



UNIVERSIDADE FEDERAL DE SANTA CATARINA
CENTRO DE CIÊNCIAS BIOLÓGICAS
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOTECNOLOGIA E BIOCÊNCIAS

Gustavo Rocha

Prospecção e caracterização molecular de novas defensinas do tipo CSa β (*Cg*-Defs) e big defensinas (*Cg*-BigDefs) na ostra *Crassostrea gigas*

Florianópolis

2020

Gustavo Rocha

Prospecção e caracterização molecular de novas defensinas do tipo CSa β (*Cg*-Defs) e big defensinas (*Cg*-BigDefs) na ostra *Crassostrea gigas*

Dissertação submetida ao Programa de Pós-Graduação em Biotecnologia e Biociências da Universidade Federal de Santa Catarina para a obtenção do título de Mestre em Biotecnologia e Biociências.

Orientador: Prof. Dr. Rafael Diego da Rosa.

Florianópolis

2020

Ficha de identificação da obra elaborada pelo autor,
através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

Rocha, Gustavo

Prospecção e caracterização molecular de novas defensinas do tipo CSaB (Cg-Defs) e big defensinas (Cg-BigDefs) na ostra *Crassostrea gigas* / Gustavo Rocha ; orientador, Rafael Diego da Rosa, 2020.

58 p.

Dissertação (mestrado) - Universidade Federal de Santa Catarina, Centro de Ciências Biológicas, Programa de Pós Graduação em Biotecnologia e Biociências, Florianópolis, 2020.

Inclui referências.

1. Biotecnologia e Biociências. 2. Imunidade de invertebrados. 3. Ostra do Pacífico. 4. Peptídeos antimicrobianos. 5. Fenômeno de presença/ausência de genes (PAV). I. Diego da Rosa, Rafael. II. Universidade Federal de Santa Catarina. Programa de Pós-Graduação em Biotecnologia e Biociências. III. Título.

Gustavo Rocha

Prospecção e caracterização molecular de novas defensinas do tipo CSa β (Cg-Defs) e big defensinas (Cg-BigDefs) na ostra *Crassostrea gigas*

O presente trabalho em nível de mestrado foi avaliado e aprovado por banca examinadora composta pelos seguintes membros:

Prof. Dr. Rafael Diego da Rosa
Universidade Federal de Santa Catarina

Profa. Dra. Patricia Hermes Stoco
Universidade Federal de Santa Catarina

Profa. Dra. Franceli Rodrigues Kulcheski
Universidade Federal de Santa Catarina

Certificamos que esta é a **versão original e final** do trabalho de conclusão que foi julgado adequado para obtenção do título de mestre em biotecnologia e biociências.

Prof. Dr. Glauber Wagner
Coordenador do Programa de Pós-Graduação

Prof. Dr. Rafael Diego da Rosa
Orientador

Florianópolis, 29 de junho de 2020.

Dedico este trabalho a todas as pessoas que se preocupam em promover o bem e a desenvolver uma sociedade melhor.

AGRADECIMENTOS

Gostaria de agradecer inicialmente ao professor Dr. Rafael Diego da Rosa, que além de ser um profissional que eu admiro muito, é também um orientador extremamente responsável e preocupado com o desenvolvimento profissional de seus alunos. Sou muito grato de fazer parte de seu grupo de pesquisa e de ter tido a oportunidade de desenvolver esta dissertação sob sua liderança.

Também sou grato à professora Dra. Luciane Maria Perazzolo, por todas as dicas durante a realização de meus experimentos e apresentações, além dos ensinamentos acerca do bom funcionamento do laboratório.

Além disso, eu não poderia deixar de agradecer à doutora Mariana Pilotto e aos mestres Cairé Barreto, Gabriel Matos e Natanael Farias pelas infinitas ajudas no decorrer do meu mestrado. Com certeza, os ensinamentos que eu aprendi com vocês me enriqueceram muito e me tornaram um profissional melhor capacitado.

Aos meus grandes amigos de laboratório Flávio, Nicolas e Talita, pois além de todo o suporte, vocês ajudaram a tornar os meus dias mais felizes.

Ao aluno de iniciação científica Leonardo Bortoli, pela parceria durante o processo coorientação. Sem dúvidas, foi um momento de muito aprendizado e responsabilidade em minha formação acadêmica.

À professora Dra. Ilíada Rainha, pelos ensinamentos durante o meu estágio de docência, pois foi uma experiência incrível que certamente vai pautar a minha atuação profissional futura.

Aos colegas de corredor, pelas inúmeras trocas de experiências, risadas e cafés na copa.

A todos os professores que participaram de maneira positiva em minha formação.

Aos laboratórios multiusuários LAMEB e CEBIME e todo o seu corpo técnico, pelo auxílio na realização dos experimentos.

Aos funcionários da limpeza e segurança da UFSC, por toda a colaboração em manter a estrutura de nosso ambiente de trabalho.

À Pós-Graduação em Biotecnologia e Biociências e às agências financiadoras CAPES e CNPq, pelo suporte financeiro e técnico para o êxito de meu trabalho.

E por fim, à minha família, por todo o apoio e amor envolvido. Sou muito grato por tudo o que vocês representam para mim. Obrigado!

RESUMO

As defensinas são um dos peptídeos antimicrobianos (AMPs) mais abundantes e bem distribuídos entre os seres vivos. Elas compreendem famílias multigênicas de peptídeos estabilizados por ligações dissulfeto e que apresentam atividade antimicrobiana contra um amplo espectro de microrganismos. Ostras *Crassostrea gigas* são organismos extremamente polimórficos e que possuem uma vasta diversidade de AMPs, incluindo defensinas do tipo CS $\alpha\beta$ e big defensinas. Embora possuam funções análogas, essas defensinas não são filogeneticamente relacionadas. Nesse contexto, o presente estudo investigou a diversidade molecular de defensinas do tipo CS $\alpha\beta$ (*Cg-Defs*) e big defensinas (*Cg-BigDefs*) em *C. gigas*. Por meio de buscas *in silico*, duas novas defensinas do tipo CS $\alpha\beta$ (nomeadas *Cg-Defh3* e *Cg-Defh4*) e três big defensinas (nomeadas *Cg-BigDef4*, *Cg-BigDef5* e *Cg-BigDef6*) foram encontradas em *C. gigas*. As duas novas *Cg-Defs* apresentam um peptídeo sinal, seguido de um pró-domínio e de um peptídeo maduro estabilizado por três ligações dissulfeto, possuindo uma alta identidade aminoacídica com defensinas do tipo CS $\alpha\beta$ de insetos. Além disso, as *Cg-Defh3* e *Cg-Defh4* são as primeiras defensinas de caráter aniônico identificadas em *C. gigas*. As três novas sequências de big defensinas são compostas por um peptídeo sinal, seguido de um pró-domínio e de um peptídeo maduro, possuindo a mesma assinatura molecular encontrada nas demais *Cg-BigDefs*. Ademais, os resultados também revelaram que as novas *Cg-Defs* e *Cg-BigDefs* apresentaram uma expressão muito baixa, ou até mesmo ausente, ainda que após a infecção experimental por patógenos de interesse aquícola. Esse resultado está possivelmente relacionado ao fato dessas defensinas estarem sujeitas ao fenômeno genético de PAV (*gene presence/absence variation*), no qual um gene está presente em alguns, mas não em todos os indivíduos de uma mesma espécie. Em conjunto, os resultados demonstram que as defensinas de *C. gigas* são altamente diversas em nível de sequências, variedade de assinaturas moleculares e características bioquímicas, abrindo assim perspectivas para estudos futuros que elucidem a participação desses efetores na imunidade de moluscos bivalves.

Palavras-chave: imunidade de invertebrados, ostra do Pacífico, peptídeos antimicrobianos, fenômeno de presença/ausência de genes (PAV).

ABSTRACT

Defensins are one of the most abundant and well-distributed antimicrobial peptides (AMPs) among living organisms. They comprise multigenic families of disulfide-stabilized peptides that exhibit antimicrobial activity against a wide spectrum of microorganisms. *Crassostrea gigas* oysters are highly polymorphic organisms showing a broad diversity of AMPs, including both CS $\alpha\beta$ -containing defensins and big defensins. Although they have analogous functions, these defensins are not phylogenetically related. In this regard, the present study investigated the molecular diversity of CS $\alpha\beta$ -containing defensins (*Cg*-Defs) and big defensins (*Cg*-BigDefs) in *C. gigas*. By using *in silico* searches, two novel CS $\alpha\beta$ -containing defensins (named *Cg*-Defh3 and *Cg*-Defh4) and three big defensins (named *Cg*-BigDef4, *Cg*-BigDef5 and *Cg*-BigDef6) were found in *C. gigas*. The two *Cg*-Defs possess a signal peptide, followed by a pro-domain and a mature peptide stabilized by three disulfide bonds, and showed a high amino acid identity with insect CS $\alpha\beta$ -containing defensins. Moreover, *Cg*-Defh3 and *Cg*-Defh4 are the first anionic defensins to be identified in *C. gigas*. The three novel big defensins have a signal peptide, followed by a pro-domain and a mature peptide, and show the same molecular signature found in the other *Cg*-BigDefs. In addition, results revealed that the novel *Cg*-Defs and *Cg*-BigDefs showed a very low or, even absent, gene expression, even after an experimental infection with pathogens of aquaculture interest. This result is probably related to the fact that these defensins can be affected by the genetic phenomenon of PAV (gene presence/absence variation), in which a gene is present in some, but not in all individuals belonging to the same species. Altogether, results showed that defensins from *C. gigas* are highly diverse in sequences, molecular signatures and biochemical features, opening new perspectives for further studies on the participation of these effectors in the immunity of bivalve mollusks.

