

BREVETOXIN: HOW IS IT MADE AND WHY?

A Thesis

by

NATALIE JEANETTE THOMPSON

Submitted to the Office of Graduate Studies of  
Texas A&M University  
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

August 2011

Major Subject: Microbiology

Brevetoxin: How Is It Made and Why?

Copyright 2011 Natalie Jeanette Thompson

BREVETOXIN: HOW IS IT MADE AND WHY?

A Thesis

by

NATALIE JEANETTE THOMPSON

Submitted to the Office of Graduate Studies of  
Texas A&M University  
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Approved by:

Chair of Committee,  
Committee Members,

Head of Department,

Lisa Campbell  
Paul Straight  
Alan Pepper  
U.J. McMahan

August 2011

Major Subject: Microbiology

## ABSTRACT

Brevetoxin: How Is It Made and Why? (August 2011)

Natalie Jeanette Thompson, B.S., Texas A&M University

Chair of Advisory Committee: Dr. Lisa Campbell

*Karenia brevis* is the major harmful algal bloom-forming species in the Gulf of Mexico, and produces neurotoxins, known as brevetoxins, that cause large fish kills, neurotoxic shellfish poisoning, and human respiratory distress. Brevetoxins are polyethers that bind voltage-sensitive sodium channels, opening them for prolonged periods of time. Clonal cultures of *K. brevis* exhibit unique brevetoxin profiles, which not only differ from one another, but also change when subjected to different environmental conditions. The brevetoxin structures were elucidated 30 years ago without any breakthroughs for the biosynthetic pathway. These unique ladder-like polyethers have 10 (PbTx-1) or 11 (PbTx-2) rings, indicating that they are synthesized as secondary metabolites by polyketide synthases. The extensive size of the genome and the lack of histones and nucleosomes combined with the additional regulatory step of a trans-splicing spliced leader sequence make normal molecular techniques ineffective in determining the genes involved in toxin synthesis. The goal of this project is to identify a potential link between toxin, gene, and function. One objective is to take the next step towards identifying the genes associated with the synthesis and regulation of brevetoxins and to help elucidate the hypothesized gene clusters of multi-protein enzymatic

complexes involved in brevetoxin production, one for each backbone. The second objective is to make an effort to determine the *in vivo* function of the costly brevetoxins by identifying possible ion channels, which could be osmotically regulated by the toxins.

Genes for polyketide synthases (PKS) were identified in *K. brevis*, obtained from Expressed Sequence Tag (EST) libraries. In this work, reverse transcription polymerase chain reactions (RT-PCR) were used to generate pools of complementary DNA (cDNA), which was used in real-time quantitative polymerase chain reactions (qPCR) to give relative amounts of PKS transcripts. *K. brevis* clones have shown a significant increase in toxin production after a rapid shift from high salinity to low salinity, indicating a regulation of brevetoxin synthesis. To gain a better understanding of regulation of toxin production during algal blooms, we compared the toxin levels under different conditions to the transcript levels of PKS genes, as determined by quantitative RT-PCR. In a separate line of investigation, an *in silico* analysis of the EST library was performed to identify ion channel genes expressed by *K. brevis*, which may be the *in vivo* binding site of brevetoxin. The information generated from this project will help to elucidate the effects of environmental variations on toxin production and the biological function of toxin production -- valuable information for the shellfish industries and public health.

## DEDICATION

To my family and friends, for being there.

## TABLE OF CONTENTS

	Page
ABSTRACT .....	iii
DEDICATION .....	v
TABLE OF CONTENTS .....	vi
LIST OF FIGURES .....	viii
LIST OF TABLES .....	x
1. INTRODUCTION .....	1
1.1 <i>Karenia brevis</i> and Brevetoxin.....	1
1.2 Spliced Leader Sequence.....	6
1.3 Polyketide Synthases.....	8
1.4 Ion Channels.....	10
1.5 Summary.....	17
2. DIFFERENTIAL EXPRESSION OF PKS TRANSCRIPTS .....	20
2.1 Overview.....	20
2.2 Materials and Methods.....	20
2.2.1 Clonal Cultures.....	20
2.2.1.1 Media and Cultures.....	20
2.2.1.2 Growth Rates.....	21
2.2.1.3 Constant Salinity and Salinity Change.....	22
2.2.2 Polymerase Chain Reaction .....	23
2.2.2.1 DNA and RNA Extraction.....	23
2.2.2.2 Reverse Transcription.....	23
2.2.2.3 PCR and Quantitative PCR.....	24
2.2.2.4 Sequencing.....	26
2.3 Results.....	26
2.3.1 Growth Rates.....	26
2.3.2 Sequencing.....	28
2.3.3 Quantitative PCR.....	28
3. ION CHANNEL IDENTIFICATION .....	34

	Page
3.1 Overview.....	34
3.2 Materials and Methods.....	34
3.2.1 EST Analysis.....	34
3.2.2 Nucleotide Sequence Analysis.....	35
3.2.3 Sequencing.....	36
3.2.4 Amino Acid Sequence Analysis.....	37
3.3 Results.....	38
3.3.1 In Silico EST Analysis.....	38
3.3.2 Sequence Analysis.....	41
4. DISCUSSION AND CONCLUSIONS .....	46
4.1 Polyketide Synthases.....	46
4.2 Ion Channels.....	49
REFERENCES .....	53
APPENDIX .....	66
VITA .....	77



## LIST OF FIGURES

		Page
Figure 1	Structures of brevetoxin .....	3
Figure 2	Differential production of total brevetoxin .....	5
Figure 3	Differential toxin profiles at salinity 27 (A) and 35 (B) .....	6
Figure 4	Basic mechanisms of fatty acid synthesis (A) and polyketide synthesis (B) .....	9
Figure 5	Structure of a sodium channel with alpha subunit and beta subunits .....	12
Figure 6	Common structural motif of voltage gated potassium channels .....	14
Figure 7	Pfam seed alignment of ion channel clan .....	15
Figure 8	Growth curves of the three <i>Karenia brevis</i> clones .....	27
Figure 9	Two control experiments .....	29
Figure 10	Standard curves generated for qPCR .....	31
Figure 11	Relative expression of PKS genes in three clonal cultures .....	32
Figure 12	Profile of EST sequences with protein similarity.....	40
Figure 13	Profile of species with closest match to EST sequences .....	41
Figure 14	Ion channel protein domains .....	42
Figure 15	Multiple sequence alignment of EST 48705280 .....	66
Figure 16	Multiple sequence alignment of EST 48706986 .....	67
Figure 17	Multiple sequence alignment of EST 158897096 .....	68
Figure 18	Multiple sequence alignment of EST 158895975 .....	71

	Page
Figure 19	Multiple sequence alignment of EST 158893679 ..... 72
Figure 20	Multiple sequence alignment of EST 158891827 ..... 73
Figure 21	The predicted transmembrane domain structure of 48705280 ..... 74
Figure 22	The predicted transmembrane domain structure of 48706986 ..... 74
Figure 23	The predicted transmembrane domain structure of 158897096 ..... 75
Figure 24	The predicted transmembrane domain structure of 158895975 ..... 75
Figure 25	The predicted transmembrane domain structure of 158893679 ..... 76
Figure 26	The predicted transmembrane domain structure of 158891827..... 76

## LIST OF TABLES

		Page
Table 1	Clonal cultures of <i>Karenia brevis</i> and first collection location.....	21
Table 2	Primers designed from PKS genes for qPCR.....	25
Table 3	Slopes and R <sup>2</sup> values for standard curves .....	30
Table 4	Primers designed for ion channel sequence elongation.....	37

## 1. INTRODUCTION

### 1.1. *Karenia brevis* and Brevetoxin

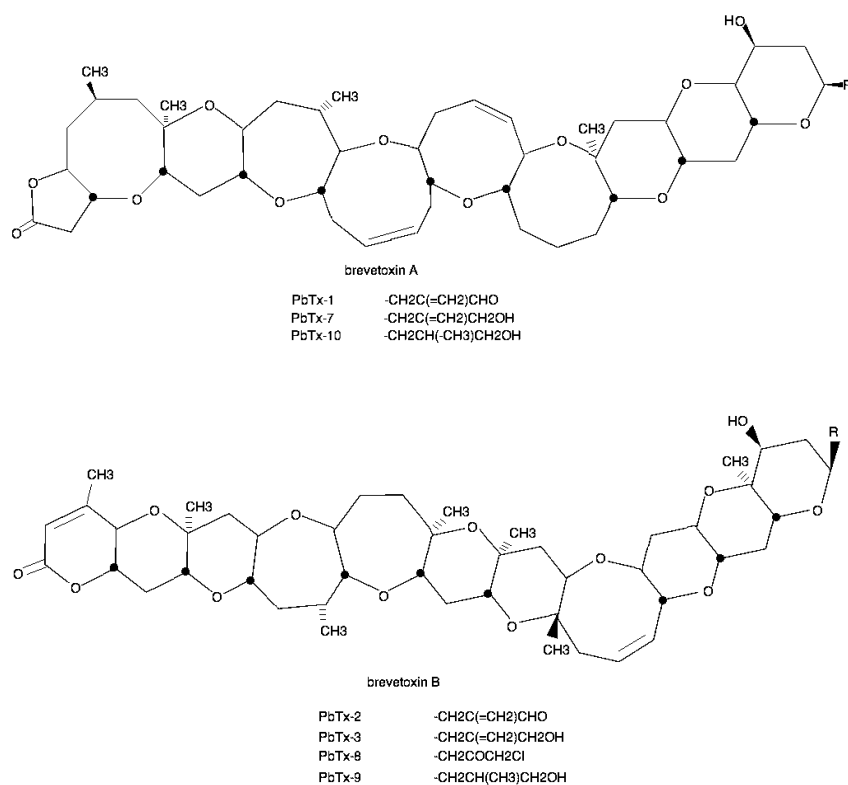
*Karenia brevis* is an unarmored dinoflagellate that blooms frequently in the Gulf of Mexico and occasionally in the southeast Atlantic coast of the United States (Steidinger et al., 1998). These harmful algal blooms, also known as “red tides,” are defined as cell abundance in seawater that is far greater than background. Typical background cell abundance for *K. brevis* is 1 – 10 cells per liter L, where as the legal limit for shellfish bed closure is 5000 cells per L (Heil, 2009; Tester et al., 2008). Blooms of *K. brevis*, observed in records dating back to 1528, discolor the water a reddish-brown and cause large-scale fish kills, marine mammal mortalities, toxic shellfish poisoning, and respiratory irritation due to aerosolized toxins (Landsberg, 2002; Magana et al., 2003; Steidinger, 2009). This small (18 - 40  $\mu\text{m}$  wide) marine eukaryote is photosynthetic and contains chlorophyll, beta-carotene, and fucoxanthin pigments. Possessing both longitudinal and transverse flagella, the single cell moves vertically in the water column and undergoes phased, synchronous cell division throughout the diel cycle. The life cycle of *K. brevis* has yet to be completely determined, but consists of a planktonic vegetative asexual stage and hypothetical planktonic-planozygote and benthic-cyst stages (Persson et al., 2008). *K. brevis* can survive at a range of salinities,

---

This thesis follows the style of Harmful Algae.

from oceanic salinity 35 to coastal salinity 27, and has also been reported in estuarine environments (Brown et al., 2006).

*Karenia brevis* produces a specific group of toxins composed of two main backbones, type A (PbTx-1) and type B (PbTx-2), which degrade into other congeners (PbTx-3, -5, -6, -7, -8, -9, -10) that are primarily extracellular, released and degraded after cell lysis (Figure 1) (Baden et al., 2005; Bourdelais et al., 2005; Pierce et al., 2001). These brevetoxins are neurotoxic polyethers that depolarize voltage-sensitive sodium channels VSSC, opening them for prolonged periods of time, which leads to neurotoxic shellfish poisoning and death (Backer, 2009; Kirkpatrick et al., 2004; Landsberg et al., 2009). PbTx-2 is typically the most abundant form produced in the cell; PbTx-1 is less abundant, but more toxic, requiring a lower concentration to have the same toxic effect as PbTx-2 (Steidinger, 2009). Other polycyclic ethers are also synthesized by *K brevis*, such as brevenal and brevisin that are competitive inhibitors of brevetoxin for the VSSC (Bourdelais et al., 2004; Fuwa et al., 2006; Satake et al., 2008; Satake et al., 2009).

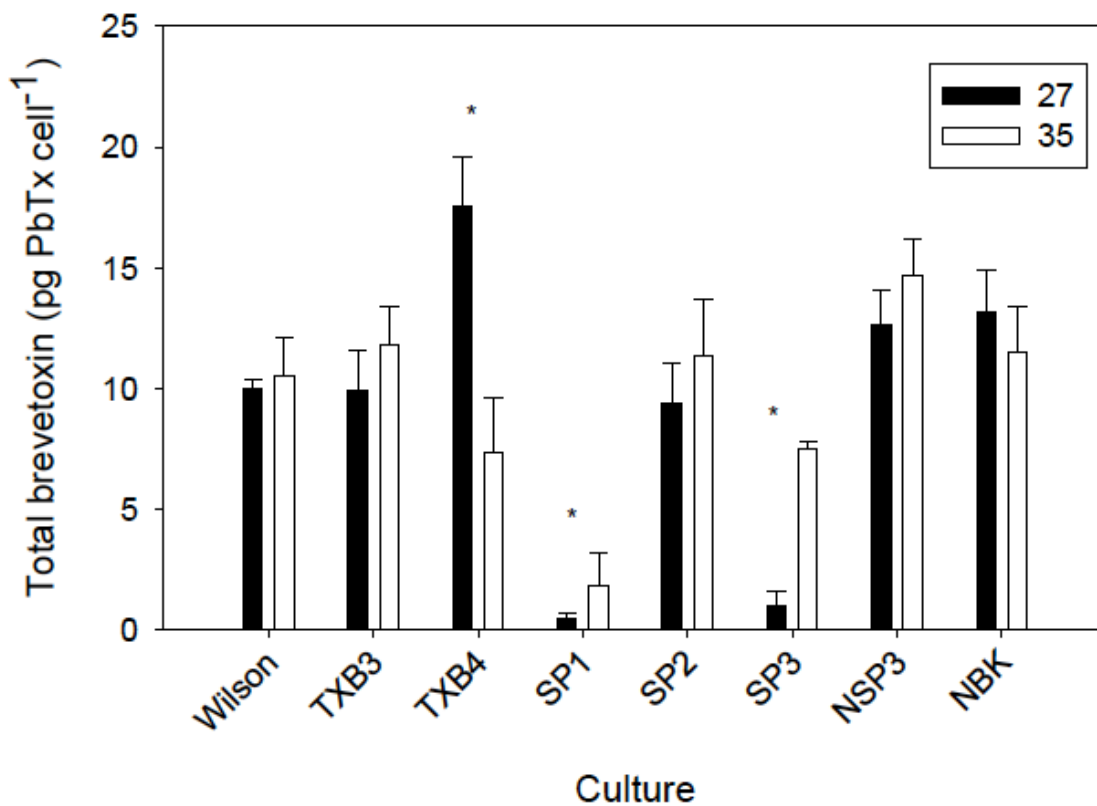


**Fig. 1.** Structures of brevetoxin. Redrawn from (Snyder et al., 2005).

The nearly annual blooms, in the Gulf of Mexico on the coasts of Florida and Texas, are composed of *Karenia brevis* predominately; five other *Karenia* species have been observed in the Gulf, however none are known to produce brevetoxins. Considerable physiological variability has been observed among clonal cultures of *K. brevis* (McKay et al., 2006; Schaeffer et al., 2007; Walsh et al., 2006). Clonal cultures display variation in both toxin production and growth rate when grown under different environmental conditions, such as salinity, light, temperature, and nutrients (Magana and Villareal, 2006; Vargo, 2009; Vargo et al., 2008). Even when grown under identical conditions, clonal cultures display unique toxin profiles with different amounts of total

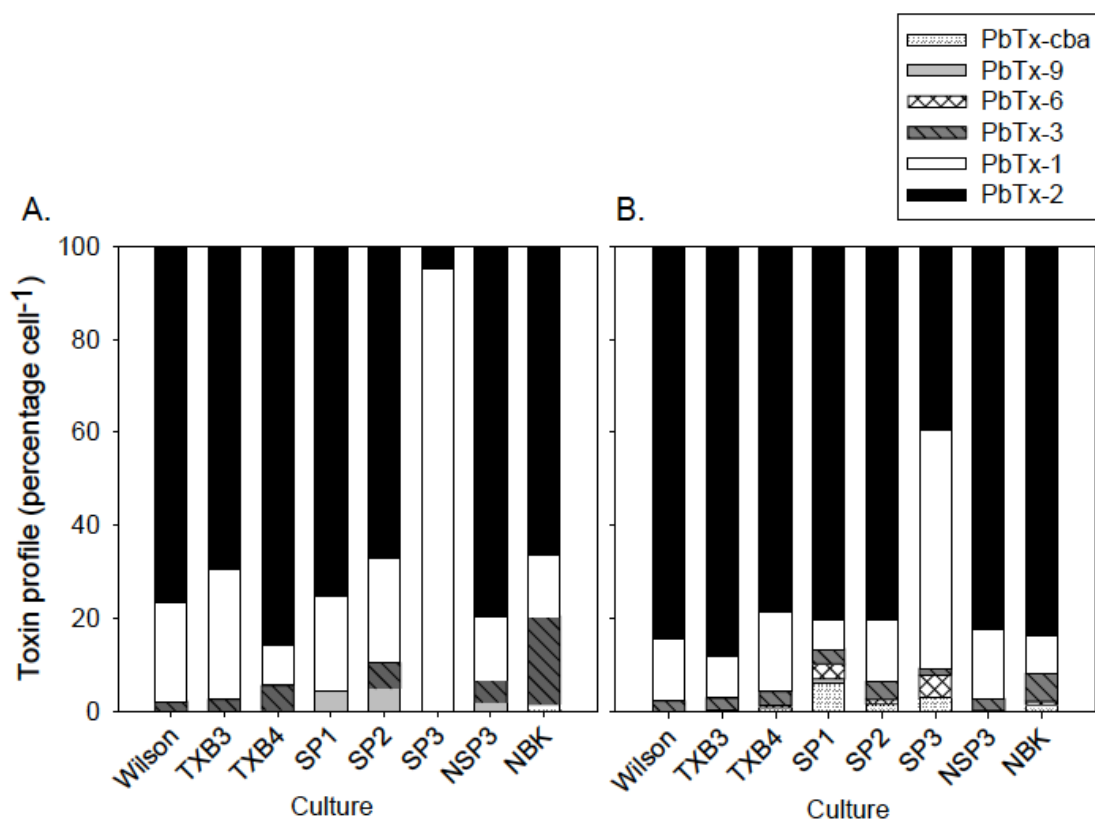
toxin and different compositions, synthesizing more of either PbTx-1 or PbTx-2 (Figures 2 and 3). Variation in amounts of brevisin and brevenal produced are also observed (Errera et al., 2010; Lekan and Tomas, 2010). The differential toxin profiles may be due to changes in the regulation of the genes involved in producing the toxins (Miller-Morey and Van Dolah, 2004).

With a genome size approximately 30 times the size of the human genome  $1 \times 10^{11}$  base pairs, sequencing the genome of *Karenia brevis* is not possible at this time (Rizzo et al., 1982). This necessitates a novel way to determine the genes involved in the brevetoxin biosynthetic pathways. Expressed Sequence Tag EST libraries for *K. brevis* have led to the development of microarrays that can be used to measure expression of genes involved in cell cycle and circadian rhythms in *K. brevis* (Barbier et al., 2003; Lidie et al., 2005; McLean and Pirooznia, 2008; Richardson et al., 2007; Van Dolah et al., 2007). The study by Van Dolah et al. showed differential gene expression in three clusters of genes, a total of 9.8% of the 4629 genes on the microarray. Of the previously annotated genes, 33% were identified as genes involved in post-transcriptional or translational processing (Van Dolah et al., 2007). Their work concluded that a lack of typically observed transcriptional regulation in circadian controlled genes, due to their lack of expression change over a diel cycle, implied the use of a post-transcriptional regulatory step, specifically by a mechanism involving the spliced leader SL sequence.



**Fig. 2.** Differential production of total brevetoxin. Taken from (Errera et al., 2010), showing total amount of brevetoxin produces by eight different clonal cultures of *Karenia brevis* at salinity 27 and salinity 35.





**Fig. 3.** Differential toxin profiles at salinity 27 (A) and 35 (B). Taken from (Errera et al., 2010), showing the types of toxins produced by eight different clonal cultures.

## 1.2. Spliced Leader Sequence

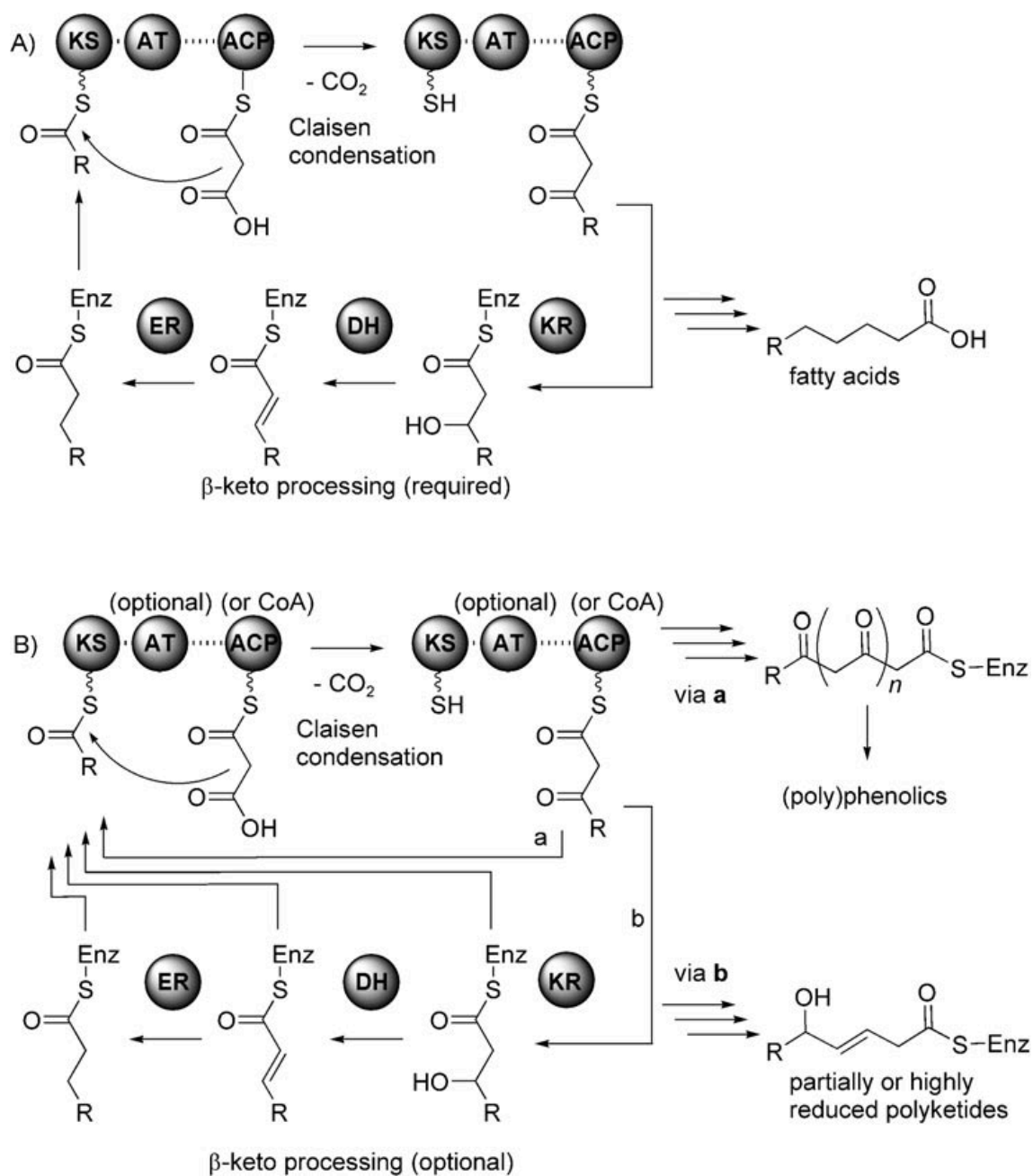
There are other complications that make it difficult to undertake typical molecular studies in *Karenia brevis*: the nucleus contains permanently condensed chromosomes with a lack of histones and nucleosomes, and the nuclear encoded pre-mRNA is additionally modified from poly-cistronic to mono-cistronic messages by a trans-spliced SL sequence (Lidie and Van Dolah, 2007; Plumley, 1997; Van Dolah et al., 2009; Zhang et al., 2007). This novel post-transcriptional regulatory step is seen in few

eukaryotes: Alveolates, Euglenozoans, Bilaterans Nematoda, Rotifera, Platyhelminthes, Cnidarians, and Urochordates Ascidiidae, Appendicularia (Zhang et al., 2007). The model used for observing this *trans*-splicing event is in *Trypanosomes*, where *trans*-splicing occurs in a similar fashion to typical *cis*-splicing (Lustig et al., 2007). Typical splicing, as modeled in *Saccharomyces cerevisiae*, occurs via small nuclear ribonucleoproteins snRNPs U1, U2, U4, U5, U6 in conjunction with Sm proteins, the ring-forming RNA-binding proteins involved in assembly of the spliceosome (Alverca et al., 2006). In *Trypanosomes*, the post-transcriptional mRNA capping occurs by the SL RNA binding the Sm proteins to form a spliceosomal RNP. The SL RNA then *trans*-splices onto the 5' end of each open reading frame ORF, which converts the polycistronic message into multiple mono-cistronic messages (Ambrosio et al., 2009).

