](https://www.elveflow.com/)


69. Cepin, U. Understanding qPCR Efficiency and Why It Can Exceed 100%. [https://biosistemika.com/blog/qpcr-efficiency-over-100/](https://biosistemika.com/blog/qpcr-efficiency-over-100/).

70. Peery, C. A bioinformatic approach to finding geranyl diphosphate synthase genes in the *Pseudo-nitzschia multiseries* genome. California State University, Sacramento, 2018.