Keywords: invertebrate immunity, Pacific oyster, antimicrobial peptides, gene presence/absence variation phenomenon (PAV).

LISTA DE FIGURAS

| | |
|--|----|
| Figura 1 – Orientação das ligações dissulfeto nas estruturas de <i>cis</i> - e <i>trans</i> -defensinas e sua diversidade estrutural | 16 |
| Figura 2 – Distribuição das <i>cis</i> - e <i>trans</i> -defensinas nos Eucariotos | 17 |
| Figura 3 – Estrutura do precursor, alinhamento dos peptídeos maduros e estrutura gênica das defensinas do tipo CSa β de <i>C. gigas</i> e estrutura tridimensional da <i>Cg</i> -Defm | 20 |
| Figura 4 – Estrutura do precursor, alinhamento dos peptídeos maduros e estrutura gênica das big defensinas de <i>C. gigas</i> e estrutura tridimensional da <i>Cg</i> -BigDef1 | 23 |
| Figura 5 – Delineamento da infecção experimental | 27 |
| Figura 6 – Eletroforese em gel de agarose (1,5%) corado com brometo de etídio representando a amplificação dos transcritos para as novas <i>Cg</i> -Defs em amostras de hemócitos de ostras <i>C. gigas</i> | 35 |
| Figura 7 – Estrutura do precursor, sequências nucleotídica e aminoacídica deduzidas e estruturas gênicas das novas <i>Cg</i> -Defs | 36 |
| Figura 8 – Alinhamento das defensinas do tipo CSa β de <i>Crassostrea gigas</i> : membros contendo oito resíduos de cisteínas (<i>Cg</i> -Defm; <i>Cg</i> -Defh1; <i>Cg</i> -Defh2 e <i>Cg</i> -Defdg) e as novas defensinas identificadas neste estudo contendo seis resíduos conservados de cisteínas (<i>Cg</i> -Defh3 e <i>Cg</i> -Defh4) | 37 |
| Figura 9 – Modelamento tridimensional das <i>Cg</i> -Defs | 38 |
| Figura 10 – Alinhamento dos peptídeos maduros de defensinas do tipo CSa β presentes em diferentes táxons | 39 |
| Figura 11 – Estrutura do precursor, sequências nucleotídica e aminoacídica deduzidas e estruturas gênicas das novas <i>Cg</i> -BigDefs | 41 |
| Figura 12 – Alinhamento do peptídeo maduro e esquema ilustrativo dos domínios das big defensinas de <i>Crassostrea gigas</i> (<i>Cg</i> -BigDef1, <i>Cg</i> -BigDef2, <i>Cg</i> -BigDef3, <i>Cg</i> -BigDef4, <i>Cg</i> -BigDef5 e <i>Cg</i> -BigDef6) | 42 |
| Figura 13 – Modelamento tridimensional das <i>Cg</i> -BigDefs | 42 |
| Figura 14 – Alinhamento dos peptídeos maduros de big defensinas presentes em diferentes táxons | 43 |
| Figura 15 – Análise da expressão de <i>Cg</i> -BP1 nos hemócitos de ostras experimentalmente infectadas com a bactéria <i>Vibrio tasmaniensis</i> LGP32 e com o herpesvírus de ostra (OsHV-1 μ Var) | 45 |
| Figura 16 – Padrão da variação por presença/ausência (PAV) dos genes relacionados as novas defensinas em indivíduos de <i>C. gigas</i> provenientes do Brasil e da França | 46 |

LISTA DE TABELAS

| | |
|--|----|
| Tabela 1 – Sequências nucleotídicas dos iniciadores utilizados neste trabalho | 30 |
| Tabela 2 – Análise da expressão das novas defensinas em ostras experimentalmente infectadas com a bactéria <i>Vibrio tasmaniensis</i> LGP32 e com o herpesvírus de ostra (OsHV-1 μ Var) | 44 |

LISTA DE ABREVIATURAS E SIGLAS

AITDs – *Ancient invertebrate-type defensins*
AMPs – Peptídeos antimicrobianos
AS – Iniciador antissenso
BPI – Proteínas indutoras da permeabilidade bacteriana
cDNA – DNA complementar
CDS – Região codificante
CITDs – *classical insect-type defensins*
CNV – Variações no número de cópias gênicas
CS $\alpha\beta$ – *Cysteine-stabilized α -helix β -sheet motif*
Cq – Ciclo de quantificação
dNTPs – Deoxinucleotídeos trifosfatados
EDTA – Ácido etilenodiamino tetra-acético
EST – *Expressed Sequences Tags*
gDNA – DNA genômico
GMQE – *global model quality estimation*
IMD – *immune deficiency*
kDa – Quilodalton
LPS – Lipopolissacarídeo
MAMPs – Padrões moleculares associados a microrganismos
MM – Massa molecular
NCBI – *National Center for Biotechnology Information*
NF- κ B – *nuclear factor-kappa B*
PAV – Variação de presença/ausência gênica
pb – Pares de base
PCR – Reação em Cadeia da Polimerase
PDB – *Protein Data Bank*
pI – Ponto isoelétrico teórico
qPCR – PCR quantitativa em tempo real
RT-PCR – Transcrição reversa seguida de reação em cadeia da polimerase
S – Iniciador senso
SDS – Dodecil sulfato de sódio
T_m – Temperatura de desnaturação (*melting temperature*)
TSA – *Transcriptome Shotgun Assembly*
UFC – Unidade formadora de colônia

LISTA DE SÍMBOLOS

°C – Graus Celsius
C+ – Controle positivo
C- – Controle negativo
< – Menor
% – Porcentagem
 α – Letra grega alfa
 β – Letra grega beta
 Δ – Letra grega delta
 θ – Letra grega teta
 μ – Micro
✓ – Sinal de visto

SUMÁRIO

| | |
|--|----|
| 1. INTRODUÇÃO | 14 |
| 1.1. Peptídeos antimicrobianos | 14 |
| 1.2. Defensinas | 15 |
| 1.3. As defensinas de ostras | 18 |
| 1.3.1. Defensinas do tipo CS$\alpha$$\beta$ (Cg-Defs) | 19 |
| 1.3.2. Big defensinas (Cg-BigDefs) | 22 |
| 2. OBJETIVOS | 25 |
| 2.1. Objetivo geral | 25 |
| 2.2. Objetivos específicos | 25 |
| 3. MATERIAIS E MÉTODOS | 26 |
| 3.1. Prospecção <i>in silico</i> de sequências | 26 |
| 3.2. Animais, infecção experimental e coleta dos tecidos | 26 |
| 3.3. Extração de ácidos nucleicos e síntese de cDNA | 27 |
| 3.4. Escolha dos iniciadores | 29 |
| 3.5. Amplificação das sequências de interesse | 29 |
| 3.6. Clonagem molecular e sequenciamento nucleotídico | 31 |
| 3.7. Análise de sequências | 32 |
| 3.8. Análises de expressão gênica | 32 |
| 3.9. Análises de presença/ausência de genes (PAV) | 33 |
| 4. RESULTADOS | 34 |
| 4.1. Identificação de novas defensinas na ostra do Pacífico <i>Crassostrea gigas</i> | 34 |
| 4.2. Caracterização de novas defensinas do tipo CS$\alpha$$\beta$ (Cg-Defs) | 34 |
| 4.3. Caracterização de novas big defensinas (Cg-BigDefs) | 39 |
| 4.4. Análise transcricional das novas defensinas de <i>C. gigas</i> | 43 |
| 4.5. Investigação do fenômeno de presença/ausência de genes (PAV) | 45 |
| 5. DISCUSSÃO | 47 |
| 6. CONCLUSÃO | 52 |
| REFERÊNCIAS | 53 |
| APÊNDICE 1 – Sequências de entrada utilizadas (<i>queries</i>) | 57 |
| ANEXO 1 – Artigo científico em coautoria (relacionado à dissertação de Mestrado) | 58 |

1. INTRODUÇÃO

1.1. Peptídeos antimicrobianos

Peptídeos antimicrobianos (*Antimicrobial Peptides*; AMPs) são moléculas produzidas por praticamente todos os seres vivos, sendo considerados efetores de defesa evolutivamente conservados (HANCOCK; SAHL, 2006). Os AMPs podem ser codificados por genes, ou gerados a partir da hidrólise de moléculas precursoras (BULET; STÖCKLIN; MENIN, 2004) e são classicamente descritos como moléculas anfipáticas e catiônicas de até 10 kDa, possuindo uma região hidrofílica rica em aminoácidos carregados positivamente e outra região hidrofóbica (BROGDEN, 2005). Esses peptídeos compõem um grupo polifilético, que mesmo não possuindo correlações filogenéticas, apresentam funções análogas (DESTOUMIEUX-GARZÓN et al., 2016), atuando na inibição ou morte, de maneira direta ou indireta, de uma ampla gama de bactérias, fungos e vírus (GUANÍ-GUERRA et al., 2010). O primeiro peptídeo com atividade antimicrobiana foi descoberto e isolado em 1939, da espécie de bactéria Gram-positiva *Brevibacillus brevis*, ao apresentar atividade antimicrobiana *in vitro* contra bactérias Gram-positivas (DUBOS, 1939). Porém, foi somente após o fim da chamada “era de ouro dos antibióticos”, no final dos anos 60, que inúmeros outros AMPs foram identificados em diversos táxons, desde bactérias até plantas, fungos e animais (PHOENIX; DENNISON; HARRIS, 2013).