Every dinoflagellate studied thus far has been found to contain a similar SL RNA sequence at the start of their nuclear encoded mRNA, often in degenerative repeats (Bachvaroff and Place, 2008; Monroe and Van Dolah, 2008; Zhang et al., 2009). A consensus sequence 5'-AUUUUGG-3' is seen in the SL sequence for the Sm-binding complex site, which is where the Sm ring forms (Zhang et al., 2009; Zhang et al., 2007). The SL genes have varied genomic arrangement and have been found in clusters with 5S rRNA and U6 snRNA, a small nuclear RNA that binds Sm-like proteins (Lidie and Van Dolah, 2007; Roy and Irimia, 2009). The specific SL sequence 5'-UCCGUAGCCAUUUUGGCUCAAG-3' allows detection and sequencing of expressed genes of *K. brevis*, as well as ensuring that the transcripts are from the dinoflagellate and not the associated bacteria in non-axenic cultures.

### 1.3. Polyketide Synthases

Brevetoxins are hypothesized to be produced as secondary metabolites by polyketide synthases PKS. These PKSs synthesize a chemically and functionally diverse group of carbon skeleton natural products by stepwise condensation using simple building blocks such as acetate and propionate (Hertweck, 2009; Tae et al., 2007). The biosynthesis of polyketides is similar to fatty acid synthesis FAS, but with a much more complex chain assembly (Figure 4). During the serialized reactions of chain elongation and modification, there are specific proteins common to both polyketide synthesis and FAS: Acyl Carrier Protein ACP, Ketoreductase KR, Dehydratase DH, and Enoyl Reductase ER; some proteins are specific to only polyketide synthesis: Ketosynthase KS and Acyl Transferase AT (Khosla, 2009). The FAS occurs in a specific and consistent fashion whereas polyketide synthesis can create a broader range of products by varying the order of the modifications and the proteins involved at each elongation step (Hertweck, 2009). There are three types of PKS: types I, II, and III. These are defined as modular, non-modular, or iterative, based on the properties of the PKS domains. Modular indicates the arrangement of the enzymes in large multi-domain catalytic modules, whereas non-modular contains individual enzymes with singular function in multi-protein complexes. Iterative indicates that the same KS domain catalyzes multiple rounds of elongation (Hertweck, 2009; Perez et al., 2008; Snyder et al., 2003; Snyder et al., 2005).



**Fig. 4.** Basic mechanisms of fatty acid synthesis (A) and polyketide synthesis (B).

Reproduced with permission from (Hertweck 2009) Copyright Wiley-VCH Verlag GmbH & co. KGaA.

Studies are underway to detect PKS genes in dinoflagellates and other harmful algal-bloom forming species, due to the unknown origins of many of the toxins in that group (Gontang et al., 2010; Jones et al., 2009a; Jones et al., 2009b; Macpherson et al., 2003; Monroe and Van Dolah, 2008; Pearson et al., 2008; Perez et al., 2008; Shimizu, 2003; Snyder et al., 2003; Snyder et al., 2005; Tillett et al., 2000). The PKS transcripts used in this study were identified in *K. brevis* by the presence of the SL sequence and characterized by phylogenetic analyses. It has been suggested that dinoflagellates utilize type I PKS, the modular enzymatic protein with multiple catalytic domains, based on sequence identity (Monroe and Van Dolah, 2008). These type 1-like PKS genes with ACP, KS, and KR domains require further study to characterize the enzymes they produce and determine their link to brevetoxin production (Monroe and Van Dolah, 2008; Snyder et al., 2003; Snyder et al., 2005).

#### **1.4. Ion Channels**

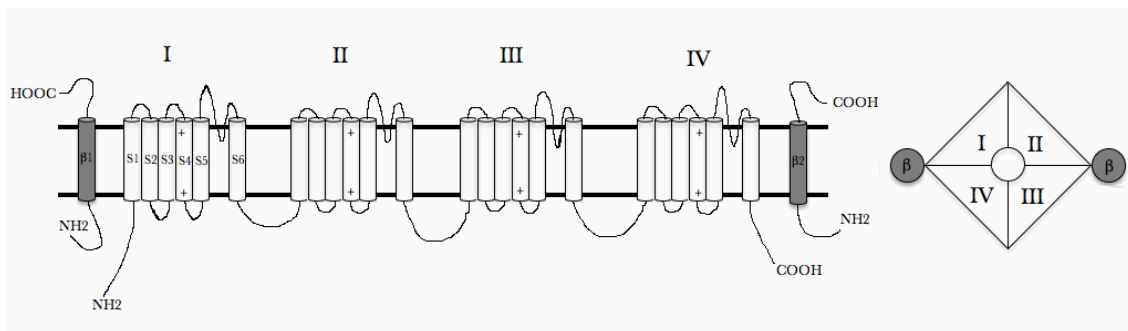
Brevetoxins bind and depolarize VSSCs in vertebrates by binding to the sodium channel site 5, which results in depolarization of the membrane, prolonged channel opening, and increased permeability of sodium through the channel (Baden et al., 1988; Landsberg, 2002; Steidinger, 2009). However, a function for brevetoxin within the cell of *K. brevis* has not been determined, though it represents a high metabolic cost for the cell to make this secondary metabolite (Schaeffer et al., 2009). Sodium channels have been studied in the dinoflagellate, *Noctiluca miliaris*, and a diatom, *Odontella sinensis*, both of which are Chromalveolates, the same supergroup as *Karenia brevis*. The studies on *N. miliaris* determined that sodium and chloride channels were involved in the ion

currents depolarizing and hyperpolarizing the membrane as a mechanism for the contraction of the tentacle for food gathering (Oami et al., 1995b). It was also concluded that external calcium concentrations inactivate the sodium current (Oami et al., 1995a). The studies on *O. sinensis* concluded that these sodium and calcium channels are important in the evolution of eukaryotic signaling mechanisms (Taylor, 2009). It was also hypothesized that these channels may be involved in environmental sensing, including osmotic shock (Taylor, 2009). These osmoregulatory strategies are employed to protect the integrity and the metabolism of the cell, and are extremely well conserved throughout the species mentioned (Kultz, 2001). Osmotic shock can also cause DNA damage, which often triggers specific DNA repair pathways. The cellular responses to osmotic shock include osmosensor proteins and phospho-proteins in the signal transduction pathways to regulate ion transport (Kultz, 2001). Phylogenetic analysis of voltage gated channel domains showed high homology of four Chromalveolate proteins, including the voltage sensors and pore selectivity filters Kitchen and Taylor, in preparation.

Environmental sensing of osmotic stress, as well as other stresses, was found in a marine diatom, *Phaeodactylum tricornutum*; internal calcium concentrations changed as an effect of hypo-osmolarity, indicating signal transduction (Bowler et al., 2008; Falciatore et al., 2000). The transport of ions, such as sodium, calcium, and potassium, through specific channels are important to the role of osmoregulation in maintaining the integrity and function of the cellular membrane (Schroeder and Hedrich, 1989). For example, using the sodium channel model in rat neurons, the osmotic pressure affected

the inactivation of the sodium current through the channels (Chizhnikov and Sorokina, 1986; Krishtal et al., 1983).

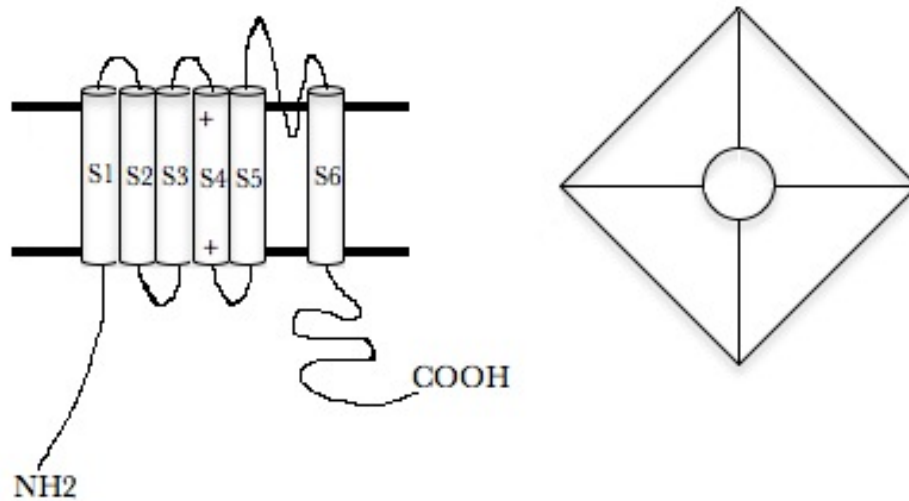
Sodium channels have a alpha subunit composed of a transmembrane protein of four repeating motifs I, II, III, and IV and associations with other subunit proteins (Figure 5) (Catterall, 1986; Marban et al., 1998). These channels convey impulses for a variety of functions and increase the sodium permeability upon fast depolarization. These ion channels are important for a variety of functions in higher eukaryotic multicellular organisms and have now been located in dinoflagellates and diatoms, as previously described (Anderson et al., 2005). It is important to determine the function of sodium channels in these more primitive eukaryotes, since neurotoxins bind to receptor sites on the sodium channel, changing the activation or inactivation of the channel (Anderson et al., 2005; Catterall, 1986).



**Fig. 5.** Structure of a sodium channel with alpha subunit and beta subunits. Redrawn from (Marban et al., 1998).

To date, neither calcium nor potassium ion channels have been identified in any dinoflagellate; however, many of the different toxins made by these harmful algal-bloom forming species cause poisoning by binding to calcium or potassium channels in higher eukaryotes (Bowler et al., 2008; Catterall, 1996; Hackett et al., 2004; Kamiyama and Suzuki, 2009; Landsberg et al., 2006; Leflaive and Ten-Hage, 2007; Macpherson et al., 2003; McElhiney and Lawton, 2005). Plant cells contain both calcium and potassium channels, which show homology to animal ion channels. A review on *Arabidopsis* determined that hyper-polarization-activated  $\text{Ca}^{2+}$ , depolarization activated  $\text{Ca}^{2+}$  and stretch-activated  $\text{Ca}^{2+}$  channels were present in plants (Harz and Hegemann, 1991). Calcium channels consist of a beta subunit 1, with one transmembrane domain, and an alpha subunit similar to the sodium channel 1, with four homologous units each made of six transmembrane domains; other than these two primary structures, there is also another alpha subunit 2, a gamma subunit, and a delta subunit (Falciatore et al., 2000). The same *Arabidopsis* review, by Harz and Hegemann, concluded that there were Shaker-Type  $\text{K}^+$  and Inward-Rectifying Shaker  $\text{K}^+$  channels in plants. Potassium channels have more diverse motifs and are homotetrameric complexes (Figure 6) (Falciatore et al., 2000; Kultz, 2001).





**Fig. 6.** Common structural motif of voltage gated potassium channels. Redrawn from (Falciatore et al., 2000).

At the genetic level, each ion channel will have specific motifs present in their amino acid sequence (Baden et al., 2005; Bouche and Bouchez, 2001; Naar et al., 2002; Pierce et al., 2001; Schroeder and Hedrich, 1989). These motifs, small functional units, can be predicted with various programs, such as ExPASy Proteomics Server (<http://expasy.org/>), to identify ion transport domains in genes that have not yet been annotated (Mullan, 2004). The Pfam database creates seed alignments of protein sequences in a family and uses that to determine the existence of a domain in a sequence (Figure 7) (Sammut et al., 2008).

## PFO3814

```

ATKA1_STAAM/1-557      MEIILFLTMVMITYVFSGLYLRVALVQS.....SRVDLIFTRFENMCFKIICT.DLEHMSAKTYVXHFLLAF
ATKA_ANASL/6-557      LQIQLTLCIVIAITFVLRGRYIARVFLER.....TILDPLMNPFIERSIYVISVVRPKDDMTGWQYIRAILYT
ATKA_HALSA/10-580     ..YAVFFTTILAVLVFVAGEYLAWVYREQANSRHPFGLSWFERLDEIFTPLENGLYRLSGINPRREMTWKGLKAVLVF
ATKA_RHILO/7-561      IQILVFCGIIQLLTKPLGFYMRVFNDR.....TFLSFIFGPLERGLYRICGTSEREQHWTAAYAVALLF
ATKA_ALIAC/7-561      LAIFLVVVALVACANPLGGYIYRTFVGER.....VRPDAVMVFFVERAIYRLIGVNFVEVMDWKAYLRAMMVS
ATKA_MYCTU/7-570      LQFAALIAVLLLTAFALGGYLAKEYDEAK.....KPGDRVFGPIERVIYQVCRVDPSEQRWSTYALSVLAF
ATKA_PSEAE/7-562      LLILAFFALVLFVAPFLGRYCFKVMGQR.....NLLSAPLAPLERVYRFLGVDPASEQDWKTYTLALLAF
ATKA_THEAC/25-579     IILILYLIASVLAIFLSFHIAKIYLDEK.....TVFSKITGRIISFFERMIGESPDHGMFTKEYFINLLF
ATKA_CAUCR/7-562      AEIALTSLAVAIQWPLGVFLSRVWNGEK.....TFLDPVMRPVEGLFYKACGVDPSKSQSHAYALALLAF
ATKA_ECOLI/7-557      LLIATFLLVLMVLARPLGSGLARLIN.....DIFLPGTTGVERVFLFRALGVSDR.EMNWKQYLCAILGL
ATKA_STRCO/9-554      LQLLALTAALALAHVFLGNYLARVYSSPK.....HLR...IEKWIYKSIQADFDTEMRWPAYLRGVLA
ATKA_SYNY3/6-557      VQIALILAILVATAPELLGRYMARVFLGQS.....TWLDKIARPLESLIFAQSGITKHPSMGATQYVSAALIS
ATKA_DEIRA/6-561      LQFVVLLGVMAALVVPVGNWLYGVATGQK.....DTALERLTYRVICINPRERMDWKRYGLAIVVS
ATKA_CLOAB/4-557     LQIAIILIVFVLLCIPIGRYMYKVEHKK.....TLLEFVLDKIDGFYKLSGIQKEEMNKQYIFALLMG

```

## PFO3814

```

ATKA1_STAAM/1-557      NCFMCFITFVLLIVQQLF-----LNPNHILNQSIDLAFNTAISFLTNSNLQHYNGESDVTYLT-
ATKA_ANASL/6-557      NLFMGLVYSLIHYQRLLP-----WNFNCFGVPRWDIILHTVVSFVINTDQOYHAGETTLTSYFS-
ATKA_HALSA/10-580     NVCIWVLLFVLMFQDQLP-----MNFVCGGSEWDLAFHTASSFTSNTNQOYHSET.LSVFT-
ATKA_RHILO/7-561      NLACFLVLYALQRLQCSLP-----YNPACMGAVDPALAFNTAASFVINTNWQYCGESTMSYLL-
ATKA_ALIAC/7-561      NLVMALEFAYAVFRLQGVLP-----LNFPAHIPMPFYLAFNTAASFVINTNWQYAGEQSLSYLS-
ATKA_MYCTU/7-570      SVMSFLLLYGIARFQGVLP-----FNPDKPAVTDHVAFAAASFMTNTNWQYSYSEATMSHFT-
ATKA_PSEAE/7-562      NLACLVLLESILMLQCSLP-----LNPQHLFGLWTLAFNTAASFVINTNWQAYSSEASLSYFS-
ATKA_THEAC/25-579     NFFACLISFLVIMFQKYLPE-----F.SYDVTVMGMSPLDNTVVSFLTNTLQHYSNPMLRSGYFS-
ATKA_CAUCR/7-562      NLIQFVVFVYAVLRQGVLP-----LNPQCFPGLSGLAFNTAISFTINTNWQYSYAGETTMSTLS-
ATKA_ECOLI/7-557      NMLGLAVLFFMLLQHYLP-----LNPQQLPGLSVDLALNTAASFVINTNWQYSYSETTLTSYFS-
ATKA_STRCO/9-554      SLACVLFVLYLQRLQGVLP-----G.SLGFASIDPDQAFNTAASFVANTNWQYSYEQAMGHVV-
ATKA_SYNY3/6-557      NLVMGVFVELILMFQCSLP-----LNPGLAASPVDLALHTAISFVINTNWQYSYSETTYTYS-
ATKA_DEIRA/6-561      NLVMLLLAYLLIRLQCALP-----WNPACLAAQSPDLAWNTVVSFMTNTNWQAYSSEQSLSYFS-
ATKA_CLOAB/4-557     NAVPAIICYIILRIQAVGI-----FNPNHVKCMEQGLTENTIISFLTNTLQOYAGETGASYSLS-

```

## PFO3814

```

ATKA1_STAAM/1-557      ---QMIVMTYLMFTSSASGYAVCIAMLRRTILT.-----NIIQNFYQDVRVIFVRVLLPLSCLISILLMTOGV
ATKA_ANASL/6-557      QVAALGFLMFTSAATGLAVGIAFIRGLTKK.-----K.LGNFYIDLTRGITRILLPLISVIGAIALVLLGV
ATKA_HALSA/10-580     ---HTFGICIAMFLTPTATGLALMPAFARAFTKED-----PRLGNFYENVVRGLVRFLPLISLLIAIILMAGEG
ATKA_RHILO/7-561      ---QMAGLTVQNFVSAATGIAIAIALIRCFARASG-----KSIQNFVDMTRCTLYVLLPLCIVLTLVYVWLG
ATKA_ALIAC/7-561      ---QMIGITFLQFTSAATGLAAMAFIRGLSRQKT-----DALGNFVVDLVKAHTRLLPLAAILAVLLLALGV
ATKA_MYCTU/7-570      ---QMTGLAVQNFVSAAGMCMVLAALIRGLARKRA-----STLGNFVVDLARTVLRIMFPLSFVVAIILVSGOV
ATKA_PSEAE/7-562      ---QMVGLTVQNFVSAAGLVCVLVALARGISRRST-----RQLGNFVVDLTRCTLYVLLPLCILLLALLVWQGV
ATKA_THEAC/25-579     ---QTFVITGLMFLSAGTGFAASMAFVRCIRDTG-----N.IQNFYHDFLVSIFDLILLPLTLVILLIILVLAGI
ATKA_CAUCR/7-562      ---QMLVITVQNFVSAATGATVAAALARAFVANRG-----EGVGNFWADLVRTTYLPLAVVAVVLAALGL
ATKA_ECOLI/7-557      ---QMACLTVQNFVSAASGIAVIFALIRAFTRQSM-----STLGNFVVDLRLITLWLVLPVALLIALFIIQOQA
ATKA_STRCO/9-554      ---QTAGLAVQNFVSAAGVIAVAVALVRGFAARSRT-----GELGNFWADLVRRVVRVLRVPIAAVGAIVLVACGV
ATKA_SYNY3/6-557      ---QACALGFLMFTSAATGIAVAIAFIRGLTQO.-----A.IQNFYQDLVLSITRILLPLISLVGAILLLVAGV
ATKA_DEIRA/6-561      ---QMAVITTFMFTSAATGFAAALAFMRGLAGRDE-----TNLGNFVVDVTRLIYRVLLPVCFGLALVFWQGM
ATKA_CLOAB/4-557     ---QMIVITTFMFFAAATGIAVALAFIRALSKK.-----KKLGNFYVDLVRITRILLPLSIIIVAFIYIQGV

```

## PFO3814

```

ATKA1_STAAM/1-557      PQTLSA.NLMIRTLSGHIQHIAFGPIASLESIKHLGTNCGGFLAAGNSATPFENFNINWSNFIEMGSMMLLPMMSMLFLGRM
ATKA_ANASL/6-557      PQTLEGG.SLTITTTLEGGTQYIARGPVASFEMIKMLGENGGFFAANSAPFFENFNATNFIEIAMIAMIAPAAIYTYGV
ATKA_HALSA/10-580     VQTILGGQLTANTFTMGIQINIRIGPHAGIEAIKMFCTNCGGINAANAATAFENPTFLSNLVLTLAMPITGFSAIYAWGAV
ATKA_RHILO/7-561      PQTLEP.YVDATTLGAKQTIALGPVASQVAIKMLGTNCGGFFNANAAPFFENPDASINLIQMVTFIPALGAALTNVFGRM
ATKA_ALIAC/7-561      PETLSG.PAVVHTLAGSMQTIARGPVATLEAIKQLGTNCGGFFNANSAPFFENFNATWTCILEIMGMLIPTALVFTAGHF
ATKA_MYCTU/7-570      IQNLHG.FIVANTLEGAPQLIPGGPVASQVAIKQLGTNCGGFFNVNSAPFFENYTFIENFVENWAILIIPFALCFAPGKM
ATKA_PSEAE/7-562      PQTFFD.VVHATTLQAEQTIPLGPAASQIAIKQLGTNCGGFFGVNSTHPPFNPSAWSNLFEVASIILIPALVFTFGHY
ATKA_THEAC/25-579     PETMORYITVNAFLTNKVYNIPLEPVATLEAIKNICTNCGGFFYCANAAYPFNPDWFTNLVEFVSFVIFIPASLISLQIV
ATKA_CAUCR/7-562      PQTLLA.GVTAHTLEGAEQKISLYAVASQEAIKMLGINGCGIFNANSAPFFENPTLNLITAIISINTLQWAAFFAGRT
ATKA_ECOLI/7-557      LQNFLP.YQAVNTVEGAQQLLEMGFPVASQEAIKMLGTNCGGFFNANSAPFFENPTALTNFVQMLAIFLIPALCFAPGEV
ATKA_STRCO/9-554      IQNFSG.IHEVQFMGGTQEWNGGAVASQEVIKELGTNCGGFFNANSAPFFENPTFTNLFEIFLILLIPVALTRTFGIM
ATKA_SYNY3/6-557     PETLAG.PAQVTTLEGATQWIARGPVVAHFEIKELGENCGGFFGINSAPFFENPNFNANLETVIMMVIIPALITTYGIM
ATKA_DEIRA/6-561     PQTLEA.YANATTLEGAQRIALGPVASLESIKHVGTCNCGGFFSMNAAPFFENPTLTLNHLISMLLLPSALTYVFGRM
ATKA_CLOAB/4-557     PQTLSA.NKTVTTIEKQLQNIPLGFPVASLEAIKQLGTNCGGFFSANSAPFFENPTLTVNSQIITLILLASMVVCFGIM

```

Fig. 7. Pfam seed alignment of ion channel clan.

