

ANÁLISE DAS SEQUÊNCIAS DE NUCLEOTÍDEOS DE NOVOS SUBTIPOS DE TOXINAS E PROTEÍNAS DE ANÊMONAS DO MAR POR SEQUENCIAMENTO DE RNA DE NOVA GERAÇÃO

Leandro Augusto Freire Boralli

Ursula Castro de Oliveira

Milton Yutaka Nishiyama Junior

Gandhi Rádis Baptista

Irina Kerkis

Álvaro Rossan de Brandão Prieto da Silva

Inácio de Loiola Meirelles Junqueira de Azevedo

Andre Junqueira Zaharenko

ANÁLISE DAS SEQUÊNCIAS DE NUCLEOTÍDEOS DE NOVOS SUBTIPOS DE TOXINAS E PROTEÍNAS DE ANÊMONAS DO MAR POR SEQUENCIAMENTO DE RNA DE NOVA GERAÇÃO

Leandro Augusto Freire Boralli
Ursula Castro de Oliveira
Milton Yutaka Nishiyama Junior
Gandhi Rádís Baptista
Irina Kerkis

Álvaro Rossan de Brandão Prieto da Silva
Inácio de Loiola Meirelles Junqueira de Azevedo
Andre Junqueira Zaharenko

Editora chefe

Profª Drª Antonella Carvalho de Oliveira

Editora executiva

Natalia Oliveira

Assistente editorial

Flávia Roberta Barão

Bibliotecária

Janaina Ramos

Projeto gráfico

Camila Alves de Cremo

Daphynny Pamplona

Gabriel Motomu Teshima

Luiza Alves Batista

Natália Sandrini de Azevedo

Imagens da capa

iStock

Edição de arte

Luiza Alves Batista

2022 by Atena Editora

Copyright © Atena Editora

Copyright do texto © 2022 Os autores

Copyright da edição © 2022 Atena Editora

Direitos para esta edição cedidos à Atena Editora pelos autores.

Open access publication by Atena Editora



Todo o conteúdo deste livro está licenciado sob uma Licença de Atribuição Creative Commons. Atribuição-Não-Comercial-NãoDerivativos 4.0 Internacional (CC BY-NC-ND 4.0).

O conteúdo do texto e seus dados em sua forma, correção e confiabilidade são de responsabilidade exclusiva dos autores, inclusive não representam necessariamente a posição oficial da Atena Editora. Permitido o *download* da obra e o compartilhamento desde que sejam atribuídos créditos aos autores, mas sem a possibilidade de alterá-la de nenhuma forma ou utilizá-la para fins comerciais.

Todos os manuscritos foram previamente submetidos à avaliação cega pelos pares, membros do Conselho Editorial desta Editora, tendo sido aprovados para a publicação com base em critérios de neutralidade e imparcialidade acadêmica.

A Atena Editora é comprometida em garantir a integridade editorial em todas as etapas do processo de publicação, evitando plágio, dados ou resultados fraudulentos e impedindo que interesses financeiros comprometam os padrões éticos da publicação. Situações suspeitas de má conduta científica serão investigadas sob o mais alto padrão de rigor acadêmico e ético.

Conselho Editorial**Ciências Biológicas e da Saúde**

Profª Drª Aline Silva da Fonte Santa Rosa de Oliveira – Hospital Federal de Bonsucesso

Profª Drª Ana Beatriz Duarte Vieira – Universidade de Brasília

Profª Drª Ana Paula Peron – Universidade Tecnológica Federal do Paraná

Prof. Dr. André Ribeiro da Silva – Universidade de Brasília

Profª Drª Anelise Levay Murari – Universidade Federal de Pelotas

Prof. Dr. Benedito Rodrigues da Silva Neto – Universidade Federal de Goiás



Prof. Dr. Cirênio de Almeida Barbosa – Universidade Federal de Ouro Preto
Prof^o Dr^a Daniela Reis Joaquim de Freitas – Universidade Federal do Piauí
Prof^o Dr^a Débora Luana Ribeiro Pessoa – Universidade Federal do Maranhão
Prof. Dr. Douglas Siqueira de Almeida Chaves – Universidade Federal Rural do Rio de Janeiro
Prof. Dr. Edson da Silva – Universidade Federal dos Vales do Jequitinhonha e Mucuri
Prof^o Dr^a Elizabeth Cordeiro Fernandes – Faculdade Integrada Medicina
Prof^o Dr^a Eleuza Rodrigues Machado – Faculdade Anhanguera de Brasília
Prof^o Dr^a Elane Schwinden Prudêncio – Universidade Federal de Santa Catarina
Prof^o Dr^a Eysler Gonçalves Maia Brasil – Universidade da Integração Internacional da Lusofonia Afro-Brasileira
Prof. Dr. Ferlando Lima Santos – Universidade Federal do Recôncavo da Bahia
Prof^o Dr^a Fernanda Miguel de Andrade – Universidade Federal de Pernambuco
Prof. Dr. Fernando Mendes – Instituto Politécnico de Coimbra – Escola Superior de Saúde de Coimbra
Prof^o Dr^a Gabriela Vieira do Amaral – Universidade de Vassouras
Prof. Dr. Gianfábio Pimentel Franco – Universidade Federal de Santa Maria
Prof. Dr. Helio Franklin Rodrigues de Almeida – Universidade Federal de Rondônia
Prof^o Dr^a Iara Lúcia Tescarollo – Universidade São Francisco
Prof. Dr. Igor Luiz Vieira de Lima Santos – Universidade Federal de Campina Grande
Prof. Dr. Jefferson Thiago Souza – Universidade Estadual do Ceará
Prof. Dr. Jesus Rodrigues Lemos – Universidade Federal do Piauí
Prof. Dr. Jônatas de França Barros – Universidade Federal do Rio Grande do Norte
Prof. Dr. José Aderval Aragão – Universidade Federal de Sergipe
Prof. Dr. José Max Barbosa de Oliveira Junior – Universidade Federal do Oeste do Pará
Prof^o Dr^a Juliana Santana de Curcio – Universidade Federal de Goiás
Prof^o Dr^a Lívia do Carmo Silva – Universidade Federal de Goiás
Prof. Dr. Luís Paulo Souza e Souza – Universidade Federal do Amazonas
Prof^o Dr^a Magnólia de Araújo Campos – Universidade Federal de Campina Grande
Prof. Dr. Marcus Fernando da Silva Praxedes – Universidade Federal do Recôncavo da Bahia
Prof^o Dr^a Maria Tatiane Gonçalves Sá – Universidade do Estado do Pará
Prof. Dr. Maurilio Antonio Varavallo – Universidade Federal do Tocantins
Prof^o Dr^a Mylena Andréa Oliveira Torres – Universidade Ceuma
Prof^o Dr^a Natiéli Piovesan – Instituto Federaci do Rio Grande do Norte
Prof. Dr. Paulo Inada – Universidade Estadual de Maringá
Prof. Dr. Rafael Henrique Silva – Hospital Universitário da Universidade Federal da Grande Dourados
Prof^o Dr^a Regiane Luz Carvalho – Centro Universitário das Faculdades Associadas de Ensino
Prof^o Dr^a Renata Mendes de Freitas – Universidade Federal de Juiz de Fora
Prof^o Dr^a Sheyla Mara Silva de Oliveira – Universidade do Estado do Pará
Prof^o Dr^a Suely Lopes de Azevedo – Universidade Federal Fluminense
Prof^o Dr^a Vanessa da Fontoura Custódio Monteiro – Universidade do Vale do Sapucaí
Prof^o Dr^a Vanessa Lima Gonçalves – Universidade Estadual de Ponta Grossa
Prof^o Dr^a Vanessa Bordin Viera – Universidade Federal de Campina Grande
Prof^o Dr^a Welma Emídio da Silva – Universidade Federal Rural de Pernambuco



Análise das sequências de nucleotídeos de novos subtipos de toxinas e proteínas de anêmonas do mar por sequenciamento de RNA de nova geração

Diagramação: Natália Sandrini de Azevedo
Correção: Maiara Ferreira
Indexação: Amanda Kelly da Costa Veiga
Revisão: Os autores

Dados Internacionais de Catalogação na Publicação (CIP)

A532 Análise das sequências de nucleotídeos de novos subtipos de toxinas e proteínas de anêmonas do mar por sequenciamento de RNA de nova geração / Leandro Augusto Freire Boralli, Ursula Castro de Oliveira, Milton Yutaka Nishiyama Junior, et al. – Ponta Grossa - PR: Atena, 2022.

Outros autores

Gandhi Rádis Baptista

Irina Kerkis

Álvaro Rossan de Brandão Prieto da Silva

Inácio de Lóiola Meirelles Junqueira de Azevedo

Andre Junqueira Zaharenko

Formato: PDF

Requisitos de sistema: Adobe Acrobat Reader

Modo de acesso: World Wide Web

Inclui bibliografia

ISBN 978-65-5983-902-5

DOI: <https://doi.org/10.22533/at.ed.025221702>

1. Invertebrados marinhos - Anêmona-do-mar. 2. Toxinas. 3. Transcriptoma. 4. Bunodossoma caissarum. I. Boralli, Leandro Augusto Freire. II. Oliveira, Ursula Castro de. III. Nishiyama Junior, Milton Yutaka. IV. Título.

CDD 593

Elaborado por Bibliotecária Janaina Ramos – CRB-8/9166

Atena Editora

Ponta Grossa – Paraná – Brasil

Telefone: +55 (42) 3323-5493

www.atenaeditora.com.br

contato@atenaeditora.com.br



Atena
Editora
Ano 2022

DECLARAÇÃO DOS AUTORES

Os autores desta obra: 1. Atestam não possuir qualquer interesse comercial que constitua um conflito de interesses em relação ao artigo científico publicado; 2. Declaram que participaram ativamente da construção dos respectivos manuscritos, preferencialmente na: a) Concepção do estudo, e/ou aquisição de dados, e/ou análise e interpretação de dados; b) Elaboração do artigo ou revisão com vistas a tornar o material intelectualmente relevante; c) Aprovação final do manuscrito para submissão.; 3. Certificam que o texto publicado está completamente isento de dados e/ou resultados fraudulentos; 4. Confirmam a citação e a referência correta de todos os dados e de interpretações de dados de outras pesquisas; 5. Reconhecem terem informado todas as fontes de financiamento recebidas para a consecução da pesquisa; 6. Autorizam a edição da obra, que incluem os registros de ficha catalográfica, ISBN, DOI e demais indexadores, projeto visual e criação de capa, diagramação de miolo, assim como lançamento e divulgação da mesma conforme critérios da Atena Editora.



DECLARAÇÃO DA EDITORA

A Atena Editora declara, para os devidos fins de direito, que: 1. A presente publicação constitui apenas transferência temporária dos direitos autorais, direito sobre a publicação, inclusive não constitui responsabilidade solidária na criação dos manuscritos publicados, nos termos previstos na Lei sobre direitos autorais (Lei 9610/98), no art. 184 do Código Penal e no art. 927 do Código Civil; 2. Autoriza e incentiva os autores a assinarem contratos com repositórios institucionais, com fins exclusivos de divulgação da obra, desde que com o devido reconhecimento de autoria e edição e sem qualquer finalidade comercial; 3. Todos os e-book são *open access*, *desta forma* não os comercializa em seu site, sites parceiros, plataformas de *e-commerce*, ou qualquer outro meio virtual ou físico, portanto, está isenta de repasses de direitos autorais aos autores; 4. Todos os membros do conselho editorial são doutores e vinculados a instituições de ensino superior públicas, conforme recomendação da CAPES para obtenção do Qualis livro; 5. Não cede, comercializa ou autoriza a utilização dos nomes e e-mails dos autores, bem como nenhum outro dado dos mesmos, para qualquer finalidade que não o escopo da divulgação desta obra.



Dedico esse trabalho ao meu pai, que embora separado de nós por 17 anos ainda é fonte de sabedoria e alegrias para a minha família.

AGRADECIMENTOS

A Deus primeiramente, pois não cai um fio de nosso cabelo sem que Ele permita.

A minha família, que sem o seu apoio eu não estaria terminando esse trabalho.

A Débora, minha namorada, melhor amiga, futura esposa e com quem pretendo passar a vida junto. Seu amparo foi fundamental.

A CAPES, que sem o apoio monetário eu não teria conseguido me manter durante os anos de trabalho dessa tese.

A Dra. Irina Kerkis, que abriu as portas do seu laboratório para esse aluno que aqui escreve, anos atrás antes de decidir sobre o possível tema da pós-graduação.

Ao Dr. Alvaro Rossan de Brandão Prieto-da-Silva, que foi imprescindível para esse trabalho, tanto na questão científica quanto nos problemas normais da vida.

Ao Dr. André Junqueira Zaharenko, a cabeça que planejou esse trabalho cujo conhecimento em toxinas de cnidários tornou essa tese possível.

Ao Laboratório Especial de Toxinologia Aplicada (LETA), do Instituto Butantan, que sem a colaboração técnica dos pesquisadores Dra. Ursula Castro de Oliveira e Dr. Milton Nishiyama esse trabalho não teria saído do papel.

Aos meus amigos, cujos nomes são muitos e que sempre me apoiaram de todas formas possíveis, nos bons e maus momentos.

SUMÁRIO

LISTA DE ABREVIATURAS	1
RESUMO	2
ABSTRACT	3
INTRODUÇÃO.....	4
REVISÃO BIBLIOGRÁFICA.....	6
O Filo Cnidaria	6
Estudos das toxinas.....	6
Actinoporinas: Toxinas citolísticas de anêmonas do mar	7
Toxinas bloqueadoras de canais de potássio depende de voltagem do tipo I – SHKT	9
MATERIAL E MÉTODOS.....	15
Origem das Amostras	15
Preparo do tecido.....	15
Extração de RNA total.....	15
Construção das bibliotecas de cDNA	15
Purificação do mRNA	15
Preparo do mRNA	16
Construção da biblioteca de cDNA	16
Sequenciamento	17
Processamentos do sequenciamento, montagem das leituras e análises computacionais	18
Processamento.....	18
Montagem das leituras	18
Análises computacionais.....	19
RESULTADOS E DISCUSSÃO	21
Extração de RNA	21
Purificação do mRNA.....	21
Síntese do cDNA e validação das bibliotecas	24
Sequenciamento	27
Análise do sequenciamento.....	29

Análise dos transcritos similares a Actinoporinas	36
Análise dos transcritos similares a toxinas bloqueadoras de canais de potássio dependente de voltagem do tipo I – ShKT.....	44
CONCLUSÃO.....	56
REFERÊNCIAS	57
SOBRE OS AUTORES	61

LISTA DE ABREVIATURAS

BCS-SP	Bunodosoma caissarum proveniente da costa de São Paulo
BCS-Aq	Bunodosoma caissarum proveniente do Arquipélago São Pedro e São Paulo
ShK	Toxina bloqueadora de canais de potássio do tipo Ia, proveniente da anêmona <i>Stichodactyla helianthus</i>
ShKT	Domínio conservado da toxina ShK presente em vários tipos diferentes de proteínas em diversos filos
BgK	Toxina bloqueadora de canais de potássio do tipo Ib, proveniente da anêmona <i>Bunodosoma granulifera</i>
CRISP	Domínio proteico presente em vários organismos, como eucariotos vertebrados e invertebrados. Significa Cysteine-rich Secretory Proteins
PCR	Sigla em inglês para Reação em cadeia da Polimerase (Polymerase Chain Reaction)
FPKM	Sigla em inglês para Fragments Per Kilobase Of Exon Per Million
TMM	Sigla em inglês para Trimmed Mean of M-values

RESUMO

As anêmonas do mar são organismos sésseis que empregam um aparato peçonhento (localizado preferencialmente em seus tentáculos) para produzir e inocular toxinas paralisantes em suas presas ou para defender-se. Nosso grupo contribuiu recentemente na caracterização do peptidoma de 3 espécies de anêmonas do mar, mostrando que a diversidade molecular das toxinas desses organismos é muito maior que o esperado. Do ponto de vista da informação nucleotídica obtida das anêmonas, poucas espécies foram investigadas em detalhe, revelando que no momento apenas uma pequena fração do arsenal tóxico dos cnidários foi investigado. Adicionalmente, novas proteínas estruturais que ocorrem em anêmonas apresentam uma similaridade gênica e estrutural impressionante com proteínas de organismos derivados, tal como vertebrados. Neste sentido, pretendemos analisar e comparar o transcriptoma de duas populações da anêmona brasileira *Bunodosoma caissarum*, uma população da costa de São Paulo e outra população localizada no arquipélago isolado São Pedro e São Paulo. Essa análise será feita através do sequenciamento de RNA de nova geração, pelo sistema Illumina. Com esse sequenciamento e com análises de bioinformática esperamos encontrar informações referentes a novas toxinas nunca antes estudadas nessa espécie e também analisar as diferenças no perfil de expressão entre as duas populações.

ABSTRACT

Sea anemones are sessile organisms that employ a venomous apparatus (preferably located in their tentacles) to produce and inoculate paralyzing toxins in their prey or to defend themselves. Our group recently contributed to the characterization of the peptidoma of 3 species of sea anemones, showing that the molecular diversity of the toxins of these organisms is much higher than expected. From the point of view of the nucleotide information obtained from anemones, few species were investigated in detail, revealing that at the moment only a small fraction of the toxic arsenal of the cnidarians was investigated. In addition, new structural proteins that occur in anemones present structural and genetic similarity to proteins of derived organisms, such as vertebrates. In this sense, we intend to analyze and compare the transcriptome of two populations of the Brazilian anemone *Bunodosoma caissarum*, a population of the São Paulo coast and another population located in the isolated São Pedro and São Paulo archipelago. This analysis will be done through the sequencing of new generation RNA, by the Illumina system. With this sequencing and with bioinformatics analyzes we hope to find information regarding new toxins never before studied in this species and also to analyze the differences in the expression profile between the two populations

INTRODUÇÃO

As anêmonas-do-mar são animais sésseis do filo Cnidaria, o mesmo filo das águas-vivas, e são divididos em quatro classes, no qual as anêmonas fazem parte da classe Anthozoa, junto com os corais verdadeiros. Esses animais possuem uma simetria radial e apresentam apenas duas camadas celulares: a ectoderme e gastroderme, com uma matriz extracelular entre elas, chamada mesogleia. São essencialmente predadores de crustáceos e peixes (1,2).

Esse filo recebeu esse nome por causa de uma estrutura específica que esses animais possuem: o cnidócito. Essas células especializadas servem tanto para defesa quanto para predação e conseguem inocular sua toxina, que contém uma ampla variedade de componentes, através de uma estrutura chamada nematocisto (3), que lembra uma espécie de arpão que perfura o tecido do alvo após ser estimulada fisicamente. Embora existam aproximadamente 30 tipos de cnidócitos distintos os cnidários contém apenas uma pequena fração dessas diferentes células (2).

Desde a década de 70 que as anêmonas começaram a serem estudadas para se entender o mecanismo de ação de suas toxinas. Nessa década tivemos alguns trabalhos muito elucidativos sobre as características das toxinas desses animais (4,5). Um desses trabalhos foi o trabalho pioneiro de Beress, no qual conseguiram purificar três pólipéptidos com atividade neuronal e cardiológica, da anêmona *Anemonia sulcata*, onde foi investigado a ação e mecanismos das toxinas ATX-I, II e III. (6).

O estudo aprofundado das toxinas das anêmonas do mar tem se mostrado um campo bem proveitoso para identificação de possíveis novos fármacos. Um exemplo é a toxina ShK, da anêmona caribenha *Stichodactyla helianthus*. Essa toxina é um potente bloqueador de canais de potássio dependente de voltagem, que atualmente foi usada como modelo para o desenvolvimento de uma droga capaz de bloquear especificamente o canal de potássio dependente de voltagem 1.3, presentes em linfócitos T, que são mediadores de doenças autoimunes, como lúpus (7,8).

Conforme as tecnologias vão evoluindo, mais ferramentas são utilizadas na investigação em busca de novas toxinas, e uma dessas novas tecnologias mais promissoras é o sequenciamento de nova geração. No entanto, poucos trabalhos com anêmonas do mar foram publicados utilizando esses novos métodos. A técnica de sequenciamento de mRNA que os sistemas de sequenciamento de nova geração proporciona uma grande ferramenta no trabalho de prospecção de novas toxinas, pois é possível analisar todo o material que está sendo expresso em proteínas no animal (9,10).

O objetivo do trabalho é utilizar-se da capacidade de sequenciamento do sistema Illumina para analisarmos todo o transcriptoma de duas populações distintas da anêmona *Bunodosoma caissarum*, uma população da costa de São Paulo (denominada Bcs-SP) e outra população isolada no Arquipélago São Pedro e São Paulo (denominada Bcs-Aq),

no Oceano Atlântico. Com os dados do transcriptoma das populações de *Bunodosoma caissarum* temos também como objetivo analisar esses dados em busca de possíveis novas toxinas e também compararmos o perfil de transcriptoma das diferentes populações.

REVISÃO BIBLIOGRÁFICA

O FILO CNIDARIA

O filo Cnidaria compreende o grupo de seres vivos que primeiramente apresentaram tecidos diferenciados na evolução em nosso globo. Esses animais são divididos em quatro classes: Hydrozoa, Scyphozoa, Cubozoa e Anthozoa, esse último o qual as anêmonas fazem parte, junto com os corais verdadeiros. Esses animais possuem uma simetria radial e possuem apenas duas camadas celulares, a ectoderme e gastroderme, com um matriz extracelular entre elas chamada mesogleia (1,2).

Esse filo compartilha uma estrutura específica que todos esses animais possuem: o cnidócito. Essas células urticantes tem a função de defesa, ajudar na competição espacial interespecie e também de predação das presas que servem de alimento para os cnidários (1). Essas células especializadas conseguem inocular as toxinas através de uma estrutura chamada nematocisto, que lembra uma espécie de arpão que perfura o tecido do alvo após ser estimulada fisicamente (2).

ESTUDOS DAS TOXINAS

Atualmente são conhecidas 255 toxinas de cnidários, divididas em 19 diferentes famílias, nos quais encontramos, por exemplo, neurotoxinas que atuam em canais de sódio e canais de potássio, com peso molecular por volta de 5 kDa. Também são conhecidas toxinas enzimáticas, com ações conhecidas como formadoras de poros em corpos celulares (conhecidas como citolisinas, com 18-20 kDa), fosfolipases e metaloproteases (1,2,11).

Os estudos das toxinas das anêmonas são relativamente recentes, com o primeiro trabalho sobre toxinas de cnidários sendo publicado em 1968, num estudo sobre purificação da toxina da anêmona *Condylactis gigantea* (12) sendo seguido em 1975 de outro trabalho, de Beress, sobre a purificação de polipeptídeos com atividade de neurotóxicas presentes na peçonha da anêmona *Anemonia sulcata* (6).

Em seu trabalho de 1975, Laszlo Béress e colaboradores purificaram e isolaram inicialmente 3 neurotoxinas peptídicas onde duas (com tamanho entre 45 e 50 aminoácidos), homólogas entre si, tiveram demonstrado os seus efeitos na inativação de canais de sódio dependentes de voltagem. Em trabalho posterior Caterral e Béress demonstraram que um desses peptídeos purificados no trabalho de 1975 (Toxin II) possui um sítio receptor similar ao encontrado em toxinas inibidores de ionóforos de sódio nos axônios neurais e músculos dos escorpiões *Leiurus quinquestriatus* e *Buthus tamulus* (13).