Devido a sua ampla distribuição, os AMPs possuem uma imensa diversidade bioquímica e estrutural, o que permite a eles apresentarem mecanismos de ação bastante diversificados (BULET; STÖCKLIN; MENIN, 2004). Alguns AMPs interagem diretamente com a membrana celular dos microrganismos por meio de interações eletrostáticas entre a sua região catiônica e a porção aniônica dos fosfolipídeos microbianos. Essa interação permite a inserção da região hidrofóbica do peptídeo na membrana, gerando desestabilização ou até mesmo a formação de poros, causando o extravasamento do conteúdo intracelular (BROGDEN, 2005). Ainda existem outros tipos de AMPs que são interiorizados e que podem se ligar a biomoléculas (DNA, RNA, proteínas) ou inibir vias metabólicas fundamentais do microrganismo (LEI et al., 2019; WILMES et al., 2011). Além disso, os AMPs apresentam funções biológicas, como a regulação da microbiota de animais, modulação de efetores imunológicos, indução de repostas pró-inflamatórias, indução da secreção de citocinas, entre outras (LEI et al., 2019).

Diversas características dos AMPs os tornam candidatos ao desenvolvimento de novos antibióticos de origem natural. Em se tratando de moléculas de origem proteica, o modo de ação dos AMPs, ao interagirem com componentes fundamentais de microrganismos, faz com que o

surgimento de cepas resistentes seja minimizado (BROGDEN, 2005). Além disso, a especificidade da interação dos AMPs a componentes moleculares exclusivos de microrganismos (SCHMITT et al., 2010a) faz com que essas moléculas sejam, na maioria das vezes, inertes em relação a células animais. Essa característica lhes confere um grande potencial para o uso e o desenvolvimento de novos fármacos humanos e veterinários. Até o ano de 2019, mais de 5.000 AMPs haviam sido registrados na plataforma DRAMP (*Data repository of antimicrobial peptides*) Database (<http://dramp.cpu-bioinfor.org/>), dos quais aproximadamente 70 desses peptídeos estão em estágios de testes clínicos para o desenvolvimento de novos fármacos (KOO; SEO, 2019). Devido a características proteicas dos AMPs, a maioria dos estudos clínicos enfocam na utilização tópica dessas moléculas, para o tratamento de lesões cutâneas, queimaduras, dermatites, acnes e infecções fúngicas em humanos (KOO; SEO, 2019). Entretanto, peptídeos como a ramoplanina, isolado de bactérias do gênero *Actinoplanes*, encontram-se em estágio III de estudos clínicos para o tratamento de infecções gastrointestinais por *Enterococcus* sp. e *Clostridium difficile*, por meio de administração oral (KOO; SEO, 2019).

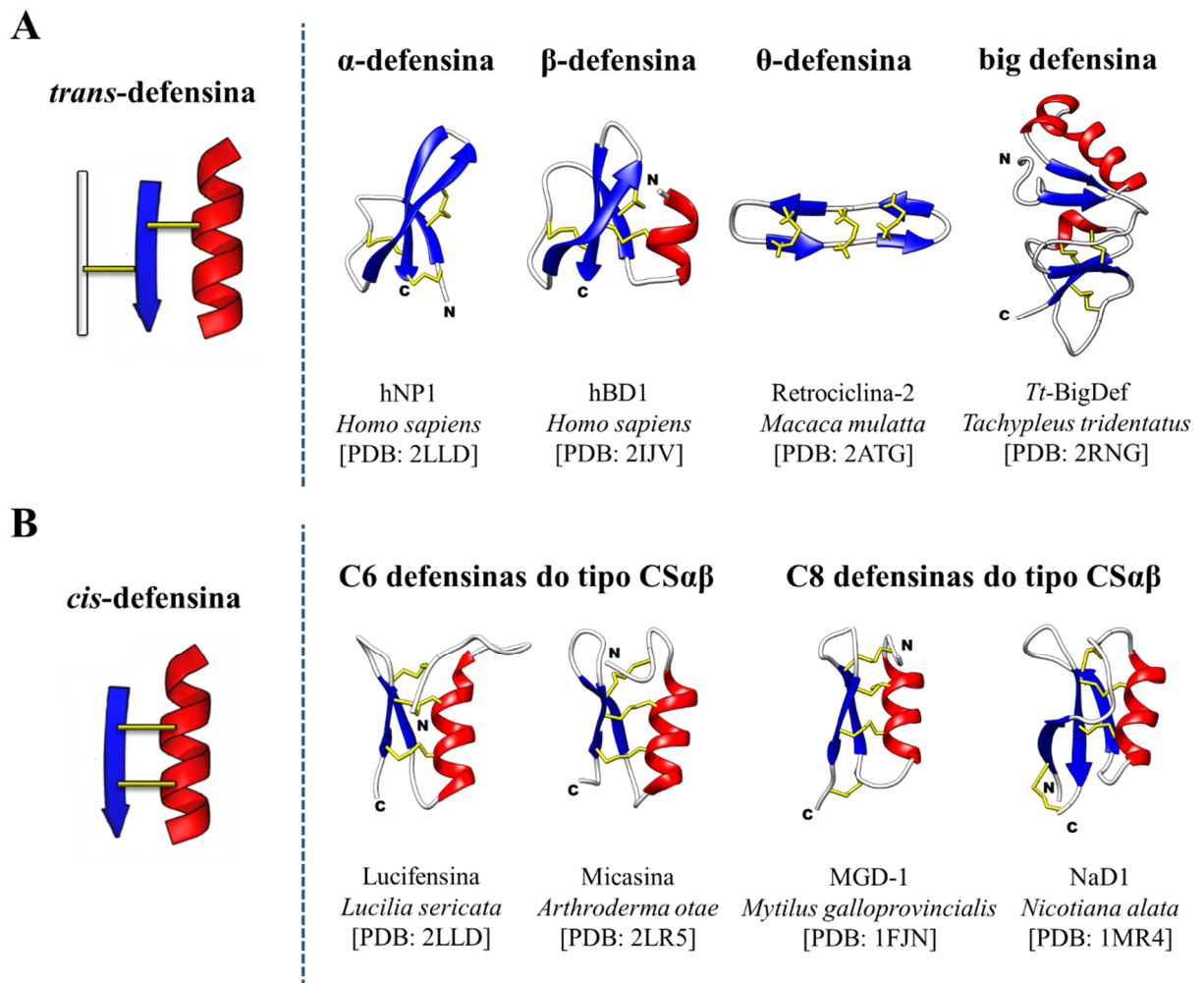
1.2. Defensinas

As defensinas compreendem um dos mais abundantes e amplamente distribuídos AMPs descritos até o momento, sendo identificadas em fungos, plantas e animais vertebrados e invertebrados (SCHMITT et al., 2012). As defensinas são peptídeos geralmente catiônicos e anfipáticos que contêm de seis a doze resíduos de cisteínas, formando de três a seis pontes dissulfeto, que estabilizam a sua conformação tridimensional (SCHMITT; ROSA; DESTOUMIEUX-GARZÓN, 2016; SHAFEE et al., 2016). A distribuição e o pareamento dos resíduos de cisteínas permitem classificá-las em duas superfamílias: *cis*-defensinas e *trans*-defensinas (Figura 1) (SHAFEE et al., 2016).