PF03814

```

ATKA1_STAAM/1-557      LSRHGRKVRHRLILFVAMFFIFIAILLTLMWSEYRGN..PILANLGIY.....PNMEKKEVRFAGLSALFIVIT
ATKA_ANASL/6-557      AKNI....KQAWLLFMWVIVFVILVWVAATBELQGNP.LVNGTLGIEK.....PNLECKELRFQWAETALWAVMT
ATKA_HALSA/10-580     VGNR....SHGVAIVAFAFFVIYMALTVAVVGETGTNACMVVTNGLHVDQT...VGNMEKKEVRFQPTASAIWGLST
ATKA_RHILO/7-561      VGNQ....RQGWAILASMGALFIAAGVAVCYWAEAAAGN..PLVHALGLD.....CGNMEKKEVRFQIALSALFAVIT
ATKA_ALIAC/7-561      LRSR....RLAIVLCTLLGAILLAGAYIVYAYEAAGNP..ILAHALGIH.....GPNMEKKEVRFGLPLTSLFVAAT
ATKA_MYCTU/7-570      VHDR....RQGWAVLAIMGIIWIGMSVAAMSFEAKGN..PRLDALGVTTQQTVDQSGGNLEKKEVRFQVGSGLWAAST
ATKA_PSEAE/7-562      VKDL....RQSRAILGCMILLFCLGLLGLSLWAEYQFN..PALAQLPIEQ.....GAFMEKKEVRFQTAASMLWAVIT
ATKA_THEAC/25-579     FQDR....KFRMLYVWVWVFFIFDALFAFFEFAGV..PFLHLGYT.....GNMVKETAIGISQSTIFAVGA
ATKA_CAUCR/7-562      VLAK....KDVRLVIAAFVLLFACAVGVYATETQAP..PAQVAAQVD.....TSANMEKKEVRFQAPATAAWVAMT
ATKA_ECOLI/7-557      MCDR....ROGRMLLWAMSVIFVICVGVVWMAEVCGN..PHLLALGTD.....SINMEKKEVRFQVLSVSLFAVVT
ATKA_STRCO/9-554      TGSV....RQGYAILGTMAAIWAGFVALMMWTEFAHHC.PALQAA.....CGAMEKKEVRFQICGSSSLFAVTT
ATKA_SYNY3/6-557      AGNP....EQGWLIFWVVFILYILIAIAAVEFQGNP..LINQILGETQ.....PNLEKKEVRFQVWLTALWAVST
ATKA_DEIRA/6-561     LGNL....KQGVVIFCGMLVMFVFLQITVYSFEQAGN..PILSKLGAEQAITATQAGGNMEKKEVRFQIAQTSLFAVTT
ATKA_CLOAB/4-557     IKKK....KQAVAIFAAMVLLLAGAAICFSAEKAGN..PALSRIQLSQS.....MGNLEKKEVRFQIASLFLTTVT

```

PF03814

```

ATKA1_STAAM/1-557     TAFTTGVNMMHDSLTPICGLGPMVLMMLNVV..FGGECVGLMNLILFVLLTVE-----ICSLMV-----
ATKA_ANASL/6-557     TATMCGAVNGMHDALMPGCLFATLNLFLQII..WGGQGTGTAYLFYLLITVE-----LTGLMV-----
ATKA_HALSA/10-580     TGTITNGGVNSMHNSTALGAFSLLFATNNI..SNGVGTGLLNILMFVILTAF-----IGALMI-----
ATKA_RHILO/7-561     TAASCGAVNAMHDSFTALCGMTPLINMOLGEVIVCGVCGAGFYGILMFIVVAVF-----VAGLMV-----
ATKA_ALIAC/7-561     TAYTTGAVNAMHDSLMPLSGCVPLLEFMMFNLI..FGGKGVGLLNILMFLITAVF-----ISGLMV-----
ATKA_MYCTU/7-570     TGTNSGVNSMHSYTPICGMVPLAHMMLGEVSPGGTCVGLNGLLVMAILAVF-----IAGLMV-----
ATKA_PSEAE/7-562     TAASNGSVNAMHDSLSPLGGMPLLNMMMLGEVIFCGVCGAGLYGMLLFLVITAVF-----IAGLMV-----
ATKA_THEAC/25-579     TITSTGASNAALVSYTPAGIIGVLIQLLNDP..LGGVGTGVNLNIFMYIIFTVE-----IASLMV-----
ATKA_CAUCR/7-562     TCGASNGSVNGKHSSIMPLGGGIAMFLMQLGELPCCGICGSVAIMVVMALLSVE-----VAGLMV-----
ATKA_ECOLI/7-557     TAASCGAVIAMHDSFTALCGMVPMWMOIGEVVFCGVGSGLYGMLLFLVITAVF-----IAGLMI-----
ATKA_STRCO/9-554     TLTSTGAVDSFHSSYTLCCGITMGLMMLGEIAPCGVGSGLYGLMVMVAVAVF-----IAGLMV-----
ATKA_SYNY3/6-557     TGTMCQAVNGMHDLSLMPFCGGFVLSDFLQIIL..WGGQGTGTAYLFVLLITVE-----LTGLMV-----
ATKA_DEIRA/6-561     TAATTSVDSMHSYTPICGGVPLMQLMMLNVV..FGGKGVGFINFVQYLVLVGF-----IAGLMV-----
ATKA_CLOAB/4-557     TDTSCGAVNMHDSLTPICGAVPLINMMLNVI..FGGCVGVGMNIMYAILTVE-----LCGLMV-----

```

PF03814

```

ATKA1_STAAM/1-557     SITNPSFHGISQVMYEMTSAAAANNSSGFEGLKDDTTFWNISTCIIMLLSRYIPIILQLMIASSLVNKKSYHQDKYTIKID
ATKA_ANASL/6-557     GINTNPGFHGISQVVEYASASSNNSSGFEGLDCTLFFNLVNGVQVLLARYVPIIAMLILLADSMRKTQVPTPCTLKTD
ATKA_HALSA/10-580     SMNNPQFRGFSEVLYEFFSASANNSSGFEGLDCTLFFNLVNGVQVLLARYVPIIAMLILLADSMRKTQVPTPCTLKTD
ATKA_RHILO/7-561     SIANGCPHGFSEVLYAYTSAAAANNSSGFEGLDCTLFFNLVNGVQVLLARYVPIIAMLILLADSMRKTQVPTPCTLKTD
ATKA_ALIAC/7-561     SMNPGHGFTEVLYAYTSAAAANNSSGFEGLDCTLFFNLVNGVQVLLARYVPIIAMLILLADSMRKTQVPTPCTLKTD
ATKA_MYCTU/7-570     SRNNPQPHGLSEILLYAYTSANNSSGFEGLDCTLFFNLVNGVQVLLARYVPIIAMLILLADSMRKTQVPTPCTLKTD
ATKA_PSEAE/7-562     AVTNPGPHGFSEVLYAYTSANNSSGFEGLDCTLFFNLVNGVQVLLARYVPIIAMLILLADSMRKTQVPTPCTLKTD
ATKA_THEAC/25-579     SFVNPESRITELLYEFASAAAANNSSGFEGLDCTLFFNLVNGVQVLLARYVPIIAMLILLADSMRKTQVPTPCTLKTD
ATKA_CAUCR/7-562     GLMHSQPHGLSEILLYAYVSGTANNSSGFEGLDNTMFWNVSAGVVMFLGRYLSIILLVAVASSFAAKRAVPATQGTKTD
ATKA_ECOLI/7-557     AMLNPGPHGFSEVLYAVSSAANNSSGFEGLDNTMFWNVSAGVVMFLGRYLSIILLVAVASSFAAKRAVPATQGTKTD
ATKA_STRCO/9-554     SMTNSGAGHGFSEVLYAYTSANNSSGFEGLDNTMFWNVSAGVVMFLGRYLSIILLVAVASSFAAKRAVPATQGTKTD
ATKA_SYNY3/6-557     GISPNGFHGISQVAYEYASAAAANNSSGFEGLDNTMFWNVSAGVVMFLGRYLSIILLVAVASSFAAKRAVPATQGTKTD
ATKA_DEIRA/6-561     SLNPNPGHGFSEVLYAYTSANNSSGFEGLDNTMFWNVSAGVVMFLGRYLSIILLVAVASSFAAKRAVPATQGTKTD
ATKA_CLOAB/4-557     GISPNGFHGLTQVLYQFTSSAANNSSGFEGLDNTMFWNVSAGVVMFLGRYLSIILLVAVASSFAAKRAVPATQGTKTD

```

Fig. 7. Continued.

**PF03814**

```

ATKA1_STAAM/1-557 -----GKTPEYLNMPICAREMKC---IVLVFLHPILLIVFSALAFMIFGASE
ATKA_ANASL/6-557 -----GRTPEIFGRKIEKREIVL---ASLILLVHPIVVLIIPSAIALAYPYSLS
ATKA_HALSA/10-580 -----CRRPQYLGKKLEWQEMRY---VFVVILVLPILVLIPOAAAIVYQCAID
ATKA_RHILO/7-561 -----GRTPEYLGKKIEAKEVKM---AMLAILCLPLAMLIPTAIAVVLPSAVA
ATKA_ALIAC/7-561 -----GRTPEIFGKKIEAREMKL---ATAAMLVHPFVILVPTAIAAMALPSARA
ATKA_MYCTU/7-570 -----GRTPEYLGKKIQATEMKL---VTLYILAMPIALLSFAAASVLISSALA
ATKA_PSEAE/7-562 -----GRTPEYLGKKLEAKEVRL---LVFTLLVMPVGVVLVGAIAASLPCPAS
ATKA_THEAC/25-579 -----GKLFPEIMSLRISSEKIKY---STLSLITHPLLVVIPLGITLMIPIHLS
ATKA_CAUCR/7-562 -----GRTPEYLGKKVEAREIQF---SILAVVIIPLSMLGFSGIAAVLPEALK
ATKA_ECOLI/7-557 -----GRTPEYLGKKIDVREMKL---TALAIVTPTLVLMGAALAMMTDAGRS
ATKA_STRCO/9-554 -----GRTPEYLGKKITREIKF---AACYILITPALVIVFTAAMALETPGD
ATKA_SYNY3/6-557 -----GRTPEFLGRKIEKREIVL---ASLILLHPIALIIPTAITLAFPTTLA
ATKA_DEIRA/6-561 -----GRTPEFLGKKIEAKEVKL---TMLAVLAHPLSLGFTALACVLPALN
ATKA_CLOAB/4-557 -----GRTPEFLNKKIECKEIKL---VAFAILVHPFLLIMSSALALTTQGLA

```

**PF03814**

```

ATKA1_STAAM/1-557 KPYFGVSLIVFIVLLSGLTFIPVLLLGPIG...EFLTL
ATKA_ANASL/6-557 SLLFTTIVTAGIVLILGVLTFPPVLAGPIA...EAFKL
ATKA_HALSA/10-580 TPAFVGLLIGVIIIVSALVFLPALVFGPIG...ELLSG
ATKA_RHILO/7-561 GPLFVGLLIGVIVIVGGLTFFPALAVGPII...EHLAM
ATKA_ALIAC/7-561 TFAPGYVFAVFIIVGALTFPPYLAGPIG...EQLQL
ATKA_MYCTU/7-570 KPLFVGLVIGVVLIVGGLTFFPALALGPIV...EQLST
ATKA_PSEAE/7-562 GPLFVSLITVTILLVGGTLFPLALGPIA...EQLSL
ATKA_THEAC/25-579 NWIFVLLIAAMILIGLLSYFPIIVLGPLLSWAHDFNL
ATKA_CAUCR/7-562 GGLFVGLLIGVILLIGGLQFFPAMALGPIV...EHFQA
ATKA_ECOLI/7-557 GPLFVGLLIGTVLLVGCALTFIPALALGPVA...EYLS
ATKA_STRCO/9-554 KPLFAGLLACAVLIITGLTYFPALALGPIA...EGLAA
ATKA_SYNY3/6-557 TTLETCITACAIILGALTFPLPILVLPVA...EAFNL
ATKA_DEIRA/6-561 TALFGALTWFVILIVGALTFPLPALTLGPVA...DHLQM
ATKA_CLOAB/4-557 NTTFVTTLIVIIIVIGALTFPLPAVALGPS...EYLTTL

```

This alignment is coloured according to the ClustalX colouring scheme:

- Glycine (G)
- Proline (P)
- Small or hydrophobic (A,V,L,I,M,F,W)
- Hydroxyl or amine amino acids (S,T,N,Q)
- Charged amino-acids (D,E,R,K)
- Histidine or tyrosine (H,Y)

Fig. 7. Continued.

## 1.5. Summary

*Karenia brevis* produces toxins that negatively affect the shellfish industries and public health. This study took a genetic approach to determine the effects of environmental variations on toxin production and to elucidate the biosynthetic pathway of the toxins. Polyketide synthases look promising as candidates for the genes involved in the brevetoxin biosynthetic pathway; however, with the complications of the *K. brevis*



genome and the post-transcriptional regulation, a method to evaluate the function of individual PKS genes must be employed.

To aid in identifying the genes associated with the synthesis of brevetoxin, three putative PKS transcripts from (Monroe and Van Dolah, 2008) were identified from the *K. brevis* EST library. They were used in this study to determine the effectiveness of using qPCR as a method for observing a direct link between the expression of a PKS gene and the levels of one of the brevetoxin backbone. This study used the variation in the production of the toxins to make a correlation to the gene expression obtained from qPCR. Three clonal cultures of *K. brevis*, with salinity as the differential environmental condition, were used to elucidate the hypothesized gene clusters of multi-protein enzymatic complexes involved in brevetoxin production, one for each backbone.

To determine the *in vivo* function of the brevetoxins, which could be osmotically regulating ion channels, the EST library and the SL sequence were used to identify ion channels produced by *K. brevis*. There was evidence of ion channels in other micro-eukaryotic organisms with hypothesized roles such as osmoregulation and signaling. With what is known about the structure of brevetoxin and its function in mammalian neurons, along with toxin profile data displaying a significant increase in production of brevetoxin after a salinity shock, it seems highly plausible that there are ion channels that are affected by brevetoxin present in *K. brevis*.

This study made progress towards gleaning valuable information of the role of polyketide synthases in the synthesis of the many polyether metabolites of *Karenia brevis*. This study also makes the first step in characterizing ion channels from the *K.*

*brevis* EST library, which is necessary to determine of the role brevetoxin plays in the cell.

## 2. DIFFERENTIAL EXPRESSION OF PKS TRANSCRIPTS

### 2.1. Overview

Described in this section are the methods used to extract, reverse transcribe, and study the expression of polyketide synthase genes from *Karenia brevis* clonal cultures. The unique toxin profiles for each clone under different environmental conditions will be used to compare to the results of the relative amounts of gene expression obtained by qPCR. These cultures were acclimated to two different salinity levels that relate to oceanic or coastal waters and a culture that undergoes a rapid shift from high to low salinity. These clonal cultures experiencing identical conditions were used to determine a correlation of the toxin production to the polyketide synthase genes.

### 2.2. Materials and Methods

#### 2.2.1. Clonal Cultures

##### *2.2.1.1. Media and Cultures*

Three clones of *Karenia brevis* were used in this study: one Wilson that produces more total toxin at higher salinity and serves as a baseline; one SP3 that produces more PbTx-1 than the other clonal cultures; and one SP1 that produces less total toxin at all environmental conditions relative to the baseline clone (Table 1). All cultures were grown in L1-Si enriched natural seawater collected from the Gulf of Mexico (Guillard and Hargreaves, 1994). Seawater was filtered and sterilized, then enriched with sterile-filtered nutrients. Two salinities were used to represent offshore oceanic water and inshore coastal water (35 and 27). Salinities were measured using a Vista refractometer Model #A366ATC. For salinity experiments at 27, the salinity 35 water was diluted with

MilliQ water to reduce salinity. All experiments were conducted at 25°C under irradiance levels of 70  $\mu\text{Ein m}^{-2} \text{s}^{-1}$  and at 12:12 hour light:dark cycle using cool white bulbs. Two other *Karenia* species were used as negative controls, *Karenia papilionacea* and *Karenia mikimotoi*, and were grown at salinities corresponding to their natural environment, salinities 35 and 30, respectively. Both species are closely related to *K. brevis*, but do not synthesize brevetoxin.

**Table 1**

Clonal cultures of *Karenia brevis* and first collection location

Isolate Synonym	Collection Location date
Wilson CCFWC268	John's Pass, Florida 1953
SP3	South Padre Island, Texas 1999
SP1	South Padre Island, Texas 1999
<i>Karenia papilionacea</i>	Hawkes Bay, NZ, Cawthron Institute
<i>Karenia mikimotoi</i>	Corpus Christi Bay, Texas 2001

#### 2.2.1.2. Growth Rates

For each salinity treatment, cultures were grown in three replicate 50 ml glass tubes for approximately 7 to 18 days, depending on individual growth rates. Culture growth was monitored by daily in vivo fluorescence measurements recorded at the same time each day using a Turner Design TD-700 fluorometer Sunnyvale, CA (Campbell et al., 2004). Growth rates were calculated based on a Model 1 linear regression of three or more logarithmically transformed data points during logarithmic growth phase (Brand et al., 1981). For each treatment, cell concentrations were



determined using a Sedwick-rafter cell and an Olympus BX60 transmitted light microscope 100X.

#### *2.2.1.3. Constant Salinity and Salinity Change*

For the constant salinity experiments, clones were acclimated to culture conditions for at least 10 generations prior to experimentation. For the salinity change experiments, clones were acclimated to culture conditions for at least 10 generations prior to experimentation at the higher oceanic salinity, and then rapidly shifted to the lower coastal salinity over the course of one hour. The experiment to test for differences among the clones was performed by growing concurrent cultures, as previously described, at salinities 35 and 27. The salinity shift experiment was performed by growing the cultures at salinity 35. Aliquots of MilliQ water were added an hour after reaching late log phase to the cultures growing at 35 salinity until a salinity of 27 was reached. Each experiment was performed twice.

Two control experiments were performed to ensure the gene expression differences were not due to any variation in culture volume or cell number. Both controls used three Wilson cultures growing concurrently at 35 salinity, and in both experiments the three cultures were pelleted for RNA extraction over the course of three days, after reaching late logarithmic growth stage. However, in one control experiment, the equal culture volumes were harvested, whereas in the other, equal cell concentrations of different culture volumes were harvested.

## 2.2.2. Polymerase Chain Reaction

### *2.2.2.1. DNA and RNA Extraction*

DNA and RNA were extracted when cultures reached the late logarithmic growth stage. The culture volume used for each extraction was based on the required number of cells to produce sufficient material for RNA analysis and PCR 100 ng per RT reaction, resulting in an approximately equal number of cells in each culture extraction.

Extractions were performed in triplicate for each treatment. Cell cultures for RNA and DNA extraction were pelleted at 600 x g for 30 minutes in a centrifuge. Supernatant was aspirated and cell pellets were frozen at -80°C. Total RNA was extracted using the Qiagen RNEasy Kit; any remaining trace genomic DNA was removed using Qiagen RNase-free DNase kit Qiagen, Valencia, CA. DNA was extracted by thawing the frozen cell pellet and performing a CTAB extraction (Wang and Stegemann, 2010). Extracted RNA and DNA were stored at -20°C. Total DNA and RNA were quantified using a NanoDrop 1000 Spectrophotometer; the rRNA band pattern was used to determine the quality and integrity of the total RNA, by observing sharp, clear bands for 28S and 18S in a 2:1 ratio (Rodriguez and McDaniel, 2001).

### *2.2.2.2. Reverse Transcription*

Reverse transcription of the RNA into a cDNA pool was performed using DyNAmo SYBR Green 2-step qRT-PCR kit NEB, Ipswich, MA. Reverse transcription RT was carried out in 20 microliter ( $\mu$ L) reactions according to manufacturer's protocol. The RT reaction was performed using M-MuLV RNase H<sup>+</sup> reverse transcriptase, RT buffer, oligo d-T primers (5 nM final concentration), nuclease free

water, and 100 nanograms (ng) of total RNA. After the cDNA was generated, the 20  $\mu$ L reaction was further diluted in nuclease free water for a 1:10 dilution. The following protocol was used to perform reverse transcription reactions: 25°C for 10 min, 37°C for 30 min, 85°C for 5 min.

#### 2.2.2.3. PCR and Quantitative PCR

Using sequence information from the *Karenia brevis* EST libraries, gene specific primers were developed to produce amplicons sized between 100 and 200 base pairs (Table 2). Transcripts annotated as polyketide synthases were verified by repeating the BLASTx procedure (Lidie et al., 2005; Monroe and Van Dolah, 2008; Snyder et al., 2005). Primers were designed for each PKS sequence and the control gene, actin, using Primer3 software and synthesized by IDT Carolville, IA (Rozen and Skaletsky, 1998). Actin was chosen as the control due to its nature as a housekeeping gene; its constitutive and endogenous expression allows assumption of constant levels between the clones, enabling the relative expression of PKS genes to be normalized (Heil, 2009). PCR reactions were performed to verify primer efficiency and compare genomic DNA levels to cDNA transcript levels to rule out copy number effects and contamination of the total RNA with DNA. These reactions were performed in 10  $\mu$ L volumes with GoTaq Green Master Mix, nuclease-free water, gene-specific primers (250 nM final concentration) and 1  $\mu$ L of reverse transcribed template Promega, Madison, WI. The following PCR protocol was used to perform amplification reactions: 95°C for 5 min, then 30 cycles of 95°C for 35 sec, 57°C for 35 sec, 72°C for 35 sec, and a final extension of 72°C for 20 min. Sequences that were detected positive

in the control species were disregarded for future study, to focus on the stronger candidates for genes involved in brevetoxin synthesis.