Atualmente já se conhece várias outras toxinas de cnidários, com os mais diferentes alvos farmacológicos, sendo responsáveis pelo desenvolvimento de drogas, tanto terapêuticas como também inseticidas. Em suas toxinas as anêmonas possuem um armamento para os mais diversos alvos: peptídeos formadores de poros em membranas

celulares, neurotoxinas bloqueadoras de canais de potássio e de sódio, inibidores de proteases e fosfolipases. Esse trabalho apresentará uma revisão sobre as famílias de toxina actinoporina e as neurotoxinas bloqueadoras de canal de potássio do tipo I (1,2,14,15), sendo essas famílias o foco de uma análise mais profunda na sessão de discussão.

Actinoporinas: Toxinas citolísticas de anêmonas do mar

Várias espécies de anêmonas do mar apresentam toxinas citolísticas denominadas Actinoporinas, que exibem uma ampla diversidade de atividades biológicas, incluindo hemólise de glóbulos vermelhos, agregação de plaquetas, efeitos citotóxicos e citostáticos em fibroblastos, atividade cardiotrópica, despolarização de membranas e bloqueio de canais de potássio (16). As actinoporinas são toxinas protéicas que danificam membranas celulares formando poros. Essas citolisinas em geral, são moléculas grandes se comparadas com as neurotoxinas de anêmonas e tem por volta de 20 kDa com uma estrutura β -sandwich flanqueada por duas α -hélices (ver a estrutura tridimensional na Figura 01), outra característica distintiva em relação à maioria das neurotoxinas de anêmona é que as actinoporinas não contém pontes dissulfeto. Letalidade em mamíferos tem sido atribuída a vasoespasmos graves dos vasos coronários, arritmia cardíaca e efeitos inotrópicos (17,18).

Essas toxinas formam uma família bem caracterizada chamada de *Actinoporin family*. Entretanto, toxinas similares foram encontradas no ducto de peçonha de caramujos predadores do gênero *Conus* por análise transcriptômica e proteômica (19,20), sendo denominadas teroporinas ou conoporinas. Sequências similares também foram encontradas nos gametófitos de briófitas (21). Assim a família das actinoporinas atualmente é composta por 35 sequências que se subdividem em três subfamílias conforme sejam originárias de anêmonas do mar (Actinoporinas subfamília Actiniaria), ou de sequências similares presentes nas peçonhas de moluscos predadores do gênero *Conus* (Actinoporinas subfamília Conoidea), ou de plantas (Actinoporinas subfamília Plantas).

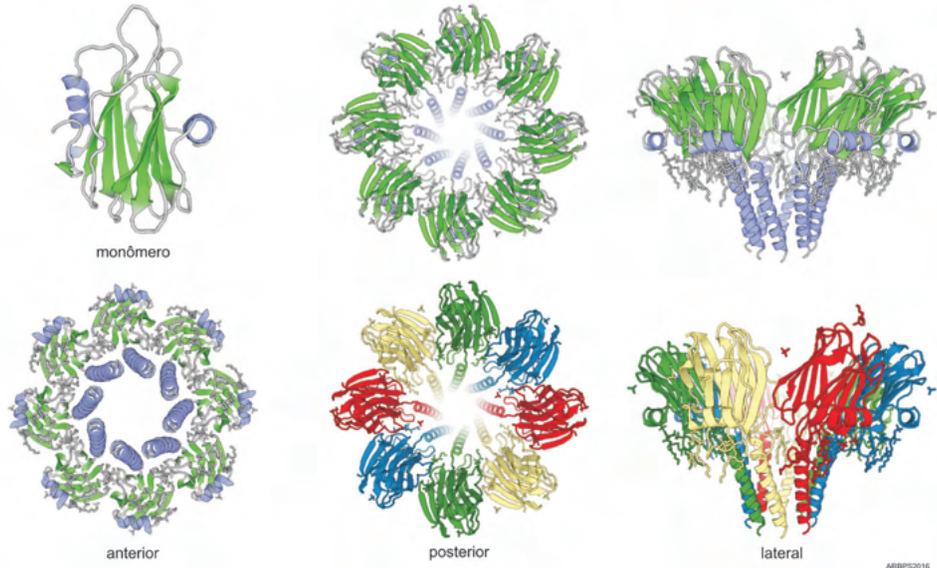


Figura 1:- Estrutura de uma Actinoporina. A figura mostra a estrutura tridimensional do monômero da Delta-Actitoxin-Afr1a (UniProtKB - B9W5G6) (PDB - 4WDC) em verde de β -sandwich, em roxo as α -hélices. Montagem da unidade biológica de simetria cíclica (C8) utilizou o modelo PDB - 4TSY. A figura mostra as vistas anterior, posterior e lateral. A α -hélice N-terminal atravessa a bicamada lipídica sendo mais bem percebida na vista anterior e lateral. Modelagem feita com o software Chimera (chimera-1.11.2) e editado pelo software Corel Drawn (versão X3)

A principal característica destas toxinas é que a partir da interação inicial do monômero com a membrana celular, ocorre um processo de automontagem em quatro estágios, até formar um poro hidrofílico seletivo para íons catiônicos que culmina na estimulação cardíaca ou hemólise. O primeiro estágio é chamado de estágio hidrossolúvel ou hidrofílico e corresponde à toxina em sua forma monomérica livre. O segundo estágio envolve o reconhecimento específico de esfingomiéline da membrana pelo monômero da toxina (mas não o colesterol, nem a fosfatidilcolina) através de vários sítios de ligação específicos, usando uma região rica em resíduos aromáticos e com colaboração de lipídios fosfocolina adjacente ao local de ligação na membrana, que ocorre principalmente através de interações hidrofóbicas. O terceiro estágio envolve uma oligomerização parcial (dimerização) acompanhada pela transferência da região N-terminal da toxina para a interface lípido-água até finalmente a total oligomerização resultando na imersão das α -hélices N-terminais na membrana envolvendo 8 subunidades da toxina (Figura 01) com a formação de um poro toroidal seletivo para cátions (22–25).

Interessantemente, as anêmonas do mar não sofrem dano, pois suas membranas celulares possuem um análogo fosfonoesfingolípido no lugar de esfingomiéline que é o receptor inicial de reconhecimento (26). Por sua atividade lítica as actinoporinas são toxinas

letais para crustáceos, moluscos e pequenos peixes. Também devido à sua atividade citolítica, as actinoporinas de anêmonas têm potencial de aplicação farmacológica na terapia de câncer, especialmente em associação com anticorpos específicos criando imunotoxinas capazes de lisar seletivamente células tumorais (27,28).

De forma geral as actinoporinas da subfamília das anêmonas são expressas na forma de um precursor, normalmente entre 212 aminoácidos de tamanho, composto por um peptídeo sinal de aproximadamente 19 aminoácidos, seguido por um pró-peptídeo de tamanho variável (16 a 20 aminoácidos) rico em aminoácidos ácidos correspondendo à sequência consenso de endereçamento para o nematocisto (29) incluindo a díade básica que indica o ponto de clivagem para a toxina madura com aproximadamente 179 aminoácidos. Os sítios de ligação com fosfocolina e com lipídios de membrana são identificados por ensaios de ligação e mutagênese-sítio dirigida também são mostrados em toxinas de referência no alinhamento da figura 03.

Toxinas bloqueadoras de canais de potássio depende de voltagem do tipo I – SHKT

Canais de potássio são uma parte muito importante para a vida, pois regulam o potencial de membrana celular dos seres vivos, portanto toxinas que atuam sobre esses canais são de grande interesse para a ciência. A primeira toxina bloqueadora de canais de potássio encontrada em anêmonas do mar foi a ShKT, da *Stichodactyla helianthus* [P29187] (30), formando uma nova família de toxinas bloqueadoras de canais de potássio do tipo I (sea anemone type 1 potassium channel toxin family). Essa toxina causou interesse pois bloqueava o canal Kv1.3, que estão envolvidos na proliferação de Linfócitos-T, tornando esse peptídeo um potencial imunossupressor (31,32). Também foi isolada da anêmona *Bunodosoma granulifera* uma toxina similar à ShKT, que recebeu o nome BgK (P29186) (33). As toxinas similares à ShKT receberam a classificação de toxinas bloqueadoras de canal de potássio do tipo 1a, enquanto as similares à BgK receberam a classificação tipo 1b.

dissulfeto (ShK 17-32, BgK 20-34) aproxima os dois loops que interconectam os trechos α -hélice, puxando-se em direção ao centro molecular conferindo uma forma globular às toxinas. Existem em anêmonas do mar outras toxinas bloqueadoras de canais de potássio dependentes de voltagem (k-neurotoxinas) importantes denominadas k-neurotoxinas tipo II ((1)), que apresentam um domínio Kunitz-type (p.ex. APEKTx1 [P86862] ou a-dendrotoxin [P00980]). Existem também k-neurotoxinas (BDS) tipo 3 de anêmonas do mar (37) que apresentam um domínio β -defensina símile totalmente diferente formado exclusivamente por folhas- β pregueadas estabilizadas por 3 pontes dissulfeto. E por fim também existem as k-neurotoxinas de escorpião CSa β (p.ex. charybdotoxin) que também tem 3 pontes dissulfeto (38), não é possível alinhar a sequência da ShK e da BgK com essas toxinas, pois a estrutura tridimensional é muito diferente entre estes domínios.

Atualmente existe uma superfamília de proteínas cuja principal característica é a presença do domínio ShKT, em decorrência da sua alta similaridade com as toxinas ShK e BgK. Existem hoje descritas 21 toxinas de anêmonas com domínio ShKT. Não existem outras neurotoxinas com esse domínio em outros grupos animais exceto as toxinas do grupo CRISPs (*Cysteine-rich secretory proteins*) (39), que são formadas por dois domínios ou um domínio SCP seguido pelo ShKT. Além disso ao domínio ShKT está presente em mais de 776 proteínas (35) listadas no UNIPROT incluindo não toxinas distribuídas entre diversas famílias (40) incluindo proteínas de adesão do tipo Cadherin-like, proteases e outras enzimas (p.ex Astacin-like metalloproteinase M12A) ou modulando fluxos de cálcio via receptor de ryanodina (RyR) (40). Não se conhece o papel fisiológico do domínio ShKT ou se há uma possível relação com a regulação de canais de potássio na célula.

É característica dessa família de k-neurotoxinas tipo I (Shk/Bgk) a presença de um par de aminoácidos geralmente aromático ou hidrofóbico/básico denominada díade funcional que nas k-neurotoxinas tipo I de anêmonas é formada pelos aminoácidos Lisina-Tirosina (KY), sendo essa díade responsável pela ação bloqueadora de canal de potássio dessas toxinas (35,41), embora atualmente seja conhecida a importância de outros aminoácidos da cadeia madura para a função bloqueadora não apenas do canal Kv1.3, mas também de outros canais.(1,42,43). Esse bloqueio funciona com a díade KY ocluindo a passagem de íons pelo poro do canal de potássio. A díade funcional está melhor representada na figura abaixo.

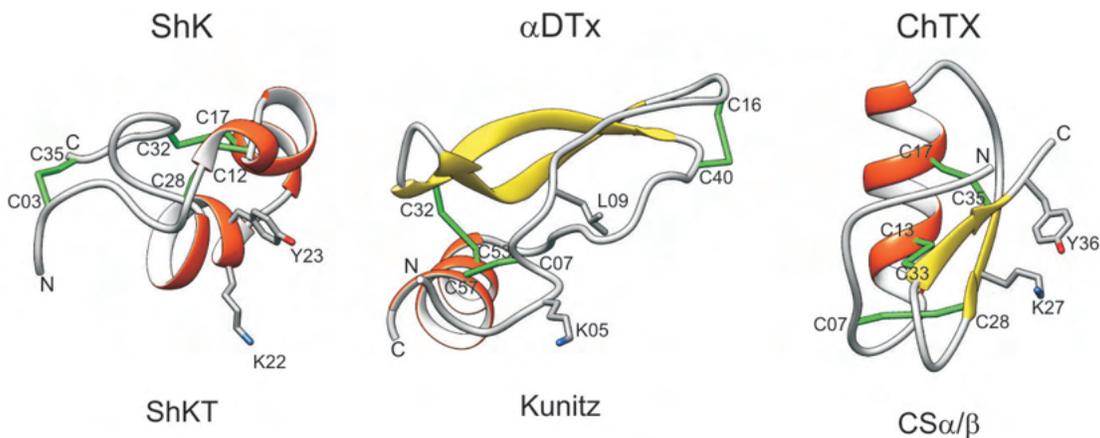


Figura 4: Díade funcional de toxinas bloqueadoras de canais de potássio dependentes de voltagem. A figura mostra o par de aminoácidos importantes para a oclusão do poro do canal presente em toxinas dominios com domínios estruturais diferentes. A figura mostra acima o nome da toxina e em baixo o domínio estrutural correspondente: ShK: Kappa-stichotoxin-She3a de *Stichodactyla helianthus* (P29187) domínio ShKT. α DTX: alfa-dendrotoxina de *Dendroaspis angusticeps* (P00980) domínio Kunitz. ChTX: Charybdotoxin de *Leiurus quinquestriatus hebraeus* (P13487) scorpion short toxin CS α / β . As cisteínas participantes de pontes dissulfeto estão numeradas segundo a posição e as pontes são mostradas em verde. A estrutura secundária da proteína está indicada em vermelho, as alfa-hélices e em amarelo as folhas-beta pragueadas. O N-terminal (N) e o C-terminal (C) de cada cadeia também são indicadas. A díade funcional é mostrada pelo par de resíduos evidenciados em cada molécula numerados segundo sua posição na cadeia.

O sitio de ligação para a família de canais de potássio dependente de voltagem (Kv1.x) da BgK de *B. granulifera* está bem estabelecido e diversos aminoácidos além da díade funcional contribuem para a constante de dissociação e para a seletividade: Kv1.1 (Kd=6 nM), Kv1.2 (Kd=15 nM), Kv1.3 (Kd=10 nM) (44,45). Além da díade funcional alguns aminoácidos discriminam entre os canais Kv1.1, Kv1.2 e Kv1.3 sendo que o núcleo de ligação essencial para o bloqueio do canal é composto pelos aminoácidos S23; K25, Y26 que, portanto, inclui a díade funcional, como se pode ver na figura abaixo:

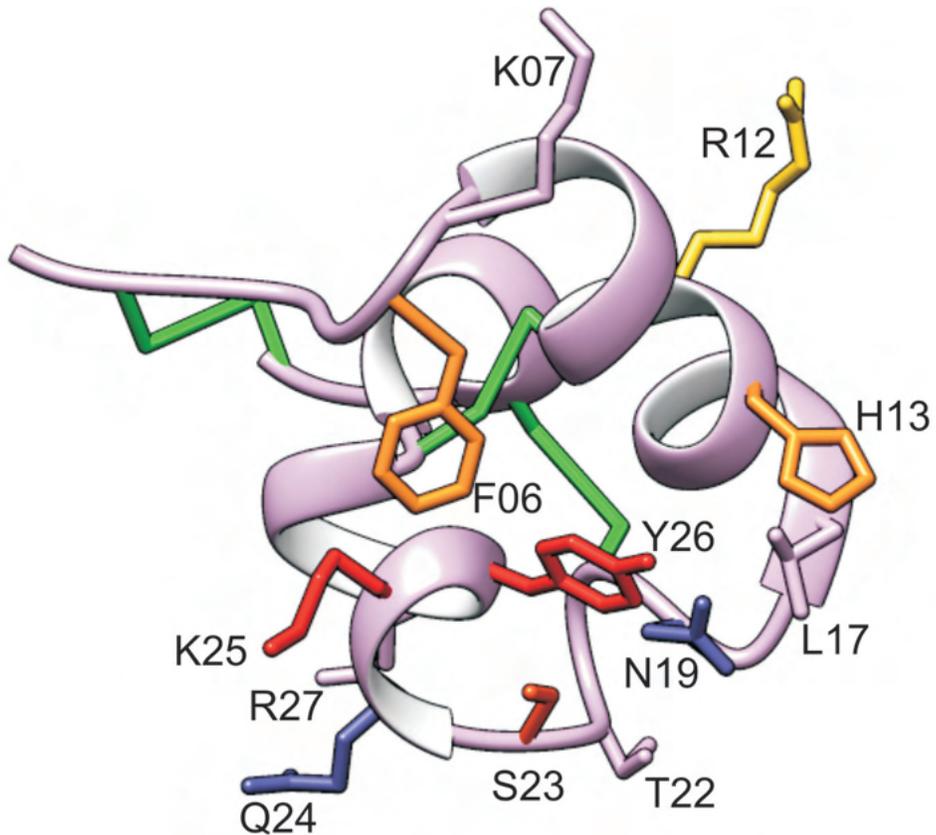


Figura 5: Sítio de ligação ao canal de potássio família Kv1.x da BgK. A figura mostra os aminoácidos da toxina que contribuem para a ligação e seletividade pelo canal de potássio. Os aminoácidos importantes para o bloqueio, incluindo a díade funcional e compartilhados pelos canais Kv1.3, Kv1.2 e Kv1.1 são mostrados em vermelho (S23; K25; Y26). Aminoácidos que contribuem exclusivamente para a ligação ao Kv1.3 são mostrados em rosa (K07; L17; T22; R27). Resíduos compartilhados entre o Kv1.3 e Kv1.1 são mostrados em roxo (N19, Q24). Resíduos importantes para a ligação aos canis Kv1.3 e Kv1.2 são mostrados em laranja (F06 e H13) e por fim o aminoácido R12 mostrado em amarelo só é importante para ligação ao Kv1.2. As pontes dissulfeto estão mostradas em verde.

Os resíduos mais importantes da BgK para o bloqueio do canal Kv1.3 ordenados de acordo com a potência de bloqueio são: K25(IC₅₀>300)>F06(IC₅₀=167)>Y26(IC₅₀=100)>S23(IC₅₀=30)>H13(IC₅₀=20) segundo determinado por Dauplais (35).

A ShK também usa um núcleo de resíduos similares à BgK mas alguns são específicos lembrando que as duas moléculas tem marcantes diferenças estruturais devido à inserção de resíduos extras na BgK após a posição 10. Os resíduos mais importantes para o bloqueio de Kv1.3 são K07, L17, T22, R27 são importantes para ligação a Kv1.3 (8,32) como mostrado na imagem abaixo:

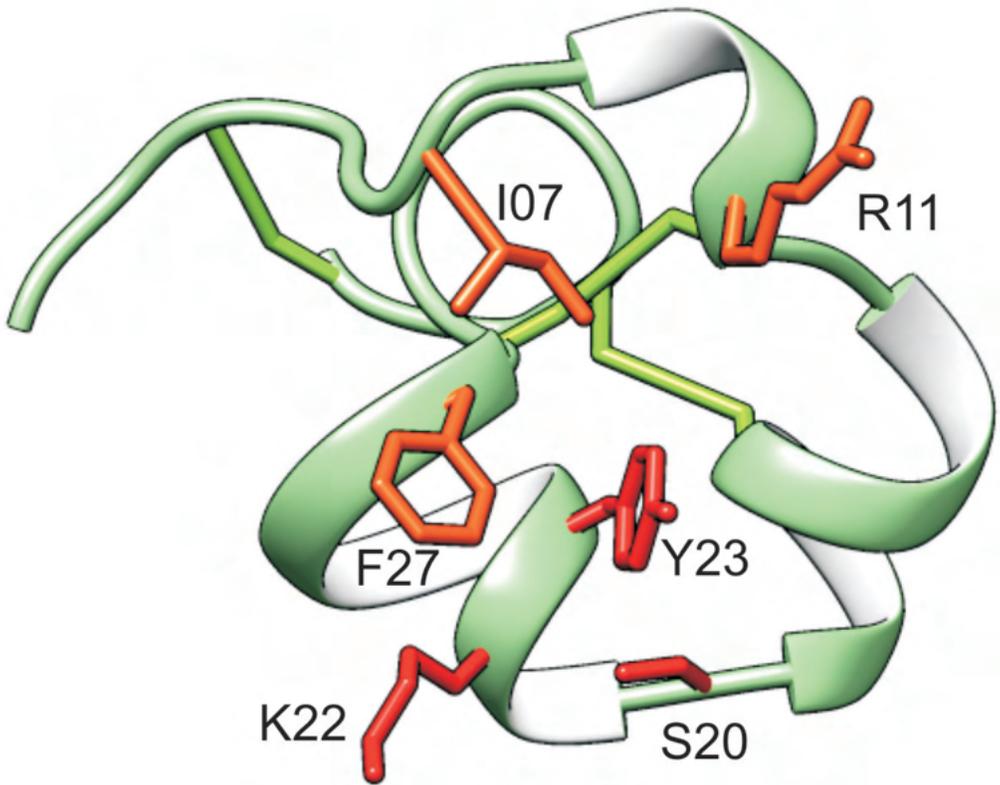


Figura 6: Sítio de ligação da ShK ao canal Kv1.3. São mostrados os aminoácidos importantes para o bloqueio do canal I07, R11, S20, F27 em laranja e a díade funcional K22, Y23 é mostrada em vermelho. As pontes dissulfeto são mostradas em verde.

Para o nosso trabalho demos ênfase apenas aos contigs que apresentaram um domínio ShK único, tal como ocorre nas k-neurotoxinas do tipo I. Entretanto, este domínio é encontrado em diversas proteínas no reino animal, inclusive algumas toxinas, como é o caso das CRISPs, que constituem de um domínio CAP e outro domínio ShK. Casos de toxinas que apresentam outros domínios junto com a ShK serão discutidos em capítulos específicos (46)..

MATERIAL E MÉTODOS

ORIGEM DAS AMOSTRAS

Os exemplares foram coletados em diferentes locais e datas. Os espécimes da *Bunodosoma caissarum* da costa de São Paulo foram coletados na latitude -23.72622967 e longitude -45.41930974 em abril de 2014 e os exemplares da *Bunodosoma caissarum* de origem do Arquipélago São Pedro e São Paulo foram coletadas na latitude 0.9168648 e longitude -29.3452703, no Oceano Atlântico, em agosto de 2007.

PREPARO DO TECIDO

O material utilizado para os experimentos foi o tecido dos tentáculos das anêmonas das duas populações, devido a sua maior concentração de nematocistos. Foi extraído cerca de 200mg de tecido, o qual foi embebido de RNA Later (Ambion, EUA) e armazenado em 4°C por 24 horas. Após esse período foi dispensado o RNA Later e o material congelado em -20°C e depois de 24 horas em -80°C, de modo a se escalonar gradativamente a temperatura e tornar mais efetiva a preservação do tecido e evitar a formação de cristais intracelularmente, o que pode causar extravasamento de conteúdo citoplasmático e perda de ácidos nucleicos, como sugerido no manual do fabricante.

EXTRAÇÃO DE RNA TOTAL

Após o preparo inicial das amostras iniciamos o processo de extração de RNA dos tecidos através do protocolo que utiliza o TRIzol (Invitrogen, EUA) como reagente, seguindo o protocolo padrão do fabricante. Foram utilizados 3 mL do reagente para os 200mg de tecido, que foi homogeneizado totalmente em um aparelho Polytron Homogenizer PT 3100, utilizando por volta de 10.000 rotações por minuto, por 1 minuto. Por fim o total extraído foi eluído em 50 μ L de água e mensurado através de absorbância de 260nm e 280nm, em um microespectrofotômetro modelo Nanodrop (Thermo Fisher Scientific, EUA).