Apesar de bastante diversas, todas as defensinas possuem uma estrutura secundária contendo folhas- β estabilizadas por ligações dissulfeto, além de possuírem funções análogas. No entanto, as *cis*- e *trans*-defensinas não apresentam qualquer correlação filogenética (GERDOL et al., 2020a; SHAFEE et al., 2016). As *trans*-defensinas apresentam duas pontes dissulfeto que apontam para lados opostos, ligando-se a elementos diferentes da estrutura secundária do peptídeo (SHAFEE et al., 2016). Essas defensinas são classificadas através de suas assinaturas aminoacídicas, além dos arranjos formados pelas pontes dissulfeto entre cisteínas (DAS et al., 2010). Dentre os representantes das *trans*-defensinas estão as defensinas cíclicas dimerizadas (θ -defensinas: Cys_{1a-3a}Cys_{2a-2b}Cys_{1b-3b}) e as cíclicas com extremidades

abertas (α -defensinas: Cys₁₋₆Cys₂₋₄Cys₃₋₅; β -defensinas e big defensinas: Cys₁₋₅Cys₂₋₄Cys₃₋₆) (SHAFEE et al., 2016). Por outro lado, as *cis*-defensinas apresentam duas pontes dissulfeto ligadas paralelamente entre uma folha- β e uma α -hélice, sendo conhecidas como “defensinas do tipo CS $\alpha\beta$ ” (*cysteine-stabilized α -helix β -sheet motif*) (Figura 1) (SCHMITT; ROSA; DESTOUMIEUX-GARZÓN, 2016; SHAFEE et al., 2016).

Figura 1. Orientação das ligações dissulfeto nas estruturas de *cis*- e *trans*-defensinas e sua diversidade estrutural.



Em azul destaca-se a orientação das folhas- β , em vermelho a estrutura de α -hélices, em branco a região do *random-coil* do peptídeo e em amarelo as ligações dissulfeto. **(A)** As *trans*-defensinas apresentam pontes dissulfeto que apontam para direções opostas, conectando regiões diferentes da estrutura do peptídeo, estabilizando a folha- β à α -hélice e ao *random-coil*. As *trans*-defensinas são classificadas de acordo com suas assinaturas moleculares e pareamento entre cisteínas na disposição Cys₁₋₆Cys₂₋₄Cys₃₋₅ nas α -defensinas, Cys₁₋₅Cys₂₋₄Cys₃₋₆ nas β -defensinas e big defensinas e Cys_{1a-3a}Cys_{2a-2b}Cys_{1b-3b} nas θ -defensinas. **(B)** Nas *cis*-defensinas, duas ligações dissulfeto paralelas conectam a folha- β à α -hélice. As defensinas do tipo CS $\alpha\beta$ contendo seis resíduos de cisteínas (C6) apresentam assinaturas moleculares e pareamento entre cisteínas na conformação Cys₁₋₄Cys₂₋₅Cys₃₋₆, enquanto que aquelas contendo oito resíduos de cisteínas (C8) distribuem-se na disposição Cys₁₋₈Cys₂₋₅Cys₃₋₆Cys₄₋₇ ou Cys₁₋₅Cys₂₋₆Cys₃₋₇Cys₄₋₈ (Adaptado SHAFEE et al., 2016; ZHU; GAO; TYTGAT, 2005).

Algumas famílias de defensinas distribuem-se apenas em pequenos grupos relacionados filogeneticamente, como no caso das α -defensinas que estão presentes em alguns grupos de mamíferos e das θ -defensinas, restritas apenas a “primatas do Velho Mundo” (Cercopithecidae). Por outro lado, as β -defensinas possuem distribuição ubíqua entre os vertebrados (ZHU; GAO; TYTGAT, 2005), além de estarem presentes em alguns artrópodes (MONTERO-ALEJO et al., 2017), enquanto que as big defensinas distribuem-se entre os táxons de cefalocordados, moluscos e alguns artrópodes marinhos (SHAFEE et al., 2017). Em relação às defensinas do tipo CS $\alpha\beta$, estas são as defensinas mais diversamente abundantes e bem distribuídas, estando presente em fungos filamentosos, plantas vasculares e em diferentes grupos de invertebrados, como em cefalocordados, moluscos e artrópodes (Figura 2) (SHAFEE et al., 2017).

Figura 2. Distribuição das *cis*- e *trans*-defensinas nos Eucariotos.

| Taxon | <i>cis</i> -defensinas | | <i>trans</i> -defensinas | | |
|-----------------|------------------------------------|---------------------|--------------------------|---------------------|---------------|
| | Defensina do tipo CS $\alpha\beta$ | α -defensina | β -defensina | θ -defensina | big defensina |
| Vertebrata | | ✓ | ✓ | ✓ | |
| Cephalochordata | ✓ | | | | ✓ |
| Mollusca | ✓ | | | | ✓ |
| Arthropoda | ✓ | | ✓ | | ✓ |
| Ascomycota | ✓ | | | | |
| Basidiomycota | ✓ | | | | |
| Angiospermae | ✓ | | | | |
| Gymnospermae | ✓ | | | | |

Ao lado esquerdo está representada uma filogenia simplificada dos eucariotos. Em azul estão representados alguns táxons do Reino Animalia, em vermelho alguns do Reino Fungi e em roxo alguns do Reino Plantae. O sinal de visto (✓), em verde, indica a presença das defensinas num determinado táxon (SHAFEE et al., 2017).

Nos invertebrados, diversos estudos demonstraram que as defensinas podem ser expressas de maneira constitutiva ou induzida, por meio de padrões moleculares associados a microrganismos (*Microbe-Associated Molecular Patterns*, ou MAMPs), pelas vias IMD (*immune deficiency*) e NF- κ B (*nuclear factor-kappa B*) (KAISER; DIAMOND, 2000; RAMIREZ et al., 2019). Dessa maneira, é possível haver um controle fino da microbiota, através de uma ação antimicrobiana contra um amplo espectro de microrganismos como bactérias Gram-positivas e Gram-negativas, além de fungos e vírus envelopados (GANZ, 2003).

O mecanismo de ação das defensinas é bastante variado. A atividade antimicrobiana pode se dar através da formação de poros ou pelo efeito detergente na membrana celular dos microrganismos (SAHL et al., 2005), pois a propriedade anfipática desses peptídeos permite que eles insiram-se na membrana plasmática fosfolipídica dos microrganismos, gerando sua desestabilização (MEADE; O'FARRELLY, 2019). Porém, estudos mais recentes demonstram a interferência desses peptídeos no fluxo de íons e síntese da parede celular dos procariotos, além da sinalização de vias imunológicas ou indução da geração de espécies reativas de oxigênio, no caso dos eucariotos (WILMES et al., 2011).

A primeira defensina foi inicialmente descoberta em granulócitos de coelhos e com atividade antimicrobiana contra um amplo espectro de bactérias Gram-positivas e Gram-negativas (SELSTED; SZKLAREK; LEHRER, 1984). O termo “defensina” foi cunhado devido ao contexto de defesa imunológica que esses peptídeos desempenham no combate a microrganismos patogênicos (GANZ et al., 1985). Desde então, a prospecção de defensinas em outros organismos revelou uma grande importância desses peptídeos na imunidade dos animais, principalmente daqueles que não possuem sistema imune adaptativo (DEFER et al., 2013). Um exemplo desses animais são as ostras, que são moluscos sésseis, osmoconformantes, que vivem em ambientes marinhos repletos de microrganismos e que, ao longo dos milhões de anos de evolução desenvolveram eficientes formas de defesa contra diferentes microrganismos patogênicos. Tais defesas incluem, entre outros, a produção de uma grande diversidade de AMPs, como defensinas da superfamília *cis* e da superfamília *trans* (DESTOUMIEUX-GARZÓN et al., 2016).

1.3. As defensinas de ostras

A partir do sequenciamento do genoma da ostra do Pacífico *Crassostrea gigas* (WANG et al., 2012) e da publicação de outros trabalhos transcritômicos (GUEGUEN et al., 2003; TARIS et al., 2009), as ostras e particularmente a espécie *C. gigas*, vem se tornando cada vez mais um modelo animal de estudos genéticos, devido especialmente, ao extremo grau de polimorfismo desses animais (BACHÈRE et al., 2015). Ostras *C. gigas* possuem um dos genomas mais polimórficos dentre as espécies animais (KIM; JUNG; GAFFNEY, 2011), possuindo um polimorfismo de nucleotídeo único (*Single Nucleotide Polymorphism*; SNP) a cada 60 pares de bases em regiões codificantes do genoma e a cada 40 pares de bases em regiões não codificantes (SAUVAGE et al., 2007). Diferentes estudos indicam que genes associados à imunidade de *C. gigas*, em especial aqueles codificantes para AMPs, sofrem de um elevado grau de polimorfismo, sendo submetidos a diferentes fenômenos genéticos, como a variações

no número de cópias gênicas (*Copy Number Variation*; CNV), assim como a variação de presença/ausência gênica (*Presense/Absense Variation*; PAV) (BACHÈRE et al., 2015; SCHMITT et al., 2010b).