**Table 2**

Primers designed from PKS genes for qPCR

Sequence Name	Primer Sequence 5'--3'
1008 F 1008 R	AAAACCGCTTGATGGAGATG GGTCATAGATTGCGCTGTGA
2006 F 2006 R	ACTCAGGCAAATCAGGATGG CCGTGGCAGATATTCGACTT
5299 F 5299 R	GCTGGATCTTGTTACAGCA TGATCCCAGGTCATCATCAA
actin F actin R	CATCACAGTTGGTGGTGAGC CATGACGGACTGGAATGTTG

Quantitative PCR was carried out using DyNAmo SYBR Green 2-step qRT-PCR kit NEB, Ipswich, MA. Protocol was optimized for the Bio-Rad CFX96 real-time PCR detection system Bio-Rad, Hercules, CA for 20  $\mu$ L reactions using the Master Mix, nuclease free water, 1  $\mu$ L diluted template, and gene specific primers (1  $\mu$ M final concentration). Cycling parameters used were according to the manufacturer's protocol: 95°C for 15 min, the 45 cycles of 94°C for 10 sec, 55°C for 25 sec, 72°C for 30 sec, a final extension of 72°C for 5 min and a melting curve of 72°C to 95°C. The melting curve was used to measure efficiency of primers to ensure correct quantification, to detect if a single amplicon was produced. Relative quantification of the transcripts was accomplished using a standard curve for the relative concentrations using two-fold

dilution series of RNA; this was used in each qPCR run to quantify the transcripts and the internal standard, actin. Linear regression analysis of the standard plot was carried out to determine efficiency of amplification; the actin primer pair was used as the reference to normalize the data (Rutledge and Cote, 2003).

#### *2.2.2.4. Sequencing*

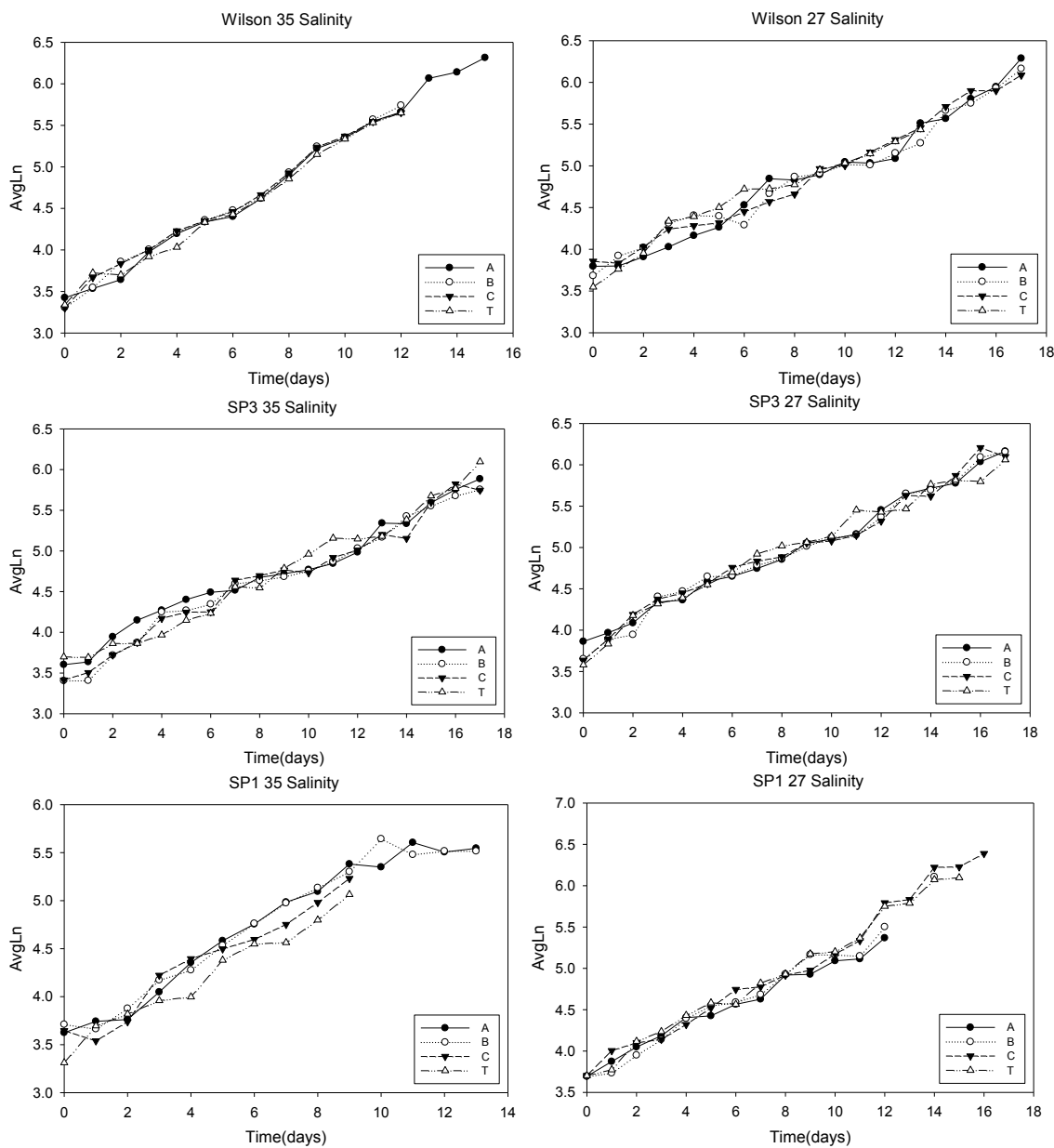
Primers were designed for sequencing of the three polyketide synthase genes from genomic DNA template, using Primer3 software and synthesized by IDT Carolville, IA (Rozen and Skaletsky, 1998). These reactions were performed in 10  $\mu$ L reactions with GoTaq Green Master Mix, nuclease-free water, gene-specific primers 250 nM final concentration, and 1  $\mu$ L of reverse transcribed template Promega, Madison, WI. The following PCR protocol was used to perform amplification reactions: 95°C for 5 min, then 30 cycles of 95°C for 35 sec, 57°C for 35 sec, 72°C for 35 sec, and a final extension of 72°C for 20 min. Sequencing was performed using ABI BigDye Terminator Sequencing kit on the ABI PRISM<sup>®</sup> DNA Sequencer, according to the manufacturer's protocol. Sequence analyses and comparisons were performed using the Sequencher 4.2.2 software.

### **2.3. Results**

#### 2.3.1. Growth rates

The growth rates of the three cultures at the two salinity levels were used to determine the approximate day to harvest cells in late logarithmic phase. For each experiment, there was a culture growing at salinity 35, one growing at salinity 27, and one growing at 35 that was shifted to 27 with MilliQ water. Growth curves were plotted

by days of growth against the average of the natural log of the fluorescence intensity (Figure 8).



**Fig. 8.** Growth curves of the three *Karenia brevis* clones. Growth at the two salinity levels 35 and 27.

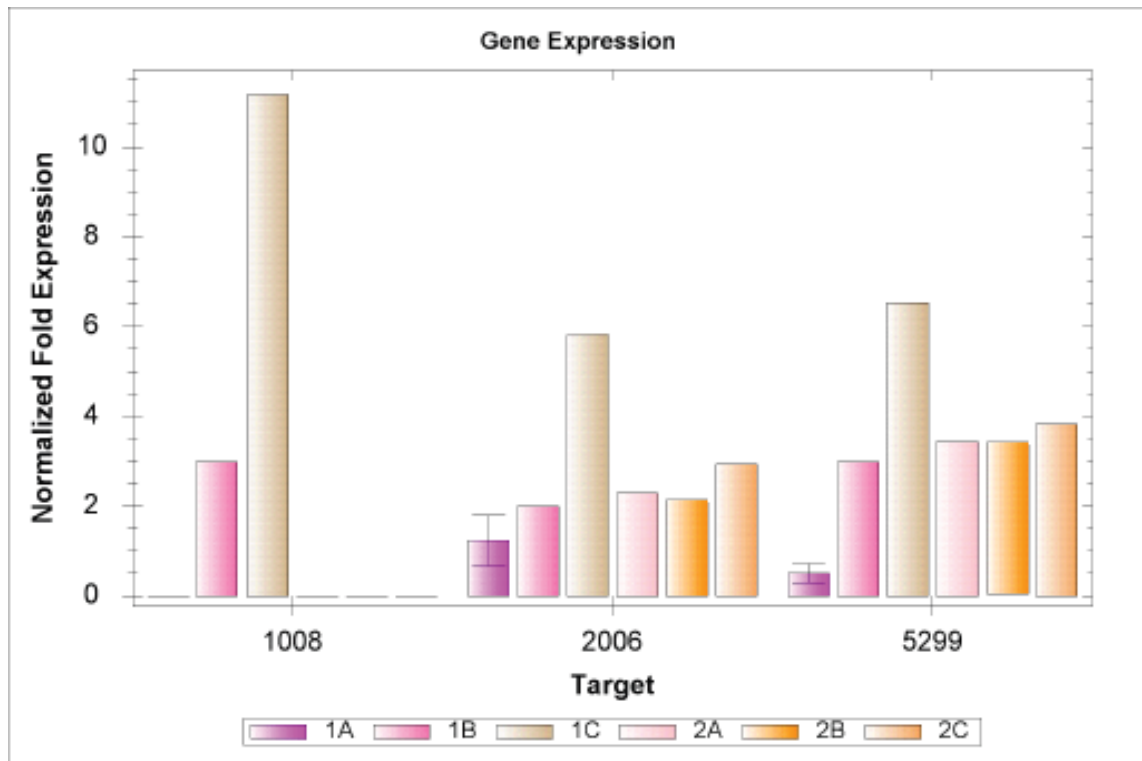
### 2.3.2. Sequencing

The three polyketide synthase genes were sequenced in each of the three clonal cultures. This was to ensure that any brevetoxin concentration differences were due to regulatory differences and not from sequence differences. There was no significant difference among the three PKS gene sequences in the three clonal cultures.

### 2.3.3. Quantitative PCR

The control experiment was performed to ensure the variation in PKS production resulted from RNA synthesis level differences and not as an artifact of RNA level differences from differential cell counts. These concurrent cultures of Wilson growing at salinity 35 were harvested on days 12, 13, and 14 of the growth curve. The three cultures of the same volume (1A, 1B, 1C) showed that, over the three days extractions occurred, an increase in the cell count caused an increase in the overall transcript abundance, as shown in Figure 9. These cultures contained approximately 3040 cells per ml (1C), 3520 cells per ml (1B), and 4360 cells per ml (1A). The three cultures with different volumes, but containing the same cell count, display the same amount of gene expression for the three polyketide synthase genes. These cultures contained approximately 3320 cells per ml (2A), 3840 cells per ml (2B), and 3600 cells per ml (2C); the culture volumes were adjusted to contain approximately 146,080 cells total for an average of 3571 cells per ml. This data shows that when cell counts are normalized, there is no statistical difference in the RNA levels. The change in normalized fold expression shows an approximately three fold increase from culture 1B to 1C, however there is only an approximately 1.5 fold

increase in cell number. This could indicate that a stress of high concentration of cells is affecting the gene expression, increasing the amount of RNA per cell.



**Fig. 9.** Two control experiments. Equal culture volumes harvested over three days (1), or equal cell concentrations of different culture volumes harvested over three days (2).

Standard curves were generated to relatively quantitate the cycle threshold (Ct) value, which is used to calculate the starting template amount, for each of the 9 samples and the 3 PKS genes (Figure 10). Seven of two-fold dilutions were made for the standards, and the entire curve was run in duplicate or triplicate. Four standard curves were obtained, one for each of the three PKS genes and one for the control gene, actin.

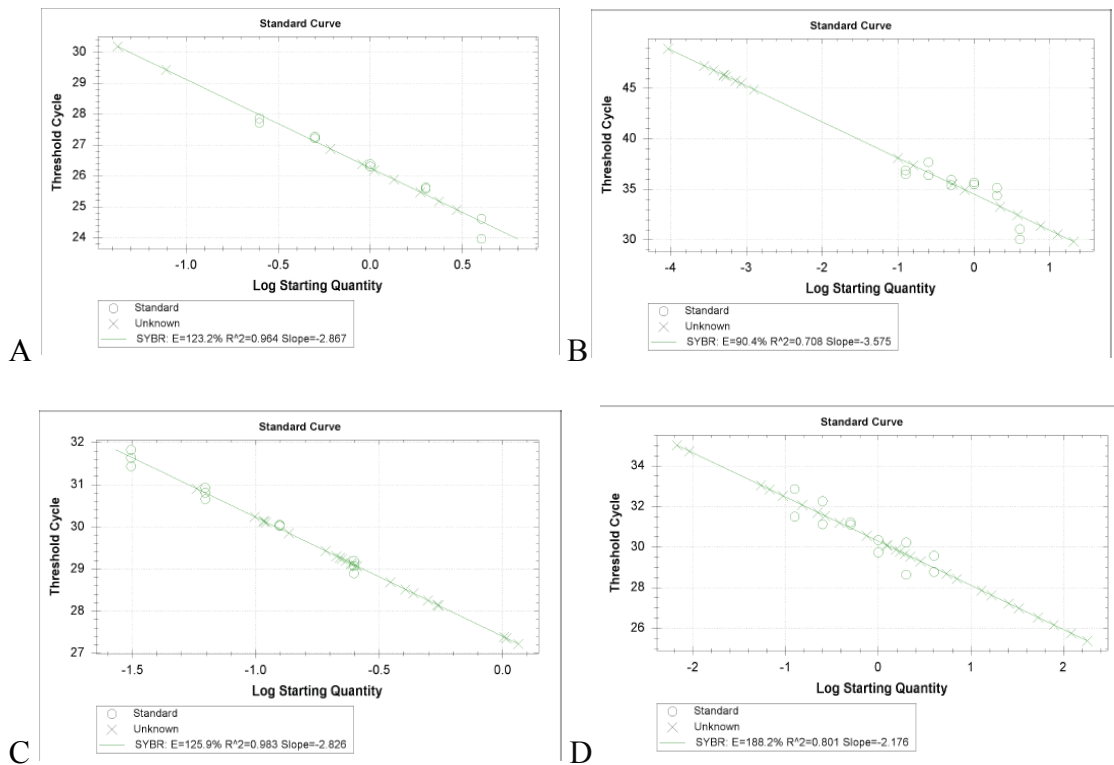


To determine the efficiency of the standard curves, the slope and the  $R^2$  value are analyzed. The results for the standard curves are acceptable (Table 3).

**Table 3**  
Slopes and  $R^2$  Values for Standard Curves

Primer	Slope	$R^2$
Actin	-2.867	.964
1008	-3.575	.708
2006	-2.826	.983
5299	-2.176	.801

To determine the gene expression of PKS genes, qPCR was run with the cDNA from clonal cultures Wilson, SP3, and SP1 at three environmental conditions: 35 salinity, 27 salinity, and a shift in salinity from 35 to 27. For all three PKS genes (1008, 2006, 5299) SP3 and Wilson showed a similar pattern of increased expression at salinity 35, whereas SP1 displayed no pattern among the three genes (Figure 11). However, the error bars make the patterns seen in SP3 and Wilson not significant.

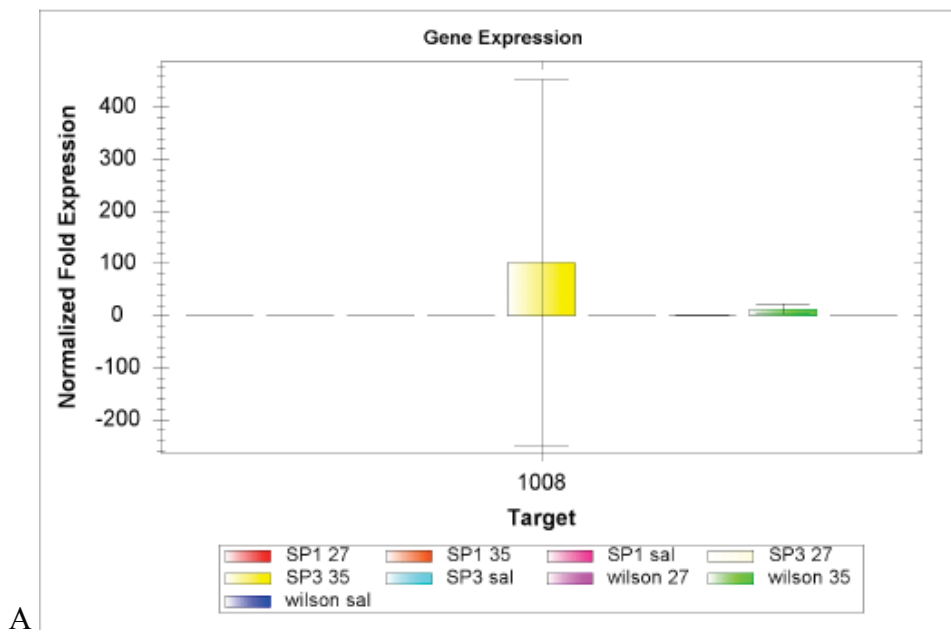


**Fig. 10.** Standard curves generated for qPCR. PKS primer sets (A) actin, (B) 1008, (C) 2006, and (D) 5299.

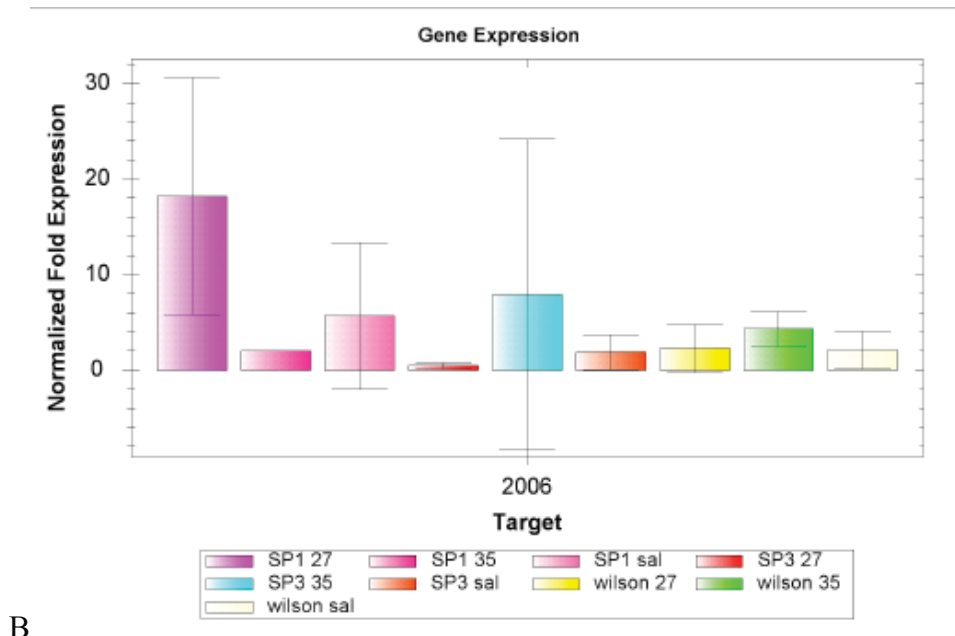
The expression of PKS gene 1008 shows an increase for clonal cultures SP3 and Wilson at the 35 salinity. The expression of PKS gene 2006 shows a trend in clonal cultures SP3 and Wilson of a maximum expression at 35 salinity, then at 27 salinity, and finally at the salinity change experiment. The expression of PKS 2006 for clonal culture SP1 shows maximum expression at 27 salinity, then for the salinity change, and finally for 35 salinity.

The expression of PKS gene 5299 expression also shows a trend in clonal cultures SP3 and Wilson of a maximum expression at 35 salinity, then at 27 salinity, and

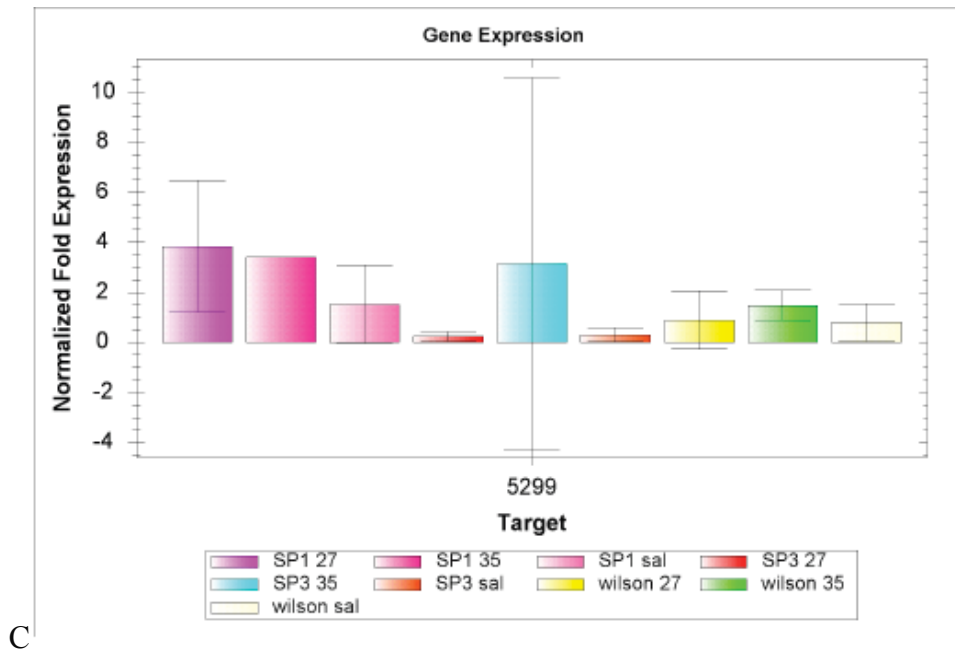
finally at the salinity change experiment. The expression of PKS 2006 for clonal culture SP1 shows maximum expression at 27 salinity, then for the salinity change, and finally for 35 salinity. The expression of clonal culture SP1 shows maximum expression at 27 salinity, then 35 salinity, and then the salinity change experiment.



**Fig. 11.** Relative expression of PKS genes in three clonal cultures. (A) 1008, (B) 2006, and (C) 5299.



B



C

Fig. 11. Continued.

### 3. ION CHANNEL IDENTIFICATION

#### 3.1. Overview

Described in this section is the genetic and bioinformatic analysis of ion channels obtained from the *Karenia brevis* EST library. The methods and programs for analyzing the nucleotide and amino acid sequences and annotating the sequences are explained. The seed alignments for ion channels were used to detect possible transcripts in *K. brevis*, and ion channel motifs were used to further characterize these sequences using proteomic software. This information will be used to better understand ion channels and the function of brevetoxin in the cell.

#### 3.2. Materials and Methods

##### 3.2.1. EST Analysis

The Expressed Sequence Tag EST library of *Karenia brevis* contains 65,266 entries, almost all of which have no annotation list of library names: *Karenia brevis* Multi-strain Library, Karbr Stressed, Karbr Dark-phase, *Karenia brevis* EST Library L99-05 (Lidie et al., 2005; McLean and Pirooznia, 2008; Richardson et al., 2007). This set of EST entries was downloaded from the public GenBank EST\_others database to an on-site computer (<http://www.ncbi.nlm.nih.gov.lib-ezproxy.tamu.edu:2048/>). These ESTs were prepared for analysis by an automated program to compare each sequence against a seed alignment set of ion channel sequences. A program was written first to translate each EST into the six open reading frames. Each open reading frame was

compared to three seed alignments: Ion Channel Library, 2TransMembrane Channel Library, and Inward Rectifying Channel Library Lockless, personal communication. Any match that was made to the EST was analyzed for a potential ion channel being expressed in *K. brevis*.