CONSTRUÇÃO DAS BIBLIOTECAS DE cDNA

Purificação do mRNA

Para a extração de mRNA foi usado o kit Dynabeads (Life Technologies, EUA), usando o seu protocolo padrão para extração de mRNA a partir de RNA total. O protocolo consiste em adicionar esferas magnéticas (beads) que possuem cauda poli-T covalentemente ligadas, no qual o mRNA se ligará por complementaridade através do poli-A e será separado do meio por ação magnética. Após esse processo as esferas são lavadas duas vezes com o tampão de lavagem próprio do kit, sendo repetido o processo de atração

magnética das esferas em cada vez. Ao final é adicionado H₂O Molecular Grade (Sigma) ao meio e em seguida aquecido à 65°C para a desnaturação do mRNA das sondas poli-T, o que possibilita que o mRNA seja recuperado junto com o sobrenadante e adicionado a outro tubo livre de RNases.

Preparo do mRNA

O mRNA resultante foi mensurado através do método de absorvância de 260nm e 280nm e também por eletroforese automatizada e por eletroferograma pelo sistema Bioanalyzer 2100 (Agilent Technologies, EUA). As amostras foram armazenadas em -80°C para uso posterior. O material a ser utilizado foi processado seguindo-se o protocolo recomendado pela Illumina (EUA), através do kit TruSeq RNA Sample Preparation (Illumina, EUA). Para o preparo das amostras foi efetuada uma segunda purificação de mRNA das mesmas. O produto final contendo apenas o mRNA teve incorporado uma mistura de reagentes da fabricante que tem como função eluir, fragmentar e incorporar nos fragmentos de mRNA iniciadores aleatórios (random primers).

Construção da biblioteca de cDNA

Em seguida as amostras passaram pelo processo de síntese da primeira fita de cDNA pelo método de transcrição reversa, na qual os iniciadores amplificam a primeira fita de cDNA. A seguir realizamos o processo de síntese da segunda fita do cDNA, o que elimina o mRNA molde da síntese da primeira fita do cDNA das amostras. O procedimento seguinte consiste em reparar as extremidades não coesivas que o cDNA apresenta. Isso permite a equalização do comprimento nas duas fitas do cDNA para logo a seguir ter incorporado no final 3' destas um nucleotídeo "A", que tem como função prevenir que esses fragmentos de cDNA se liguem a outros durante a reação de ligação dos adaptadores, cujo objetivo é incorporar sequências de nucleotídeos pré-definidas no final dos fragmentos gerados. Esses adaptadores têm como objetivo criar identificadores para cada amostra, pois os cDNAs independentes de cada uma terão o seu adaptador definido, possibilitando que o sequenciamento seja realizado por multiplexação (multiplexing), aplicando diferentes amostras simultaneamente ao aparelho.

Por fim os fragmentos de cDNA contendo os adaptadores são amplificados através da reação em cadeia da polimerase (PCR) contendo iniciadores que são específicos aos adaptadores, possibilitando um enriquecimento da quantidade total dos fragmentos. A qualidade das bibliotecas foi aferida pelo sistema de eletroforese capilar em Bioanalyzer (Agilent Technologies, EUA) e quantificação em PCR em tempo real (Kapa Sybr Library Quantification kit, KapaBiosystems).

SEQUENCIAMENTO

Após a preparação das bibliotecas, os cDNAs são adicionados à célula de fluxo do aparelho (flow cell), onde ocorrerá o processo chamado agrupamento (clusterization), no qual cada fragmento é isotermicamente amplificado. Após a geração de grupos de cDNAs (cluster generation) ocorrerá o sequenciamento das amostras, na mesma célula de fluxo. Tal célula é uma placa que tem incorporado em toda a sua extensão dois tipos de adaptadores de nucleotídeos complementares aos adaptadores incluídos os fragmentos de cDNA. Os adaptadores incorporados aos fragmentos de cDNA são ligados por complementaridade a outros adaptadores que já se encontram na célula de fluxo. Quando as amostras são injetadas no sistema os fragmentos passam por um processo de desnaturação, separando suas duas fitas, o que possibilita que um dos adaptadores das cadeias possa se ligar aos adaptadores complementares da célula de fluxo, permitindo que através de uma reação em cadeia da polimerase possa criar uma fita complementar à fita original, incorporada a célula de fluxo. Terminando o processo de polimerização o fragmento passa por uma desnaturação e o molde original da fita é lavado e excluído do processo. Em seguida a fita ligada tem o seu outro adaptador, na sua outra extremidade, também ligado por complementaridade a outro adaptador na célula de fluxo, formando assim um tipo de ponte com o fragmento de cDNA, que passam por outro processo de amplificação por polimerização seguida de outra desnaturação, criando uma série de aglomerados de fragmentos semelhantes a si mesmos. A este processo de amplificação é dado o nome de amplificação em ponte (bridge amplification).

O nosso sequenciamento utilizou-se do método de “paired-end”, o qual tem como objetivo sequenciar os dois sentidos da dupla fita do cDNA das amostras, para maior abrangência e segurança. O aparelho utilizado para o sequenciamento foi o HiSeq 1500 (Illumina, EUA) e a corrida foi realizada em 300 ciclos de 150 para cada fita, levando cerca de 72 horas para a sua finalização.

Brevemente, após o agrupamento dos fragmentos começa o sequenciamento propriamente dito, que ocorre através da síntese da cadeia de DNA complementar à cadeia molde acoplada na célula de fluxo. O processo ocorre por inserção das quatro bases nucleotídicas (adenina, guanina, citosina e timina), sendo que cada uma delas é identificada por um marcador fluorescente que será excitado através de um laser. Com a fluorescência gerada o sistema consegue identificar qual base foi incorporada simultaneamente em cada um dos cDNAs gerados e agrupados na célula de fluxo, durante o sequenciamento.

PROCESSAMENTOS DO SEQUENCIAMENTO, MONTAGEM DAS LEITURAS E ANÁLISES COMPUTACIONAIS

Processamento

Finalizado o sequenciamento os dados brutos foram processados através de uma série de programas, cujo objetivo foi filtrar os dados brutos para que a montagem dos transcritos fosse executada apenas com reads de alta qualidade. Dando início às análises de bioinformática, os arquivos fastq foram gerados para cada amostra utilizando o software CASAVA (versão 1.8.2), contendo apenas sequências aprovadas no controle de qualidade de bases específico de índice Q maior que 30. Posteriormente foi utilizado um pipeline para análise das sequências, e pré-processamento dos dados brutos, com filtro das sequências por qualidade, tamanho da leitura, sequências com regiões de homopolímeros, e de baixa complexidade, caudas poli A/T/N e adaptadores, e pontas de baixa qualidade, com os softwares fastq-mcf, versão 1.04.662 e bowtie2, versão 2.2.5 (47). Os critérios empregados para a filtragem foram: regiões de homopolímeros e baixa complexidade acima de 90% da sequência foram eliminadas, e mantidas regiões com média de qualidade acima de 25, eliminando aquelas de baixa qualidade, e após a remoção das pontas de baixa qualidade, tamanho mínimo das sequências de 40 pares de bases.

Montagem das leituras

Após o pré-processamento e aplicação de filtros para obtenção de sequências de alta qualidade, iniciou-se a montagem feita através do pacote de programas chamado Trinity (versão: r20140413p1, de 13/04/2014). O Trinity é a ferramenta padrão para montagem de novo de sequências de transcriptomas que não tenham genoma de referência. Brevemente, o Trinity utiliza-se de uma combinação de três algoritmos para a sua função: Inchworm, Chrysalis e Butterfly. No primeiro passo, o Inchworm monta as leituras em sequências únicas dos transcritos, usando um “algoritmo guloso” para tal. Esse algoritmo utiliza o algoritmo guloso, baseado no k-mer para uma montagem mais rápida e eficiente, utilizando apenas um único (melhor) representante, representando um conjunto de variantes, ou seja, o programa segue montando os transcritos seguindo as leituras que melhor se enquadram para a construção daquele transcrito dentro do k-mer estabelecido. Em seguida, o Chrysalis, a partir dos dados gerados no Inchworm, cria uma série de aglomerados de transcritos semelhantes entre si, para então utilizar o algoritmo de grafos de Bruijn para cada aglomerado, cada um refletindo a complexidade das repetições entre as variações das sequências consenso. Por fim o Butterfly analisa os caminhos percorridos pelas leituras e o seu pareamento no contexto dos consensos para determinar quais sequências são plausíveis. Através desse programa os dados obtidos foram agrupados em milhares de sequências que se sobrepõe, formando sequências contíguas.

Nessa etapa foi definido que usaríamos os resultados do sequenciamento das

duas espécies de *Bunodosoma caissarum* como uma única espécie, pois é esperado que boa parte dos seus genes sejam compartilhados entre elas, sendo apenas alguns transcritos específicos, portanto o Trinity acaba por considerar transcritos com alta homologia entre as espécies como um único transcrito (contig). Não foi utilizado nenhum genoma de referência para a montagem.

Análises computacionais

Esses contigs foram analisados e anotados através de softwares do pacote Blastall, softwares esses que foram usados para fazer uma localização de sequências que compartilham alguma similaridade entre o banco de dados escolhido e os contigs sequenciados. Como procedimento padrão fizemos uma busca local contra o banco de dados geral do NCBI (NR) e do Projeto de Anotação de Toxinas Animais, do Uniprot (<http://www.uniprot.org/program/Toxins>). Para essas buscas utilizamos o critério de corte com “e-value” no valor de $1e-06$, que restringe os resultados para sequências mais similares às encontradas nos bancos de dados. O “e-value” é um parâmetro que descreve a chance de encontrarmos aleatoriamente resultados de uma busca de uma determinada sequência dentro de um banco de dados, ou seja, quanto menor o resultado desse valor menor a chance de ter ocorrido por coincidência.

Resumidamente, as etapas da análise de expressão envolveram o mapeamento dos reads nas referências com bowtie2 (47), em seguida, a estimativa do nível de abundância foi feita a partir das contagens de leituras por transcrito, normalizando-se em função de seus tamanhos, gerando um índice de expressão FPKM (Fragments Per Kilobase of exon model per Million mapped reads) com o programa RSEM (RNA-Seq by Expectation Maximization) que não requer um genoma de referência (Li e Dewey, 2011). A normalização do nível de expressão entre as amostras foi feita com o método TMM (Trimmed Mean of M-values) (48) e a análise de expressão diferencial par a par entre todas as amostras foi feita com o software edgeR (49). Para estudar os padrões de expressão dos transcritos entre as amostras, foram utilizados os transcritos diferencialmente expressos normalizados obtidos com edgeR, que apresentaram diferença de intensidade de expressão entre cada par de amostras de pelo menos 4x mais, com significância estatística corrigida para falsa descoberta de $p=0.001$.

A fim de minerarmos a maior quantidade possível de informações sobre as sequências por nós obtidas efetuamos um “BLAST” contra todas as toxinas conhecidas de anêmonas do mar, usando os mesmos parâmetros utilizado no “BLAST” contra o banco NR do NCBI e o Uniprot. Para tal criamos um banco de dados contendo todas as toxinas anotadas do filo cnidária presentes no Uniprot para a execução do algoritmo.

O último BLAST que fizemos foi contra famílias individuais de toxinas, famílias essas selecionadas conforme a sua importância segundo nosso conhecimento atual em toxinas

de cnidários e outros animais, para com isso maximizar a eficiência da nossa mineração de dados contra nosso sequenciamento.

RESULTADOS E DISCUSSÃO

EXTRAÇÃO DE RNA

A seguir apresentamos a tabela com os resultados da extração de RNA total correspondentes aos 200mg de tecido dos tentáculos das anêmonas. O RNA total obtido foi diluído em 50 μ l de água livre de RNAses.

População	Concentração de RNA (ng/ μ l)	Abs260/280
<i>Bcs-SP</i>	1623	2,03
<i>Bcs-Aq</i>	1560	1,95

Tabela 1: Resultados da absorvância do RNA total em comprimentos de ondas de 260 e 280nm.

A Coluna Concentração de RNA mostra os valores obtidos na extração em valores de ng por μ l. Abs 260/280 indica o cálculo efetuado para estimar a possível contaminação por proteínas no RNA total extraído do tecido das anêmonas através da absorvância de luz no comprimento de onda de 260nm e 280nm.

Nossa extração de RNA total conseguiu um total de 1,6 μ g/ μ l para a população Bcs-SP e 1,5 μ g/ μ l para a população Bcs-Aq, consideramos valores ideais para a seguinte purificação do mRNA do RNA total. Os valores de 2,03 para Bcs-SP e 1,95 para Bcs-Aq nos valores de absorvância 260/280 indica que conseguimos extrair RNA puro das anêmonas, sem contaminação de DNA ou proteínas.

PURIFICAÇÃO DO MRNA

Abaixo segue a tabela com os resultados da purificação de mRNA, mensurados pelo método de absorvância em comprimentos de ondas de 260 e 280nm:

Espécie	Concentração de mRNA (ng/ μ l)	Volume final (μ l)	Quantidade final (ng)
<i>Bcs-SP</i>	40,7	50	2035
<i>Bcs-Aq</i>	36,4	50	1820

Tabela 2: Resultados da purificação do mRNA, mensurados por absorvância de ondas 260nm e 280nm

A Coluna Concentração de RNA mostra os valores obtidos na extração em valores de ng por μ l. Volume final representa o volume em que o material foi diluído. Quantidade final é o quanto de material purificado foi conseguido, em ng.

Também utilizamos da eletroforese automatizada, feita pelo aparelho Bioanalyser

(Agilent Technologies, EUA), para validação dos resultados da nossa purificação de mRNA das amostras. Para tal foi utilizado o “chip” próprio do sistema para RNA. Através desse procedimento pudemos verificar a distribuição dos fragmentos de mRNA conforme a sua faixa de peso molecular. A imagem abaixo é o resultado da eletroforese automatizada do RNA mensageiro.

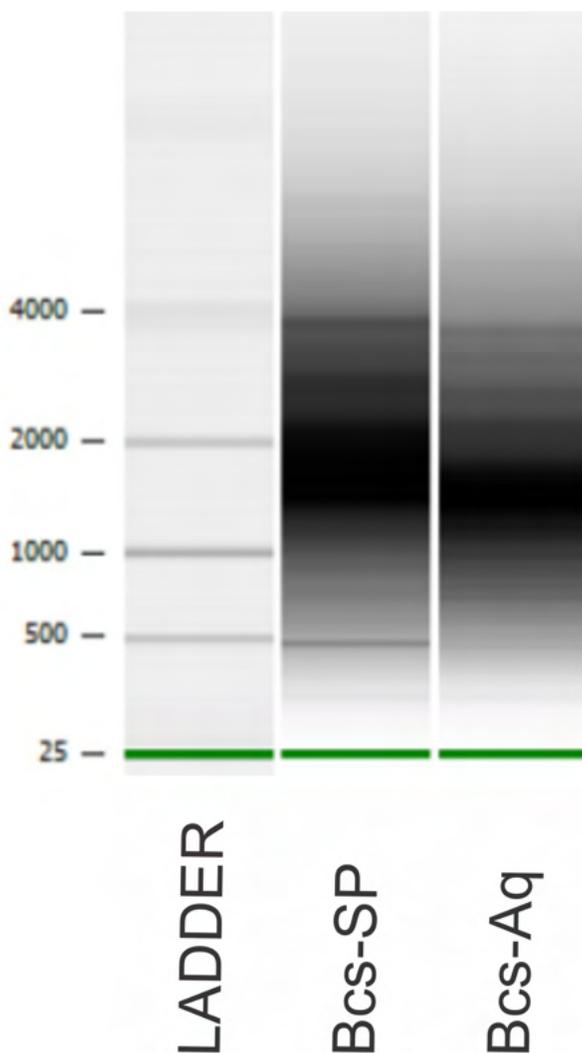


Figura 7: Imagem do resultado da eletroforese automatizada do RNA mensageiro das 2 amostras das populações de *Bunodosoma caissarum*. O marcador (*ladder*) para o “chip” específico do aparelho para RNA contém fragmentos de tamanho de 500, 1000, 2000 e 4000 pares de bases. A linha verde representa a normalização do sistema para o Ladder, Bcs-SP e Bcs-Aq em fragmentos de 25 bases.

Na figura 7 podemos observar que na amostra da população do litoral de São Paulo obtivemos maior concentração de fragmentos entre 1500 e 2100 de bases de mRNA, enquanto que na população do arquipélago obtivemos maior concentração de fragmentos por volta de 1400 e 1800 bases de mRNA. Valores levemente menores para a Bcs-Aq mas ainda assim os resultados são consistentes entre as populações.

Para melhor visualização da distribuição dos fragmentos no mRNA obtido das populações e sua qualidade fizemos um eletroferograma do material, cujo resultado se encontra abaixo.

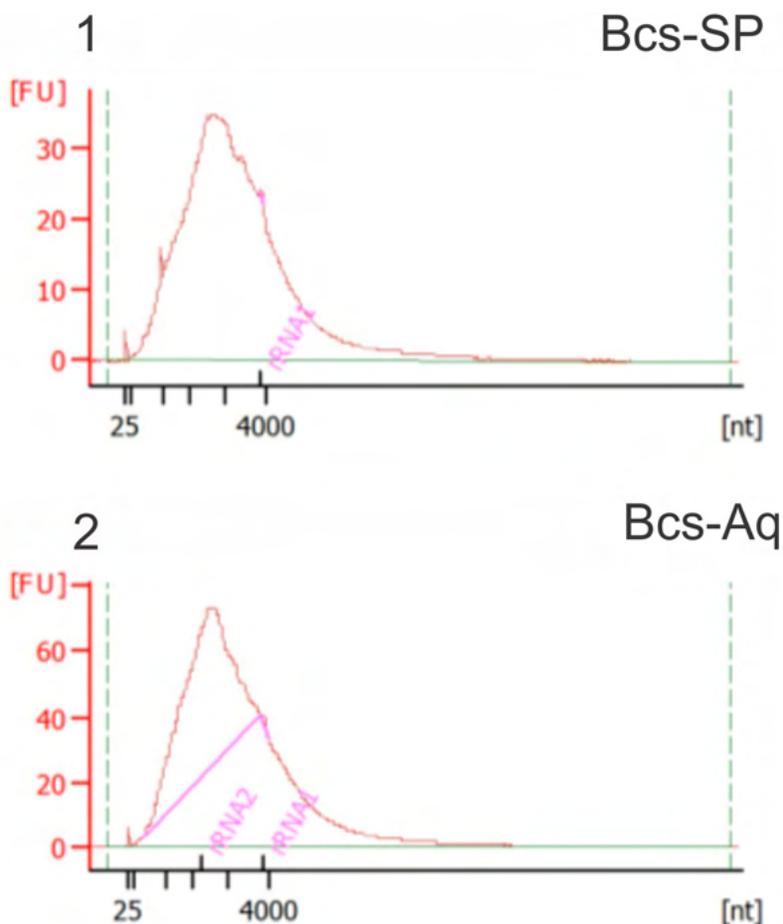


Figura 8: Eletroferograma das amostras de mRNA. FU, em vermelho é uma sigla para *Fluorescent Units*, unidades de fluorescência utilizadas pelo sistema. As linhas em magenta simbolizam a área delimitada pela quantidade de contaminação de rRNA na amostra de mRNA. A numeração no eixo inferior representa o número de bases nucleotídicas dos fragmentos. Os detalhes em magenta das imagens representam linhas, nas quais as áreas delimitadas pelas mesmas correspondem a contaminação de RNA ribossomal ainda presente nas amostras. O número 1 indica o gráfico correspondente a população Bcs-SP e o número 2 corresponde a população Bcs-Aq.

Na figura apresentada do eletroferograma observamos uma área magenta maior na Bcs-Aq, enquanto que na amostra Bcs-SP o tamanho dessa área magenta foi praticamente inexistente. Isso demonstra que a amostra que representa maior nível de contaminação com o rRNA foi Bcs-Aq.

SÍNTESE DO CDNA E VALIDAÇÃO DAS BIBLIOTECAS

O passo seguinte do trabalho foi fazer uma nova purificação do mRNA das amostras, seguindo o protocolo da Illumina, e a construção da biblioteca de cDNA para o sequenciamento. Após a construção da biblioteca foi feita uma nova validação pelo sistema de eletroforese automatizada (Figura 9) e eletroferograma (Figuras 10), com a diferença que para esse procedimento utilizamos o “chip” para DNA do aparelho Bioanalyser (Agilent Technologies, EUA).

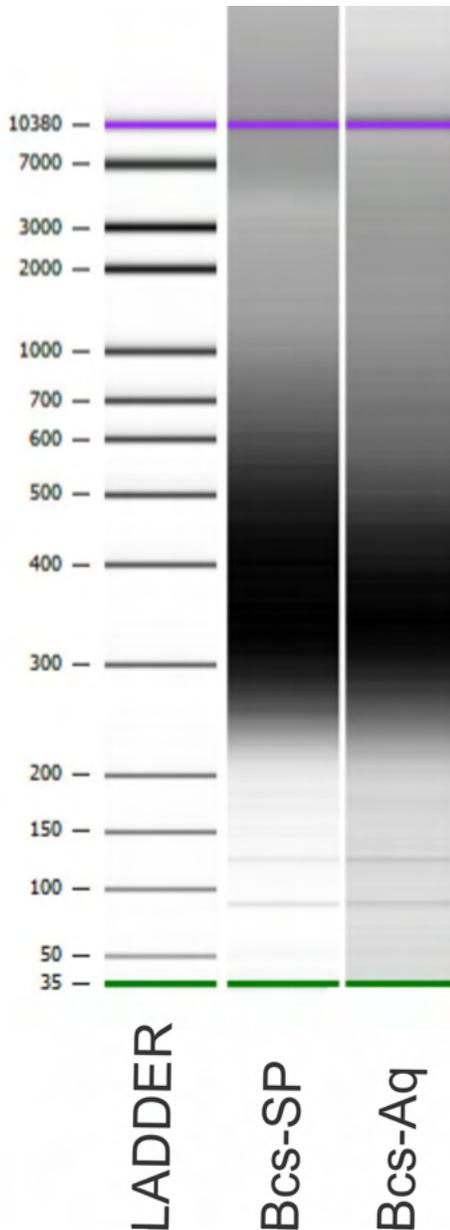


Figura 9: Imagem do resultado da eletroforese automatizada da biblioteca de cDNA das 2 amostras das populações de *Bunodosoma caissarum*. O marcador (*ladder*) para o “chip” específico do aparelho para DNA contém fragmentos de tamanhos entre 50, 7000 ses. A numeração verde na base representa fragmentos de 35 pares de bases incorporado às amostras, para fins de normalização dos fragmentos, assim como a numeração roxa representa fragmentos de 10380 pares de bases incorporado às amostras, para fins de normalização.

Na figura 9 podemos observar que os picos de fragmentos nas duas amostras ficaram semelhantes entre si, com a maior concentração de fragmentos variando na faixa

entre 400 e 500 pares de bases, demonstrando a eficiência na construção da biblioteca de cDNA para o sequenciamento. Também é possível constatar uma maior concentração de fragmentos menores na Bcs-Aq do que em comparação a Bcs-SP, o que não inviabiliza o sequenciamento das amostras.