Diferentes famílias de AMPs codificados por genes foram identificadas e caracterizadas na espécie *C. gigas*, dentre elas as proteínas indutoras da permeabilidade bacteriana (*Cg-BPIs*) (GONZALEZ et al., 2007a), peptídeos ricos em prolina (*Cg-PRPs*) (GUEGUEN et al., 2009), *macrophage-expressed gene 1 protein* (*Cg-Mpeg1*) (HE; ZHANG; YU, 2011), moluscidas (*Cg-Molluscidin*) (SEO et al., 2013), defensinas do tipo CS $\alpha\beta$ (*cis*-defensinas *Cg-Defs*) (GONZALEZ et al., 2007b; GUEGUEN et al., 2006; ZHANG et al., 2018) e big defensinas (*trans*-defensinas *Cg-BigDefs*) (ROSA et al., 2011), o que demonstra que as ostras são excelentes organismos para a prospecção de novos AMPs. Além disso, por abranger dentro de um mesmo genoma exemplares de defensinas do tipo CS $\alpha\beta$ (*cis*-defensinas) e big defensinas (*trans*-defensinas), a espécie *C. gigas* torna-se um importante foco de estudo para se entender a evolução e diversificação das defensinas, pois embora apresentem funções análogas, as *cis*- e *trans*-defensinas não possuem qualquer correlação filogenética (GERDOL et al., 2020a; SHAFEE et al., 2016).

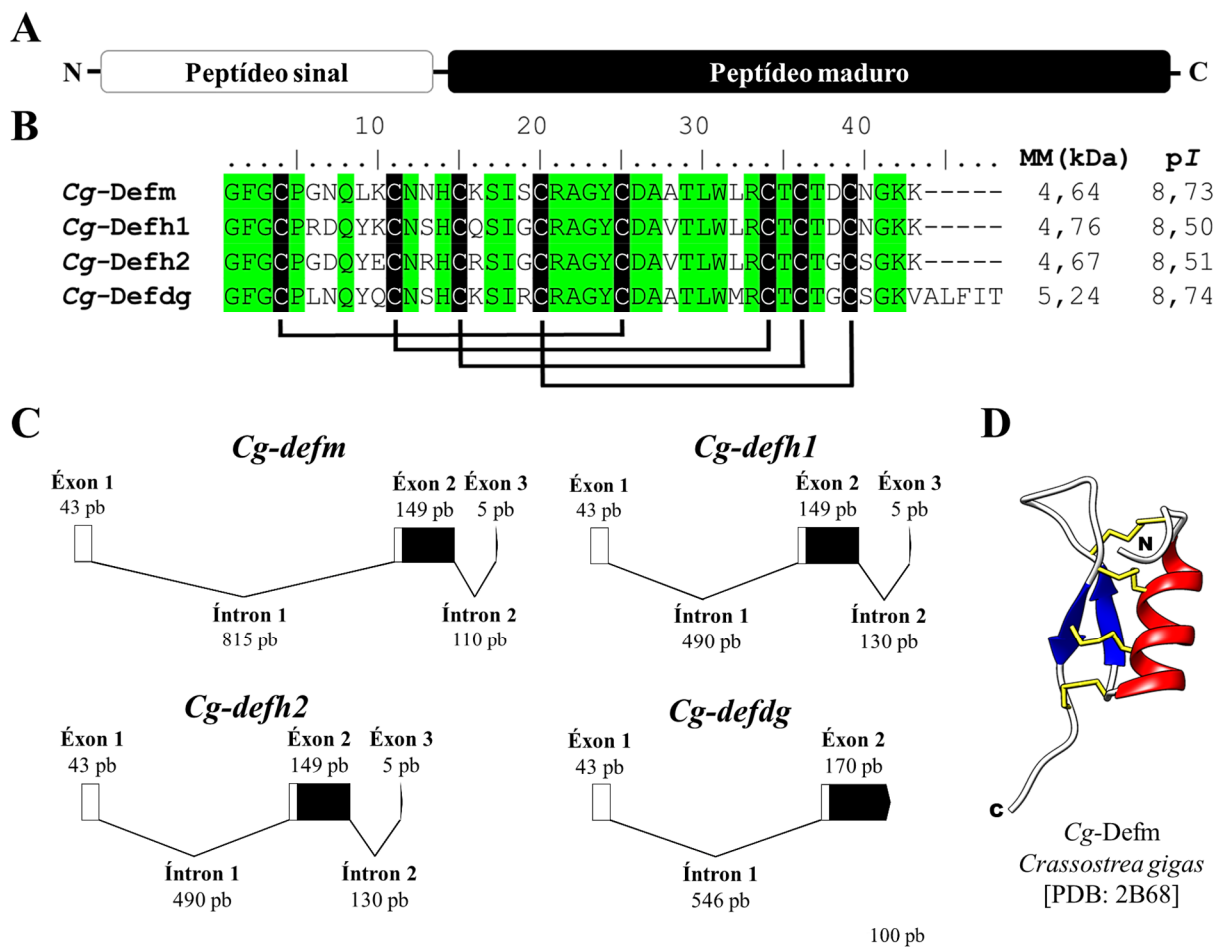
1.3.1 Defensinas do tipo CS $\alpha\beta$ (*Cg-Defs*)

As defensinas do tipo CS $\alpha\beta$ (*cysteine-stabilized α -helix β -sheet motif*) são os representantes da *cis*-defensinas. Esse é o grupo de defensinas mais diverso e bem distribuído, apresentando uma grande variedade de assinaturas moleculares e padrões de pareamentos entre cisteínas (SHAFEE et al., 2016), estando presente em diversos táxons, como em cefalocordados, moluscos, anelídeos, artrópodes, nematóides, cnidários, além de fungos e plantas superiores (SHAFEE et al., 2017). Em ostras, são peptídeos anfifílicos e catiônicos, que possuem uma estrutura secundária conservada, composta por uma α -hélice e duas folhas- β antiparalelas, estabilizadas por três (Cys₁₋₄Cys₂₋₅Cys₃₋₆) ou quatro (Cys₁₋₅Cys₂₋₆Cys₃₋₇Cys₄₋₈) pontes dissulfeto (SCHMITT; ROSA; DESTOUMIEUX-GARZÓN, 2016).

A primeira defensina do tipo CS $\alpha\beta$ descrita em ostras foi identificada na espécie americana, *Crassostrea virginica*. O peptídeo identificado como “*American oyster defensin*” (AOD) possui 38 aminoácidos e é estabilizado por três ligações dissulfeto, possuindo atividade antimicrobiana contra bactérias Gram-positivas e Gram-negativas (SEO et al., 2005). Posteriormente, na espécie *C. gigas*, as primeiras defensinas do tipo CS $\alpha\beta$ foram identificadas em diferentes tecidos, sendo constitutivamente expressas no manto (*Cg-Defm*) (GUEGUEN et al., 2006), nos hemócitos (*Cg-Defh1* e *Cg-Defh2*) (GONZALEZ et al., 2007b) e na glândula

digestiva (*Cg-Defdg*) (ZHANG et al., 2018). Essas defensinas de *C. gigas* são expressas como precursores contendo um peptídeo sinal de 22 aminoácidos, seguido de um peptídeo maduro catiônico de 4,6 a 5,2 kDa, contendo oito resíduos conservados de cisteínas ($Cys_{1-5}Cys_{2-6}Cys_{3-7}Cys_{4-8}$) (Figura 3) (SCHMITT; ROSA; DESTOUMIEUX-GARZÓN, 2016; ZHANG et al., 2018). Esses peptídeos são codificados por diferentes sequências gênicas e apresentam uma estrutura gênica composta de três éxons separados por dois íntrons, para as *Cg-defm*, *Cg-defh1* e *Cg-defh2* e dois éxons separados por um íntron, para a *Cg-defdg* (Figura 3) (ZHANG et al., 2018).

Figura 3. Estrutura do precursor, alinhamento dos peptídeos maduros e estrutura gênica das defensinas do tipo CS $\alpha\beta$ de *C. gigas* e estrutura tridimensional da *Cg-Defm*.



(A) Estrutura esquemática do precursor das *Cg-Defm*s composta por um peptídeo sinal na região N-terminal, seguido por um peptídeo maduro catiônico na porção C-terminal. (B) No alinhamento das *Cg-Defm*s, em verde, destacam-se os resíduos aminoácidos conservados em todas as sequências e, em preto, os resíduos conservados de cisteínas. Embaixo das sequências, as barras pretas indicam a posição das ligações dissulfeto entre as cisteínas ($Cys_{1-5}Cys_{2-6}Cys_{3-7}Cys_{4-8}$). Ao lado de cada sequência está indicado a massa molecular (MM) em quilodaltons (kDa) e o ponto isoelétrico teórico (pI). (C) Os éxons estão indicados pelos retângulos, enquanto que os íntrons estão representados pelas linhas pretas. Nos retângulos, em branco, está indicado a porção referente ao peptídeo sinal, enquanto que, em preto, está o peptídeo maduro. A escala dos genes está representada no canto inferior direito (100 pares de bases). (D) Em vermelho, destaca-se a estrutura de uma α -hélice e, em azul, as duas folhas- β antiparalelas. O

arranjo tridimensional do peptídeo é estabilizado através de quatro pontes dissulfeto (Cys₁₋₅Cys₂₋₆Cys₃₋₇Cys₄₋₈) destacadas em amarelo. As regiões N- e C-terminal do peptídeo são indicadas pelas letras “N” e “C”, respectivamente.