### 3.2.2. Nucleotide Sequence Analysis

With the EST sequences that were determined to be good matches for the ion channel seed sequences  $E$  value  $< e-05$ , further analysis was carried out. Each sequence was used in a BLASTn search against the 65,266 EST entries of *Karenia brevis* to determine the number of EST matches; singletons were identified as sequences with no other similar sequences in the EST database. Identical sequences were determined to be those with more than 95% identity  $E$  value  $< e-10$  over a 150 base pair region; Contigs were developed with those sequences that had high query coverage  $> 75\%$  and identity  $> 85\%$  and low  $E$  values  $< e-10$ . These contigs were used to assemble the EST sequences into transcript assemblies TAs using the Sequencher 4.2.2 software and were clustered into groups. TAs and singletons were then used in a BLASTx search protein database using translated nucleotide sequence using the GenBank non-redundant nucleotide database and the Alveolata (taxid:33630) database to evaluate their possible function as ion channels; they were also used in a tBLASTx search translated nucleotide database using translated nucleotide sequence using the Alveolata database. Those with highest sequence similarity to other ion channel sequences, including those of other related micro-eukaryotes Alveolates, were further studied. These sequences were analyzed and

annotated based on the closest matches in both the nr database and the Alveolata database.

### 3.2.3. Sequencing

TAs and singletons that were narrowed down to be further studied were used to design primers for a nested PCR of the 5' end, using Primer3 software and synthesized by IDT Carolville, IA (Rozen and Skaletsky, 1998). The forward primer that was used for each sequence was the spliced leader sequence; this is to ensure that the gene from the EST library is a nuclear encoded gene from *K. brevis* (Table 4). The nested PCR reaction was run for the detection of the spliced leader sequence at the 5' end. Two successive reactions were run: the first with the SL forward primer and the first gene specific reverse primer; the second with the SL forward primer and the internal gene specific reverse primer (Guillou et al., 2002; Hart et al., 2007). These reactions were performed in 10  $\mu$ L volumes with GoTaq Green Master Mix, nuclease-free water, gene-specific primers 250 nM final concentration, and 1  $\mu$ L of reverse transcribed template Promega, Madison, WI. The following PCR protocol was used to perform amplification reactions: 95°C for 5 min, then 30 cycles of 95°C for 35 sec, 57°C for 35 sec, 72°C for 35 sec, and a final extension of 72°C for 20 min. Sequencing was performed using ABI BigDye Terminator Sequencing kit on the ABI PRISM<sup>®</sup> DNA Sequencer, according to the manufacturer's protocol. Sequence analyses and comparisons were performed using the Sequencher 4.2.2 software.

**Table 4**  
Primers designed for ion channel sequence elongation

Sequence Name	Primer Sequence 5'--3'
<b>Spliced Leader</b>	TCCGTAGCCATTTTGGCTCAAG
<b>48705280 R1</b>	GAAGAGCTCGATTTGCGTTT
<b>48705280 R2</b>	CGAATTGAGGATGGAAGACC
<b>48706986 R1</b>	AACGTTCCAAGCATCAAACC
<b>48706986 R2</b>	CAGGGTCATCTCCCAGAGAA
<b>158891827 R1</b>	CACGCAGGAGCAACTCTACA
<b>158891827 R2</b>	TCTAAGCCCATGGTGACACA
<b>158893679 R1</b>	CAAGATACGCATCGCTTTGA
<b>158893679 R2</b>	AATCTTTCGTCGGCCCTTT
<b>158893749 R1</b>	ATGAACTTCGGCAGACTCGT
<b>158893749 R2</b>	AGATGGGAACACTCCGTCAA
<b>158893834 R1</b>	ATAATCCTGATTGCGGCTTG
<b>158893834 R2</b>	CAACGAAAACAGCAAATGGA
<b>158895975 R1</b>	AATCCATCCCGTGAAAACAA
<b>158895975 R2</b>	CCCATCATCTGTTGCTGTTG
<b>158897096 R1</b>	TTGGGAAGTCTTTGGTGAGG
<b>158897096 R2</b>	CGTCTCCCAAAGGACTTGAC
<b>158901885 R1</b>	ATGAGCTTGGAGAAGGACGA
<b>158901885 R2</b>	GCTGAGAACAATTGCCAGGT
<b>159007202 R1</b>	TCCATTGGGTTTTTGCTTTC
<b>159007202 R2</b>	CACTGGTGCGCATACTTCTT

#### 3.2.4. Animo Acid Sequence Analysis

With the sequenced nucleotides, the amino acid sequence could then be determined using the Translate Tool (<http://expasy.org/>). The translated amino acid sequences were then analyzed to determine the putative open reading frame. Using the likely open reading frame for each of the ion channels, the sequences were analyzed for channel motifs. Each sequence was analyzed for ion channel domains using Pfam (<http://pfam.sanger.ac.uk/>) and was compared to homologous sequences using Blastp against GenBank sequence databases (Mullan, 2004; Sammut et al., 2008). Multiple



sequence alignments were made with other Alveolata species sequences and the closest homology match from the nr database threshold ( $E$  value  $< e^{-4}$ ) to understand and develop a fuller picture of the structure and function of each ion channel using ClustalW (<http://www.ebi.ac.uk/Tools/msa/clustalw2/>).

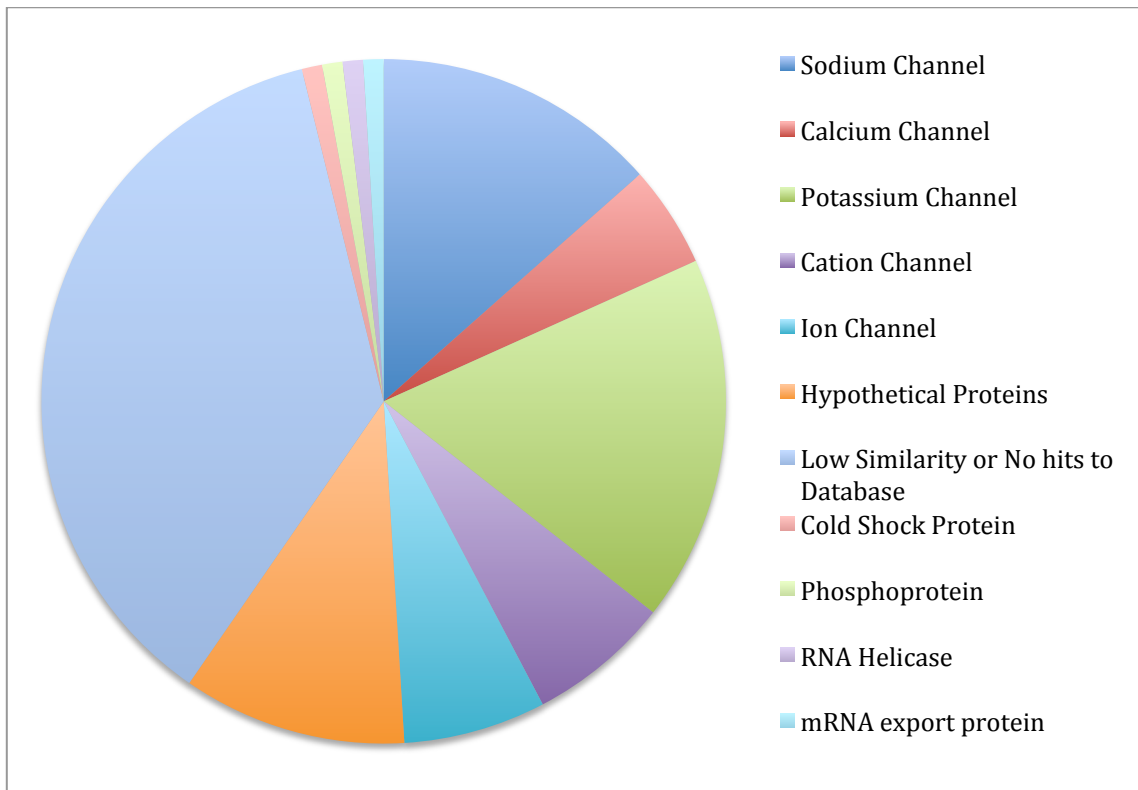
The amino acid sequences were also analyzed for common structural motifs of ion channels. Hydrophobicity plots were created to find the transmembrane domains and were compared to what would be expected for sodium channels; these were obtained using the Transmembrane Hidden Markov Model (<http://www.cbs.dtu.dk/services/TMHMM-2.0/>).

### **3.3. Results**

#### **3.3.1. In Silico EST Analysis**

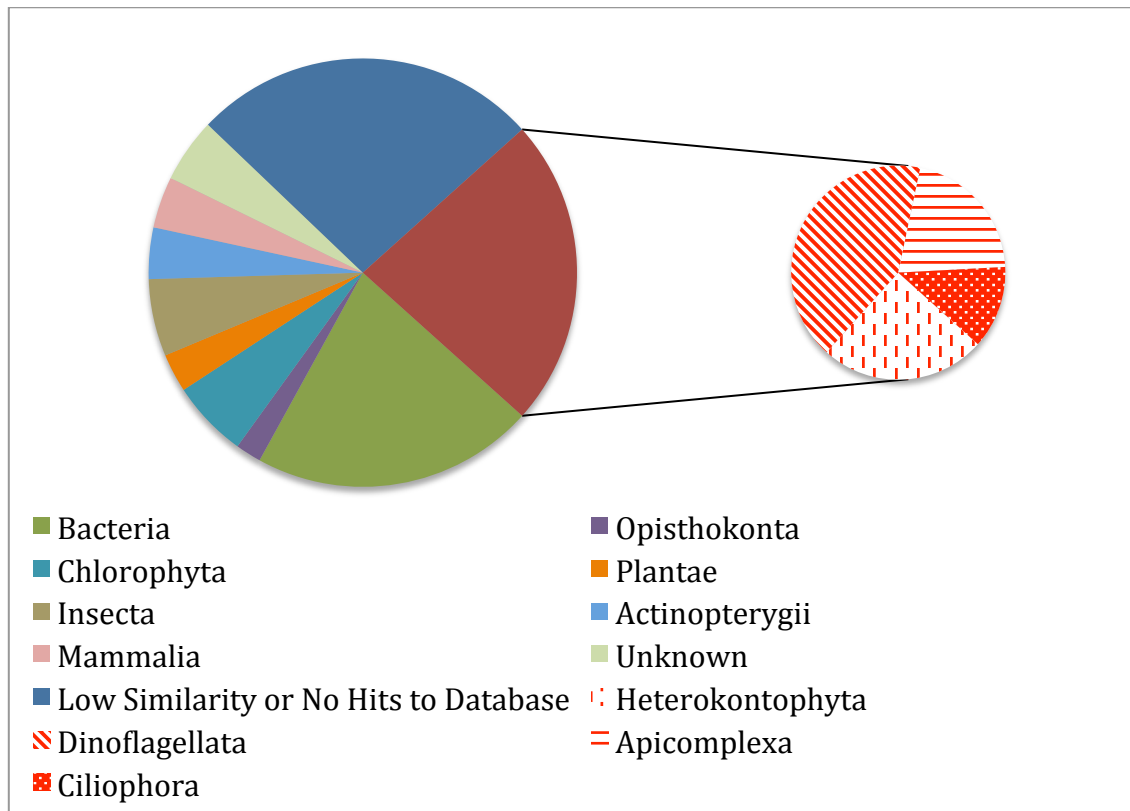
There were 500 oligonucleotide matches out of the 65266 entries from the comparison to the three seed alignment sets of ion channel sequences .766% of the EST library; out of those 500 sequences, 365 were discarded due to misleading hits on short common sequences and high  $E$  values  $> e^{-05}$ . With the 135 sequences remaining .206% of the EST library, further analysis was carried out. It was determined that there were duplicate sequences or sequences from the same gene in those 135 EST sequences, when showing more than 95% homology  $E$  value  $< e^{-150}$ . Focusing on the longest matches to ion channel sequences with the lowest  $E$  values threshold  $E$  value  $< e^{-15}$ , ten sequences were selected for further study. Six of these sequences (48705280, 48706986, 158891827, 158895975, 158893679, 158897096) were determined to contain the spliced leader sequence after sequencing the 5' end.

The remaining 129 EST sequences were also analyzed to create a fuller picture of what may be occurring inside a *Karenia brevis* cell. The EST set was assembled into 50 TAs and 54 singletons with a total set of 104 unique gene fragments using BLAST search and Sequencher 4.2.2 software; these were divided again into putative functional groups based on the BLASTX searches against the GenBank non-redundant database  $E$  value  $< e^{-4}$ . Of the ion channel matches to the EST sequences, there was evidence of voltage-gated channels, ligand gated channels, and cyclic nucleotide gated channels. There were also four matches to non-ion channel sequences; specifically, one cold shock protein, one phosphoprotein, one RNA helicase and one mRNA export protein. Of these four sequences, the most significant is the match to the RNA helicase, EST 159003625; this *K. brevis* sequence significantly matched four sequences with high query coverage  $>80\%$  and a low  $E$ -value  $< 4e^{-9}$  to DEAD/DEAH box helicases in *Perkinsus marinus*, *Neospora caninum*, and *Toxoplasma gondii*. The other three non ion channel sequences had low  $E$  values, but the query coverage were not significant enough to draw any conclusions. Cumulative database matches to different proteins of the sequences detected in this group of EST sequences are shown in Figure 12.



**Fig. 12.** Profile of EST sequences with protein similarity. Based on GenBank sequence database, figure labeled starting at the top moving clockwise.

The matches were also analyzed to help confirm that they were being expressed by *Karenia brevis* and not by any associated bacteria or other contaminant. Each sequence was used to identify its highest match, based on *E*-value, to a species; this gives confidence that the sequences with high similarity to other Chromalveolate species are being expressed by *Karenia brevis*. These species were put into groups and the cumulative database matches of the different proteins detected in this group of EST sequences to the highest matched species groups are shown in Figure 13.



**Fig. 13.** Profile of species with closest match to EST sequences. Based on BLAST search with subset pie of chromalveolata groups, figure labeled starting in the lower right pie section and moving clockwise.

### 3.3.2. Sequence Analysis

Within the six ion channel sequences, the voltage dependent motif (RxxRxxRxxR) was found in all of the amino acid sequences (48705280, 48706986, 158891827, 158895975, 158893679, 158897096). All six ion channels were matched with ion transport protein motifs in Pfam with significant *E* values (Figure 14). The multiple sequence alignments show that five of the six amino acid sequences have close homology to other Alveolata organisms; the sixth sequence (158897096) has close



channel family proteins of *Tetrahymena thermophila*, a Chromalveolate Ciliophora. It also contains the spliced leader and the voltage dependent motif. The hydrophobicity plot for transcript 48705280 shows six transmembrane domains, as expected for a sodium channel structural motif (see first figure on p. 74).

The alignment of EST 48706986 in Figure 16 shows the closest homology match to the ion transport protein of *Polymorphum gilvum* (326414624), an alphaproteobacterium, and two voltage gated cation channels of *Perkinsus marinus*, a Chromalveolate species. In the search against the Alveolata group, there were matches to conserved hypothetical proteins of dinoflagellates and voltage gated cation channels of *Perkinsus marinus*, a Chromalveolate species. It also contains the spliced leader and the voltage dependent motif. The hydrophobicity plot for transcript 48706986 shows four transmembrane domains, indicating that it either is not the full length sequence, or that it has a structural motif more similar to the four transmembrane domain, two pore channels (see second figure on p. 74).

The alignment of EST 158897096 in Figure 17 shows the closest homology match to a cation channel of *Bos taurus* (300794851), five cation channels of *T. thermophila* (118363637, 118364415, 118377893, 118377903, 118395994), and two voltage gated cation channels of *P. marinus* (294898822, 294955854). In the search against the Alveolata group, there were matches to conserved hypothetical proteins of dinoflagellates, cation channels of *T. thermophila* and voltage gated cation channels of *P. marinus*. It also contains the spliced leader and the voltage dependent motif. The hydrophobicity plot for transcript 158897096 shows six transmembrane domains,

indicating that it either is not the full length sequence, or that it has a structural motif more similar to the four transmembrane domain, two pore channels (see first figure on p. 75).

The alignment of EST 158895975 in Figure 18 shows the closest homology match to the voltage gated sodium channel of *S. sputigena* (260888220), a voltage gated sodium channel of *Hydra magnipapillata* (221126494), a Cnidarian, and a cation channel of *T. thermophila* (118377903). In the search against the Alveolata group, there was one match to a conserved hypothetical protein of a Chromalveolate, *P. marinus*. It also contains the spliced leader and the voltage dependent motif. The hydrophobicity plot for transcript 158895975 shows six transmembrane domains, as expected for a sodium channel structural motif (see second figure on p. 75).

The alignment of EST 158893679 in Figure 19 shows the closest homology match to an ion transport protein of a *Silicibacter* species (159418114), an alphaproteobacterium, and a conserved hypothetical protein of *P. marinus*. It also contains the spliced leader and the voltage dependent motif. The hydrophobicity plot for transcript 158893679 shows five transmembrane domains, which may indicate that it is not the full-length sequence (see first figure on p. 76).

The alignment of EST 158891827 in Figure 20 shows the closet homology match to a cation channel family protein of *Rhodospirillum centenum* (209966108), an alphaproteobacterium, and a conserved hypothetical protein of *P. marinus*. The EST sequence 158891827 also has closet homology to sodium channels. The search against the Alveolata group shows many hits to conserved hypothetical proteins of

dinoflagellates. It also contains the spliced leader and the voltage dependent motif. The hydrophobicity plot for transcript 158891827 shows six transmembrane domains as expected for a sodium channel structural motif (see second figure on p. 76).



## 4. DISCUSSION AND CONCLUSIONS

### 4.1. Polyketide Synthases

Due to negative effects of brevetoxin on the shellfish industry and public health, the determination of the biosynthetic pathway of the brevetoxins produced by *Karenia brevis* is necessary. Having more information about the genes involved in the synthesis of brevetoxins will play a major role in the effort to know more about the precursors to the toxins, more about the breakdown products, more about the interactions with other polyethers synthesized by *K. brevis*, and more about the environmental conditions that promote toxin production. This information will be used to understand why brevetoxins are being made and to develop models for predicting and monitoring algal blooms. This knowledge may also be helpful in understanding the fate of the toxins in shellfish, the possible enzymatic interactions in vivo, the gene regulation process in the cell, and the possibility of combinatorial biosynthesis, which is used in bacterial studies for novel antibiotics and other therapeutic agents (Rodriguez and McDaniel, 2001).

Polyketide synthases are the most promising candidates for the genes involved in producing the two brevetoxin backbones. PKS genes are currently being studied in *K. brevis* and other dinoflagellates that produce toxin, since they synthesize a physically and functionally diverse group of carbon skeleton natural products and use simple building blocks, creating a broad range of products by the variety of the modifications and the proteins involved at each elongation step. The type 1-like PKS genes, localized

to *K. brevis* by SL sequence and phylogenetic analysis, were used in this study to determine the effectiveness of using qPCR as a method for creating a direct link between the expression of a PKS gene and the levels of one of the brevetoxin backbone. This study used the variation in the toxin production to make a correlation to the gene expression of the PKS genes from three different clonal cultures and three environmental conditions.

The control experiments, harvested over three days, with either same culture volume or same cell count, show the correlation of expression levels of PKS genes resulting from RNA level differences in cells and not from differences in cell counts Figure 9. The sequencing of the coding sequences of the PKS genes 1008, 2006, and 5299 show no single nucleotide polymorphisms, suggesting that any differences in toxin production and composition may be from differences in the regulation of genes.

The toxin production differences among the three clones studied were used to compare to the expression levels obtained from the clonal cultures. In most clones, more total toxin is produced in cultures acclimated to 35 salinity than 27 salinity, but there is a fifteen fold increase after a rapid salinity shift (Errera & Campbell, in review). Typically, more PbTx-2 is produced ~90% (Errera et al., 2010); however SP3 is unusual in that it produces predominately PbTx-1 (Figures 2 &3).

We did not a significant correlation from the results obtained from the qPCR between toxin production and expression of PKS genes 1008, 2006, and 5299 (Figure 11). The overlapping error bars and lack of a consistent trend in the expression levels between the clonal cultures compared to the toxin production did not enable us to make a

prediction for a correlation between PKS 1008, 2006 and 5299 and the brevetoxin production. The lack of expression of PKS 1008 by any other clonal culture other than SP3 and Wilson at 35 salinity does not indicate that it is involved in the toxin synthesis, due to the presence of PbTx-1 and PbTx-2 in Wilson, SP1, and SP3 in all environmental conditions.

Since there was a fifteen-fold increase in total toxin production when a clonal culture underwent the salinity change from 35 to 27, it was expected that a high expression of a PKS gene involved in toxin production would be seen. The lack of induction of PKS transcript levels in the salinity change experiment for all three clonal cultures indicates that there may be a more involved regulation process in the biosynthesis of the toxin. There is more total toxin produced at salinity 35 than at 27; SP3 and Wilson show a trend of more PKS expression at 35 salinity. SP3 produces more PbTx-1 than the other clonal cultures, yet it followed a similar trend as the Wilson clonal culture. These data indicate that the biosynthesis of the brevetoxins may involve many genes to make up multi-enzymatic domains as well as other regulatory products. This will make determining the genes involved in the production extremely difficult.

While using qPCR worked as a method to determine expression in *K. brevis*, it did not provide concrete evidence whether a PKS gene was involved in the biosynthetic pathway. Other methods to aid in this process, in association with the qPCR, need to be developed. Knockouts of the PKS genes should be produced to conclusively determine its role in toxin synthesis. This can be performed by transfection of *K. brevis* cells with homologous sequences to the flanking regions of the PKS genes for recombination to

occur. Another method to create gene knockouts would be to use a RNA interference RNAi system by using small interfering RNA homologous to the PKS RNA for post-transcriptional gene silencing (Bouche and Bouchez, 2001). Another method that can be used to determine the role of PKS genes in toxin biosynthesis is using western blotting by developing antibodies to the candidate genes (Naar et al., 2002).

Once it can be determined which PKS genes are involved in biosynthetic pathway of the brevetoxins, the qPCR method will be useful for quickly characterizing clonal cultures to understand the differences between the toxin compositions of the clonal cultures. Samples of *K. brevis* algal blooms collected from the field can be used with the previously described experimental procedure to observe differential expression occurring within a given timeframe or between locations during a bloom. This will be useful for characterizing algal blooms and understanding more to develop models for predicting and monitoring algal blooms.

#### **4.2. Ion Channels**

The two brevetoxin backbones have an unknown in vivo function; considering it is a large costly carbon molecule, it must have an important function in *K. brevis*. There was a fifty-fold increase in brevetoxin production after a shock from high salinity to low salinity; this possibly indicates a function involving osmoregulation. Since brevetoxins affect voltage sensitive sodium channels in vertebrates, by binding sodium channel site 5 and depolarizing the membrane, it is possible that it has similar function of binding ion channels in the *K. brevis* cell. This study made an effort to identify ion channels in the

EST library and determine that they were real ion channel sequences and were nuclear encoded genes of *K. brevis*.

Two Chromalveolata species, *N. miliaris* and *O. sinensis*, have known ion channels that are hypothesized to have function in signaling, environmental sensing, osmoregulation, and contractile motion. Another Chromalveolata species, *Phaeodactylum tricornerutum*, showed evidence of osmoregulation. Gaining more information about the genes producing ion channels in *K. brevis* will play a role in the effort to know more about the interactions with brevetoxins and other polyethers synthesized in the cell. This information will be used to understand why brevetoxins are being made, what their functions are, how they are regulated, and to determine the possibility of using molecular techniques to disrupt the biosynthetic pathway of the brevetoxins.