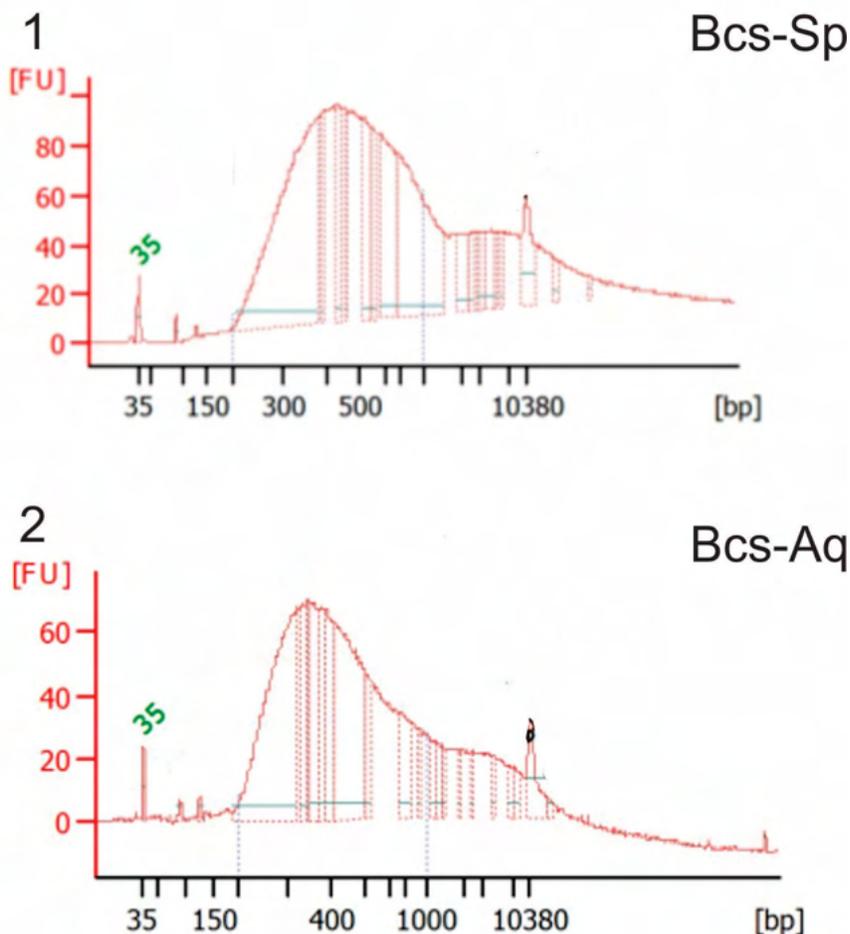


Figura 10: Eletroferograma das bibliotecas de cDNA das amostras. FU é sigla para *Fluorescent Units* e bp para pares de bases. Os números acima das curvas mostram os tamanhos em pares de bases dos fragmentos de cDNA. O pico verde na base representa fragmentos de 35 pares de bases incorporado às amostras, para fins de normalização dos fragmentos, assim como o ultimo pico representa fragmentos de 10380 pares de bases incorporado às amostras, para fins de normalização.

No eletroferograma confirmamos o que observado na eletroforese automatizada apresentado na figura n, com a maior parte dos fragmentos de cDNA presentes na faixa de 400 e 500 pares de bases nas duas populações de BCS-Sp e entre 300 e 400 pares de

bases na população Bcs-Aq.

SEQUENCIAMENTO

Após a construção das bibliotecas o material foi sequenciado no aparelho Illumina Hi-Seq 1500, processo que demorou 72 horas. Com o sequenciamento finalizado os dados brutos gerados foram convertidos na forma de arquivos com a extensão “fastq”. Tal tipo de extensão fastq é um formato de conversão de dados gerados pelos sequenciamentos efetuados pelo sistema Illumina™ e que é lido pela maioria dos algoritmos de análise e processamento de dados gerados por plataformas de sequenciamento de nova geração. Na tabela 3 apresentamos os resultados iniciais do sequenciamento, com o valor total de leituras obtidas em cada amostra e os valores dos dados que foram filtrados pelo pré-processamento dos mesmos:

Bcs-SP	Resultados	%
Leituras totais	120,537,468	100
Total das leituras boas	117,224,693	97.25
Total das leituras filtradas	3,312,775	2.75
• Leitura muito curta	2,816,563	2.34
• Filtrada na qualidade	451,625	0.37
• Filtrada no homopolímero	42,545	0.04
• Filtrada na baixa complexidade	2,042	0.00
Leituras aparadas	14,143,715	11.73
Bcs-Aq		
Leituras totais	44,790,462	100
Total das leituras boas	43,914,424	98.04
Total das leituras filtradas	876,038	1.96
• Leitura muito curta	732,191	1.63
• Filtrada na qualidade	119,315	0.27
• Filtrada no homopolímero	23,181	0.05
• Filtrada na baixa complexidade	1,351	0.00
Leituras aparadas	4,326,737	9.66

Tabela 3: Resultados iniciais do sequenciamento das amostras

O total das leituras foi o total de leitura que foram obtidas no sequenciamento das populações de *Bunodosoma caissarum*. Desse total foram excluídas as leituras que não cumpriram o mínimo esperado de qualidade. Leituras muito curtas: leituras filtradas devido ao seu tamanho reduzido. Filtrada na qualidade: leituras que tiveram baixa qualidade no

sinal durante o sequenciamento. Filtrada no homopolímero: leituras que foram filtradas por causa da presença predominante de homopolímero na sua sequência. Filtrada na baixa complexidade: leituras filtradas devido à baixa complexidade das suas bases nucleotídicas, como predominância de repetições ao longo de toda leitura. As leituras aparadas foram aquelas que tiveram diferenças no tamanho das leituras forward e reverse no sequenciamento, sendo assim necessária o procedimento de aparar as pontas para que tenham leituras compatíveis no tamanho.

Conforme visto na tabela anterior conseguimos uma grande quantidade de leituras boas em relação ao número de leituras totais, onde tivemos um aproveitamento de mais de 97% em cada amostra. Porém é importante destacar que embora as amostras tenham sido preparadas de formas idênticas em todo o protocolo os números dos resultados de leituras totais nas duas populações possuem grande diferença, por volta de 3 vezes mais leituras para a Bcs-SP do que em comparação a Bcs-Aq, mas a qualidade das leituras aproveitadas foi muito similar, por volta de 98%. Isso pode ocorrer devido a qualidade do mRNA que foi purificado das amostras, p.ex. a amostra de RNA de Bcs-Aq apresenta maior conteúdo de rRNA. O RNA é muito susceptível a degradações pela sua própria característica de fita simples mesmo estando embebido em solução preservante de *RNAlater* (Ambion, EUA). Ademais, a amostra Bcs-Aq foi obtida a partir de tecidos que estavam preservados desde 2007, sendo suscetível a degradação ao longo do tempo.

Após essa filtragem inicial das leituras obtidas no sequenciamento os dados foram processados no software Trinity, cujos detalhes do seu funcionamento se encontram no Material e Métodos (item 3.6.2). Os resultados dessa montagem se encontra na tabela 4 a seguir.

Leituras	Transcritos	Genes	Tamanho médio	Ribossomal
161.139.117	199.517	124.790	950,22	13

Tabela 4: Estatística preliminar da montagem das leituras do sequenciamento pelo pacote Trinity. As duas populações foram consideradas como uma única população

A coluna Leituras corresponde ao total de leituras que foi processado pelo Trinity. Transcritos foi o total de sequencias que o programa conseguiu montar. Genes é o número de genes individuais que o software conseguiu identificar. Tamanho médio representa o tamanho médio de todos os transcritos montados. Ribossomal é a quantidade de transcritos ribossomais identificados após a montagem.

Na tabela 4 a seguir estão os resultados da montagem do sequenciamento do mRNA das anêmonas, que foi feito com o programa Trinity, detalhes do seu funcionamento se encontram no Material e Métodos. Nessa parte do processamento as duas populações

de BCS do litoral sudeste de São Paulo e a do arquipélago foram consideradas como uma, para fins de análises, já que como as duas populações são da mesma espécie então esperávamos encontrar grande parte do seu transcriptoma idêntico entre si, portanto seria mais sensato considerar as duas populações como uma única amostra. Em uma etapa posterior das análises será feito um levantamento da expressão gênica de cada transcrito identificado nas amostras e, com isso, teremos os contíguos expressos diferentemente em cada dado amostral da espécie *Bunodosoma caissarum*:

A quantidade de transcritos que o programa Trinity resultou para as duas populações foi de quase 200 mil. Mas é observado uma diferença na contagem entre as colunas de transcritos e genes. Essa diferença ocorre porque o programa identifica em suas análises que alguns genes possuem diferentes isoformas entre si, ou seja, são sequências semelhantes entre si que o algoritmo classifica como sendo um mesmo gene, mas em diferentes isoformas.

Na coluna Ribossomal temos a informação de 13 genes ribossomais identificados após a montagem das leituras. Isso demonstra que mesmo após nosso procedimento de purificação do mRNA, afim de excluirmos todo o restante do material gênico ainda ficamos com transcritos ribossomais na nossa amostra.

ANÁLISE DO SEQUENCIAMENTO

O passo seguinte do trabalho foi analisar nossas sequências em busca por similaridades em bancos de dados com informações proteicas, tanto de componentes celulares quanto de toxinas. Nossa primeira opção foi o banco de dados de proteínas Non-Redundant (NR), do NCBI, que inclui dados dos bancos GenBank, PDB, SwissProt, PIR e PRF. Essa busca foi feita através do software BLAST, como detalhado no Material e Métodos. Nessa parte do trabalho as análises de expressão das sequências adquiridas foram separadas entre as duas populações. Esse método de montagem como uma única população e analisar a expressão de cada sequências entre as duas populações permite identificarmos os transcritos comuns entre as populações, com as diferentes expressões entre elas. Essa primeira análise gerou a tabela abaixo.

Categorias	Contigs	FPKM Bcs-SP	% Bcs-SP	FPKM Bcs-Aq	% Bcs-Aq
Toxinas	315	14756,52	1,63	57966,98	4,65
Não-caracterizadas	912	2567,1	0,28	4434,96	0,36
Preditas	43098	242548,63	26,77	253803,7	20,37
Sem resultado	134010	542761,28	59,91	809432,21	64,97
Celulares	13872	74868,12	8,26	80157,07	6,43
Hipotéticas	7310	28525,81	3,15	40151,23	3,22
Total	199517	906027,46	100	1245946,15	100

Tabela 5. Resultados do BLAST contra os transcritos, divididos em diferentes categorias.

As categorias estão divididas entre: Toxinas – as possíveis toxinas; Não-caracterizadas – proteínas identificadas presentes no banco de dados, mas sem caracterização; Preditas – proteínas preditas, sem maiores informações; Sem resultado – transcritos que não tiveram resultado no BLAST; Celulares – transcritos identificados como possíveis proteínas celulares. Hipotéticas – proteínas anotadas como hipotéticas nos bancos de dados. Coluna Contigs mostra a quantidade de transcritos identificado em cada categoria. FPKM Bcs-SP representa o nível de expressão de cada categoria na população Bcs-SP, segundo o método FPKM. % Bcs-SP representa o nível de expressão de cada categoria para a população Bcs-SP, em porcentagem. FPKM Bcs-Aq representa o nível de expressão de cada categoria na população Bcs-Aq, segundo o método FPKM. % Bcs-Aq representa o nível de expressão de cada categoria para a população Bcs-Aq, em porcentagem. FPKM é a sigla para Fragments Per Kilobase of transcript per Million.

Um detalhe importante que foi identificado nessa primeira análise foi a diferença de expressão entre as duas populações, principalmente nas toxinas, onde tivemos uma diferença 3 vezes maior para a população do Arquipélago. Tal diferença pode ter ocorrido por causa da qualidade do material sequenciado da Bcs-Aq, que estava mais degradado do que em relação à Bcs-SP, ou simplesmente a diferença seja real, já que a população do Arquipélago está completamente isolada e isso pode ter influenciado a expressão dos transcritos.

Na tabela a maior porcentagem de expressão identificada nas duas populações foi da categoria “Sem resultado”, que significa que esses contigs não possuem resultado identificado em bancos de dados. Esse fato é explicado pelo fato que esses animais são poucos estudados, então consideramos normal não encontrarmos muitas informações sobre seus dados gênicos.

Os transcritos dentro da categoria Celulares foram aqueles que tiveram como resultado no BLAST proteínas que são conhecidas sua função não-tóxica no organismo de eucariotos. Essa categoria teve um total de 8,26% de expressão para a Bcs-SP e 6,43% para a Bcs-Aq, porém como grande parte da anotação dos nossos contigs se encontram na

categoria “Sem resultado”, “Não caracterizadas”, “Preditas” e “Hipotéticas”, esse número pode não representar a real expressão nesses animais enquanto vivos.

Na categoria “Toxinas” obtivemos apenas 315 contigs identificados como toxinas, com um total de 1,63% de expressão na Bcs-SP e 4,65% de expressão na Bcs-Aq. Esses valores foram significativamente pequenos em relação ao todo transcriptoma das anêmonas, talvez não representando um valor real da expressão de toxinas dos animais. Por isso decidimos fazer uma mineração dos dados, afim de extrairmos o máximo possível de informações de transcritos similares à toxinas.

Para essa mineração criamos um banco de dados pessoal, contendo todas as toxinas revisadas existentes no banco de dados do Uniprot, mais precisamente do *Animal toxin annotation project*. Esse banco foi escolhido pois se trata de um programa de anotação específico para a anotação de toxinas. Nesse banco estão presentes toxinas de vários animais diferentes, como serpentes, aranhas, escorpiões, anêmonas do mar, caramujos e etc. Todos os dados presentes foram curados por revisores não-automatizados, o que aumenta sua credibilidade.

Como resultado do BLAST das nossas sequencias contra o banco de dados de toxinas do Uniprot obtivemos 3613 possíveis toxinas. Um resultado significativamente superior do que o obtido anteriormente, apenas com o resultado do BLAST contra o banco de dados NR, NCBI. Porém como forma de confirmarmos nossos resultados decidimos fazer um último BLAST, contra todas toxinas conhecidas de cnidários depositados no banco de dados Uniprot. Essa análise gerou um total de 30 sequencias com similaridade com toxinas de cnidários que não tínhamos identificado anteriormente. Esse resultado nos causou estranheza, já que achávamos que com o BLAST contra todas toxinas depositadas usando os mesmos parâmetros já conseguiríamos recuperar todas as toxinas possíveis, o que não se mostrou verdadeiro.

Embora essas sequencias adicionais que identificamos com o BLAST contra toxinas de cnidários tenha causado certo espanto, não foi algo que ilógico, pois o algoritmo do BLAST trabalha com buscas por similaridades em bancos de dados, portanto quanto menor o banco maior sua similaridade. Já que o banco contendo apenas toxinas de cnidários (com 255 toxinas) seja menor que o banco total de toxinas (com 6622 toxinas), a similaridade será bem melhor para o BLAST contendo apenas toxinas de cnidários.

Já que foi possível encontrar mais similares nas nossas sequencias à toxinas que passariam despercebidos com um BLAST contra um banco contendo apenas toxinas de cnidários nós decidimos fazer análises com cada família de toxina do nosso interesse contra nossas sequencias. Essas toxinas que escolhemos para esses BLASTs individuais foram todas as famílias conhecidas em cnidários e também foram escolhidas toxinas com base em trabalhos publicados recentemente sobre possíveis novas toxinas.

Obtivemos um total de 308 novas sequencias similares a toxinas nessa nova análise, demonstrando quantas possíveis toxinas iriam ser ignoradas no nosso trabalho, ficando com um total de 3951 sequencias. Porém, percebemos que nosso trabalho poderia ter uma grande quantidade de falsos positivos, pois muitas proteínas já conhecidas e não sendo toxinas possuem domínios em sua constituição que compartilham características com toxinas, como é o caso do domínio ShKT, que foi primeiro descoberto na toxina ShK (30), da *Stichodactyla helianthus*, e que também está presente em proteínas em praticamente todos seres vivos.

Para resolver esse problema verificamos cada sequência individualmente, analisando quais domínios estavam presentes em sua constituição e quais características eram normais apenas em toxinas, esperando assim separar quais sequencias eram proteínas não-tóxicas e as quais estavam mais próximas de possíveis toxinas. Desse procedimento resultou a figura abaixo, representando a expressão de cada família identificada perante o total expresso como toxina.

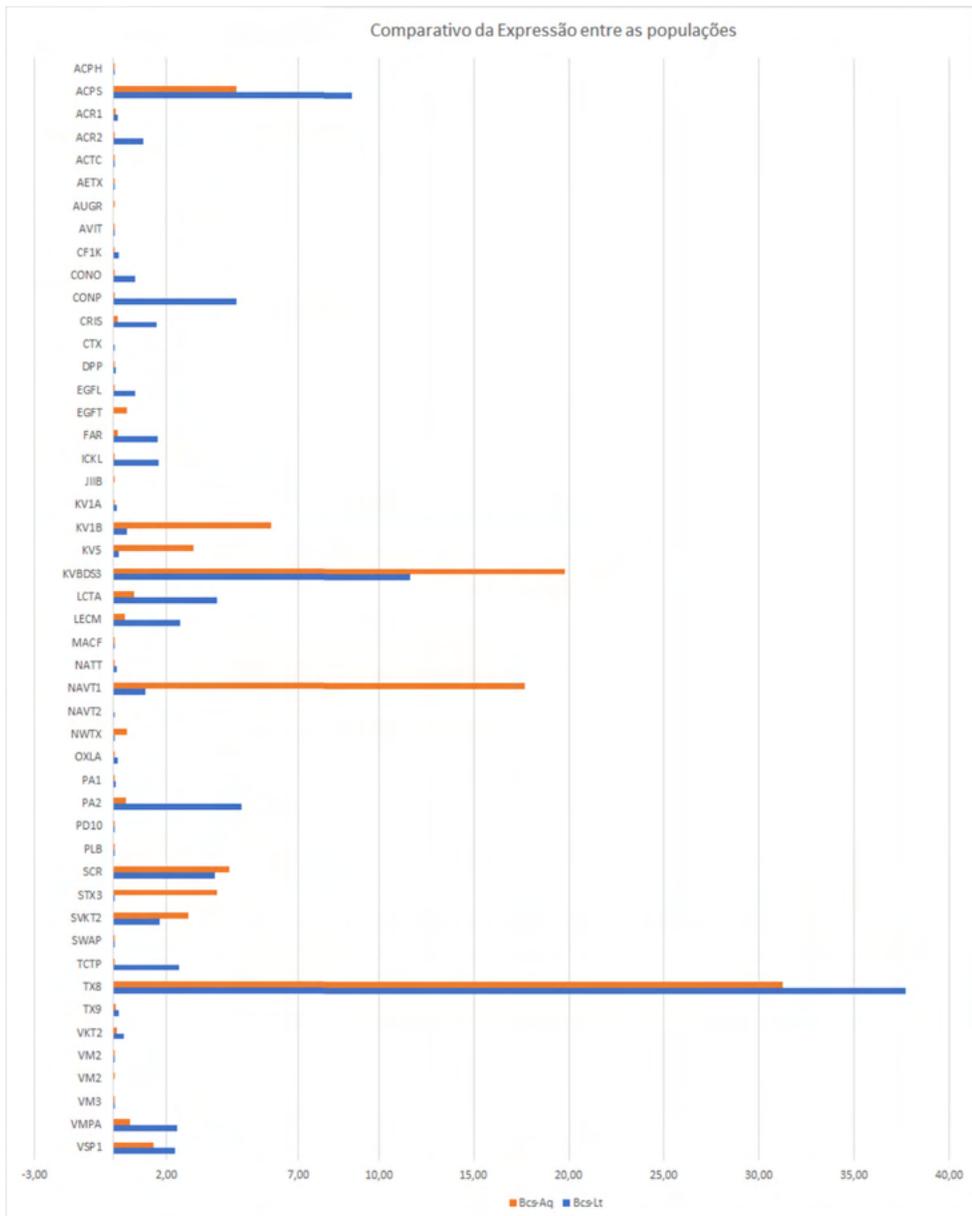


Figura 11: Famílias de toxinas expressas nas duas populações e seu nível de expressão.

As siglas em ordem alfabética representam as seguintes famílias de toxinas: ACES: Venom Acetylcholinesterase ; ACPH - Venom acid phosphatase; ACR1 - Acorrhagin-1; ACR2 - Acorrhagin-2; ACPS - Actinoporin (Sea Anemone); ACTC - Actinoporin (Conoidea); AETX – AETX; AUGR – Augerpeptide; AVIT – AVIT; CF1K - Cf1 (Possible Knottin); CONO - Conotoxin O2 (ICK-LIKE); CONP - Conopeptide P-like (kazal type); CRIS – CRISP; CTX - Cytolisin (jellyfish toxin); DPP - PLA2; EGFL - EGF-Like; EGFT - EGF Toxin; FAR – FARP; ICKL - ICK-like (BgrTx1); JIBB - Possible new Toxins (Jiibo3); KV1A - Sea anemone Type 1a Kv (ShK-Like); KV1B - Sea anemone Type 1b Kv (BgK-like); KV5 - Sea anemone Kv Type 5 (ICK-like); KVBDS3 - Sea anemone B-defensin (Kv type 3); LCTA - Latrotoxin Superfamily; LECM - Lectin/C-Type; MACF - MACPF; NATT - Natterin (Aerolysin-Like); NAV1 - Sea anemone B-defensin (Nav type I); NAV2 - Sea anemone B-defensin (Nav type II); NW1X - Possible new Toxins (Av1931/Bcs1802); OXLA - L-amino-acid oxidase; PA1 – Lipase/PLA1;

PA2 - PLA2; PD10 - Peptidase S10; PLB – PLB-Like; SCR - SCRiP; STX3 - Sea anemone Short toxin Type 3; SVKT2 - Sea Anemone Kunitz Type 2 Kv Toxin; SWAP - Waprin-Like; TCTP - TCTP Family; TX8 - Sea anemone 8 toxin family; TX9 - Sea anemone structural class 9a Family; VKT2 - Kunitz-type Family; VM2 - Peptidase M12B (PII); VM1 - Peptidase M12B (PI); VM3 - Peptidase M12B (PIII); VMPA - Peptidase M12A (Astacin-like); VSP1 - Peptidase S1.

No gráfico acima temos uma visão geral das toxinas encontradas e a sua expressão nas duas populações, em porcentagem. Vemos que a possível toxina mais expressa nas populações é a Sea anemone 8 toxin Family, que não possui função definida e que só tinha sido encontrada na anêmona *Anemonia viridis*. As famílias que também são bem expressas pertencem as famílias de toxinas de cnidários clássicas, como as toxinas bloqueadoras de canais de potássio e sódio. As famílias mais conhecidas serão discutidas posteriormente no trabalho, em capítulos específicos para cada uma.

O último passo para finalizarmos a análise geral do transcriptoma foi acrescentarmos as novas sequencias identificadas através dos BLASTs individuais como possíveis toxinas a tabela total das duas populações. Assim esperamos termos nos aproximado mais da porcentagem real de toxinas e outros transcritos que são expressos pelas duas populações de Bcs. Segue abaixo os dois gráficos correspondentes à essa análise.