Análises da expressão gênica das *Cg*-Defs foram realizadas após desafios com bactérias Gram-negativas do gênero *Vibrio* (GONZALEZ et al., 2007b; GUEGUEN et al., 2006). Embora a *Cg*-Defm não tenha apresentado nenhuma modulação da expressão no tecido de manto (GUEGUEN et al., 2006), a presença dos transcritos da *Cg*-Defh1 e *Cg*-Defh2 aumentou em tecidos de manto e brânquias, devido à infiltração hemocitária nesses tecidos, sugerindo assim, que as defensinas possam atuar como uma primeira barreira de defesa nesses organismos (GONZALEZ et al., 2007b). Foi descrito na espécie de besouro *Rhodnius prolixus*, que as defensinas do tipo CSαβ apresentam uma expressão induzida pela via de sinalização IMD que, por meio de algum estímulo externo, como a presença de MAMPs ou citocinas, ativam a via NF-κB controlando a transcrição de genes associados à imunidade, como as defensinas (VIEIRA et al., 2018).

Testes antimicrobianos demonstraram que as *Cg*-Defs apresentam atividade antimicrobiana contra bactérias Gram-positivas, apresentando uma concentração inibitória mínima (MIC) entre 0,01 a 6 μM (BACHÈRE et al., 2015). Essa capacidade antimicrobiana se dá através da aptidão das defensinas de se ligarem ao lipídeo II, molécula precursora da produção da peptidoglicana (SCHMITT et al., 2010a). No entanto, atividade antimicrobiana, em menor grau, contra bactérias Gram-negativas e fungos filamentosos, também foram descritas (GUEGUEN et al., 2006; SCHMITT et al., 2012; ZHANG et al., 2018).

Por se tratarem de organismos extremamente polimórficos, foi reportado em ostras *C. gigas* o fenômeno genético de CNV (*copy number variation*) (BACHÈRE et al., 2015). Essa ocorrência genética consiste em um tipo de polimorfismo em que um gene pode estar em diferentes números de cópias nos genomas de indivíduos de uma mesma população (BACHÈRE et al., 2015). O fenômeno genético de CNV está associado a uma expressão diferenciada de genes, levando a uma alta variabilidade fenotípica. Toda essa plasticidade genética induz uma maior resiliência a fatores ambientais, como por exemplo, a uma maior resistência a enfermidades (ROSA et al., 2012; SPRINGER et al., 2009). Em *C. gigas*, a primeira evidência de CNV foi descrita em um estudo relacionado a genes associados à imunidade (SCHMITT et al., 2010b). Enquanto que alguns genes como o da *Cg*-*bpi* apresentam-se no genoma de *C. gigas* sempre como uma simples cópia, foi reportado nesse estudo que os genes das defensinas do tipo CSαβ *Cg*-*defm*, *Cg*-*defh1* e *Cg*-*defh2* estavam distribuídos em múltiplas cópias (SCHMITT et al., 2010b).

1.3.2 Big defensinas (Cg-BigDefs)

As big defensinas são peptídeos multidomínio, compostos por uma região N-terminal hidrofóbica e por uma região C-terminal (β -*defensin-like*), estabilizada por três ligações dissulfeto (Cys₁₋₅Cys₂₋₄Cys₃₋₆), conformação idêntica ao das β -defensinas de vertebrados (ROSA et al., 2011). Estudos recentes sugerem que as β -defensinas são evolutivamente derivadas das big defensinas (GERDOL et al., 2020a).

A primeira sequência de uma big defensina foi identificada na espécie de quelicerado marinho *Tachypleus tridentatus* (SAITO et al., 1995). Sua distribuição se dá principalmente em espécies marinhas dos táxons Lophotrochozoa, Arthropoda e Cephalochordata, sendo conhecida apenas uma sequência de big defensina em uma espécie não marinha, representada pelo mexilhão de água doce, *Hyriopsis cumingii* (LOTH et al., 2019).

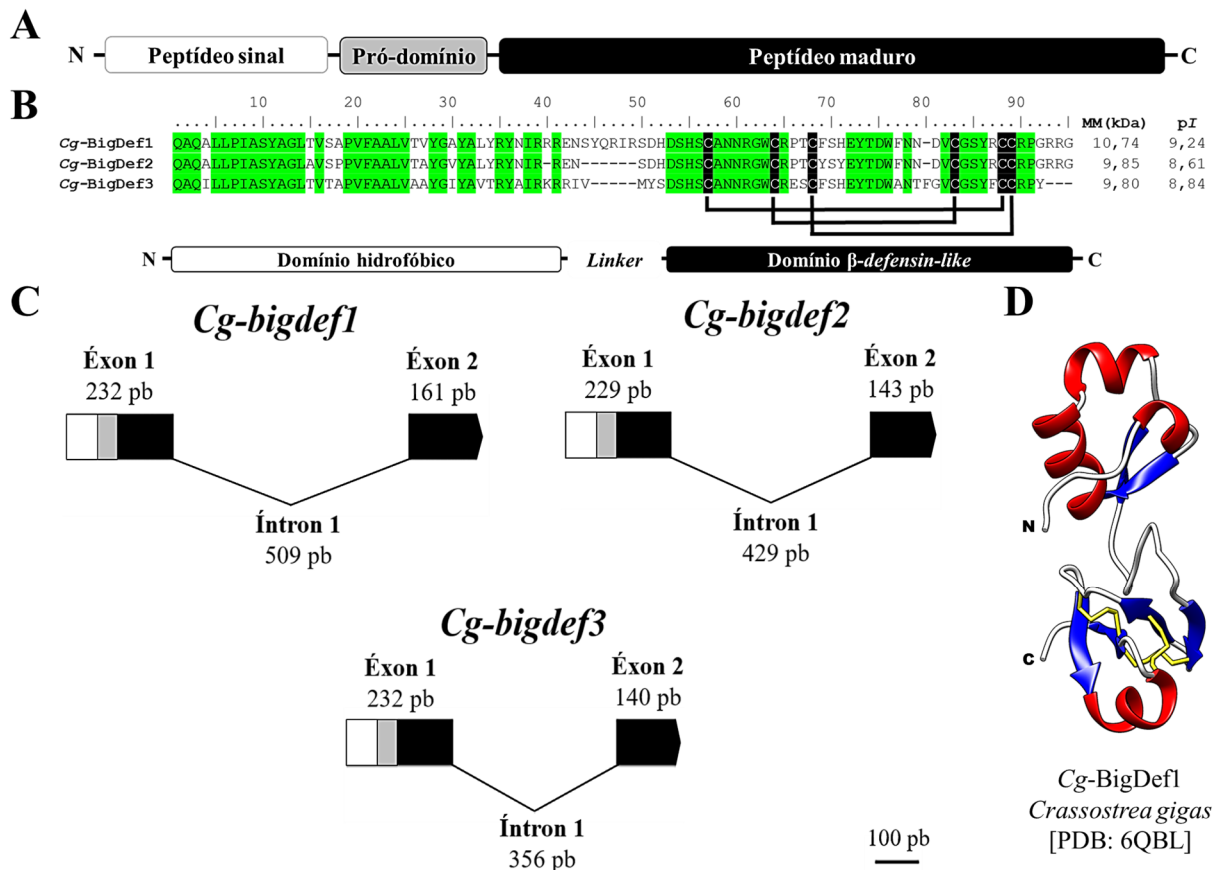
Em ostras *C. gigas* foram descritas big defensinas sendo expressas nos hemócitos (Cg-BigDef1, Cg-BigDef2, Cg-BigDef3) (ROSA et al., 2011). São peptídeos catiônicos de 9,8 a 10,7 kDa (SCHMITT et al., 2012), compostos por um peptídeo sinal, seguido de um pró-domínio na região N-terminal e de um peptídeo maduro contendo uma região hidrofóbica seguida de uma região β -*defensin like*, conectadas através de uma região *linker*, na porção C-terminal (GERDOL et al., 2020a) (Figura 4). As Cg-BigDefs são codificadas por diferentes sequências gênicas organizadas em dois éxons separados por um íntron (GERDOL et al., 2020a) (Figura 4).

Enquanto que as big defensinas Cg-BigDef1 e Cg-BigDef2 são altamente induzidas frente a infecções por bactérias Gram-negativas do gênero *Vibrio*, a Cg-BigDef3 é expressa constitutivamente nos hemócitos (ROSA et al., 2011). Um estudo recente sugere que a regulação da expressão da Cg-BigDef1 é feita através da via de sinalização NF- κ B/Rel, na qual uma variedade de MAMPs, como lipopolissacarídeos bacterianos, provocam a ativação dessa via, induzindo a translocação de fatores de transcrição para o núcleo das células e conduzindo a expressão de genes associados à imunidade, dentre eles as big defensinas (LI et al., 2020).