It was determined that 49% of the 104 unique sequences were matches to ion, cation, sodium, calcium, potassium channels in the GenBank nr database Figure 12. There were four sequences that did not match any ion channels: cold shock protein 158898942 YP\_002129794.1, ACU45003.1, phosphoprotein 159018956 CBZ51262.1, RNA helicase 159003625 CBZ49647.1, mRNA export protein/het-R 158895135 XP\_002404672.1, CBZ50532.1. These matches to both the non redundant database and the Alveolata database, identified in the parentheses, may indicate that transcriptional regulation occurs.

It was determined that 62 of 104 sequences had significant matches to genes from the Alveolata group; this gives confidence that these ion channels and other

proteins are being expressed by *K. brevis*. Of the 104 unique sequences, 23% had a highest homology match to a Chromalveolate channel protein; of those 24 sequences, 42% had a highest homology to other dinoflagellates Figure 13. This give significant evidence that these sequences are ion channels functioning in the *K. brevis* cell. It was determined that of the 104 sequences, 16% were matches to organisms of Plantae, Insecta, Actinopterygii, or Mammalia; the matches to Plantae may be indicative of chloroplast localization of the ion channels. The other matches may indicate contamination; 20% of the sequences had highest homology matches to bacterial sequences, which may be from contamination or from the non-axenic culture. These matches may also be indicating the high conservation of ion channel domains such as voltage sensors or selectivity pores. Further analysis of the homology between the amino acid sequences and the structures of the folded ion channels should be conducted.

Ten sequences from the EST library were sequenced on the 5' end to determine the presence of the spliced leader to ensure they were nuclear encoded genes. There were six sequences that were further studied after the SL was found. Depending on procedure of the generation of the EST library, which were 3' ESTs, the SL could be present in the remaining 4 sequences of the 10 oligonucleotides selected to be sequenced, but were not detected in this study. There was also evidence of degenerative SL sequences, whose function is not completely understood.

EST sequences 48705280, 48706986, and 158897096 all had significant matches  $E$  value  $< e^{-11}$  to channel proteins of Alveolates; these should be further studied to understand their role in *K. brevis* ion channel formation. All of the six sequences had

significant homology to sodium channel proteins, except for 158893679, which had a higher  $E$  value than could be considered significant. Most of the six sequences had the structural motif of the six transmembrane domains of a sodium channel; the few that did not are either not completely sequenced or have more similar structure to a four transmembrane, two pore channel. While these sequences have sequence similarity to sodium channels in eukaryotes, they have structure more similar to either potassium channels, or the bacterial sodium channel. This study successfully showed the presence of ion channels in the *K. brevis* EST library, and that they are nuclear encoded genes with a spliced leader sequence. These results will help determine the next directions to take to determine not only how brevetoxin is synthesized, but also its function within the cell.

## REFERENCES

- Alverca, E., Franca, S., and de la Espina, S.M.D. 2006. Topology of splicing and snRNP biogenesis in dinoflagellate nuclei. *Biology of the Cell* 98, 709-720.
- Ambrosio, D.L., Lee, J.H., Panigrahi, A.K., Nguyen, T.N., Cicarelli, R.M.B., and Gunzl, A. 2009. Spliceosomal proteomics in *Trypanosoma brucei* reveal new RNA splicing factors. *Eukaryotic Cell* 8, 990-1000.
- Anderson, P.A.V., Roberts-Misterly, J., and Greenberg, R.M. 2005. The evolution of voltage-gated sodium channels: Were algal toxins involved? *Harmful Algae* 4, 95-107.
- Bachvaroff, T.R., and Place, A.R. 2008. From stop to start: Tandem gene arrangement, copy number and trans-splicing sites in the dinoflagellate *Amphidinium carterae*. *PLoS ONE* 3(8): e2929.
- Backer, L.C. 2009. Impacts of Florida red tides on coastal communities. *Harmful Algae* 8, 618-622.
- Baden, D.G., Bourdelais, A.J., Jacocks, H., Michelliza, S., and Naar, J. 2005. Natural and derivative PbTx: Historical background, multiplicity, and efforts. *Environmental Health Perspectives* 113, 621-625.
- Baden, D.G., Mende, T.J., Szmant, A.M., Trainer, V.L., Edwards, R.A., and Roszell, L.E. 1988. Brevetoxin binding - molecular pharmacology versus immunoassay. *Toxicon* 26, 97-103.



- Barbier, M., Leighfield, T.A., Soyer-Gobillard, M.O., and Van Dolah, F.M. 2003. Permanent expression of a cyclin B homologue in the cell cycle of the dinoflagellate *Karenia brevis*. *Journal of Eukaryotic Microbiology* 50, 123-131.
- Bouche, N., and Bouchez, D. 2001. Arabidopsis gene knockout: Phenotypes wanted. *Current Opinion in Plant Biology* 4, 111-117.
- Bourdelaïs, A.J., Campbell, S., Jacocks, H., Naar, J., Wright, J.L.C., Carsi, J., and Baden, D.G. 2004. Brevenal is a natural inhibitor of brevetoxin action in sodium channel receptor binding assays. *Cellular and Molecular Neurobiology* 24, 553-563.
- Bourdelaïs, A.J., Jacocks, H.M., Wright, J.L.C., Bigwarfe, P.M., and Baden, D.G. 2005. A new polyether ladder compound produced by the dinoflagellate *Karenia brevis*. *Journal of Natural Products* 68, 2-6.
- Bowler, C., Allen, A.E., Badger, J.H., Grimwood, J., Jabbari, K., Kuo, A., Maheswari, U., Martens, C., Maumus, F., Otilar, R.P., et al. 2008. The *Phaeodactylum* genome reveals the evolutionary history of diatom genomes. *Nature* 456, 239-244.
- Brand, L.E., Murphy, L.S., Guillard, R.R.L., and Lee, H.T. 1981. Genetic variability and differentiation in the temperature niche component of the diatom *Thalassiosira pseudonana*. *Marine Biology* 62, 103-110.
- Brown, A.F.M., Dortch, Q., Van Dolah, F.M., Leighfield, T.A., Morrison, W., Thessen, A.E., Steidinger, K., Richardson, B., Moncreiff, C.A., and Pennock, J.R. 2006.

Effect of salinity on the distribution, growth, and toxicity of *Karenia* spp.

Harmful Algae 5, 199-212.

Campbell, L., Loret, I.P., Villareal, T.A., Soltysiak, K., and Gold, J.R. 2004. Linking population and physiological diversity in *Karenia brevis* from the Texas coast. In: Steidinger, K.A. et al. (Eds.), Harmful Algae 2002. Proceedings of the Xth International Conference on Harmful Algae, Florida Fish and Wildlife Conservation Commission and Intergovernmental Oceanographic Commission of UNESCO.

Catterall, W. 1986. Molecular properties of voltage-sensitive sodium channels. Annual Review of Biochemistry 55, 953-985.

Catterall, W. 1996. Structure and function of voltage gated ion channels. Annual Review of Biochemistry 64, 493-531.

Chizhnikov, I., and Sorokina, Z. 1986. Non-temperature related effects of osmotic pressure on inward sodium current in the membrane of isolate rat spinal ganglia neurons. Neurophysiology 18, 380-385.

Errera, R.M., Bourdelais, A., Drennan, M.A., Dodd, E.B., Henrichs, D.W., and Campbell, L. 2010. Variation in brevetoxin and brevenal content among clonal cultures of *Karenia brevis* may influence bloom toxicity. Toxicon 55, 195-203.

Falciatore, A., d'Alcal`a, M.R., Croot, P., and Bowler, C. 2000. Perception of environmental signal by a marine diatom. Science 288, 2363-2366.

- Fuwa, H., Ebine, M., Bourdelais, A.J., Baden, D.G., and Sasaki, M. 2006. Total synthesis, structure revision, and absolute configuration of brevenal. *Journal of the American Chemical Society* 128, 16989-16999.
- Gontang, E.A., Gaudencio, S.P., Fenical, W., and Jensen, P.R. 2010. Sequence-based analysis of secondary-metabolite biosynthesis in marine actinobacteria. *Applied and Environmental Microbiology* 76, 2487-2499.
- Guillard, R.R.L., and Hargreaves, P.E. 1994. *Stichochrysis immobilis* is a diatom, Not a chrysophyte. *Phycologia* 33, 66-66.
- Guillou, L., Nezan, E., Cueff, V., Denn, E.E.L., Cambon-Bonavita, M.A., Gentien, P., and Barbier, G. 2002. Genetic diversity and molecular detection of three toxic dinoflagellate genera *Alexandrium*, *Dinophysis*, and *Karenia* from French coasts. *Protist* 153, 223-238.
- Hackett, J.D., Anderson, D.M., Erdner, D.L., and Bhattacharya, D. 2004. Dinoflagellates: A remarkable evolutionary experiment. *American Journal of Botany* 91, 1523-1534.
- Hart, M.C., Green, D.H., Bresnan, E., and Bolch, C.J. 2007. Large subunit ribosomal RNA gene variation and sequence heterogeneity of *Dinophysis* Dinophyceae species from Scottish coastal waters. *Harmful Algae* 6, 271-287.
- Harz, H., and Hegemann, P. 1991. Rhodopsin-regulated calcium currents in *Chlamydomonas*. *Nature* 351, 489-491.
- Heil, D.C. 2009. *Karenia brevis* monitoring, management, and mitigation for Florida shellfish harvesting areas. *Harmful Algae* 8, 608-610.

- Hertweck, C. 2009. The biosynthetic logic of polyketide diversity. *Angewandte Chemie-International Edition* 48, 4688-4716.
- Jones, A.C., Gerwick, L., Gonzalez, D., Dorrestein, P.C., and Gerwick, W.H. 2009a. Transcriptional analysis of the jamaicamide gene cluster from the marine cyanobacterium *Lyngbya majuscula* and identification of possible regulatory proteins. *Bmc Microbiology* 9:247.
- Jones, A.C., Gu, L.C., Sorrels, C.M., Sherman, D.H., and Gerwick, W.H. 2009b. New tricks from ancient algae: Natural products biosynthesis in marine cyanobacteria. *Current Opinion in Chemical Biology* 13, 216-223.
- Kamiyama, T., and Suzuki, T. 2009. Production of dinophysistoxin-1 and pectenotoxin-2 by a culture of *Dinophysis acuminata* Dinophyceae. *Harmful Algae* 8, 312-317.
- Khosla, C. 2009. Structures and mechanisms of polyketide synthases. *Journal of Organic Chemistry* 74, 6416-6420.
- Kirkpatrick, B., Fleming, L.E., Squicciarini, D., Backer, L.C., Clark, R., Abraham, W., Benson, J., Cheng, Y.S., Johnson, D., Pierce, R., et al. 2004. Literature review of Florida red tide: Implications for human health effects. *Harmful Algae* 3, 99-115.
- Krishtal, O., Osipchuk, Y., and Pidoplichko, V. 1983. The transmembrane gradient of osmotic pressure modifies the kinetics of sodium currents in perfuse neurons. *Cellular and Molecular Life Sciences* 39, 494-495.
- Kultz, D. 2001. Cellular osmoregulation: Beyond ion transport and cell volume. *Zoology* 104, 198-208.

- Landsberg, J.H. 2002. The effects of harmful algal blooms on aquatic organisms. *Reviews in Fisheries Science* 10, 113-390.
- Landsberg, J.H., Flewelling, L.J., and Naar, J. 2009. *Karenia brevis* red tides, brevetoxins in the food web, and impacts on natural resources: Decadal advancements. *Harmful Algae* 8, 598-607.
- Landsberg, J.H., Hall, S., Johannessen, J.N., White, K.D., Conrad, S.M., Abbott, J.P., Flewelling, L.J., Richardson, R.W., Dickey, R.W., Jester, E.L.E., et al. 2006. Saxitoxin puffer fish poisoning in the United States, with the first report of *Pyrodinium bahamense* as the putative toxin source. *Environmental Health Perspectives* 114, 1502-1507.
- Leflaive, J., and Ten-Hage, L. 2007. Algal and cyanobacterial secondary metabolites in freshwaters: A comparison of allelopathic compounds and toxins. *Freshwater Biology* 52, 199-214.
- Lekan, D.K., and Tomas, C.R. 2010. The brevetoxin and brevenal composition of three *Karenia brevis* clones at different salinities and nutrient conditions. *Harmful Algae* 9, 39-47.
- Lidie, K.B., Ryan, J.C., Barbier, M., and Van Dolah, F.M. 2005. Gene expression in Florida red tide dinoflagellate *Karenia brevis*: Analysis of an expressed sequence tag library and development of DNA microarray. *Marine Biotechnology* 7, 481-493.

- Lidie, K.B., and Van Dolah, F.M. 2007. Spliced leader RNA-Mediated trans-splicing in a dinoflagellate, *Karenia brevis*. *Journal of Eukaryotic Microbiology* 54, 427-435.
- Lustig, Y., Sheiner, L., Vagima, Y., Goldshmidt, H., Das, A., Bellofatto, V., and Michaeli, S. 2007. Spliced-leader RNA silencing: A novel stress-induced mechanism in *Trypanosoma brucei*. *Embo Reports* 8, 408-413.
- Macpherson, G.R., Burton, I.W., LeBlanc, P., Walter, J.A., and Wright, J.L.C. 2003. Studies of the biosynthesis of DTX-5a and DTX-5b by the dinoflagellate *Prorocentrum maculosum*: Regiospecificity of the putative Baeyer-Villigerase and insertion of a single amino acid in a polyketide chain. *Journal of Organic Chemistry* 68, 1659-1664.
- Magana, H.A., Contreras, C., and Villareal, T.A. 2003. A historical assessment of *Karenia brevis* in the western Gulf of Mexico. *Harmful Algae* 2, 163-171.
- Magana, H.A., and Villareal, T.A. 2006. The effect of environmental factors on the growth rate of *Karenia brevis* (Davis G. Hansen and Moestrup). *Harmful Algae* 5, 192-198.
- Marban, E., Yamagishi, T., and Tomaselli, G. 1998. Structure and function of voltage-gated sodium channels. *Journal of Physiology* 508, 647-657.
- McElhiney, J., and Lawton, L.A. 2005. Detection of the cyanobacterial hepatotoxins microcystins. *Toxicology and Applied Pharmacology* 203, 219-230.
- McKay, L., Kamykowski, D., Milligan, E., Schaeffer, B., and Sinclair, G. 2006. Comparison of swimming speed and photophysiological responses to different

external conditions among three *Karenia brevis* strains. *Harmful Algae* 5, 623-636.

McLean, T.I., and Pirooznia, M. 2008. Functional genomics and molecular analysis of a subtropical harmful algal bloom species, *Karenia brevis*. Unpublished.

Miller-Morey, J.S., and Van Dolah, F.M. 2004. Differential responses of stress proteins, antioxidant enzymes, and photosynthetic efficiency to physiological stresses in the Florida red tide dinoflagellate, *Karenia brevis*. *Comparative Biochemistry and Physiology C-Toxicology & Pharmacology* 138, 493-505.

Monroe, E.A., and Van Dolah, F.M. 2008. The toxic dinoflagellate *Karenia brevis* encodes novel type I-like polyketide synthases containing discrete catalytic domains. *Protist* 159, 471-482.

Mullan, L. 2004. Domains and motifs -- proteins in bite-sized chunks. *Briefings in Bioinformatics* 5, 71-74.

Naar, J., Bourdelais, A., Tomas, C., Kubanek, J., Whitney, P.L., Flewelling, L., Steidinger, K., Lancaster, J., and Baden, D.G. 2002. A competitive ELISA to detect brevetoxins from *Karenia brevis* formerly *Gymnodinium breve* in seawater, shellfish, and mammalian body fluid. *Environmental Health Perspectives* 110, 179-185.

Oami, K., Naitoh, Y., and Sibaoka, T. 1995a. Modification of voltage-sensitive inactivation of Na<sup>+</sup> current by external Ca<sup>2+</sup> in the marine dinoflagellate *Noctiluca miliaris*. *J Comp Physiol A-Sens Neural Behav Physiol* 176, 635-640.

- Oami, K., Naitoh, Y., and Sibaoka, T. 1995b. Voltage-gated ion conductances corresponding to regenerative positive and negative spikes in the dinoflagellate *Noctiluca miliaris*. *J Comp Physiol A-Sens Neural Behav Physiol* 176, 625-633.
- Pearson, L.A., Moffitt, M.C., Ginn, H.P., and Neilan, B.A. 2008. The molecular genetics and regulation of cyanobacterial peptide hepatotoxin biosynthesis. *Critical Reviews in Toxicology* 38, 847-856.
- Perez, R., Liu, L., Lopez, J., An, T.Y., and Rein, K.S. 2008. Diverse bacterial PKS sequences derived from okadaic acid-producing dinoflagellates. *Marine Drugs* 6, 164-179.
- Persson, A., Smith, B.C., Wikfors, G.H., and Alix, J.H. 2008. Dinoflagellate gamete formation and environmental cues: Observations, theory, and synthesis. *Harmful Algae* 7, 798-801.
- Pierce, R.H., Henry, M.S., Blum, P., and Payne, S. 2001. Gymnodinium breve toxins without cells: intra-cellular and extra-cellular toxins. In: Hallegraeff, G.M, Blackburn, S.I., Bolch, C.J., and R.J. Lewis, R.J. (Eds.), *Harmful Algal Blooms*, IOC of UNESCO, Paris, pp. 421-424.
- Plumley, F.G. 1997. Marine algal toxins: Biochemistry, genetics, and molecular biology. *Limnology and Oceanography* 42, 1252-1264.
- Richardson, P., Lucas, S., Rokhsar, D., Wang, M., Lindquist, E.A., Nosenko, T., Lidie, K.L., Bhattacharya, D., and Van Dolah, F.M. 2007. DOE joint genome institute *Karenia brevis* EST project. Unpublished.



- Rizzo, P.J., Jones, M., and Ray, S.M. 1982. Isolation and properties of isolated nuclei from the Florida red tide dinoflagellate *Gymnodinium breve* (Davis). *Journal of Protozoology* 29, 217-222.
- Rodriguez, E., and McDaniel, R. 2001. Combinatorial biosynthesis of antimicrobials and other natural products. *Current Opinion in Microbiology* 4, 526-534.
- Roy, S.W., and Irimia, M. 2009. Splicing in the eukaryotic ancestor: Form, function and dysfunction. *Trends in Ecology & Evolution* 24, 447-455.
- Rozen, S., and Skaletsky, H. 2000. Primer3 on the WWW for general users and for biologist programmers. *Methods in molecular biology* 132, 365-86.
- Rutledge, R.G., and Cote, C. 2003. Mathematics of quantitative kinetic PCR and the application of standard curves. *Nucleic Acids Research* 31, 93.
- Sammut, S., Finn, R., and Bateman, A. 2008. Pfam 10 years on: 10000 families and still growing. *Briefings in Bioinformatics* 9, 210-219.
- Satake, M., Bourdelais, A.J., Van Wagoner, R.M., Baden, D.G., and Wright, J.L.C. 2008. Brevisamide: An unprecedented monocyclic ether alkaloid from the dinoflagellate *Karenia brevis* that provides a potential model for ladder-frame initiation. *Organic Letters* 10, 3465-3468.
- Satake, M., Campbell, A., Van Wagoner, R.M., Bourdelais, A.J., Jacocks, H., Baden, D.G., and Wright, J.L.C. 2009. Brevisin: An aberrant polycyclic ether structure from the dinoflagellate *Karenia brevis* and its implications for polyether assembly. *Journal of Organic Chemistry* 74, 989-994.

- Schaeffer, B.A., Kamykowski, D., McKay, L., Sinclair, G., and Milligan, E.J. 2007. A comparison of photoresponse among ten different *Karenia brevis* (Dinophyceae) isolates. *Journal of Phycology* 43, 702-714.
- Schaeffer, B.A., Kamykowski, D., Sinclair, G., McKay, L., and Milligan, E.J. 2009. Diel vertical migration thresholds of *Karenia brevis* (Dinophyceae). *Harmful Algae* 8, 692-698.
- Schroeder, J., and Hedrich, R. 1989. Involvement of ion channels and active transport in osmoregulation and signaling of higher plant cells. *Trends in Biochemical Science* 14, 187-192.
- Shimizu, Y. 2003. Microalgal metabolites. *Current Opinion in Microbiology* 6, 236-243.
- Snyder, R.V., Gibbs, P.D.L., Palacios, A., Abiy, L., Dickey, R., Lopez, J.V., and Rein, K.S. 2003. Polyketide synthase genes from marine dinoflagellates. *Marine Biotechnology* 5, 1-12.
- Snyder, R.V., Guerrero, M.A., Sinigalliano, C.D., Winshell, J., Perez, R., Lopez, J.V., and Rein, K.S. 2005. Localization of polyketide synthase encoding genes to the toxic dinoflagellate *Karenia brevis*. *Phytochemistry* 66, 1767-1780.
- Steidinger, K.A. 2009. Historical perspective on *Karenia brevis* red tide research in the Gulf of Mexico. *Harmful Algae* 8, 549-561.
- Steidinger, K.A., Landsberg, J.H., Truby, E.W., and Roberts, B.S. 1998. First report of *Gymnodinium pulchellum* (Dinophyceae) in North America and associated fish kills in the Indian River, Florida. *Journal of Phycology* 34, 431-437.

- Tae, H., Kong, E.B., and Park, K. 2007. ASMPKS: an analysis system for modular polyketide synthases. *Bmc Bioinformatics* 8, 327.
- Taylor, A. 2009. A fast Na<sup>+</sup>/Ca<sup>2+</sup>-based action potential in a marine diatom. *PLos ONE* 4, E4966.
- Tester, P.A., Shea, D., Kibler, S.R., Varnam, S.M., Black, M.D., and Litaker, R.W. 2008. Relationships among water column toxins, cell abundance and chlorophyll concentrations during *Karenia brevis* blooms. *Continental Shelf Research* 28, 59-72.
- Tillett, D., Dittmann, E., Erhard, M., von Dohren, H., Borner, T., and Neilan, B.A. 2000. Structural organization of microcystin biosynthesis in *Microcystis aeruginosa* PCC7806: An integrated peptide-polyketide synthetase system. *Chemistry & Biology* 7, 753-764.
- Van Dolah, F.M., Lidie, K.B., Monroe, E.A., Bhattacharya, D., Campbell, L., Doucette, G.J., and Kamykowski, D. 2009. The Florida red tide dinoflagellate *Karenia brevis*: New insights into cellular and molecular processes underlying bloom dynamics. *Harmful Algae* 8, 562-572.
- Van Dolah, F.M., Lidie, K.B., Morey, J.S., Brunelle, S.A., Ryan, J.C., Monroe, E.A., and Haynes, B.L. 2007. Microarray analysis of diurnal- and circadian-regulated genes in the Florida red-tide dinoflagellate *Karenia brevis* (Dinophyceae). *Journal of Phycology* 43, 741-752.
- Vargo, G.A. 2009. A brief summary of the physiology and ecology of *Karenia brevis* Davis (G. Hansen and Moestrup comb. nov.) red tides on the West Florida Shelf

and of hypotheses posed for their initiation, growth, maintenance, and termination. *Harmful Algae* 8, 573-584.

Vargo, G.A., Heila, C.A., Fanning, K.A., Dixon, L.K., Neely, M.B., Lester, K., Ault, D., Murasko, S., Havens, J., Walsh, J., et al. 2008. Nutrient availability in support of *Karenia brevis* blooms on the central West Florida Shelf: What keeps *Karenia* blooming? *Continental Shelf Research* 28, 73-98.