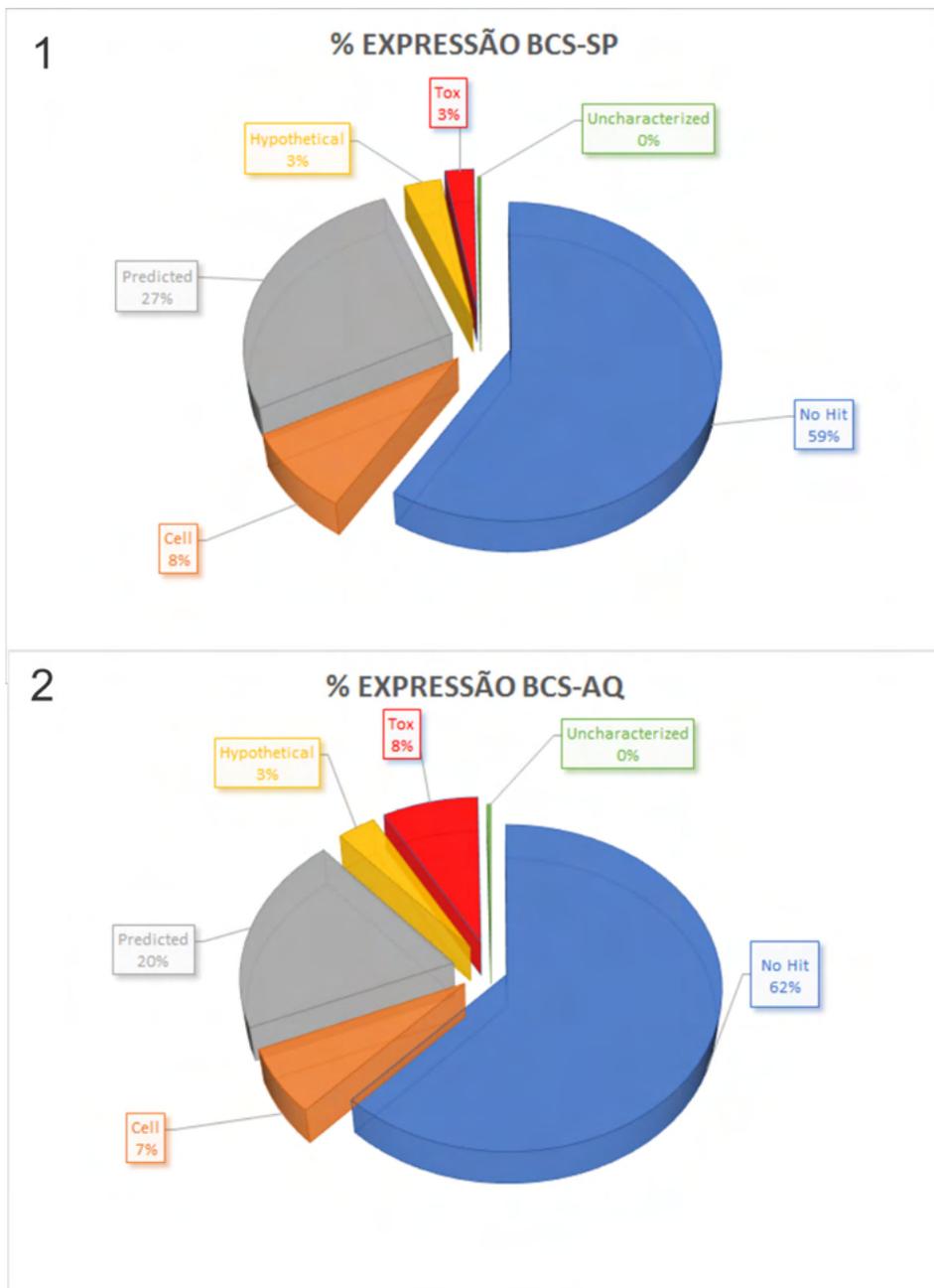


Figura 12: Gráfico do total de expressão nas duas populações de *Bunodosoma caissarum*. O gráfico 1 representa a população Bcs-SP e o gráfico 2 representa a população Bcs-Aq. Em azul está a fatia das sequências que não tiveram similaridades identificadas (No Hit). Na fatia laranja estão as sequências similares a proteínas celulares (Cell). Cinza são as sequências que foram preditas em possíveis proteínas (Predicted). As proteínas hipotéticas estão em amarelo (Hypothetical). Em verde as não caracterizadas (Uncharacterized). A fatia em vermelho são as sequências que foram identificadas como possíveis toxinas (Tox). O BLAST foi feito com o banco de dados NR, do Genbank.

Com as análises finalizadas dos BLASTs específicos para diferentes classes de toxinas conseguimos enriquecer o total expresso de toxinas perante o toda a expressão das populações, passando de 1,63% para 3% expresso na Bcs-SP e 4,65% para 8% de expressão na Bcs-Aq, praticamente o dobro nas duas populações, mostrando a efetividade do nosso método de mineração de dados com BLASTs individuais.

ANÁLISE DOS TRANSCRITOS SIMILARES A ACTINOPORINAS

Foram encontrados doze *contigs* com similaridade com citolisinas pertencentes à família das actinoporinas conforme mostrado na tabela abaixo.

Contig ID	Acerto no Blast: Identificador (Accession code)	E-value	Actinoporinas (subfamílias)
c28100_g1_i1	ACTP2_ACTEQ: Equinatoxin II de <i>Actinia equina</i> (P61914)	9.09E-50	Actiniaria
c61681_g1_i1	ACTP5_ACTEQ: DELTA-actitoxin-Aeq1b de <i>Actinia equina</i> (Q93109)	3.71E-118	Actiniaria
c64268_g1_i1	ACTP5_ACTEQ: DELTA-actitoxin-Aeq1b de <i>Actinia equina</i> (Q93109)	1.91E-111	Actiniaria
c64268_g1_i2	ACTP4_ACTEQ: DELTA-actitoxin-Aeq1c de <i>Actinia equina</i> (Q9Y1U9)	3.93E-119	Actiniaria
c64268_g1_i3	ACTP5_ACTEQ: DELTA-actitoxin-Aeq1b de <i>Actinia equina</i> (Q93109)	3.07E-110	Actiniaria
c64268_g1_i4	ACTP5_ACTEQ: DELTA-actitoxin-Aeq1b de <i>Actinia equina</i> (Q93109)	1.63E-103	Actiniaria
c123467_g1_i1	ACTP1_CINAN: Tereporin-Ca1 de <i>Cinguloterebra anilis</i> (P0DN66)	1.24E-07	Conoidea
c83042_g1_i1 (incompleto)	ACTP1_CINAN: Tereporin-Ca1 de <i>Cinguloterebra anilis</i> (P0DN66)	1.13E-09	Conoidea
c60312_g2_i2 (incompleto)	ACTP1_CINAN: Tereporin-Ca1 de <i>Cinguloterebra anilis</i> (P0DN66)	1.32E-07	Conoidea
c60312_g2_i3	ACTP1_CINAN: Tereporin-Ca1 de <i>Cinguloterebra anilis</i> (P0DN66)	1.72E-07	Conoidea
c64761_g2_i2	ACTP2_MONPT: Echotoxin-2 de <i>Monoplex parthenopeus</i> (Q76CA2)	9.57E-07	Actiniaria
c94977_g1_i1 (Incompleto)	ACTPC_ACTTE: DELTA-actitoxin-Ate1a de <i>Actinia tenebrosa</i> (P61915)	4.42E-25	Actiniaria

Tabela 6: Actinoporinas de *Bunodosoma caissarum*.

A tabela mostra os *contigs* identificados na biblioteca de *B. caissarum* com similaridade com Actinoporinas das subfamílias de anêmona do mar e de moluscos indicando os respectivos E-values e os números de acesso do UNIPROT. Dois *contigs* representam sequencias parciais e não foram utilizadas na figura do alinhamento.

Os *contigs* classificados como pertencentes às actinoporinas da subfamília actiniaria apresentaram uma organização similar, contendo regiões 5'UTR, seguida pela região codificante e 3'UTR, variando em tamanhos de 999 a 1656 nucleotídeos. A maioria destas sequências apresentou sinais de poliadenilação clássicos (AATAAA) no final da região 3'UTR. Entretanto, o contig c94977_g1_i1 está incompleto. Ainda assim, este contig tem 74% de similaridade (calculada pelo EMBOSS Needle) com a região C-terminal de DELTA-actitoxin-Ate1a de *Actinia tenebrosa* (P61915) sendo claramente uma actinoporina.

Para melhor entender as relações de similaridade das sequências identificadas foi feito um alinhamento dos *contigs* completos com os seus respectivos acertos (hits) encontrados pelo BLAST e com algumas actinoporinas cuja estrutura tridimensional e dados funcionais são bem conhecidos, o que permitiu construir um dendrograma mostrado na figura abaixo:

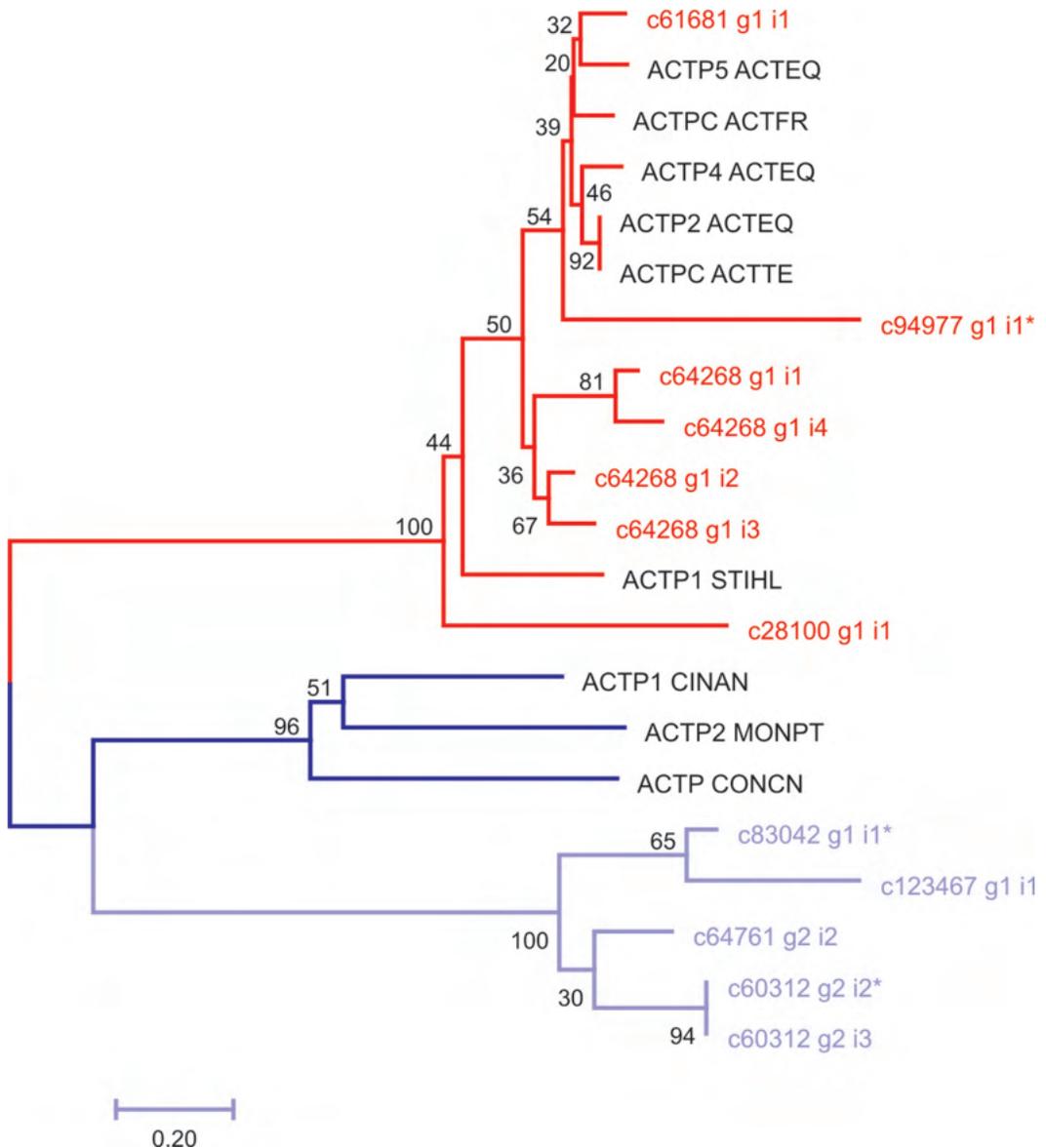


Figura 13: Dendrograma dos contigs similares às toxinas da família das Actinoporinas. O dendrograma foi montado por neighbour joining utilizando um modelo de substituição JTT de acordo com um alinhamento feito por Muscle. A árvore mostra em vermelho sequências de actinoporinas da subfamília actinoriária: ACTP2_ACTEQ: DELTA-actitoxin-Aeq1a (Equinatoxin II) de *Actinia equina* (P61914). ACTPC_ACTTE: DELTA-actitoxin-Ate1a (Tenebrosin-C) de *Actinia tenebrosa* (P61915). ACTP4_ACTEQ: DELTA-actitoxin-Aeq1c de *Actinia equina* (Q9Y1U9). ACTPC_ACTFR: DELTA-actitoxin-Afr1a (Fragaceatoxin C) de *Actinia fragacea* (B9W5G6). ACTP5_ACTEQ: DELTA-actitoxin-Aeq1b de *Actinia equina* (Q93109). ACTP1_STIHL: DELTA-stichotoxin-She4a (Sticholysin I) de *Stichodactyla helianthus* (P81662). Em azul sequências da subfamília conoidea: ACTP2_MONPT: Echotoxin-2 de *Monoplex parthenopeus* (Q76CA2). ACTP_CONCN: Conoporin-Cn1 de *Conus consors* (P0DKQ8). ACTP1_CINAN: Tereporin-Ca1 de *Cinguloterebra anilis* (P0DN66). Em roxo claro se destaca um grupo de contigs que tiveram similaridade com sequências da subfamília conoidea. Os contigs incompletos estão indicados por asterisco. Os números indicam os valores de bootstrap de 500 replicatas.

Sete *contigs* (c28100_g1_i1; c61681_g1_i1; c64268_g1_i1 e suas isoformas; i2; i3 e i4) foram mais similares com sequências pertencentes à subfamília actiniária. Os melhores E-values foram com dois precursores de equinatoxinas de *Actinia equina*: ACTP4_ACTEQ: DELTA-actitoxin-Aeq1c (Q9Y1U9) com o qual o *contig* c64268_g1_i2 mostrou 89.3% de similaridade e ACTP5_ACTEQ: DELTA-actitoxin-Aeq1b (Q93109) com o qual o c61681_g1_i1 mostrou 89.4% de similaridade. De forma geral todos os *contigs* pertencentes à família Actiniária tem similaridades acima de 80%.

Todos os *contigs* mostrados em vermelho no dendrograma, exceto o *contig* parcial c94977_g1_i1, apresentaram ORFs completos com *start* e *stop* códons, codificando precursores variando em tamanho de 212 a 215 aminoácidos e peptídeos sinais de 19 aminoácidos foram identificados pelo SignalP, como está mostrado no alinhamento da figura 03. Seguidos por um pró-peptídeo contendo a sequência de endereçamento para o nematocisto indicando claramente o ponto de clivagem da toxina madura. Algumas sequências apresentam uma substituição conservativa na díade de clivagem com KK no lugar de KR, ver figura 03.

No dendrograma o *contig* c28100_g1_i1 foi o mais divergente entre as sequências pertencentes às actinoporinas da subfamília actiniária. Este *contig* apresenta também um *start* códon e um *stop* códon, além de um sinal de poliadenilação ao fim da região 3'UTR, mas codifica um precursor maior com 244 aminoácidos. Entre o peptídeo sinal e a região madura há uma grande inserção no lugar do pró-peptídeo e a sequência consenso (DEDEDIEKR) não é claramente identificada, ver figura 9 abaixo. Entretanto, este possível pró-peptídeo também apresenta inúmeros aminoácidos negativos e um sítio de clivagem KR a partir da posição 74 predizendo uma região da toxina madura um pouco menor com 170 aminoácidos que apresenta 64.2% de similaridade com a toxina madura de ACTP2_ACTEQ: Equinatoxin II de *Actinia equina* (P61914) ou 65.9% de similaridade com ACTP1_STIHL: DELTA-stichotoxin-She4a (Sticholysin I) de *Stichodactyla helianthus* (P81662).

Apesar de esse precursor ser o maior, com 244 aminoácidos, em compensação a parte da toxina madura é 9 resíduos menor que as outras actinoporinas, que tem em geral 179 aminoácidos.

É possível observar no alinhamento da figura 9 que a maioria dos sítios para ligação fosfololina, a região de interação inicial rica em aromáticos outros sítios importantes como o de adesão elular se mostraram igualmente conservadas nas sequências de *B. caissarum*. Entretanto, algumas substituições podem indicar seletividade por diferentes análogos de fosfolipídeos. As actinoporinas de anêmonas têm diferentes capacidades de hemólise, mas ainda assim existe bastante conservação na região correspondente à a-hélice N-terminal, que em geral é rica em aminoácidos hidrofóbicos.

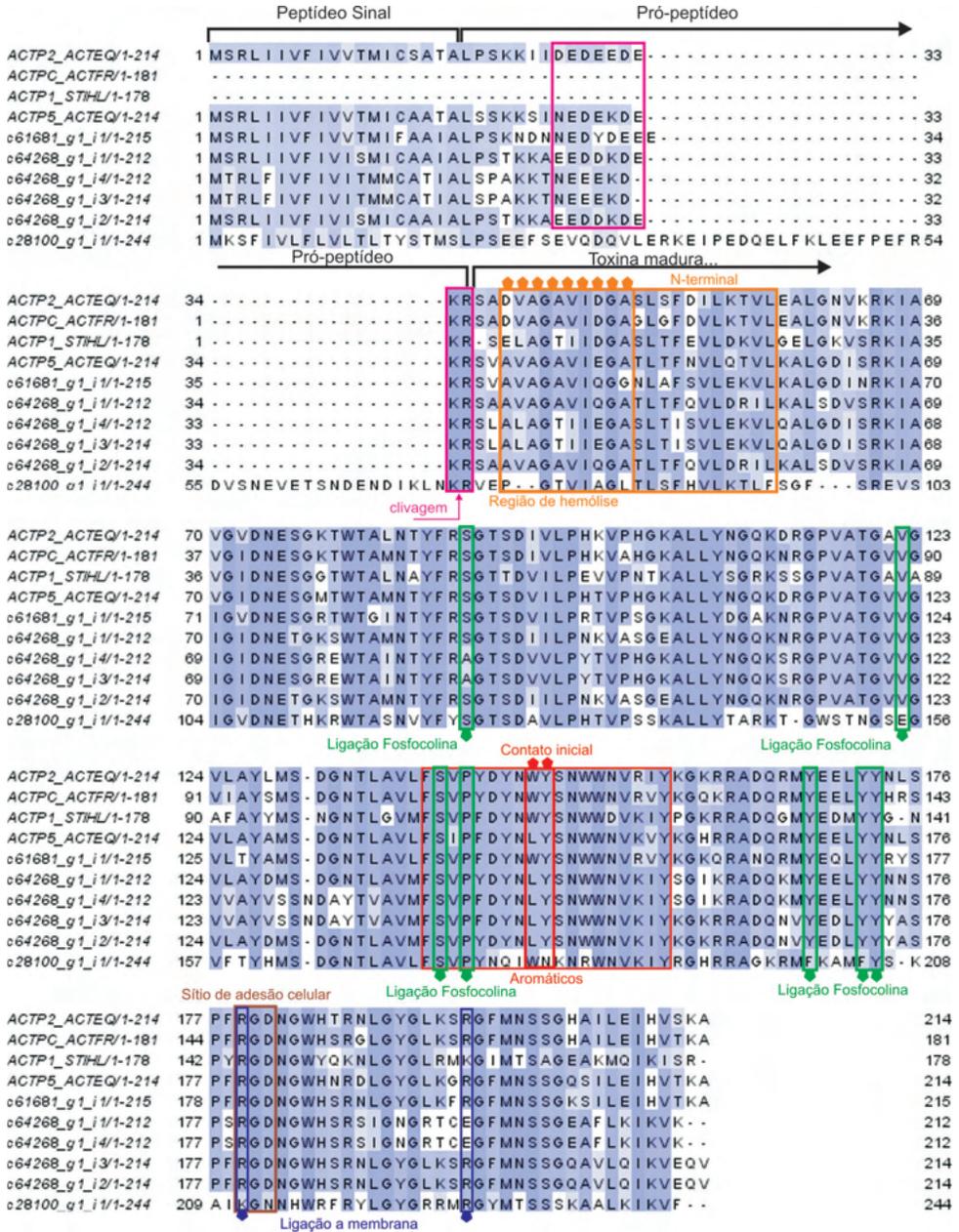


Figura 14: Alinhamento de Actinoporinas. O alinhamento dos *contigs* completos de *B. caissarum* (mostrados em dendrograma) com actinoporinas de anêmonas do mar bem estudadas. ACTP2_ACTEQ: Precursor da Equinatoxin-II de *Actinia equina* (P61914); ACTPC_ACTFR: a sequência madura da Fragaceatoxin C de *Actinia fragacea* (B9W5G6) e ACTP1_STIHL: toxina madura Sticholysin-I de *Stichodactyla helianthus* (P81662). A região do peptídeo sinal e propéptídeo estão indicadas. A região correspondente ao sítio de endereçamento da secreção para o nematocisto está indicada por uma caixa cor de rosa e o sítio de clivagem (KR). A região da toxina madura indicada por uma seta da mesma cor indicando o par de aminoácidos básicos (KR). A região da a-hélice N-terminal está indicada por uma caixa laranja e a parte importante para hemólise é indicada por pentágonos da mesma cor. A região rica em aromáticos está indicada por uma caixa vermelha e os sítios de contato inicial com

esfingolipídios de membrana estão indicados por caixas marcadas com pentágonos vermelhos. Os sítios de ligação com fosfololina estão indicados com caixas marcadas com pentágonos verdes. Outros sítios de interação com lipídios da membrana celular estão indicados por caixas marcadas com pentágonos roxos e um provável sítio de adesão celular (RGD) está marcado com uma caixa marrom. Os números indicam a posição do resíduo no alinhamento. A escala de cores de conservação está indicada segundo a matriz BLOSUM62.

Além da alta similaridade e da conservação geral dos sítios de ligação, as actinoporinas de *B. caissarum* apresentaram na tabela 7 massas moleculares e pontos isoelétricos teóricos muito semelhantes aos descritos para outras actinoporinas conhecidas.

Contig ID (toxina madura)	Tamanho (aa)	Massa molecular Teórica (Da)	pI teórico
c28100_g1_i1	170	19392.26	10.51
c61681_g1_i1	179	19827.56	9.95
c64268_g1_i1	177	19369.88	9.12
c64268_g1_i2	179	19710.28	9.34
c64268_g1_i3	180	19815.43	9.20
c64268_g1_i4	178	19475.03	8.90

Tabela 7: Massa molecular e pI teórico da região madura das actinoporinas de *B. caissarum*.

Até o momento não existem descrições de actinoporinas em anêmonas do gênero *Bunodosoma*. Entretanto, de OLIVEIRA (50) relata que o perfil cromatográfico da peçonha *B. caissarum* apresenta dois picos majoritários com atividade hemolítica. Um pico denominado C1 corresponde a uma proteína com 15495 Da e tem atividade fosfolipásica que teve o seu N-terminal parcialmente sequenciado identificando essa toxina como uma PLA2. O outro maior pico denominado C3 apresentou uma massa de 19757 Da, sem atividade enzimática e a toxina purificada foi reconhecida por anticorpos anti-equitoxina II (ACTP2_ACTEQ: P61914). Esta citolisina foi denominada de Caissarolysin I (Bcs I) e que não foi sequenciada. Suas características de massa são compatíveis com as massas teóricas encontradas neste seqüenciamento.