A atividade antimicrobiana da Cg-BigDef1 foi testada contra diferentes bactérias Gram-positivas e Gram-negativas de interesse clínico e ambiental. Os resultados demonstraram atividade bactericida em uma concentração de peptídeo entre 0,6 a 20 μ M (LOTH et al., 2019). A capacidade antimicrobiana das Cg-BigDefs está associada à capacidade desses peptídeos de formarem estruturas microscópicas de *nanonets*, na qual as big defensinas se agregam formando uma trama de fibras, que aprisionam e inativam os microrganismos (GERDOL et al., 2020a; LOTH et al., 2019). Estudos sugerem que a porção N-terminal hidrofóbica das big defensinas apresenta a capacidade de formar a estrutura de fibras das *nanonets*, mantendo o contato

próximo entre o peptídeo e a membrana do microrganismo, enquanto que fica a cargo da porção β -defensin-like a atividade antimicrobiana capaz de inativar o microrganismo (LOTH et al., 2019).

Figura 4. Estrutura do precursor, alinhamento dos peptídeos maduros e estrutura gênica das big defensinas de *C. gigas* e estrutura tridimensional da *Cg-BigDef1*.



(A) Estrutura esquemática do precursor das *Cg-BigDefs*, que é composta por um peptídeo sinal na região N-terminal, seguido de um pró-domínio e de um peptídeo maduro catiônico na porção C-terminal. (B) No alinhamento das *Cg-BigDefs*, em verde destaca-se os resíduos aminoacídicos conservados em todas as sequências e, em preto, os resíduos conservados de cisteínas. Embaixo das sequências, as barras pretas indicam a posição das ligações dissulfeto (Cys₁₋₅Cys₂₋₄Cys₃₋₆). Abaixo do alinhamento, em escala, o retângulo branco destaca o domínio hidrofóbico, que se conecta ao retângulo preto, representando o domínio β -defensin-like, através de uma região *linker*. Ao lado de cada sequência está indicado a massa molecular (MM) em quilodaltons (kDa) e o ponto isoelétrico teórico (pI). (C) Os éxons estão indicados pelos retângulos, enquanto que os íntrons estão representados pelas linhas pretas. Nos retângulos, em branco, está indicado o peptídeo sinal, em cinza o pró-domínio, enquanto que, em preto, está o peptídeo maduro. A escala dos genes está representada no canto inferior direito (100 pares de bases). (D) Em vermelho destacam-se as estruturas de α -hélices e, em azul, as folhas- β . O arranjo tridimensional do peptídeo é estabilizado através de três pontes dissulfeto (Cys₁₋₅Cys₂₋₄Cys₃₋₆) destacadas em amarelo. As regiões N- e C-terminal do peptídeo são indicadas pelas letras “N” e “C”, respectivamente.

Uma variação extrema de CNV também foi reportado nas big defensinas de *C. gigas*, o fenômeno genético conhecido como variação de presença/ausência gênica (PAV) (ROSA et al., 2015). Esse fenômeno é retratado quando determinado gene ocorre em alguns indivíduos, porém é completamente ausente em outros. O PAV foi bastante estudado em populações de

bactérias e em espécies de plantas de interesse econômico (GABUR et al., 2020), porém foi apenas em 2015 que a ocorrência desse fenômeno foi descrita em invertebrados por Rosa e colaboradores (2015). Nesse estudo, 163 indivíduos de *C. gigas* foram diagnosticados em relação a presença dos genes das *Cg-BigDef1*, *Cg-BigDef2* e *Cg-BigDef3*, no qual ficou demonstrada a prevalência do fenômeno de PAV atuando em algum grau, sobre todos os indivíduos. Esse estudo constatou que pelo menos um desses três genes estavam ausentes nos indivíduos. Surpreendentemente, aproximadamente 24% dos animais diagnosticados não apresentaram nenhum dos três genes em seus genomas (ROSA et al., 2015).

O fenômeno de PAV foi bastante explorado em um estudo recente com a espécie de mexilhão *Mytilus galloprovincialis*. Essa espécie conta com 60.338 genes anotados em bancos de dados genômicos, no entanto foi observado que ao menos 14.820 genes codificantes (24,25% do total) estavam completamente ausentes entre indivíduos (GERDOL et al., 2020b). Esse fenômeno genético pode estar relacionado à sobrevivência e adaptação evolutiva dos organismos, pois ao enfrentarem diferentes condições ambientais, determinados conjuntos de genes podem ser selecionados em um curto espaço de tempo em uma linhagem, porém sem perder a variabilidade genética geral de uma população (GERDOL et al., 2020b). Essa adaptação evolutiva faz com que os AMPs sejam massivamente submetidos ao fenômeno de PAV, no qual existem inúmeras isoformas de AMPs em determinada espécie, porém cada indivíduo possui uma combinação de poucos, porém distintos, AMPs (GERDOL et al., 2020b).

Nas últimas décadas vem se destacando o crescente interesse em pesquisas envolvendo AMPs, devido ao seu grande espectro de atividade biológica. Além disso, a prospecção de AMPs torna-se bastante relevante em se tratando de uma alternativa ao uso de antibióticos convencionais, pois o número limitado de antibióticos e suas similaridades de modos de ação induzem à resistência bacteriana, o que vem se tornando cada vez mais um problema de saúde pública e no controle de enfermidades.

Ostras *C. gigas* expressam em seus tecidos uma grande diversidade de moléculas efetoras, dentre elas os AMPs. Essas moléculas podem ser constitutivamente expressas ou induzidas por fatores estressores, no entanto, elas estão submetidas a um complexo polimorfismo genético, como o fenômeno de PAV, o que pode dificultar a identificação de novos AMPs nessas espécies. Portanto, a hipótese deste trabalho é a de que ostras *Crassostrea gigas* expressam em seus tecidos uma grande diversidade de defensinas do tipo CSa β e big defensinas.

2. OBJETIVOS

2.1. Objetivo geral

Identificar e caracterizar em nível molecular e transcricional novas defensinas em ostras da espécie *Crassostrea gigas*.

2.2. Objetivos específicos

- Prospectar em bancos de dados públicos novas sequências codificantes de defensinas do tipo CS $\alpha\beta$ e de big defensinas em ostras *C. gigas*;
- Caracterizar a estrutura do precursor, estrutura gênica e características bioquímicas das novas defensinas;
- Analisar o perfil de expressão das novas defensinas frente a infecções experimentais por patógenos de interesse aquícola;
- Avaliar a ocorrência e a prevalência do fenômeno de presença/ausência gênica (PAV) nos genes das novas defensinas em *C. gigas*.

3. MATERIAIS E MÉTODOS

3.1. Prospecção *in silico* de sequências

Sequências codificantes de defensinas do tipo CS $\alpha\beta$ e de big defensinas foram utilizadas como sequências de entrada (*query*) para a prospecção de novas defensinas em *C. gigas* (Apêndice 1). As buscas foram realizadas utilizando os bancos de dados não anotados EST (*Expressed Sequences Tags*) e TSA (*Transcriptome Shotgun Assembly*), disponíveis no NCBI (*National Center for Biotechnology Information*), através da ferramenta tBLASTx (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>). As sequências completas e com *e-value* inferior a 10^{-15} foram selecionadas e traduzidas através do programa ExPASy Translate Tool (<https://web.expasy.org/translate/>).

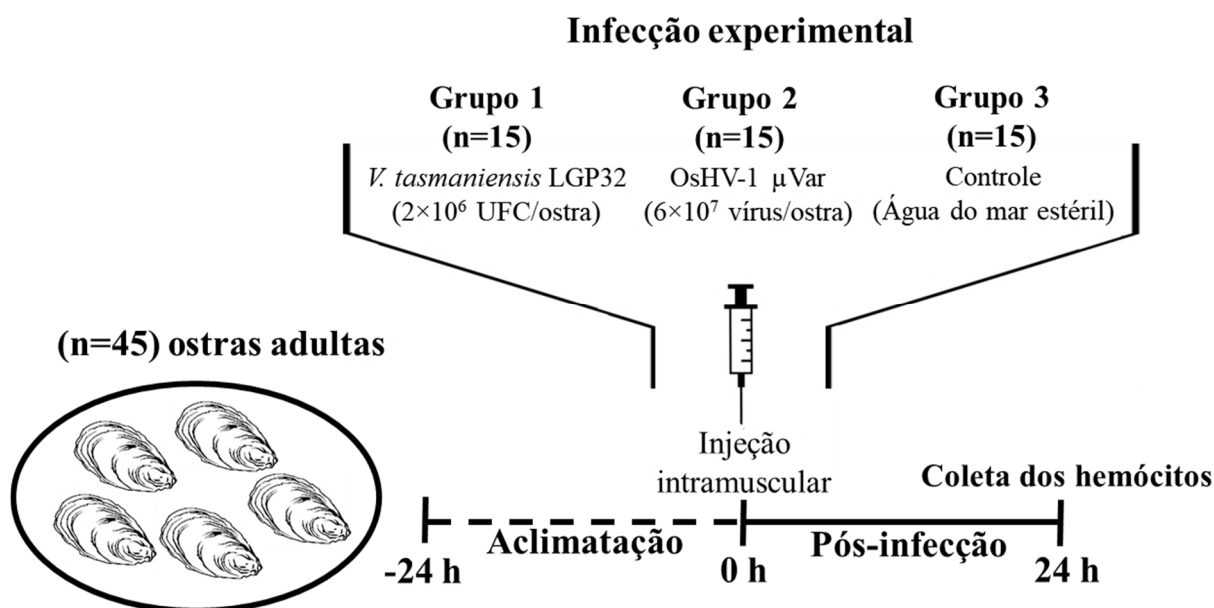
3.2. Animais, infecção experimental e coleta dos tecidos

Neste trabalho, foram utilizadas ostras adultas (8-10 cm) da espécie *Crassostrea gigas* (Thunberg, 1793) provenientes da França (Montpellier, Occitanie) e do Brasil (Florianópolis, Santa Catarina). Ostras de ambas procedências foram utilizadas nos ensaios de clonagem molecular (França: Cg-Defs; Brasil: Cg-BigDefs) e verificação do fenômeno de PAV enquanto que as ostras provenientes da França foram ainda utilizadas nos ensaios de expressão gênica.