Walsh, J.J., Jolliff, J.K., Darrow, B.P., Lenos, J.M., Milroy, S.P., Remsen, A., Dieterle, D.A., Carder, K.L., Chen, F.R., Vargo, G.A., et al. 2006. Red tides in the Gulf of Mexico: Where, when, and why? *Journal of Geophysical Research-Oceans* 111.

Wang, L., and Stegemann, J.P. 2010. Extraction of high quality RNA from polysaccharide matrices using cetyltrimethylammonium bromide. *Biomaterials* 31, 1612-1618.

Zhang, H., Campbell, D.A., Sturm, N.R., and Lin, S.J. 2009. Dinoflagellate spliced leader RNA genes display a variety of sequences and genomic arrangements. *Molecular Biology and Evolution* 26, 1757-1771.

Zhang, H., Hou, Y.B., Miranda, L., Campbell, D.A., Sturm, N.R., Gaasterland, T., and Lin, S.J. 2007. Spliced leader RNA trans-splicing in dinoflagellates. *Proceedings of the National Academy of Sciences of the United States of America* 104, 4618-4623.

## APPENDIX

```

gi|118377893| 1486 ACENKTDVSFNTFEEKPEHDYVSTRKKEILNLEKKGQNLKRVDSGDAVPSPTIAIK-
gi|118377903| 1463 FCNQQVEAFDFEQCCESAEYKYLQQRKKDIQNFEEKGICELKVQTSADDVLSISIKYK-
gi|118363637| 1522 NCDYFQKQPQTEQLKEPLEKEYVQVVKQQVNNLIQKQNKIKSYWSGDDVLAIKVYK-
gi|260888220| 1 -----MNVLEKALHWK-
gi|48705280| 1 -----CHVFVSWLRRRSPVPSVCAGVSKSFKKTTNPEVCTRVHP
consensus 1561 q f e eh yv kkei n kkg ik wsgd vl ti yk

gi|118377893| 1545 ECKQLNLIISKLLISKYDMEONEGGLIGIIQALRRRRIKYIVHSPIFEYSIMICVIINTII
gi|118377903| 1522 QCDALSKVLDTINCRQDIIEONEGGLIGIIQALRRRVRKGFVMHPVENLIMICVLLNIVV
gi|118363637| 1581 CCQAIKTLMKCLNYPNTIIEONEGGLIGIIVNAERRRVRMLVQTFLYENLIMAVVMNTVI
gi|260888220| 12 HTQKIIITFVIILN-----AAVLGVITNRTLSAEEVLFLEADKACLIVIFVE
gi|48705280| 42 CFETFGGIVVLFN-----FMVMTETHATATADDNVMRFVANRALLIYYTIE
consensus 1621 qq l ii lN eqneggligiialrrrik v pifenlimicviintvi

gi|118377893| 1605 LSLDGIIVDTQGLQTISQFNISFTFIFAIDMILKIIGDGFTEVSDKMNIFDCIIVCLSL
gi|118377903| 1582 LCMDFIIPPTSS-AYILDTFNLIFTILFAIDMGLKIIGMGLVMYASDKMNIFFAVIVLSLI
gi|118363637| 1641 LSLDGSVDSL-----FALYCADRMNIFDFGLISLI
gi|260888220| 59 LIAKLLVYRRS-----FWSGWNIEFVIVLS--
gi|48705280| 91 LARLYVFRRK-----SFAAGNIIIFTIVFTGW
consensus 1681 L ldgiv y sdkmNIFd iiv lsl

gi|118377893| 1665 FELFFIIEITGSGSSFSAFRAIRILRVFVLRVTRLVRSIQFMKLMITATGSSLSQFIYIL
gi|118377903| 1641 FELIFLNGT---SSVSALRSIRILRVFVLRVTRLLRSIQFMKILVTAISSLEQFIYIL
gi|118363637| 1673 FDIIFLGGSSGSSFSAFRSIRILRVFVLRVTRLLRSIKFMKILKVYSSILDKIFYIE
gi|260888220| 86 -SIVFISSS-----ISVIRAFRIRILKALAEFPELQILVSSMLKAPSMIAL
gi|48705280| 120 ISELEIFQVT-AGMDQIAMLSRIRILRVLRVRFVAFREYVVLVACTGRCCKTFLVAS
consensus 1741 felifit t ss sa r irilrvfvlrvtrllrsl fmkilvtai s l l yil
*****

gi|118377893| 1725 LLLFLFMFIYLLGMSLFGGSEFNMPDK--PTRMNFDSFLNSAFATILLLTIENWNLLQI
gi|118377903| 1698 LLLFLFMFIYALLGMSLFGGNMPVYNSNIPSRNNDQFINAFAVIDLLTIENWNDDLTN
gi|118363637| 1733 LLLFLYIFIYTLGMLYGGSFLSVSR--SNRMVDFSLQASFVYDLLTIENWNLLIYV
gi|260888220| 134 LLLFLYIFIAVFGSPMYGDAPPELFG-----DIGGSMFTLFOVMTFSSWATAVAR
gi|48705280| 179 GLLEFVNVNVAIISVEFLDEKVAIIP--SGAYRSCGCDIAFSNHWYANETWFQITST
consensus 1801 lLlflimfiyallgmslfgg f v rm fd flnsaFsil lltienWn ll

gi|118377893| 1783 CIRSNVSNYSVLELISWIFIGNYIFMNLVLAIVMDSFDIAEVQADRRREIENRFIEVVDI
gi|118377903| 1758 CLISNVSNYVCTAYLISMIFIGNYILLGLVLAIMLDSFNIAEVQNDRQEIENRFIEVVDL
gi|118363637| 1791 ALRQCVNVGTCYLYLISCFIGNYIFMNLVLAIVMIGAFNSSFQDFKKEEIANTVIEVDFL
gi|260888220| 185 FIMAMYPP--YAWLYFVFEHLLTAITLLNVVGIIVVEAVGISEAVKQKQAAEAAENAPPE
gi|48705280| 236 ADWTTGRPAEEFWFYAFVFNVFXDLDGTEPYCCSHXWTPFFXQGXCRRLCHCXQS
consensus 1861 i snv l lylliswifignylmnlvlaivmd f s dr ei nr evv di

gi|118377893| 1843 LFNYSSVG--TTIIAQTSSFSSGYSNS-----
gi|118377903| 1818 TFGHTSMG--TFMNNLSQSMGNQSSNN-----
gi|118363637| 1851 DLGHITLNSITFMNMSNSQSRLSSTHSFVKNRQRLMLSSGSSGQYSDEHSQNHSQH
gi|260888220| 243 ER-----
gi|48705280| 296 X-----
consensus 1921 d ss tt l s

```

Fig. 15. Multiple Sequence Alignment of EST 48705280.

```

gi|294955854| 1003 LQKQLYTKKIYLEFAHVKDLVKKRRKMYFFCTSSKKEETFIMTCILLNTAVTGMKMFPPS
gi|294898822| 1291 LQKQLYTKKIYLEFAHVKDLVKKRRKMYFFCTSSKKEETFIMTCILLNTAVTGMKMFPPS
gi|326414624| 1 -----MRDRLKELIASRRWEMTIVGVIVLNAVTLGETSPAVM
gi|48706986| 1 -----GPNTSGYPDRTIY-ERHGWNGTEEVW
consensus 1321 lqkqlytkkiylefahvkdldpvkrrkmyffctsskketfimtciillntavtgmkmfpps

gi|294955854| 1063 DAYNTTLATLNIFAFIFTAEALKLYATRWWYFDSWNCFDFVCVVATAIGILVDLL-S
gi|294898822| 1351 DAYNTTLATLNIFAFIFTAEALKLYATRWWYFDSWNCFDFVCVVATAIGILVDLL-S
gi|326414624| 39 GRIGGLTTLDTILAVFVELTLRIFARAFRDPWSLFDFTIVAIALE-----
gi|48706986| 27 KLRTDDGTVCGILLLIFLWEMTLRLYVLRMNSWEDAWNVDFIVVAVDAIMSVLGLGK
consensus 1381 daynttl tlnyl fiFtaEa LkiyatRwwyfyDsWncDFvcvV ataiGilvdlL s

gi|294955854| 1122 TVTIGTLMSATRIFRIARIFRLVRFAKGLNOLFIFVLSIPKLFNVALLLLLLLFLFTVL
gi|294898822| 1410 TVTIGTLMSATRIFRIARIFRLVRFAKGLNOLFIFVLSIPKLFNVALLLLLLLFLFTVL
gi|326414624| 90 --PASGPFAVDRALRIRLRLSVVPSLRRIGLIAALECLGSIVLALVFVFAVM
gi|48706986| 87 DIFSVSVLRVRLAKLNRSKVLRVFPELRLIACLSSVKALFWGTVLIFSLMWSMI
consensus 1441 tvtigtlms iRifriaRlfrlvr a gL ql ia vlsipklfnvaiLllllllflftvl
*****

gi|294955854| 1182 GVRLFG-----ATQLSGSHDDHANFRNFYRGMTLVR-CMTGEGWNEIMHSLAKDSEFFG
gi|294898822| 1470 GVRLFG-----ATQLSGSHDDHANFRNFYRGMTLVR-CMTGEGWNEIMHSLAKDSEFFG
gi|326414624| 148 ATKLYG-----ELFP-----DWEGDLGRSTYTLFQ-VMTLESNSMG-----
gi|48706986| 147 AVLFIHPLNKTVDYGESCRCPRASSVMOATITFWQIVTCDSNGEAS-----
consensus 1501 gvrlfg atqlsgshddhanfrnfyrg mTlv cmTge Wneimhslakdseffg

gi|294955854| 1236 LVLHRPCVNDFHITADNYAQLQEKCLIDHPVQCGSFFSFLFFISYTCLITFVVLNLVIAV
gi|294898822| 1524 LVLHRPCVNDFHITADNYAQLQEKCLIDHPVQCGSFFSFLFFISYTCLITFVVLNLVIAV
gi|326414624| 183 -----IVRPVMEVHVALLFFVPILCTAETVLNLFIGI
gi|48706986| 196 -----IPLIEKHEVALFFVMFLSIGLAVNLLGV
consensus 1561 lvlhrpcvndfhitadnyaqlqekclidhpvq g PfslFFfi yt iitfvVLNLVIAV

gi|294955854| 1296 ILEGFEDSTKSDEKNVVNKCIDVWKQDFPHDMELPIDSAVRPTEIVFMETVNSELMRDN
gi|294898822| 1584 ILEGFEDSTKSXXEYVCH-----
gi|326414624| 217 IVSMOEHEHAEADRQ-----
gi|48706986| 228 VVSVAQQAKDQEEDHR-----E
consensus 1621 il gf dstksiee r

gi|294955854| 1356 FVIRRKMMRFRERGGGTLYRSIPLKQAAFIRMQIDSKGKLMHDALAEMRLKIFGSPWSP
gi|294898822| 1602 --FRRKALVRHVR-----SSELSV
gi|326414624| 235 AHDETGILEEVR-----ALRREV
gi|48706986| 247 EVVDREEQHTLLG-----ECMBADKEGN
consensus 1681 vlrrkam hlvr 1 l asp sv

gi|294955854| 1416 PYWMKAGNHRMIGSPLSCLKKDKKYQVWADYFVRWIQAYEKKNIPIWGVTQNEPQFYF
gi|294898822| 1620 QEWMQD-----
gi|326414624| 255 EALRSSQA-----
gi|48706986| 271 GTIXHLHKM-----
consensus 1741 wmkad h m

```

Fig. 16. Multiple Sequence Alignment of EST 48706986.

```

gi|118377893| 1840 DDIDFNYSVCG--HTTFLAQSSQFSSSGYSNS-----
gi|118377903| 1815 QDLDFGHTSMG--TTFMNNLSQSMGNCSSNN-----
gi|118363637| 1848 DDLDFGHTTTLNSSTTFMNNNSQSRIGSSTHSFVSKNRQMLSSGGSSQGYSDHNSQNNH
gi|294898822| 1122 TYPFVCSLNFCASSRRLGSGEGHROQSAAAS-----
gi|294955854| 834 TYPFVCSLNFCASSRRLGSGEGHROQSAAAS-----
gi|158897096| 1 -----HAFRFPVMHEGQVLWETPG-----
gi|300794851| 24 DTFSTIEHLQCLSSQAVPRHIREILDSSRQKK-----
gi|118395994| 605 DNQNSAQPOLKHOKTLESESCFSKNVSEFALIN-----
consensus 1921 e t l g sttr l t l s

gi|118377893| 1869 -----LMNKNSNSFSNSHRKSSQDSQSMSSQTAKSGSQNQIED-----
gi|118377903| 1844 -----MSRISGIIQQSGGLQSSATLSRQVINKLSQSNDSGQLDKSKIIGGDDHSSSD
gi|118363637| 1908 SQHSFSSKKSQQTKQKESFSREYKEKASSISDSINQSSDNKSSHNLELQIVSEQEQNNDKNN
gi|294898822| 1154 -----ICGPEWLASMEKSIIMPCKDCSVHYCTGELQ-----
gi|294955854| 866 -----ICGSEWLASMEKSIIMPCKDCSVHYCTGELQ-----
gi|158897096| 20 -----TEPHDHKMKSNPKPAGGHSKAKSSP-----
gi|300794851| 56 -----LMLGDQHQLVRFISKPRIVERITHAQR-----
gi|118395994| 638 -----KSLGDKSELDROSSISIQQLKSESLMKSQIEG-----
consensus 1981 ig fl eksit h lt

gi|118377893| 1908 ----EDLEEVRKREN--INQVQFTYFAGVVKCQSSLYIFSQQNIIRKICYRALVHPLFEQV
gi|118377903| 1897 RVTSINSESKLKQDNK-KNKNKFIWEGVECEKSLFLFSKEGNFRKSIYYIYKH-----
gi|118363637| 1968 QKNMFKQQAVVKDEDEDEKEETLEFDNFKCKSLYIFSMQNKIRRAFYFVKNPLFDNI
gi|294898822| 1184 -----PSEKKTASCRDECEHHPYICADVACADNVFEGG-----
gi|294955854| 896 -----PSEKKTASCRDECEHHPYICADVACADNVFEGG-----
gi|158897096| 46 -----KTSQSETEANRTPCOR-----
gi|300794851| 83 -----LMSTLRVRCSE-----
gi|118395994| 671 -----PYMLSEAPDQAPAKSYTNLN-----
consensus 2041 k fiyf gv cqksly

gi|118377893| 1961 VLMVIVLSSIKLTMDDTYVNDQESQIYIVGLSIDKFFTAFFALESLIKIIISLGLVLDNSY
gi|118377903| 1950 -AIIILTSFKLTFDITYIGTADVATQFSNDADLFFTIFFLCECLIKVIALGFILDEGSY
gi|118363637| 2028 ILAGIILNSVKLVSDTYDTGSDS-QKQASQILDIMFTVFFSIESCLKIVSYGFFFDENSY
gi|294898822| 1216 -----ECEAC
gi|294955854| 928 -----ECEAC
gi|158897096| 62 -----
gi|300794851| 94 -----
gi|118395994| 692 -----S
consensus 2101 d

gi|118377893| 2021 LRDSWSQLDFFIVICSTIDVSVTSNLSAIKILRLLRRLRPLRLLNHNKSMKLIIVTALME
gi|118377903| 2009 LRDNWQCLDFFIVVSSVLDLSSNLNAIKILRLLRRLRPLRLLSQQNKSMLIIVTALME
gi|118363637| 2087 LRDNWQKLDFFIVVVTSTIDVSLQNGNFASLKILRLLRRLRPLRVLHNNKSMKLIIVTALLE
gi|294898822| 1221 LEACTSACCPQFCKAVIKDAANCVE-----
gi|294955854| 933 LEACTSACCPQFCKAVIKDAANCVE-----
gi|158897096| 62 LAKRVIRNPFESVMAMV-----
gi|300794851| 94 RPPLSLWAGVLERPIFFN-----
gi|118395994| 693 ITQNDISEIPQENEDINKITN-----
consensus 2161 lr s dffivi iikvs i

gi|118377893| 2081 SVGGITNMIMVLLIWLMFAILALNLMKGMHYCNGP-DDLS--LNLYAYNKQ-----
gi|118377903| 2069 SIGGIANMIIVLLIWLMFAILGLNLMQGLFYCNNTTGAALSGSYNVYNYPKKQILKYIK

```

Fig. 17. Multiple Sequence Alignment of EST 158897096.

```

gi|118363637| 2147 SIVGILNMLLVIFLVWLMFAILGVSLMGKLYQCNFPNNGIS--YNIYKYNQSQ-----
gi|294898822| 1247 -----
gi|294955854| 959 -----
gi|158897096| 80 -----
gi|300794851| 113 -----
gi|118395994| 715 -----
consensus 2221

gi|118377893| 2131 -----KCLNIPGATWEIYDINCNDNIGSSMLFLFLLSFEGWPLVYVMSFIDCGGSDG
gi|118377903| 2129 NGFSFLTHQIKCEQLPGAFNDRYPNNSDNIHFSMVMLFVLSFEGWPNYVWYYLDGSDTG
gi|118363637| 2199 -----CSRVGGSWQTYDLNFDNVPALNLFVISTFEGWPAYVENYLDGSSDG
gi|294898822| 1247 -----QGGQNLNLHQHFDNIAVSMMLTFEISTEGWVDVMYAGTDATEVY
gi|294955854| 959 -----QGGQNLNLHQHFDNIAVSMMLTFEISTEGWVDVMYAGTDATEVY
gi|158897096| 80 -----IILNFVMLIVEFDSTAERANGENPAP
gi|300794851| 113 -----FIIIFLIFLNIVVLMVETELLESANQLE
gi|118395994| 715 -----EQVYWMKKRIKPLEYLRKIQKNKKKEKPDQ
consensus 2281 ga w dnig sml lfiist egw dvvw dgsd

gi|118377893| 2181 --PILNNSAYFAEFFAVFIEVGCFFSINLFSALMSFNFDIAQKKAKNKVLTDECSQWIEL
gi|118377903| 2189 --PVWNSRYEGVFFVWFIEVGSFFSINLFTALISFNFDIATKKAKNAELTDDCSQWIEL
gi|118363637| 2247 --PVYFGSAYFSIYIVFFLIVGSFFSITLFGAVLSFHMIDIALQKIET-----
gi|294898822| 1292 GQRRDHAEELWGLFFVWFMLIGSLFILNLQVGVIVDNFNKIKAQGRSLTLTQTQKWIEL
gi|294955854| 1004 GQRRDHAEELWGLFFVWFMLIGSLFILNLQVGVIVDNFNKIKAQGRSLTLTQTQKWIEL
gi|158897096| 107 -----EWEVGCWLVLALEFVVELL
gi|300794851| 141 P-----LKIIEVVAWFILLFLEILMLLS-----
gi|118395994| 745 K-----KENPDLNELYIMSEDEETIKAYQKQVN-----
consensus 2341 p favff vfi vgsffilnlfvaiiv nf ak ylt q wiel

gi|118377893| 2239 QRMIVKSTP--DFNALKKPTNKIRFMLWSFVRSYVEIFLIFCIICNIVIMAMAVDTSPQ
gi|118377903| 2247 QRLIVKSTP--DYASMKPPDNNFRKLLMRISESPKLDIFLIFCIIGNVVVMAMAVDTSSS
gi|118363637| 2292 -RLQLYLSF--NKQIQK-----FRNIVSMALSVDTSPTS
gi|294898822| 1352 QKQLYTKKIYLEFAHVKDLFVKRRKMYEFCTSSKFETFIIMTCILINTAVTGMKMPSPSD
gi|294955854| 1064 QKQLYTKKIYLEFAHVKDLFVKRRKMYEFCTSSKFETFIIMTCILINTAVTGMKMPSPSD
gi|158897096| 127 -RYYVERVE-----
gi|300794851| 168 -SFFLFWKN-----
gi|118395994| 774 -DYIQDEEQ-----
consensus 2401 qrmiv p df lk r m w s i il vt y

gi|118377893| 2297 QYDQILNDLNIFFTSVFTVETAIKKIIFAGLQGYFNGWNOFDFVVCASIVDIIVTTSMGR
gi|118377903| 2305 NYDFILQSVNLGFTGVFIABC-----WNOFDFVVCASVVDIILSYSGK
gi|118363637| 2322 SYDTILQDINLGFSSIFITECLKLIAYGFNGYFYKSSNOFDFVVSIVSDIIFTYSGN
gi|294898822| 1412 AYNTTLATINNYIFAFITFABAAKKYATR-WWYFYDSWNCDFVQVVAATGILVDLEST
gi|294955854| 1124 AYNTTLATINNYIFAFITFABAAKKYATR-WWYFYDSWNCDFVQVVAATGILVDLEST
gi|158897096| 135 -----EFADGWNKEDFFLISIDTVCSLIQLLFG
gi|300794851| 176 -----AWNVEDFVIVVISLPEIIVVLAGV
gi|118395994| 782 -----DSNFKSNSVTKEDANQSNFQYM
consensus 2461 y l ln f vf e iki a yff wnfGdfvv asii ivm lmg

gi|118377893| 2357 QFISFLRAGPQIARVLRVLRVSRLLKLIKSFQNLKLIQTTIFAPELNLNATALLFLYF
gi|118377903| 2349 SFFSFLKAGPQIARVLRVLRVSRLLKLIKSFQGLKLIQTTVFAPEFLNATALLFLYF
gi|118363637| 2382 QFIFKFLSAGPQIARVLRVLRVSRLLKLIKQFEGLOKLIQTTIYALPSLNNATALLFLYF
gi|294898822| 1471 VTIGTLMSE-----AIRIFRIARLFLVRFARGLNCLFIAFVLSLPELNVALLLLLFL

```

Fig. 17. Continued.



```

gi|294955854| 1183 VTIGTILMS-----AQRDFRDAKLEFLVFAKGLNQLFIKFLSIFKLEFVALLLLLLLE
gi|158897096| 163 SAFFVSTL-----RVLRLCKLARMSKIFVVFPELRLLMAGLGGMRSHFWQMLLITFSL
gi|300794851| 200 TSKPVWLQ-----LRLICRVLSKLFARFHQVRVILDLVRAIKSIFLLMLLLEFFY
gi|118395994| 806 MKKALSDS-----NMDMKKLEIDYPOETYGSYFNAPHEDSFGNWCIFKEGNRNLYP
consensus 2521 i l rvlrv rv rllklik f glq li a v lphl na lllliiyf
*****

gi|118377893| 2417 IFALGCVLFSVTEGN--VIDQFNNEKDEHSFILLRRCSTGEDWMTVAFDVKG---P
gi|118377903| 2409 IFSILAWLPSDLQSGW--RLSTFWNFTNEHRMILLVRRSTGEDWLVMMYDTFNA---K
gi|118363637| 2442 IFALACMLYS-----PYQSTGEDWLLIMEDLIHH---T
gi|294898822| 1525 LFTVLGVRLFGATQLSG--SHDDHANFRNEFRFMTLVRCMTGEGWNEIMHSLAKD---S
gi|294955854| 1237 LFTVLGVRLFGATQLSG--SHDDHANFRNEFRFMTLVRCMTGEGWNEIMHSLAKD---S
gi|158897096| 218 IMAVLGVQFIHP-----LNDKMTKQVYIRECCERC---P
gi|300794851| 254 IFAVAGVVFENYTRSTRQDLYHEFESDLLNSIITVPILELHMMALQDTWKVPEV
gi|118395994| 859 CLLGQCPVTEYKVFKK-----DDIYISREFEQINQOVQLQLIH
consensus 2581 ifavlgvylf d n f f a i lvrmtgedwy im di k t

gi|118377893| 2472 SNYSIIYFVVFIIIQYMMNLFVLLIINQFENYDPDNPIQNFTESEDNFKOLIGIMT
gi|118377903| 2464 GVIGLIFVVSFYTVOSYILMNLFVLLIMNEFENYINPDNPLSNFNQEEQFKRQVQMT
gi|118363637| 2474 NMMSAFVVFVFIIQYIMNSLFIILLIQFENNYINPDNPLNDFQYSEEDFKQVQFT
gi|294898822| 1580 EFGVLVLRPCVN-DFHITADNVAQLQEKCLLHPVOCGSEFSEFFISYTCIITEVVEN
gi|294955854| 1292 EFGVLVLRPCVN-DFHITADNVAQLQEKCLLHPVOCGSEFSEFFISYTCIITEVVEN
gi|158897096| 250 EAYDSVSHALTLTPQSMAGDSVGH-----
gi|300794851| 314 RTFSSIVLWLLGSIIFRNIIAMMVTFQNIARNELNEEMTHEVQHKADIFKROIQ
gi|118395994| 899 KSYEQMRICAKCLDNDHPPNINPQKSIFDEPTICSNKNEKDVSLTSQSNKRLPDP
consensus 2641 y iiy vvfiiiq i mnlfv iii f e in npls l k wvim