Os níveis de expressão de cada contig foram comparados entre as populações de *B. caissarum* do litoral de sudeste (Bcs-SP) e do arquipélago (BCS-Aq) e estão mostradas no gráfico de expressão a seguir:

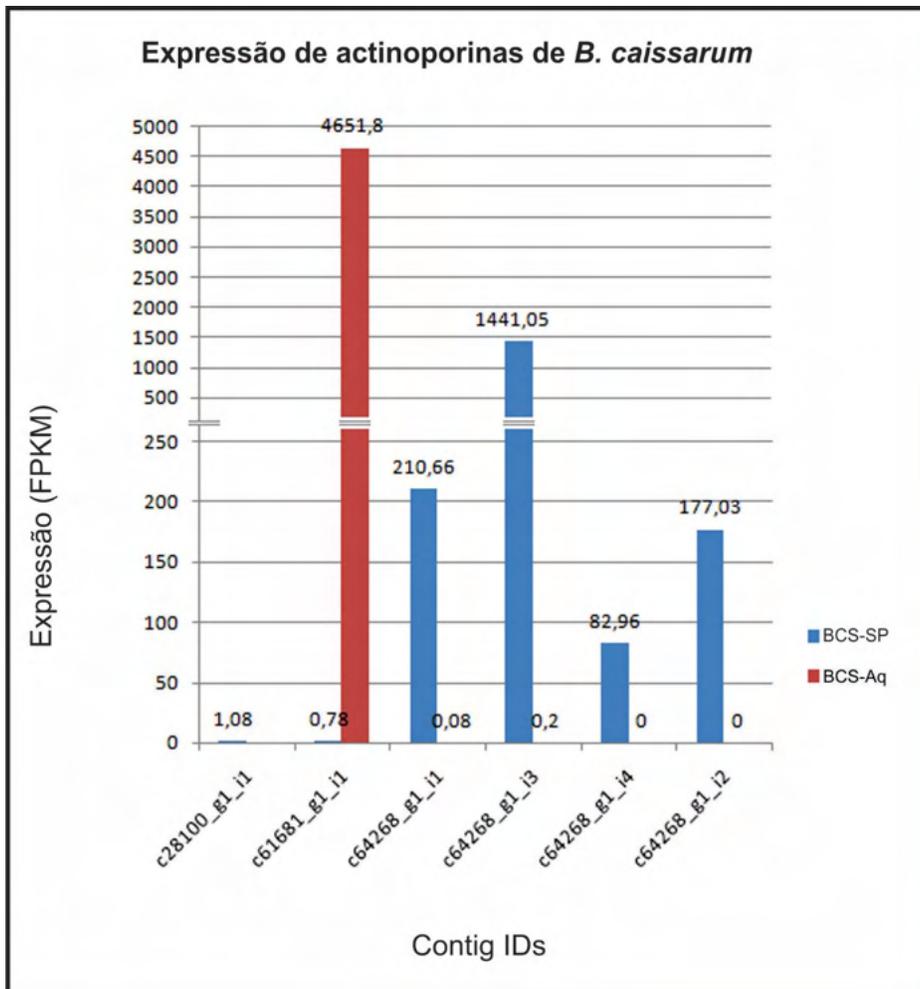


Figura 15: Gráfico da expressão das actinoporinas de *B. caissarum*. O gráfico compara o nível de expressão em FPKM (Fragments Per Kilobase Of Exon Per Million) de cada contig entre as duas populações de *B. caissarum* do Sudeste (Bcs-SP em azul) e do arquipélago (BCS-Aq em vermelho).

Na figura 10 verificamos os níveis de expressão dos *contigs* de actinoporinas nas duas populações de *B. caissarum* pertencentes à sub-família de anêmonas do mar. *Contigs* cujos acertos do BLAST foram similares a seqüências da sub-família Conoidea e seqüenciamentos incompletos parciais não foram incluídos. Dos 6 *contigs* analisados 3 foram expressos em ambas as populações, embora com níveis diferentes entre si. O contig c61681_g1_i1 com foi o mais expresso na amostra da população do arquipélago (BCS-Aq) enquanto que o contig c64268_g1_i2 é a isoforma mais expressa na população do litoral (Bcs-SP compartilhando 83.9% de similaridade entre si. Ambos os *contigs* não são exclusivos de cada população, mas as expressões do mesmo contig na população oposta é cerca de 6 a 7 mil vezes menor. Eles representam a isoforma majoritária de cada

população. Os *contigs* c28100_g1_i1, c64268_g1_i4 e c64268_g1_i2 foram expressos em diferentes níveis unicamente na Bcs-SP, o que indica uma maior diversidade de isoforma na população do litoral em comparação à do arquipélago.

Outros, cinco *contigs* mostrados em roxo claro no dendrograma da figura 8 foram identificados pelo BLAST como sequências similares a toxinas da subfamília conoidea. Notavelmente a toxina, ACTP2_MONPT: Echotoxin-2 de *Monoplex parthenopeus* (Q76CA2) a pesar de pertencer a um gastrópode, também é incluída pelo banco de dados de toxinas do UNIPROT na subfamília das actinoporinas de anêmonas do mar e foi acerto do BLAST com o *contig* c64761_g2_i2. Entretanto, estes *contigs* tiveram um E-value baixo com ordens de grandeza entre 10^{-9} a 10^{-7} , vide tabela 01. Desses apenas os *contigs* c123467_g1_i1; c64761_g2_i2; c60312_g2_i3 são possivelmente completos, contendo possível *start* e *stop* códons. A figura 05 a seguir mostra um alinhamento de actinoporinas de moluscos e os *contigs* de interesse. As actinoporinas da subfamília conoidea apresentam precursores geralmente maiores quando comparados com as de anêmona. O peptídeo sinal é maior com 24 aminoácidos e não contem pró-peptídeos entre o peptídeo sinal e toxina madura. O propetídeo, quando existe, é C-terminal. Portanto, a sequência de endereçamento consenso para o nematocisto (DEDEDIEKR) das actinoporinas de anêmona não está presente nas toxinas de moluscos.

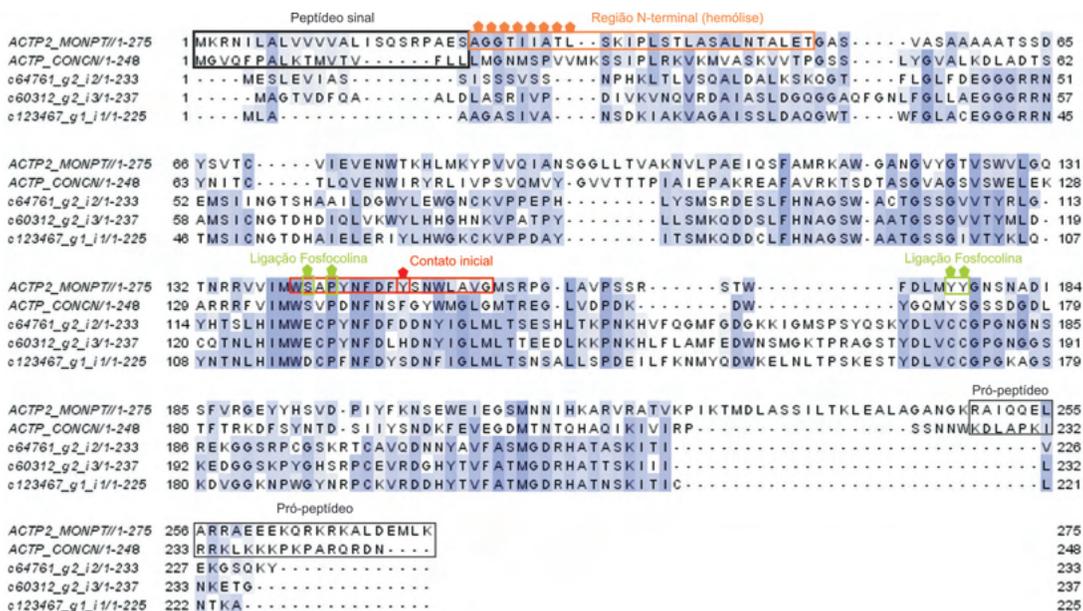


Figura 16: Alinhamento dos *contigs* similares a Actinoporinas de Conoidea. Usando como referência o *contig* c64761_g2_i2, os *contigs* mais similares a actinoporinas de moluscos (em roxo claro no dendrograma da figura 02) foram comparados com ACTP2_MONPT: precursor da Echotoxin-2 de *Monoplex parthenopeus* (Q76CA2). A região do peptídeo sinal e do pró-peptídeo estão indicados na figura em preto. A região N-terminal é indicada por uma caixa laranja e os resíduos importantes para hemólise são indicados por pentágonos laranjas. Os sítios de ligação com fosfocolina estão indicados em verde. A região rica em aromáticos e o sítio de ligação inicial a esfingomiélna estão indicados

em vermelho. Os números indicam a posição do resíduo no alinhamento. A escala de cores de conservação está indicada segundo a matriz BLOSUM62.

Caracteristicamente nenhum dos *contigs* encontrados como acertos no BLAST de toxinas do UNIPROT apresentaram peptídeos sinais preditos pelo SignalP. Entretanto, é possível encontrar o sinal consenso de Kozak eucariótico clássico (51) que determina o *start codon* nestes *contigs* originando ORFs completos que se traduzem em proteínas preditas maiores com em média 234 aminoácidos, enquanto que os precursores de actinoporinas de anêmonas normalmente tem 214 aminoácidos. Mas, nenhum peptídeo sinal foi identificado pelo SignalP. Também observamos a presença de cisteínas na sua cadeia de aminoácidos, algo incomum para actinoporinas clássicas. A similaridade das sequências alinhadas na figura 05 com Echotoxin-2 de molusco é menor que 30% e a maioria dos sítios de ligação com lipídeos não está conservada conforme o indicado por (19) para a Echotoxin-2.

Estes *contigs* também encontram acertos com outras proteínas não presentes no banco de dados de toxinas do UNIPROT. Como por exemplo, sequências preditas do genoma da anêmona *Nematostella vectensis* (XP_001633907.1) e outras sequências preditas de peixes contendo o domínio do PFAM (PF06369) de Actinoporinas de anêmonas, mas estas sequências ainda estão sobre revisão e também não apresentam peptídeos sinais.

Evidentemente estas sequências contêm um domínio similar ao das actinoporinas, mas não devem ser toxinas inclusive devido à ausência de um peptídeo sinal, falta da sequência consenso de endereçamento para o nematocisto e por serem muito divergentes, talvez elas possam ser resultado relictual de uma transferência horizontal de genes de citolisinas de microrganismos unicelulares para cnidários. Não existe evidências dessa teoria em toxinas de anêmonas, portanto isso é apenas uma teoria.

ANÁLISE DOS TRANSCRITOS SIMILARES A TOXINAS BLOQUEADORAS DE CANAIS DE POTÁSSIO DEPENDENTE DE VOLTAGEM DO TIPO I – SHKT

Encontramos 20 *contigs* no nosso sequenciamento que se assemelham às toxinas bloqueadoras de canal de potássio do tipo 1 e que possuem apenas um único domínio ShKT, apresentados na tabela abaixo:

Contig ID	Código de acesso	Espécie	Subtipo	E-value
c17614_g1_i1	K1B2_BUNCI: Kappa-actitoxin-Bcs3b (C0HJC3)	<i>Bunodosoma caissarum</i>	1b	6,26E-20
c116940_g1_i1	K1B9A_ANEVI: U-actitoxin-Avd9a (P0DN00)	<i>Anemonia viridis</i>	1b	5,32E-08
c18859_g1_i1	K1B9A_ANEVI: U-actitoxin-Avd9a (P0DN00)	<i>Anemonia viridis</i>	1b	1,47E-28
c36937_g1_i1	K1B9A_ANEVI: U-actitoxin-Avd9a (P0DN00)	<i>Anemonia viridis</i>	1b	8,61E-14
c64871_g1_i1	K1B9A_ANEVI: U-actitoxin-Avd9a (P0DN00)	<i>Anemonia viridis</i>	1b	6,10E-27
c53259_g1_i1	K1B9C_ANEVI: U-actitoxin-Avd9c (P0DN02)	<i>Anemonia viridis</i>	1b	4,23E-24
c90258_g1_i1	K1B9C_ANEVI: U-actitoxin-Avd9c (P0DN02)	<i>Anemonia viridis</i>	1b	5,61E-08
c473_g1_i1	K1BBA_ANEVI: U-actitoxin-Avd11a (P0DN05)	<i>Anemonia viridis</i>	1b	3,91E-22
c61431_g1_i1	K1BBA_ANEVI: U-actitoxin-Avd11a (P0DN05)	<i>Anemonia viridis</i>	1b	1,60E-28
c80103_g1_i1	K1BBA_ANEVI: U-actitoxin-Avd11a (P0DN05)	<i>Anemonia viridis</i>	1b	3,62E-24
c68752_g3_i1	K1B_BUNGR: Kappa-actitoxin-Bgr1a (P29186)	<i>Bunodosoma granuliferum</i>	1b	2,04E-19
c68752_g3_i2	K1B_BUNGR: Kappa-actitoxin-Bgr1a (P29186)	<i>Bunodosoma granuliferum</i>	1b	8,62E-19
c68752_g3_i3	K1B_BUNGR: Kappa-actitoxin-Bgr1a (P29186)	<i>Bunodosoma granuliferum</i>	1b	1,85E-18
c68752_g4_i1	K1B_ACTEQ: Kappa-actitoxin-Aeq4a (P81897)	<i>Actinia equina</i>	1b	5,17E-13
c103967_g1_i1	K1A_ANEER: Kappa-actitoxin-Aer3a (Q0EAE5)	<i>Anemonia erythraea</i>	1a	6,09E-31
c39229_g1_i1	K1A_ANEER: Kappa-actitoxin (Q0EAE5)	<i>Anemonia erythraea</i>	1a	2,06E-20
c39229_g1_i2	K1A_ANEER: Kappa-actitoxin (Q0EAE5)	<i>Anemonia erythraea</i>	1a	1,85E-18
c120283_g1_i1	K1B9A_ANEVI: U-actitoxin-Avd9a (P0DN00)	<i>Anemonia viridis</i>	1b	5,12E-07
c87952_g1_i1	K1B9C_ANEVI: U-actitoxin-Avd9c (P0DN02)	<i>Anemonia viridis</i>	1b	5,27E-07

Tabela 8: Toxinas bloqueadoras de canal de potássio do tipo 1 em Bcs

A tabela mostra os contigs identificados na biblioteca das populações de *B. caissarum* similares às toxinas bloqueadoras de canal de potássio do tipo I, de acordo com o Resultado do BlastX contra o banco de dados de toxinas animais do UniprotKB, o respectivo e-value e a classificação da família segundo a análise do dendrograma e do alinhamento discutido abaixo. A coluna subtipo representa o subtipo da toxina que resultou do BLAST de cada contig.

Os contigs presentes na tabela possuem o precursor semelhantes às toxinas conhecidas da família, variando de 306 a 960 nucleotídeos, porém para melhor visualizarmos a similaridade das sequências com as toxinas já conhecidas e bem descritas pertencentes à família de bloqueadores de canais de potássio dependentes de voltagem do tipo I, fizemos um alinhamento dos contigs em suas formas completas, apresentado na figura n. A única exceção dentre todas as sequências foi o contig c87952_g1_i1, com 270 aminoácidos no mínimo em sua cadeia proteica, já que não foi possível estimar o número total de aminoácidos pois não existe informação sobre start e stop códon em sua sequência. Analisando o contig na ferramenta de busca de domínios conservados do NCBI verificamos que o mesmo possui um domínio ShKT conservado em sua sequência, por isso o BLAST acusou como um possível resultado contra toxinas ShK-like. Por esse motivo esse contig foi excluído de nossas análises posteriores.

Signal Peptide

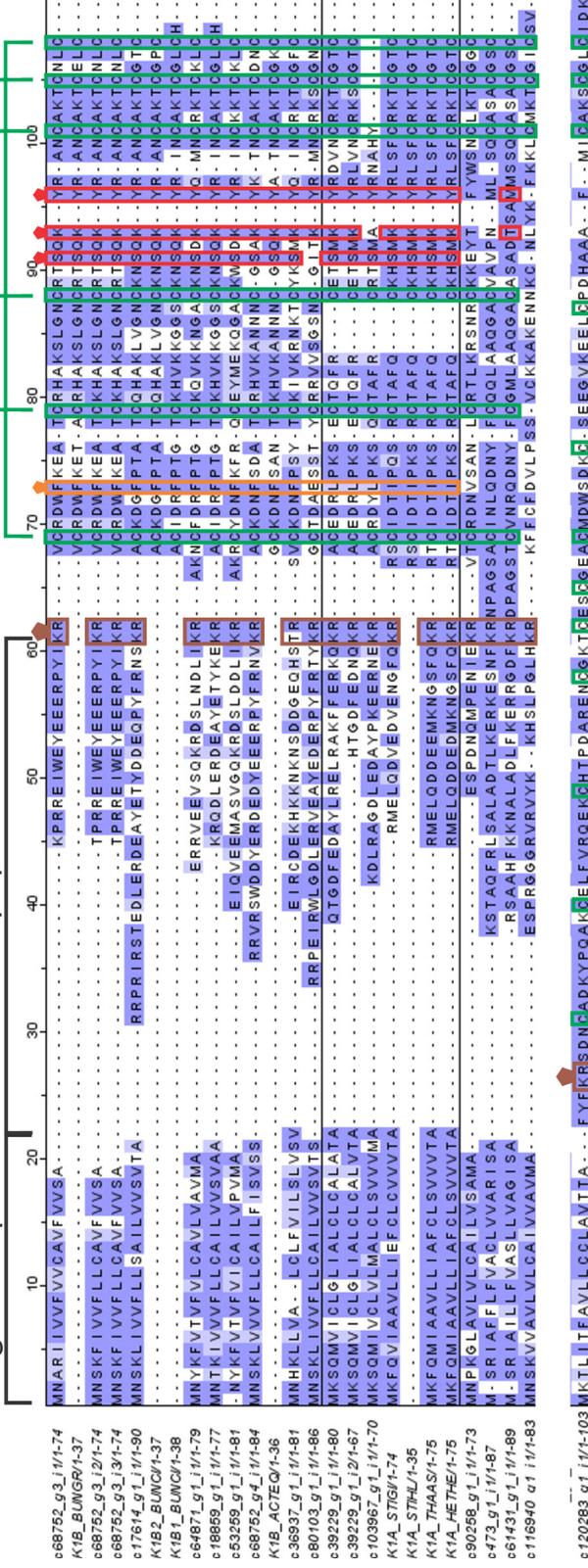


Figura 17: Alinhamento dos contigs similares às toxinas da família de bloqueadoras de canal de potássio do tipo 1. O alinhamento foi feito pelo algoritmo Muscle. Os códigos correspondem às seguintes toxinas: K1B_BUNGR: Kappa-actitoxin-Bgr1a (Bkg) de *Bunodosoma granulifera*. K1B2_BUNCI: Kappa-actitoxin-Bcs3b (BcsTx2) de *Bunodosoma caissarum*. K1B1_BUNCI: Kappa-actitoxin-Bcs3a (BcsTx1) de *Bunodosoma caissarum*. K1B_ACTEQ: Kappa-actitoxin-Aeq4a de *Actinia equina*. K1A_HETMG: Kappa-stichotoxin-Hmg1a de *Heteractis magnifica*. K1A_STHL: Kappa-stichotoxin-She3a (SHK) de *Stichodactyla helianthus*. K1A_THAAS: Kappa-thalatoxin-Tha2a de *Thalassianthus aster*. K1A_HETHE: Kappa-thalatoxin-Hhe2a de *Heterodactyla nemprichi*. Em verde estão as localizações conservadas das pontes dissulfeto nas toxinas e nos contigs. As caixas e os polígonos em vermelho indicam os resíduos mais importantes para a ação bloqueadora do canal. Em laranja estão resíduos também importantes, mas em proporção menor que as apresentadas em vermelho. Em marrom estão destacados os possíveis locais de clivagem da parte madura dos peptídeos. O contig c120283_g1_j1 foi destacado dos demais devido as suas características únicas e discordantes das demais, como a presença de 12 cisteínas em sua cadeia peptídica.

A caixa em vermelho no alinhamento, com o polígono localizado logo acima, temos posições mais importantes para a função bloqueadora dessas toxinas, que inclui a díade funcional dessa família de toxinas. Em laranja se encontra os resíduos que possuem atuação em canais de potássio, porém a sua importância para a função bloqueadora não é tão fundamental como outros resíduos. Esses resíduos não apenas indicam o canal Kv1.3 como alvo, mas também Kv1.1, Kv1.2, Kv1.6, Kv3.2 e KCa3.1.

No alinhamento apresentado na figura 17 cada sequência foi analisada em busca do seu possível peptídeo sinal através do algoritmo SignalP, que encontrou sinal de clivagem em todas sequencias. Para designarmos o possível pró-peptídeo de endereçamento para o nematocisto nos contigs todos os aminoácidos seguintes ao peptídeo sinal até a díade básica característica que indica o ponto de clivagem para a parte da toxina madura (29) foram considerados como pró-peptídeo. Porém os contigs c36937_g1_i1 e c120283_g1_i1 não apresentaram a díade característica, embora o contig c36937_g1_i1 possua uma arginina numa posição semelhante a outra sequencias, que foi considerada como possível ponto de clivagem. No contig c120283_g1_i1 foi encontrado uma díade KR 3 aminoácidos seguintes ao peptídeo sinal, característica que a difere em muito das toxinas e de outras sequencias, talvez representando que esse contig não possua características de toxina. Para melhor visualização da correlação entre as toxinas conhecidas e nossos contigs foi construído um dendograma usando como base o alinhamento discutido, que está apresentado na figura a seguir.