```

Fig. 17. Continued.

```

gi|260888220| 1 -----MNVLRRAHMKTKQKTFVITDIAALGVITNRISAEVFL--
gi|118377903| 121 NFDRIALFAVTIFEKLYKQVWYIKKESYICFCIINHNLETVISLGVLENSWMLARD
gi|158895975| 1 -----RGMGCILCTMDATATDDGDVLRVVAIAN-----
gi|221126494| 121 VIVRQFSRVIIYWFNKEKTLWITDGNPARKWAIYIITNKLEDFDFMIIITVILINCIFLAMN
consensus 121 v q l k lwwih kiii vciina vfdvltlstvian vlla

gi|260888220| 45 -----EAVDRACLVIITVELLAKLVY----KRSFVSEGWNIIDFVIVLS----
gi|118377903| 181 DPTTNSDNNFALDLSLLVIYTVEMGLKLLAMGFPSKNSFLSEGWNLDVFIITGWIIPY
gi|158895975| 29 -----SVLEVVYTAELILRLRFIM----KRFFHSEWNELDIIVITGWISE
gi|221126494| 181 N-----APVAVYVFTGIVTEEMILKFAKCLLHLHYAVLRGWNILDVWVWVLSVITL
consensus 181 eavdkvllivyTvEliikl amg i krsflsegWniIdfvivvtgwis

gi|260888220| 86 SIVFIS-----SSISVIRAPRIFRLLAAAEPELOLVSSMLKAPSMTWALLLFI
gi|118377903| 241 VINSSS-----FNLALRSLSRVLRELRSSISKVHILNQLVAFETAFLKDAFILLCF
gi|158895975| 71 IIEVFGQSVMDKVALKCDRMLRLRLVRLVAFRELYVLVAGTGCCKTTFWAGGLLFI
gi|221126494| 235 IIPGAN-----LSCISIRVLRALRTISIEGLKTYVNALNADAMSSVHITTF
consensus 241 vi v s lsairm Rvrlrlrsis fkeLqilVsallkaip mtwalilf
*****

gi|260888220| 139 VEYIFAEGSTMYG-----DAP-----
gi|118377903| 294 FGVFGLASLOLESCKKTRCLDIITGIKLGE-----EMNNIFVCAQDGDCCIQ
gi|158895975| 131 VLNWVSLLSVEFLDQVKELVPSVYR-----
gi|221126494| 286 FVCLFALFGLQLFAGVFSQKCVLNAIVYNSSSWLMQIQNSSNWHNGQICCNNSGSLNC
consensus 301 f ifaifglqlfgg lk rcidaf d n l vcaq

gi|260888220| 157 -----ELFGDIGGSMFIFCVMTFESMAFAVARPIMAMYPMAWLYFL
gi|118377903| 343 QGISYICAKMVSNSDLYNFDFGWSLENYVWVMSNGWDFIMQFVIVGFNFPAFYFF
gi|158895975| 158 -----PCSCPTAFSNIWMANLTWPHIISTADWI-MIGRPHVEEENFNYLFFV
gi|221126494| 346 PKNFTCQQNMAPNENNGFTSFDNIGWAILTSPQLINRFWEDVYIGVLSAGPYYIYFFL
consensus 361 g pnw t Fdnigw mltlfqvis e Wstivar ivamy wawlyFl

gi|260888220| 199 VFLLLTAITLNVVVGIVVEAVGT-----ISEAKQKQAAEAAENAP-----
gi|118377903| 403 IIVYLGWALLVNLCAGILAVKYESIANQSSSKGNQESNNQNLRWGIYKSNKELY
gi|158895975| 205 FNIFIVVXLVNLTIV-----AAIVDANIARVDN-----
gi|221126494| 406 VIIEFSEFLNLIILAVVVFVYQH-----EITCVDFSAQERLHQVLSPPYDDLYVLP
consensus 421 viiflgtw LvNlivgivv y issav d qaayrmen i y il

```

**Fig. 18.** Multiple Sequence Alignment of EST 158895975.

```

gi|259418114| 1 CLVIFVLEIAAKLIVR-----RLRFVNCNIFDGLVGHAVP-----G
gi|294936273| 1 FVLVYIVELIIRIYYN-----RADFFNDAENIFDCLLVFKTHVETFIVSGSALKS
gi|158893679| 1 EMHRLDSIINDHLSQSPWRLKAMRILASLFEIIVGVVTLINLVVVIETNERAQCRSAS
consensus 1 lvifvlei kl y R fF awnifd liv i lvet g

gi|259418114| 41 AQQFSVLR--ALRLRVLRLVSVAPRI--RRVVEGPTALPGMGSVFLLMALIFYIGSVM
gi|294936273| 51 LLSLRIVR--MIRLRLVRLIRFFREL--WLVITGIAMKTLAWVTLILFSLTWCALL
gi|158893679| 61 HSDPRCNEGDAIRLLNIGCLLIIVAEITFRILYVDQKFLQNWNCSDLEFVCLSGVAEYF
consensus 61 rvlr aIRllrvrlil f el rlvvegii ml mg v Llm ltyvg vm
*****

gi|259418114| 97 AKKLFQDN-----FP--EWFGLALSAYSLFQIMLESASGIVRPVMD
gi|294936273| 107 LRLVLCSTAWAFTRQIAGPYEFDPEYEPGTFWAGLDTLFPVTTQDNMMISVARPVCR
gi|158893679| 121 NSVFPSLG-----FLRMERLRL--LRIVRIITTCRELYMLVHGFST
consensus 121 tll g f ewFgtl l l tlfqittqe w m v rpv

gi|259418114| 139 VYFYWAFFVFFIMVTTFAVNLIVGLIIVNSMQDAHSEEGERTDAYRQOVLRLAIEA
gi|294936273| 167 YPEMMLFFFLPYEFTAMAILNVMLASLIN-----NAIQAAKVGLADAERV
gi|158893679| 162 YKALGWAAILMLGLLTMSIIAWEVVLHPIN-----LWVEHSNVYEECDRCPRSFASV
consensus 181 vypiaWafflpfimvT failnvlg iin e gertnayed lt ea v

gi|259418114| 199 RIRIQDEKTQK
gi|294936273| 213 EQSFPGK----
gi|158893679| 215 WEANITFFTPL-
consensus 241 r a l

```

**Fig. 19.** Multiple Sequence Alignment of EST 158893679.

```

gi|158891827| 1 KTPCCSSCPAIKS-----RTPPKKVDLLPSSDMDLETLRGTAAPAL
gi|209966108| 1 -----MTESSALPAAADAPAAPAP
gi|294932777| 1 VFFIEWILRVTSGGVWVMFDMYNFLDFLLVFITGVMANWFLPLGVDVSLRKLTVLRAL
consensus 1 e s d m d l l r a a p a l

gi|158891827| 44 RSRRHSGSFGSTLGERHFPRVWL-LVKSROEAFVIGFVIVVNCVTMGLAEELLGSAKCC
gi|209966108| 19 SARR-----LVESVGEORFIVAVILVNAVTLGLETSATAMAAAG-
gi|294932777| 61 FLVRLARAVRLIPMFEMWLLGGFLDSALLILTFIITVLLFMFSVWVEIICHNAGV
consensus 61 r lRh r lv lveS f fiv vIvN vtmglea llg aag

gi|158891827| 103 E-----QEVTAASH-----VITAVFLVELLLRGTAYGPRIFYFMPGNLWNFVAMLV
gi|209966108| 59 -----PLVTVLDR-----LALAVEVVELGLALAVY--SLSEFRQG--NNNFETIV
gi|294932777| 121 AADPVTAELPFTLHQGMFTLLQLFEVLTLETWSDSVRPLMHLFRVGLYAVLNMNLYT
consensus 121 plvt leh l laVflvel lr ay ri fpr G wn fd mlv

gi|158891827| 150 LLIGV-----VTVWLLPLLGAESSTFQTLTVLRALRLA
gi|209966108| 101 ALS-----LAPLIGWACAG---NLVLRALRL
gi|294932777| 181 AVVVENAFALARKDEEQARVKQAQRDRDILDKDLFMELDADGSELDKRFDOALGDS
consensus 181 alt l pllgaegg ltvlrALrl
*****

gi|158891827| 184 RLVRVVSQVEVWFHEVWLLRLGLMESGRVLFWTVVVFLLITYIFAVFGVVLGTKIKKKID
gi|209966108| 126 RVLRLLSVWPQMRAYQALFGALPG---MGAILAVLALIIYNSAVLATKIFGTSFPPWFG
gi|294932777| 241 KLVDKMSSGIEREDMININDLLDGDGVLTEEFTGGRMKKGARAKELVLRQLR
consensus 241 rlvrvvssv revvl lfglle g l iv viali yi av g kllgt ike h
****

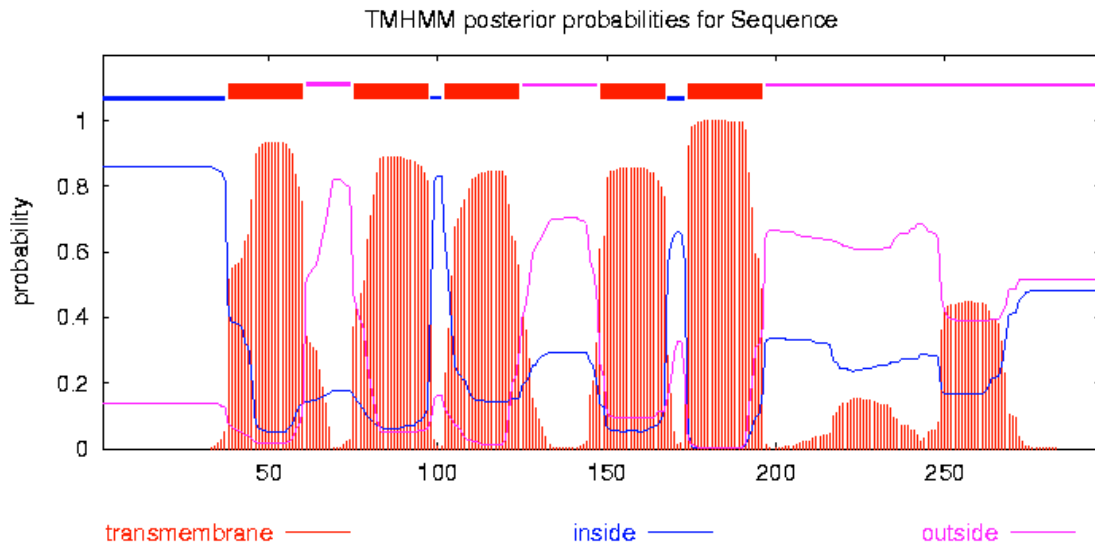
gi|158891827| 244 DEWNSTGSADQELFPLVTTVDGVLPLMYLLQVLTLDWNS-ILRPMMKHIPWSWIFFYL
gi|209966108| 183 SGRS-----MYSLEQLITLESWSMGIVREVMVYPMANVVEFVP
gi|294932777| 301 AVVDSCELLANEES-----DQVVELCESVEGTVKDLSVRLDVGQAATAIGTLRY
consensus 301 iv S qel mytL qvltldsw m ivRpm m v pwaw ff

gi|158891827| 303 YISIAVIVMMNLVTAKIVX-----NALKNSQQDAEELIA-----
gi|209966108| 222 EIVLVTFAVLNLFTAIVNSMQSLHDADTKAEAEERELEKAKHERTAMADELRALRAEIIAG
gi|294932777| 351 YFQHYDVRLEFSSRWOLRKEIQQERFKNKRAYDQPKASTTAFRVESWLYWPPATSTCCS
consensus 361 yi iv vmlnl iaiiv a k da elka g r

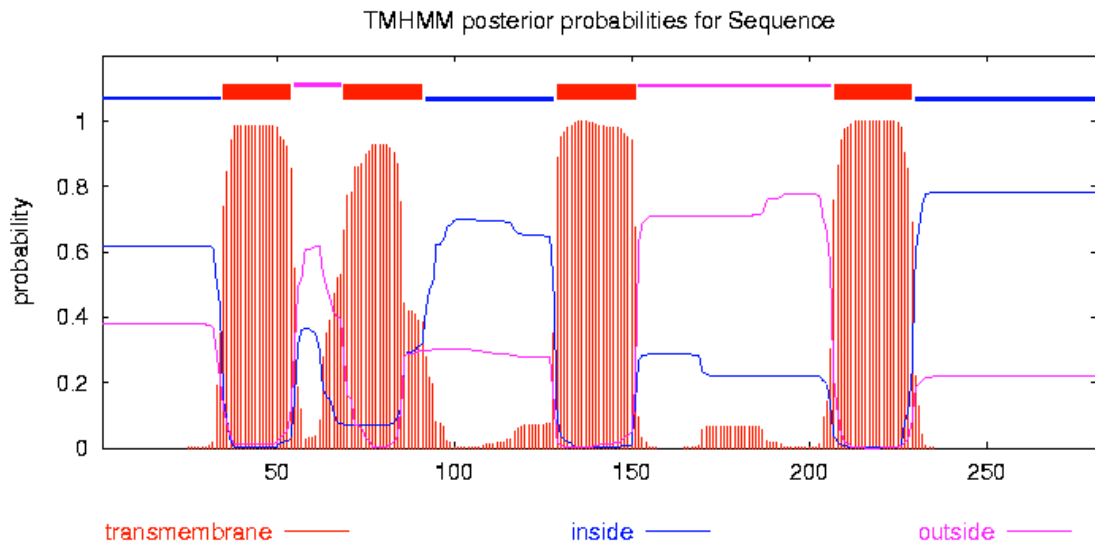
gi|158891827| -----
gi|209966108| 282 LKEAVLRDR
gi|294932777| 411 H-----
consensus 421

```

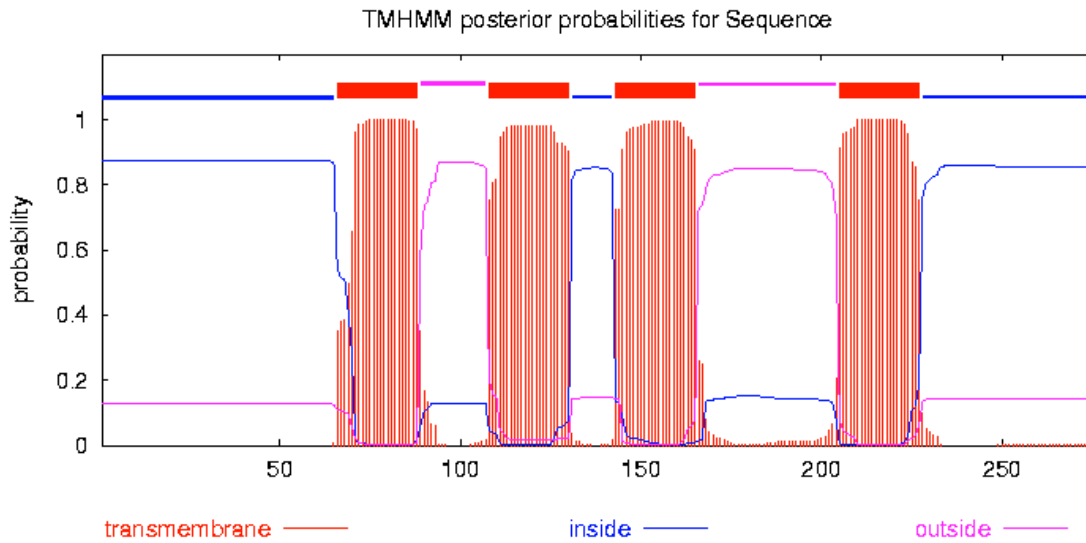
Fig. 20. Multiple Sequence Alignment of EST 158891827.



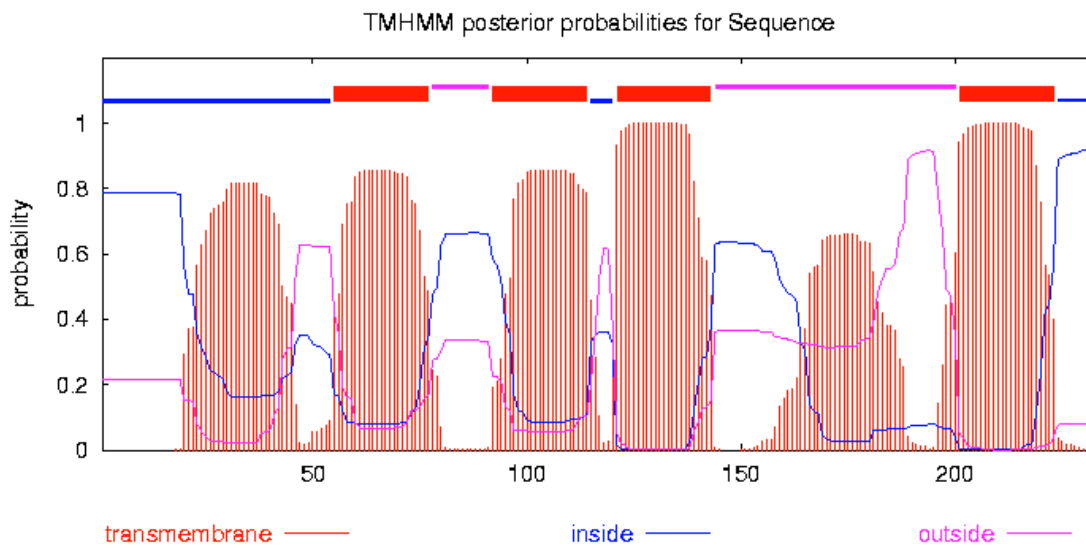
**Fig. 21.** The Predicted Transmembrane Domain Structure of 48705280.



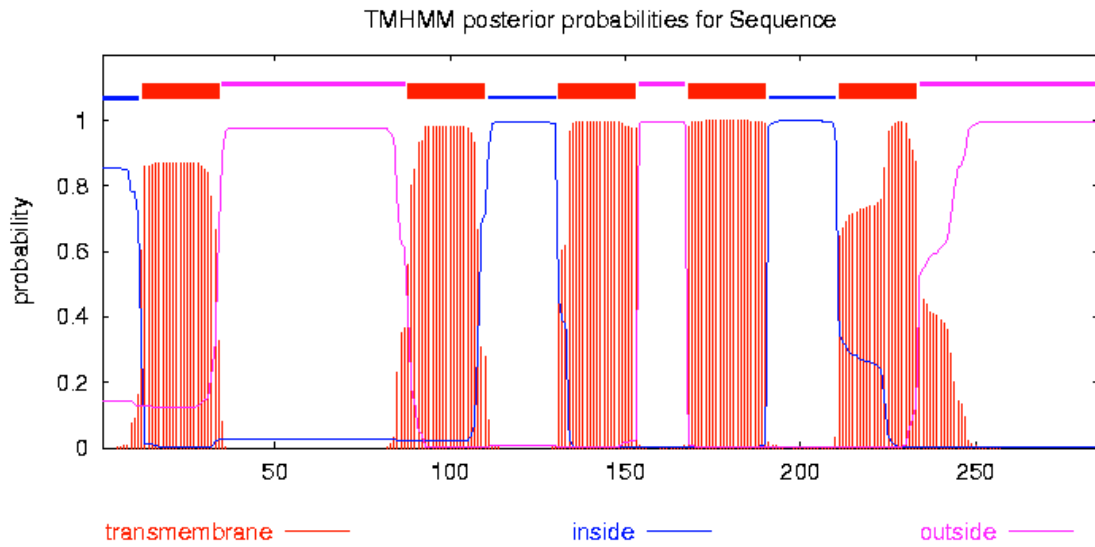
**Fig. 22.** The Predicted Transmembrane Domain Structure of 48706986.



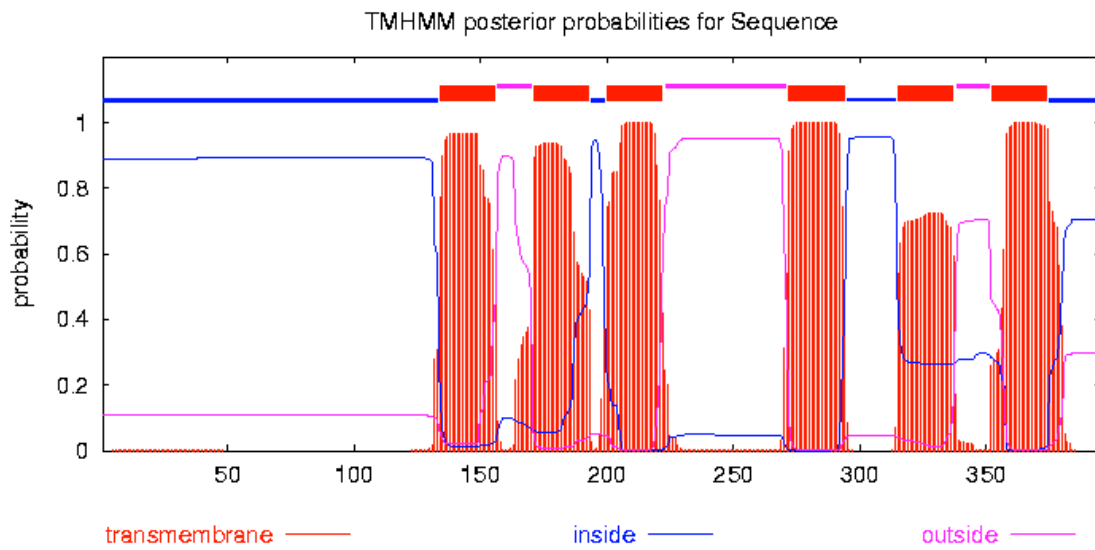
**Fig. 23.** The Predicted Transmembrane Domain Structure of 158897096.



**Fig. 24.** The Predicted Transmembrane Domain Structure of 158895975.



**Fig. 25.** The Predicted Transmembrane Domain Structure of 158893679.



**Fig. 26.** The Predicted Transmembrane Domain Structure of 158891827.

## VITA

Name: Natalie Jeanette Thompson

Address: Department of Biology, Texas A&M University  
3146 TAMU  
College Station, TX 77843

Email Address: NJThompson@cvm.tamu.edu

Education: B.S., Genetics, Texas A&M University, 2008  
M.S., Microbiology, Texas A&M University, 2011