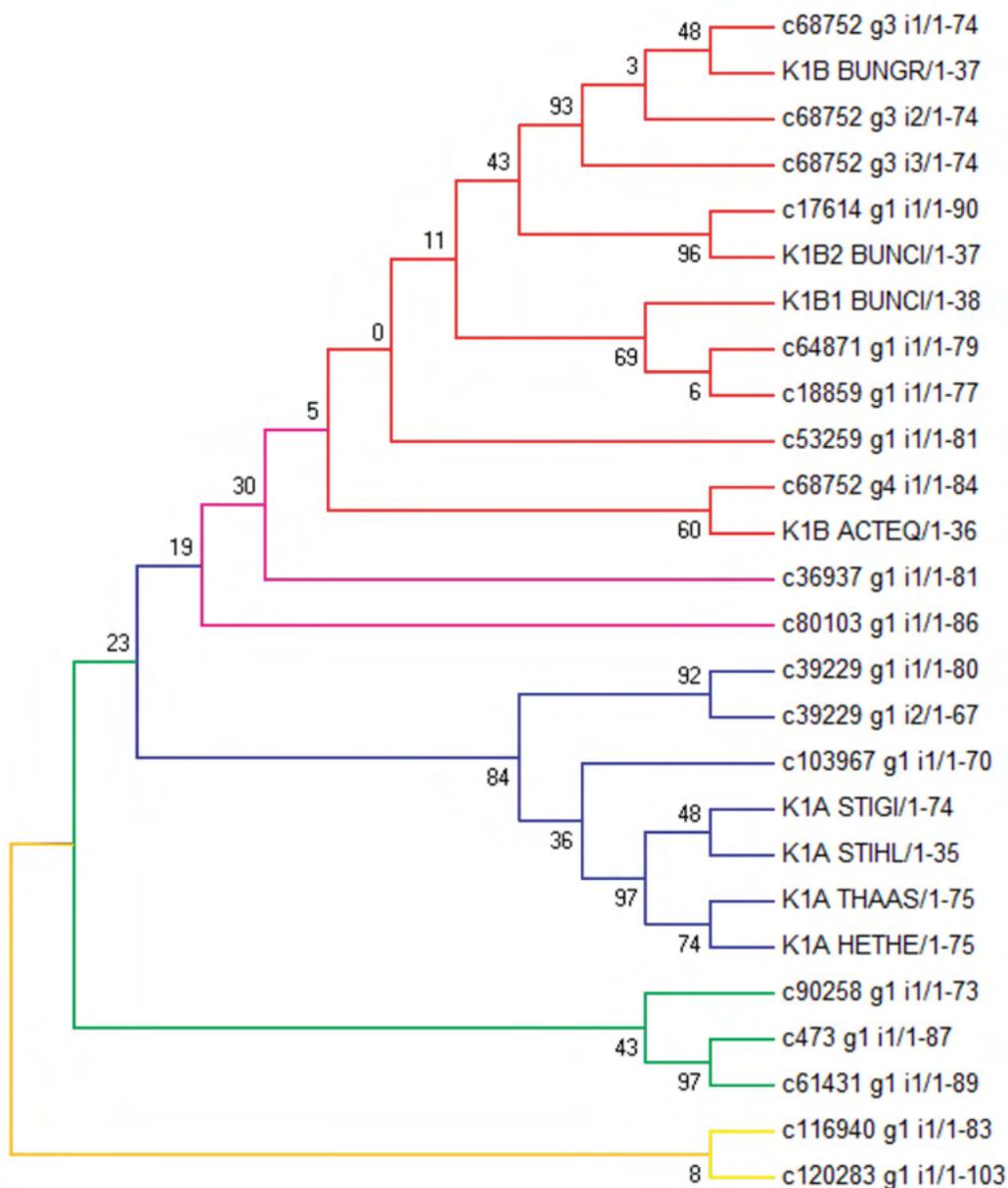


Figura 18: Dendrograma dos contigs similares às toxinas da família de bloqueadoras de canal de potássio do tipo 1. O dendrograma foi montado por neighbour joining utilizando um modelo de substituição JTT de acordo com um alinhamento feito por Muscle. Os códigos correspondem às seguintes toxinas: K1A_STIHL: Kappa-stichotoxin-She3a (ShK) de *Stichodactyla helianthus* (P29187). K1A_HETMG: Kappa-stichotoxin-Hmg1a de *Heteractis magnifica* (Q16846). K1A_STIGI: Kappa-stichotoxin-Sgt4a de *Stichodactyla gigantean* (E2S061). K1A_HETHE: Kappa-thalatoxin-Hhe2a de *Heterodactyla hemprichii* (E2S065). K1B2_BUNCI: Kappa-actitoxin-Bcs3b (BcsTx2) de *Bunodosoma caissarum* (C0HJC3). K1B_BUNGR: Kappa-actitoxin-Bgr1a (Bgk) de *Bunodosoma granulifera* (P29186). K1B1_BUNCI: Kappa-actitoxin-Bcs3a (BcsTx1) de *Bunodosoma caissarum* (C0HJC2). K1B_ACTEQ: Kappa-actitoxin-Aeq4a de *Actinia equina* (P81897). Em vermelho se destaca um grupo de contigs que tiveram similaridade às toxinas do subtipo 1b (Bgk). Em magenta se encontra um contig

que se mostrou o mais intermediário entre o tipo 1a e 1b. Em azul se destaca contigs que possuem certa semelhança com às toxinas do tipo 1a (ShK). Em verde estão os contigs do tipo 1b que possuem maior identidade com as toxinas do subtipo 1a. Em laranja representa toxinas também do subtipo 1b, porém com maior divergência do que as anteriores, por isso ficaram destacadas no dendrograma.

No grupo destacado pela cor vermelha temos 8 contigs (c68752_g3_i1, c68752_g3_i2, c68752_g3_i3, c17614_g1_i1, c64871_g1_i1, c18859_g1_i1, c53259_g1_i1 e c68752_g4_i1) temos os contigs que mais se aproximaram das toxinas do tipo 1b (BgK). O contig c17614_g1_i1 possui e-value de $5,00E-21$ e 97,3% de identidade com a toxina K1B_BUNCI, de *Bunodosoma caissarum*, que pode significar que encontramos em nossos contigs a sequência percussora dessa toxina, já que a mesma possui apenas sua parte madura conhecida e anotada. Os contigs c68752_g3_i1, c68752_g3_i2, c68752_g3_i3 são isoformas de um mesmo transcrito, possuindo e-value entre $1,00E-19$ e $4,00E-20$ para a toxina K1B_BUNGR, de *Bunodosoma granuliferum*, com identidade variando de 89,2% e 91,9%. Também é possível notar como o contig c68752_g4_i1 se destacou dos contigs anteriores, ficando num ramo isolado com a toxina K1B_ACTEQ e como o contig c53259_g1_i1 se mostrou isolado entre o c68752_g4_i1 e as outras, talvez sendo um intermediário entre essas possíveis isoformas. Essas diferenças podem indicar que diferentes isoformas de toxinas do subtipo 1b estão presentes no transcriptoma de *Bunodosoma caissarum*.

Os contigs destacados com a cor magenta no dendrograma são dois contigs (c36937_g1_i1 e c80103_g1_i1) que apresentaram características que os colocam como intermediários entre as toxinas do tipo 1b e 1a. Analisando o alinhamento desses contigs é possível observar que após a primeira cisteína os aminoácidos do contig c80103_g1_i1 não são tão conservados em relação aos outros mais similares às toxinas do subtipo 1b, enquanto que no contig c36937_g1_i1 possui maior conservação. Os dois contigs apresentam também a inserção de quatro aminoácidos, característica das toxinas do subtipo 1b.

O grupo azul contém três contigs (c39229_g1_i1, c39229_g1_i2 e c103967_g1_i1) que mais se assemelharam às toxinas do tipo 1a, o que é algo nunca antes detectado na análise de toxinas da *Bunodosoma caissarum*. O valor de e-value desses contigs foram de $2,06E-20$ para o contig c39229_g1_i1 com a toxina Kappa-actitoxin-Aer3a (Q0EAE5) da *Anemonia erythraea*, com 60% de similaridade, no contig c39229_g1_i2 temos e-value de $3,00E-19$, possuindo 76,2% de identidade com a toxina Kappa-actitoxin-Aer3a (Q0EAE5) da *Anemonia erythraea* e e-value de $3,00E-31$ para o contig e c103967_g1_i1 também com a toxina Kappa-actitoxin-Aer3a (Q0EAE5) da *Anemonia erythraea*, com 68,1% de identidade.

O grupo verde possui 3 contigs (c90258_g1_i1, c473_g1_i1 e c61431_g1_i1) que se destacaram dos grupos anteriores, ficando em ramos destacados dentro do dendrograma. Esses 3 contigs possuem a inserção característica do subtipo 1b, com as cisteínas

conservadas, porém no geral seus aminoácidos apresentam muitas diferenças em relação aos outros contigs. Também não foi encontrado os resíduos importantes para a ação bloqueadora de canais de potássio.

Por fim temos o grupo amarelo, com dois contigs (c116940_g1_i1 e c120283_g1_i1) que apresentaram as características mais divergentes do restante, ficando destacados no dendrograma. O contig c116941_g1_i1 possui certas características em comum com os contigs e toxinas anteriores, como as cisteínas e possíveis pontes dissulfeto conservadas e a díade funcional bloqueadora de canais de potássio (nesse contig a tirosina (Y) foi substituído por uma fenilalanina (F), compartilhando a propriedade de aminoácido aromático), seu e-value foi de 5,32E-08, com 41% de identidade, com a toxina K1B9A_ANEVI (U-actitoxin-Avd9a [P0DN00]), pertencente ao subtipo 1b.

O contig c120283_g1_i1 se mostrou completamente divergente das demais, com 12 cisteínas presentes na sua cadeia peptídica, totalizando 6 possíveis pontes dissulfeto, o dobro do esperado para toxinas da família aqui estudada. Outra característica divergente é o tamanho do possível propeptídeo desse contig, com apenas 5 aminoácidos até o possível ponto de clivagem na díade básica KR. Suas características indicam que provavelmente não se trata de uma toxina. Seu e-value no BLAST foi de 5,12E-07 e 38% de identidade com a toxina K1B9A_ANEVI (U-actitoxin-Avd9a [P0DN00]), pertencente ao subtipo 1b.

Os contigs c17614_g1_i1 e c18859_g1_i1 foram os contigs que mais se aproximaram das toxinas bloqueadoras de canal de potássio do tipo 1 de *Bunodosoma caissarum* já depositadas em bancos de dados e analisando o alinhamento verificamos que esses contigs são realmente as toxinas já conhecidas, porém agora temos acesso à informação do seu precursor completo.

Atualmente foram identificados na família de anêmonas *Bunodosoma* apenas toxinas bloqueadoras de canal de potássio do tipo 1, subtipo 1b, que é justamente a BgK, da anêmona *Bunodosoma granulifera*, então é natural encontrarmos maior parte dos contigs semelhantes à esse subtipo, porém encontramos três contigs que contém as características do subtipo 1a (ShKT), o contig c103967_g1_i1, c39229_g1_i1 e c39229_g1_i2. Seus precursores estão completos, com start e stop códon, e a característica do subtipo 1a em relação com o subtipo 1b também está presente, que é a falta de 4 aminoácidos logo antes da terceira cisteína. Então é possível que esses transcritos sejam realmente de uma nova toxina.

Abaixo segue a imagem com a modelagem de contigs que consideramos os mais relevantes com base em toxinas bem estudadas.

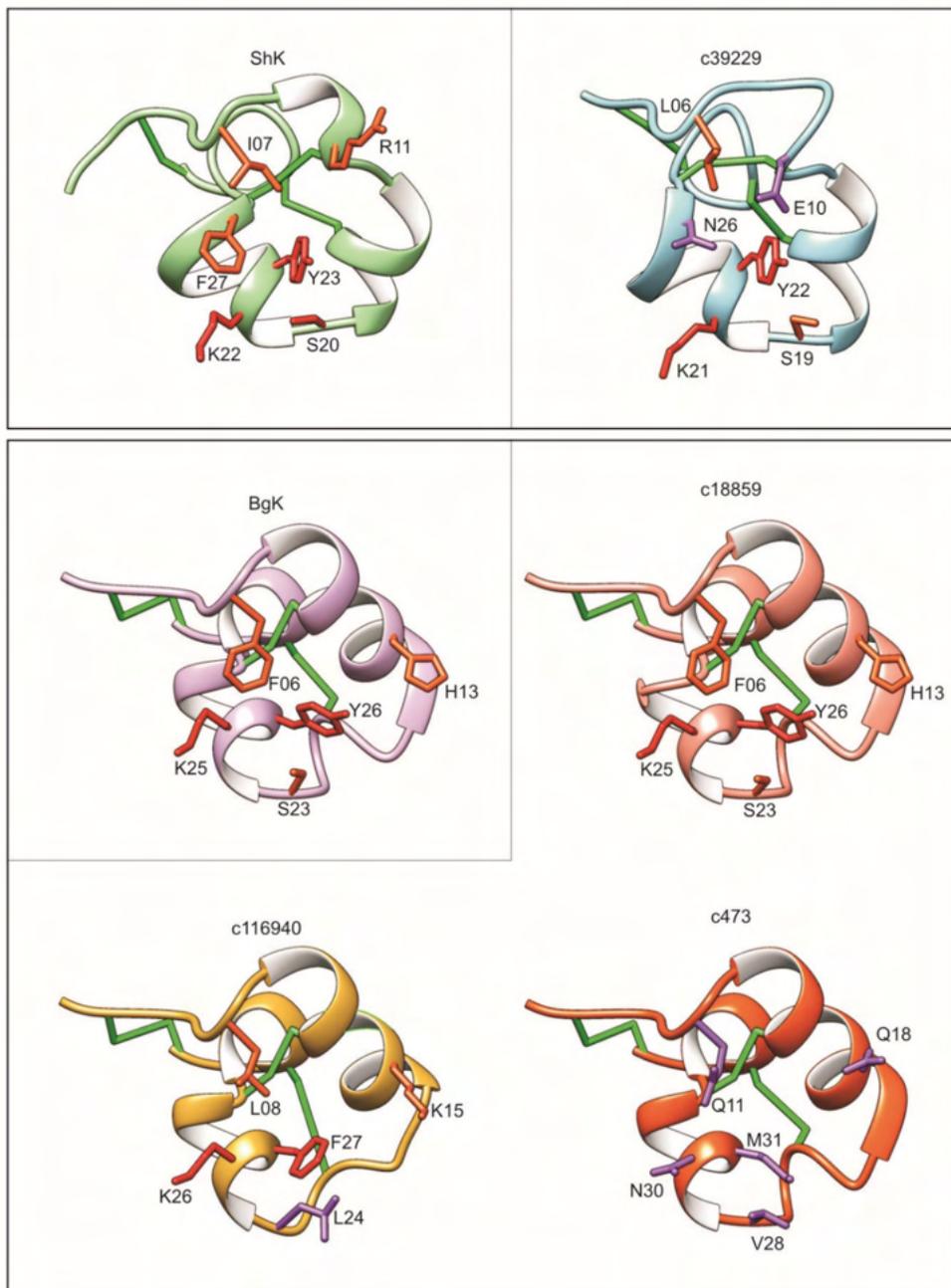


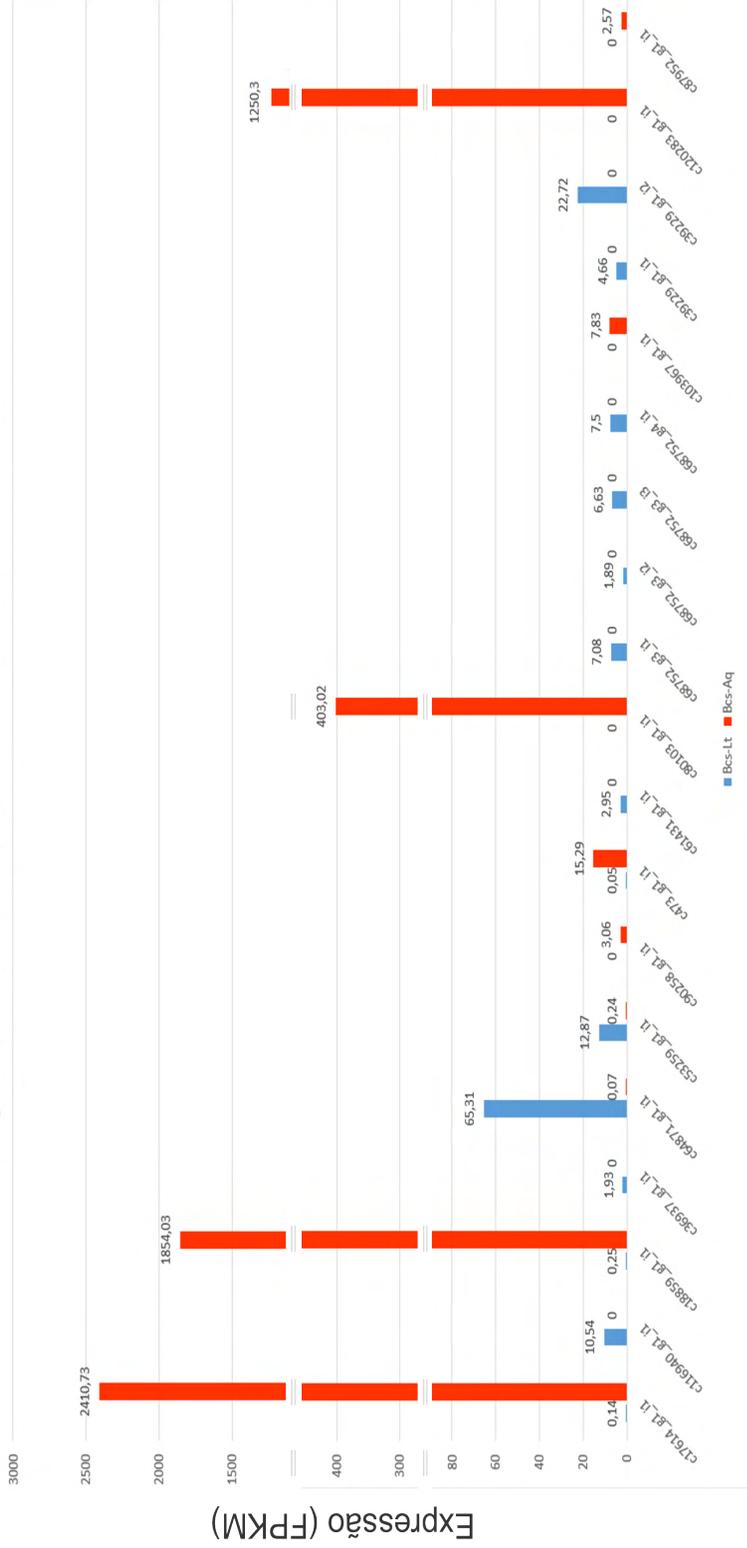
Figura 19: Modelagem de alguns contigs em relação a ShK e BgK. A toxina ShK (modelo 4LFS) está representada em verde. BgK (modelo 4Z7P) em roxo claro. O contig c39229 em azul claro, contig c18859 em salmão, c116940 em amarelo e c473 em laranja. Na modelagem em verde representa as pontes dissulfeto, em laranja os resíduos com relevância para o bloqueio e em vermelho os resíduos importantes para a função bloqueadora de canais de potássio. Em roxo se encontra os resíduos que provavelmente não compartilham das mesmas funções que os resíduos nessas posições nas toxinas.

Na figura acima temos a modelagem dos contigs c39229_g1_i1, modelado com base na ShK, c18859_g1_i1, c116940_g1_i1 e c473_g1_i1 modelados com base na BgK. No contig c39229_g1_i1 é possível identificar semelhança na estrutura geral do contig em relação à toxina, embora tenha algumas diferenças no N-terminal, por volta dos resíduos de número 10. Encontramos os resíduos L06, S19, K21 e Y22 em posições semelhantes à ShK e mantendo as mesmas características físico-químicas, porém os resíduos E10 e N26 não compartilham essas características.

Já em relação à BgK temos o contig c18859_g1_i1, que se mostrou o mais próximo da BgK, com a sua estrutura, os resíduos importantes e as pontes dissulfeto altamente semelhantes a toxina. Os contigs c166940_g1_i1 e c473_g1_i1 possuem semelhança estrutural com a BgK, porém quando analisamos os resíduos importantes para a função bloqueadora de canal de potássio encontramos discrepâncias. No Contig c166940_g1_i1 a díade funcional KY está presente, porém a Serina que está presente na posição 23 na BgK se encontra substituída por uma Lisina, que não possui as mesmas características físico-químicas, e no contig c473_g1_i1 nenhum desses resíduos importantes é encontrado, o que talvez signifique que esse contig não possua função tóxica.

Abaixo se encontra o gráfico comparativo da expressão nas duas populações de *Bunodosoma caissarum* dos contigs analisados:

Expressão de transcritos ShK-like de *B. caissarum*



Contig IDs

Figura 20: Gráfico da expressão dos contigs ShK-like de *B. caissarum*. O gráfico compara o nível de expressão em FPKM nas duas populações de *B. caissarum* do Sudeste (Bcs-SP em azul) e do arquipélago (Bcs-Aq em vermelho).

Na figura 20 podemos observar os diferentes níveis de expressão de cada população nos contigs ShK-like analisados. Primeiramente o que logo se destaca no gráfico é a diferença de expressão entre as duas populações, com a Bcs-Aq atingindo picos bem mais elevados se comparado a Bcs-SP. Isso pode ocorrer devido a necessidade da população do arquipélago de expressar mais esse tipo de toxina, talvez contra algum tipo de presa mais suscetível à essa família de toxinas. Porém isso é apenas uma teoria, não existindo fatos científicos que comprovem isso.

Dos 19 contigs analisados apenas 5 tiveram expressão identificada em ambas populações, embora com picos bem divergentes entre si. O contig com maior expressão é o c17614_g1_i1, com expressão de 2410,73 pela Bcs-Aq, enquanto o contig com maior expressão pela Bcs-SP é o c64871_g1_i1, com expressão de 65,31. Porém embora a expressão seja muito mais alta na população do arquipélago, o gráfico também indica uma maior diversidade nas isoformas dessas toxinas na Bcs-SP, 13 contigs apresentando expressão nessa população, enquanto que na Bcs-Aq temos 10 contigs.

CONCLUSÃO

Nosso trabalho teve sucesso em sequenciar as duas populações de *Bunodosoma caissarum*. Nosso objetivo de encontrar prováveis novas toxinas também foi alcançado, como demonstrado nos capítulos específicos das famílias Actinoporinas e toxinas bloqueadoras de potássio do tipo I (ShK). Também identificamos outras possíveis famílias de toxinas, portanto um estudo mais aprofundado de cada família também seria necessário.

Em relação a expressão podemos concluir que existe diferenças entre os níveis de toxinas entre as diferentes populações, o que deve ocorrer devido à grande distância existente entre a costa do estado de São Paulo e o Arquipélago São Pedro e São Paulo, o pode resultar em animais diferentes para serem predadas, assim como o próprio ambiente mais eutrofizado da região da costa de São Paulo.

REFERÊNCIAS

1. Frazão B, Vasconcelos V, Antunes A. Sea Anemone (Cnidaria, Anthozoa, Actiniaria) Toxins: An Overview. *Mar Drugs* [Internet]. 2012;10(12):1812–51. Available from: <http://www.mdpi.com/1660-3397/10/8/1812/>
2. Turk T, Kem WR. The phylum Cnidaria and investigations of its toxins and venoms until 1990. *Toxicon* [Internet]. 2009;54(8):1031–7. Available from: <http://linkinghub.elsevier.com/retrieve/pii/S0041010109003407>
3. Rachamim T, Morgenstern D, Aharonovich D, Brekhman V, Lotan T, Sher D. The Dynamically Evolving Nematocyst Content of an Anthozoan, a Scyphozoan, and a Hydrozoan. *Mol Biol Evol* [Internet]. 2015;32(3):740–53. Available from: <http://mbe.oxfordjournals.org/cgi/doi/10.1093/molbev/msu335>
4. Cariello L, de Santis a, Fiore F, Piccoli R, Spagnuolo a, Zanetti L, et al. Calitoxin, a neurotoxic peptide from the sea anemone *Calliactis parasitica*: amino acid sequence and electrophysiological properties. *Biochemistry* [Internet]. 1989;28(6):2484–9. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/2567180>
5. Calton GJ, Burnett JW, Vader W. A study of the nematocyst venoms of the sea anemone, *Bolocera tuediae*. *Toxicon* [Internet]. 1978 [cited 2017 Jul 24];16(5):443–51. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/29365>
6. Beress L, Beress R. Purification of three polypeptides with neuro- and cardiotoxic activity from the sea anemone *Anemonia sulcata*. *Toxicon* [Internet]. 1975;13(5):359–67. Available from: <http://research.bmn.com/medline/search/record?uid=MDLN.76105389>
7. Chi V, Pennington MW, Norton RS, Tarcha EJ, Londono LM, Sims-Fahey B, et al. Development of a sea anemone toxin as an immunomodulator for therapy of autoimmune diseases. Vol. 59, *Toxicon*. 2012. p. 529–46.
8. M. W. Pennington *, ‡, V. M. Mahnir §, I. Khaytin †, I. Zaydenberg †, M. E. Byrnes ‡ and, Kem§ WR. An Essential Binding Surface for ShK Toxin Interaction with Rat Brain Potassium Channels†. 1996 [cited 2017 Mar 19]; Available from: <http://pubs.acs.org/doi/abs/10.1021/bi962463g>
9. Kozlov S, Grishin E. The mining of toxin-like polypeptides from EST database by single residue distribution analysis. *BMC Genomics* [Internet]. 2011;12(1):88. Available from: <http://www.biomedcentral.com/1471-2164/12/88>
10. Honma T, Minagawa S, Nagai H, Ishida M, Nagashima Y, Shiomi K. Novel peptide toxins from acrorhagi, aggressive organs of the sea anemone *Actinia equina*. *Toxicon*. 2005;46(7):768–74.
11. Norton RS. Structures of sea anemone toxins. *Toxicon* [Internet]. 2009;54(8):1075–88. Available from: <http://linkinghub.elsevier.com/retrieve/pii/S0041010109001433>
12. Shapiro BI. Purification of a toxin from tentacles of the anemone *Condylactis gigantea*. *Toxicon* [Internet]. 1968 May [cited 2017 May 30];5(4):253–9. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/4387960>
13. Catterall WA, Beress L. Sea-Anemone Toxin and Scorpion Toxin Share a Common Receptor-Site Associated with Action Potential Sodium Ionophore. *J Biol Chem*. 1978;253(20):7393–6.

14. Bosmans F, Tytgat J. Sea anemone venom as a source of insecticidal peptides acting on voltage-gated Na⁺ channels. *Toxicon* [Internet]. 2007 Mar 15 [cited 2017 Jun 4];49(4):550–60. Available from: <http://linkinghub.elsevier.com/retrieve/pii/S0041010106004429>
15. Šuput D. In vivo effects of cnidarian toxins and venoms. *Toxicon* [Internet]. 2009 Dec 15 [cited 2017 Jun 4];54(8):1190–200. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/19281834>
16. Anderluh G, Maček P. Cytolytic peptide and protein toxins from sea anemones (Anthozoa: Actiniaria). *Toxicon*. 2002;40(2):111–24.
17. Martínez D, Campos A., Pazos F, Alvarez C, Lanio M., Casallanovo F, et al. Properties of St I and St II, two isotoxins isolated from *Stichodactyla helianthus*: a comparison. *Toxicon* [Internet]. 2001 [cited 2017 May 30];39(10):1547–60. Available from: <http://www.sciencedirect.com/science/article/pii/S0041010101001271>
18. Lanio ME, Morera V, Alvarez C, Tejuca M, Gómez T, Pazos F, et al. Purification and characterization of two hemolysins from *Stichodactyla helianthus*. *Toxicon* [Internet]. 2001 [cited 2017 May 30];39(2):187–94. Available from: <http://www.sciencedirect.com/science/article/pii/S0041010100001069>
19. Kawashima Y, Nagai H, Ishida M, Nagashima Y, Shiomi K. Primary structure of echotoxin 2, an actinoporin-like hemolytic toxin from the salivary gland of the marine gastropod *Monoplex echo*. *Toxicon* [Internet]. 2003 Oct [cited 2017 May 30];42(5):491–7. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/14529730>
20. Violette A, Biass D, Dutertre S, Koua D, Piquemal D, Pierrat F, et al. Large-scale discovery of conopeptides and conoproteins in the injectable venom of a fish-hunting cone snail using a combined proteomic and transcriptomic approach. *J Proteomics* [Internet]. 2012 Sep 18 [cited 2017 May 30];75(17):5215–25. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/22705119>
21. Hoang QT, Cho SH, McDaniel SF, Ok SH, Quatrano RS, Shin JS. An actinoporin plays a key role in water stress in the moss *Physcomitrella patens*. *New Phytol* [Internet]. 2009 Oct [cited 2017 May 30];184(2):502–10. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/19674339>
22. Álvarez C, Mancheño JM, Martínez D, Tejuca M, Pazos F, Lanio ME. Sticholysins, two pore-forming toxins produced by the Caribbean Sea anemone *Stichodactyla helianthus*: Their interaction with membranes. *Toxicon* [Internet]. 2009 Dec 15 [cited 2017 May 30];54(8):1135–47. Available from: <http://linkinghub.elsevier.com/retrieve/pii/S0041010109001330>
23. Črnigoj Kristan K, Viero G, Dalla Serra M, Maček P, Anderluh G. Molecular mechanism of pore formation by actinoporins. *Toxicon* [Internet]. 2009 Dec 15 [cited 2017 May 30];54(8):1125–34. Available from: <http://linkinghub.elsevier.com/retrieve/pii/S004101010900138X>
24. Morante K, Bellomio A, Gil-Cartón D, Redondo-Morata L, Sot J, Scheuring S, et al. Identification of a Membrane-bound Prepore Species Clarifies the Lytic Mechanism of Actinoporins. *J Biol Chem* [Internet]. 2016 Sep 9 [cited 2017 May 30];291(37):19210–9. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/27445331>
25. Rojko N, Dalla Serra M, Maček P, Anderluh G. Pore formation by actinoporins, cytolytins from sea anemones. *Biochim Biophys Acta - Biomembr* [Internet]. 2016 Mar [cited 2017 May 30];1858(3):446–56. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/26351738>
26. Meinardi E, Florin-Christensen M, Paratcha G, Azcurra JM, Florin-Christensen J. The molecular basis of the self/nonself selectivity of a coelenterate toxin. *Biochem Biophys Res Commun* [Internet].

1995 Nov 2 [cited 2017 May 30];216(1):348–54. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/7488111>

27. Tejuca M, Anderluh G, Dalla Serra M. Sea anemone cytolytins as toxic components of immunotoxins. *Toxicon* [Internet]. 2009 Dec 15 [cited 2017 May 30];54(8):1206–14. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/19268683>

28. Penton D, Perez-Barzaga V, Diaz I, Reytor ML, Campos J, Fando R, et al. Validation of a mutant of the pore-forming toxin sticholysin-I for the construction of proteinase-activated immunotoxins. *Protein Eng Des Sel* [Internet]. 2011 Jun 1 [cited 2017 May 30];24(6):485–93. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/21296830>

29. Anderluh G, Podlesek Z, Macek P, Maček P. A common motif in proparts of Cnidarian toxins and nematocyst collagens and its putative role. *Biochim Biophys Acta - Protein Struct Mol Enzymol* [Internet]. 2000 Feb 9 [cited 2017 May 30];1476(2):372–6. Available from: <http://www.scopus.com/inward/record.url?eid=2-s2.0-0033962603&partnerID=40&md5=835aa31a20b1766991623a2accc253e3>

30. Castañeda O, Sotolongo V, Amor AM, Stöcklin R, Anderson AJ, Harvey AL, et al. Characterization of a potassium channel toxin from the Caribbean sea anemone *Stichodactyla helianthus*. *Toxicon* [Internet]. 1995 May [cited 2017 Mar 19];33(5):603–13. Available from: <http://linkinghub.elsevier.com/retrieve/pii/004101019500013C>

31. Panyi G, Possani LD, Rodríguez de la Vega RC, Gáspár R, Varga Z. K⁺ channel blockers: novel tools to inhibit T cell activation leading to specific immunosuppression. *Curr Pharm Des* [Internet]. 2006 [cited 2017 Mar 19];12(18):2199–220. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/16787250>

32. Kalman K. ShK-Dap22, a Potent Kv1.3-specific Immunosuppressive Polypeptide. *J Biol Chem* [Internet]. 1998;273(49):32697–707. Available from: <http://www.jbc.org/cgi/doi/10.1074/jbc.273.49.32697>

33. Tudor JE, Pallaghy PK, Pennington MW, Norton RS. Solution structure of ShK toxin, a novel potassium channel inhibitor from a sea anemone. *Nat Struct Biol*. 1996;3:317–20.

34. Zhao Z, Tan L, Dang C, Zhang H, Wu Q, An L. Deep-sequencing transcriptome analysis of chilling tolerance mechanisms of a subnival alpine plant, *Chorispora bungeana*. *BMC Plant Biol* [Internet]. 2012;12:222. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3571968&tool=pmcentrez&rendertype=abstract>

35. Dauplais M, Lecoq A, Song J, Cotton J, Jamin N, Gilquin B, et al. On the convergent evolution of animal toxins. Conservation of a diad of functional residues in potassium channel-blocking toxins with unrelated structures. *J Biol Chem* [Internet]. 1997 Feb 14 [cited 2017 Jul 24];272(7):4302–9. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/9020148>

36. Gimenez-Gallego G, Navia MA, Reuben JP, Katz GM, Kaczorowski GJ, Garcia ML. Purification, sequence, and model structure of charybdotoxin, a potent selective inhibitor of calcium-activated potassium channels. *Proc Natl Acad Sci U S A* [Internet]. 1988 May [cited 2017 Jul 23];85(10):3329–33. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/2453055>

37. Diochot S, Schweitz H, Béress L, Lazdunski M. Sea anemone peptides with a specific blocking activity against the fast inactivating potassium channel Kv3.4. *J Biol Chem*. 1998;273(12):6744–9.

38. Masefski W, Redfield AG, Hare DR, Miller C. Molecular structure of charybdotoxin, a pore-directed inhibitor of potassium ion channels. *Science* [Internet]. 1990 Aug 3 [cited 2017 Mar 19];249(4968):521–4. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/1696395>

39. Yamazaki Y, Morita T. Structure and function of snake venom cysteine-rich secretory proteins. *Toxicon* [Internet]. 2004 Sep [cited 2017 Mar 19];44(3):227–31. Available from: <http://linkinghub.elsevier.com/retrieve/pii/S0041010104002090>
40. Gibbs GM, Scanlon MJ, Swarbrick J, Curtis S, Gallant E, Dulhunty AF, et al. The cysteine-rich secretory protein domain of Tpx-1 is related to ion channel toxins and regulates ryanodine receptor Ca²⁺ signaling. *J Biol Chem* [Internet]. 2006 Feb 17 [cited 2017 Mar 19];281(7):4156–63. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/16339766>
41. Castañeda O, Harvey AL. Discovery and characterization of cnidarian peptide toxins that affect neuronal potassium ion channels. *Toxicon* [Internet]. 2009;54(8):1119–24. Available from: <http://linkinghub.elsevier.com/retrieve/pii/S0041010109001469>
42. Kalman K, Nguyen A, Rauer H, Mahnr V, Paschetto K, Kem WR, et al. ShK-Dap 22 , a Potent Kv1.3-specific Immunosuppressive Polypeptide. *J Biol Chem*. 1998;273(49):32697–707.
43. Pennington MW, Lanigan MD, Kalman K, Mahnr VM, Rauer H, McVaugh CT, et al. Role of disulfide bonds in the structure and potassium channel blocking activity of ShK toxin. *Biochemistry*. 1999;
44. Aneiros A, García I, Martínez JR, Harvey AL, Anderson AJ, Marshall DL, et al. A potassium channel toxin from the secretion of the sea anemone *Bunodosoma granulifera*. Isolation, amino acid sequence and biological activity. *Biochim Biophys Acta* [Internet]. 1993 May 7 [cited 2017 Mar 19];1157(1):86–92. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/8098956>
45. Alessandri-Haber N, Lecoq A, Gasparini S, Grangier-Macmath G, Jacquet G, Harvey AL, et al. Mapping the functional anatomy of BgK on Kv1.1, Kv1.2, and Kv1.3. Clues to design analogs with enhanced selectivity. *J Biol Chem* [Internet]. 1999 Dec 10 [cited 2017 Mar 19];274(50):35653–61. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/10585444>
46. Modica MV, Lombardo F, Franchini P, Oliverio M. The venomous cocktail of the vampire snail *Colubraria reticulata* (Mollusca, Gastropoda). *BMC Genomics* [Internet]. 2015;16(1):441. Available from: <http://www.biomedcentral.com/1471-2164/16/441>
47. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods* [Internet]. 2012 Mar 4 [cited 2017 May 30];9(4):357–9. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/22388286>
48. Robinson MD, Oshlack A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol* [Internet]. 2010 [cited 2017 May 30];11(3):R25. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/20196867>
49. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* [Internet]. 2010 Jan 1 [cited 2017 May 30];26(1):139–40. Available from: <https://academic.oup.com/bioinformatics/article-lookup/doi/10.1093/bioinformatics/btp616>
50. de Oliveira JS, Zaharenko AJ, de Freitas JC, Konno K, de Andrade SA, Portaro FCV, et al. Caissarolysin I (Bcs I), a new hemolytic toxin from the Brazilian sea anemone *Bunodosoma caissarum*: Purification and biological characterization. *Biochim Biophys Acta - Gen Subj* [Internet]. 2006 Mar [cited 2017 May 30];1760(3):453–61. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/16458433>
51. Kozak M. An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. *Nucleic Acids Res* [Internet]. 1987 Oct 26 [cited 2017 May 30];15(20):8125–48. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/3313277>

SOBRE OS AUTORES

LEANDRO AUGUSTO FREIRE BORALLI - nascido em 1985, é biólogo desde 2009, formado pela Universidade São Camilo. Antes mesmo de ingressar na Universidade ele já se interessava por genética e o interesse apenas aumentou durante o curso. Após a formação ele se interessou em estudar no Instituto Butantan através do Programa de Aprimoramento Profissional (PAP). Através desse programa Leandro conseguiu entrar como aprimorando no Laboratório de Genética do Instituto Butantan, comandado pela Dra. Irina Kerkis. Durante esses 2 anos de PAP o aluno já começou seus estudos de toxinas de anêmonas do mar (principalmente da *Bunodosoma caissarum*) e também começou a aprender sobre bioinformática. Todos esse tempo de PAP serviu de base para o seu mestrado, onde ele foi aluno da Universidade Federal de São Paulo (pelo programa Biologia Estrutural e Funcional), embora ainda estudasse presencialmente no mesmo laboratório no Instituto Butantan.

URSULA CASTRO DE OLIVEIRA - Endereço para acessar este CV: <http://lattes.cnpq.br/0612558794411094>. ID Lattes: 0612558794411094. Graduação em Farmácia e Bioquímica pela Universidade de São Paulo (1995-2001). Iniciação Científica em Biologia Molecular no Instituto Butantan (1997-2002). Doutorado direto (2005-2011) em Ciências (Biologia da Relação Patógeno-Hospedeiro) pelo Instituto de Ciências Biomédicas da Universidade de São Paulo. Pós-doutorado no Instituto Butantan, no projeto “TOXINOLOGIA COMPARATIVA DOS ESCORPIÕES DE INTERESSE MÉDICO DO BRASIL” (2012 - atual). Em 2019 contratada pela Fundação Butantan para integrar o Controle de Qualidade Biotecnológico para o desenvolvimento e implementação de novas tecnologias. (Texto informado pelo autor)

MILTON YUTAKA NISHIYAMA JUNIOR - Endereço para acessar este CV: <http://lattes.cnpq.br/2462633063143454>. ID Lattes: 2462633063143454. Doutor em Bioinformática pela Universidade de São Paulo e pesquisador Associado do Laboratório de Toxinologia Aplicada (LETA) e Coordenador do Centro de Bioinformática e Biologia Computacional do Instituto Butantan. Atua nas áreas de aprendizagem de máquina, redes neurais, big data e descoberta do conhecimento com ênfase em Sistemas da Informação e Biologia Molecular envolvendo a predição de alvos moleculares, peptídeos bioativos e descoberta de novas toxinas, baseado na integração de dados de larga escala de multi-ômicas de *Homo sapiens* e organismos não-modelo, como animais peçonhentos e organismos patogênicos relacionados a saúde pública, para entendimento dos mecanismos moleculares que governam os sistemas biológicos. (Texto informado pelo autor)

GANDHI RÁDIS BAPTISTA - Bolsista de Produtividade em Pesquisa do CNPq - Nível 2. Endereço para acessar este CV: <http://lattes.cnpq.br/5322192969141529>. ID Lattes: 5322192969141529. Graduado em Farmácia Bioquímica pela Universidade de São Paulo (1993), mestre em Tecnologia Bioquímico-Farmacêutica pela Universidade de São Paulo (1996) e doutor em Bioquímica pela Universidade de São Paulo (2001). Professor Adjunto do Departamento de Bioquímica da Universidade Federal de Pernambuco (2005-2008).

Atualmente, Professor Associado do Instituto de Ciências do Mar da Universidade Federal do Ceará. Membro da Sociedade Americana de Química (American Chemical Society, ACS), desde 2010. Os principais interesses de investigação compreendem os peptídeos e as proteínas farmacologicamente ativos provenientes de organismos marinhos e terrestres; os receptores celulares e seus ligantes; a interação molecular; a tecnologia do DNA recombinante; o diagnóstico molecular e seus insumos; o direcionamento intracelular e a compartimentalização de polipeptídeos. (Texto informado pelo autor)

IRINA KERKIS - Bolsista de Produtividade em Pesquisa do CNPq - Nível 1D. Endereço para acessar este CV: <http://lattes.cnpq.br/4302687618153569>. ID Lattes: 4302687618153569. Possui graduação em Biologia e Química pela Universidade Federal de Tomsk (1978), mestrado em Biologia e Química pela Universidade Federal de Tomsk (1981) e doutorado em Citogenética pelo Instituto de Citologia e Genética (1989). Reconhecimento pela Universidade de São Paulo de Título de Doutor em Ciências - área Biologia (Genética) em 2004. Obteve Livre-Docência em Ciências Biológicas pela Academia de Ciências da Rússia em 1994. Atualmente ela é a diretora do Laboratório de Genética do Instituto Butantan e Pesquisadora nível VI do Instituto Butantan. Ela tem uma vasta experiência, liderança, e treinamento em Biologia Celular, especificamente na área de obtenção, caracterização e produção em alta escala de diversos tipos das células tronco embrionárias e pluripotentes induzidas, assim como adultas - mesenquimais e células tronco germinativas. Seu foco atual é a medicina translacional utilizando células tronco imaturas de polpa dentária humanas em doenças neurodegenerativas e hematopoiéticas. Além disso, ela atua na área de toxinas - peptídeos penetrantes de células, tais como a crotamina isolada do veneno das serpentes brasileiras e utilizada como o marcador de células proliferamente ativas, de ciclo celular, assim como carreador de DNA/ RNA ou outras moléculas biologicamente ativas para o interior das células. Ela publicou mais de 90 artigos científicos, 20 capítulos de livros e produziu cinco patentes, três delas já foram concedidas. Ela também atuou como palestrante principal em eventos científicos nacionais e internacionais e recebeu vários prêmios. (Texto informado pelo autor)

ÁLVARO ROSSAN DE BRANDÃO PRIETO DA SILVA - Endereço para acessar este CV: <http://lattes.cnpq.br/2277174467541698>. ID Lattes: 2277174467541698. Possui graduação em Medicina Veterinária pela Universidade de São Paulo (1991) e doutorado em Ciências Biológicas (Bioquímica) pela Universidade de São Paulo (2002). Atualmente é pesquisador científico do Instituto Butantan, onde é vice-diretor do Laboratório de Genética. Atua na área de Bioquímica, com ênfase em Biologia Molecular e Química de Proteínas. Os principais interesses de investigação compreendem: os polipeptídeos bioativos, toxinas animais e seu alvos, receptores celulares e seus ligantes; a interação molecular; a tecnologia do DNA recombinante: clonagem e expressão, bibliotecas de cDNA de organismos peçonhentos, RT-PCR. Análise e produção de antígenos e imunobiológicos. Atuando principalmente nos seguintes temas: toxinas, neurotoxinas, clonagem purificação de proteínas, antígenos recombinantes e sorologia. *Crotalus*, *Micrurus*, Cnidários, Himenópteros. (Texto informado pelo autor)

pelo autor)

INÁCIO DE LOIOLA MEIRELLES JUNQUEIRA DE AZEVEDO - Bolsista de Produtividade em Pesquisa do CNPq - Nível 2. Endereço para acessar este CV: <http://lattes.cnpq.br/9499590129608919>. ID Lattes: 9499590129608919. Possui graduação em Ciências Biológicas pelo Instituto de Biociências da Universidade de São Paulo (1998), doutorado em Biologia e Genética pela Universidade de São Paulo (2003) e pós-doutorado na área biotecnológica pelo Instituto Butantan (2004). Atualmente é Pesquisador Científico do Instituto Butantan, sendo diretor do Laboratório de Toxinologia Aplicada. Tem experiência nas áreas de genômica, transcriptômica e biologia molecular em geral, com ênfase em sequenciamento de DNA em larga escala. Atua principalmente nos seguintes temas: análises ômicas de venenos animais e de seus efeitos, evolução de venenos de serpentes, bioquímica de toxinas e biotecnologia. (Texto informado pelo autor)

ANDRE JUNQUEIRA ZAHARENKO - Endereço para acessar este CV: <http://lattes.cnpq.br/3955659940007123>. ID Lattes: 3955659940007123. Graduado em Ciências Biológicas pela USP (2000) e Doutor em Fisiologia Geral pelo Instituto de Biociências da USP em 2007. Atuação na área de Toxinologia Marinha, principalmente em neurotoxinas de anêmonas do mar. Dedicou-se ao isolamento e determinação estrutural de peptídeos e compostos de baixa massa molecular (utilizando técnicas cromatográficas e de biologia molecular, bem como espectrometria de massas) e da caracterização dos efeitos biológicos, através de experimentos eletrofisiológicos. Possui em andamento colaborações com outros grupos estrangeiros visando a determinação estrutural e análise de estrutura-função de peptídeos que interagem com canais iônicos, bem como a investigação de aspectos biofísicos que regem a atividade de canais modulados por toxinas. Realizou de 2011-2015 (pós-doc) a análise transcriptômica de toxinas destes organismos através de técnicas de sequenciamento de DNA de nova geração. Atualmente atuando na indústria farmacêutica (Novartis Corporation) na posição de MSL (Medical Scientific Liaison), na área de oncologia (tumores sólidos). Responsável pela prospecção de Centros de Pesquisa Clínica e implementação de protocolos de pesquisa clínica em diversos hospitais e serviços públicos pelo país. Bem como discussão de racional molecular de novas drogas no segmento, com stakeholders externos e internos e líderes de opinião em oncologia. (Texto informado pelo autor)

ANÁLISE DAS SEQUÊNCIAS DE NUCLEOTÍDEOS DE NOVOS SUBTIPOS DE TOXINAS E PROTEÍNAS DE ANÊMONAS DO MAR POR SEQUENCIAMENTO DE RNA DE NOVA GERAÇÃO

 www.atenaeditora.com.br
 contato@atenaeditora.com.br
 [@atenaeditora](https://www.instagram.com/atenaeditora)
 www.facebook.com/atenaeditora.com.br


Ano 2022

ANÁLISE DAS SEQUÊNCIAS DE NUCLEOTÍDEOS DE NOVOS SUBTIPOS DE TOXINAS E PROTEÍNAS DE ANÊMONAS DO MAR POR SEQUENCIAMENTO DE RNA DE NOVA GERAÇÃO

 www.atenaeditora.com.br

 contato@atenaeditora.com.br

 [@atenaeditora](https://www.instagram.com/atenaeditora)

 www.facebook.com/atenaeditora.com.br