As infecções experimentais foram realizadas com dois patógenos de interesse aquícola: a bactéria Gram-negativa *Vibrio tasmaniensis* LGP32 e o herpesvírus de ostra (OsHV-1 μ Var). Para isso, as ostras (n=45) foram aclimatadas por 24 horas e em seguida divididas em três grupos de 15 animais. O primeiro grupo foi experimentalmente infectado com 2×10^6 unidades formadoras de colônia (UFC)/animal de *V. tasmaniensis* LGP32 enquanto que o segundo grupo foi infectado com 6×10^7 partículas virais/animal de OsHV-1 μ Var. O terceiro grupo (grupo controle) foi inoculado apenas com o veículo (água do mar filtrada e autoclavada). Todas as inoculações foram feitas através de injeções (100 μ L) no músculo adutor dos animais (Figura 5).

Após 24 horas da infecção experimental, a hemolinfa de cada animal foi coletada por meio de uma punção na cavidade cardíaca, através do músculo adutor. Cada grupo foi dividido em três *pools* de cinco animais (triplicata biológica). A hemolinfa foi centrifugada (1.000 \times g por 10 min a 4°C), o sobrenadante descartado e o pellet contendo os hemócitos foi armazenado em TRIzol (Thermo Scientific) e mantido a -20°C, até o momento da extração do RNA total.

Figura 5. Delineamento da infecção experimental.



Um “n” amostral de 45 ostras *C. gigas* foi aclimatado por 24 horas e dividido em três grupos de 15 animais. O grupo 1 foi infectado com 2×10^6 UFC/ostra da bactéria Gram-negativa *Vibrio tasmaniensis* LGP32. O grupo 2 foi infectado com 6×10^7 partículas do herpesvírus de ostra (OsHV-1 μ Var)/ animal. Como controle, o grupo 3 foi injetado apenas com o veículo (água do mar filtrada e autoclavada). Após 24 horas da infecção, os hemócitos foram coletados para a extração de RNA total.

Para os testes de PAV, foram coletadas amostras de gDNA de ostras individuais. Seis ostras provenientes da França foram retiradas de suas conchas e imediatamente congeladas em nitrogênio líquido, para serem trituradas individualmente em um pó fino e armazenadas a -80°C até o momento da extração do gDNA. Por sua vez, as ostras brasileiras tiveram extraídos os tecidos de brânquia de 16 animais de maneira também individual, que foram mantidos em álcool 75% e armazenados a -20°C até o momento da extração do gDNA.

3.3. Extração de ácidos nucleicos e síntese de cDNA

Para a extração de DNA genômico (gDNA), aproximadamente 100 mg de tecido dos animais foram adicionado em 400 μL de tampão de lise [100 mM Tris-HCl (pH 8,5), 100 mM NaCl, 50 mM EDTA (pH 8,0), 1% SDS e 0,1 mg/mL de proteinase K]. Após uma incubação de 1 hora a 55°C (agitando manualmente a cada 15 min), as amostras foram centrifugadas (1.000 $\times g$ por 1 min a 4°C) e o sobrenadante foi transferido para um novo tubo, no qual foi adicionado 200 μL de acetato de potássio 3 M. As amostras foram incubadas por 30 min, centrifugadas (14.000 $\times g$ por 10 min a 4°C), o sobrenadante foi precipitado em álcool isopropílico e o precipitado contendo o gDNA foi lavado em etanol 75% gelado. O gDNA foi suspenso em 50 μL de água livre de DNases/RNases a 65°C . Para a eliminação do RNA contaminante, as

amostras de gDNA foram tratadas com a enzima RNase A (Sigma-Aldrich). Foi adicionado a cada amostra 50 µg/mL da enzima RNase A, seguida de uma incubação a 37°C por 1 hora.

Na extração de RNA total das amostras conservadas em TRIzol (Thermo Scientific), foi adicionado clorofórmio (1:5, v/v). As amostras foram agitadas vigorosamente por inversão e incubadas por três min à temperatura ambiente e então centrifugadas (12.000 ×g por 15 min a 4°C). A porção sobrenadante aquosa superior foi coletada e transferida para um novo tubo e precipitadas em álcool isopropílico (1:1, v/v) e lavado em etanol 75% gelado. O RNA foi suspenso em 40 µL de água livre de DNases/RNases a 65°C. As amostras foram então tratadas com a enzima Desoxirribonuclease I (Thermo Scientific), para a eliminação de resquícios de DNA genômico. Foi adicionado às amostras 1× do tampão da DNase I e 1 U da enzima, que então foram incubadas por 30 min a 37°C e posteriormente a 65°C por 10 min para a inativação da enzima.

Para a eliminação da RNase A e da DNase I, os ácidos nucleicos foram precipitados. Foi adicionado às amostras 300 mM de acetato de sódio (pH 5,5) e álcool isopropílico (1:1; v:v) e em seguida as amostras foram brevemente agitadas, incubadas por 10 min à temperatura ambiente e centrifugadas (12.000 ×g por 10 min a 4°C). O sobrenadante foi removido e o precipitado contendo o gDNA/RNA foi lavado em etanol 75% gelado. Após centrifugação (8.000 ×g por 5 min a 4°C), o etanol foi descartado e o pellet foi suspenso em 40 µL de água livre de DNases/RNases a 65°C. Os ácidos nucleicos foram quantificados por espectrofotometria e a integridade das amostras foi analisada por eletroforese em gel de agarose (0,8%). As amostras foram então armazenadas a -20 °C até seu uso posterior.

Para a síntese de DNA complementar (cDNA), em uma solução de 20 µL foi adicionado 400 ng de RNA, 250 ng de *random primers* e 10 mM de dNTPs, que foram aquecidos a 65°C por 5 min e imediatamente em seguida resfriados em gelo por 2 min. Foi então adicionado à solução 4 µL do tampão de reação 5× para a síntese de cDNA (Invitrogen), além de 100 mM de Ditioneitol (DTT) e 40 U do inibidor de ribonuclease RNaseOUT (Thermo Fisher Scientific). As amostras foram incubadas a 37°C por 2 min e então foi adicionado 200 U da enzima M-MLV-RT (Invitrogen). A solução foi incubada a 25°C por 10 min, 37°C por 50 min e em seguida a 70°C por 15 min para a inativação das enzimas. A qualidade das amostras foi avaliada por meio de uma reação de PCR do gene da *Cg-actin* (Tabela 1). As amostras permaneceram armazenadas a -20 °C até seu uso posterior.

3.4. Escolha dos iniciadores

Os iniciadores escolhidos para esse estudo foram baseados nas sequências obtidas através das prospecções *in silico* (*Cg-Defh3* – GenBank: GT054002; *Cg-Defh4* – GenBank: CU682678; *Cg-BigDef4* – GenBank: GHAT01130943; *Cg-BigDef5* – GenBank: GECI01014168; *Cg-BigDef6* – GenBank: GECI01022844). Os iniciadores utilizados para a identificação, clonagem molecular e sequenciamento das novas defensinas foram desenhados manualmente para que se posicionassem de maneira a flanquearem a região codificante do peptídeo (*coding sequence* ou CDS), enquanto que os iniciadores utilizados na técnica de RT-qPCR foram desenhados de maneira a serem específicos para cada sequência e que se encaixassem em parâmetros específicos, como uma proporção guanina-citosina próxima a 50%, uma temperatura de hibridização próxima de 60°C e que gerassem um produto de amplificação entre 100-200 pb. A qualidade dos iniciadores foi analisada utilizando o programa FastPCR (<https://primerdigital.com/fastpcr.html>), a fim de evitar a formação de dímeros.

3.5. Amplificação das sequências de interesse

Para a amplificação das sequências de interesse, foi realizada a técnica de reação em cadeia da polimerase em duas etapas (*second-round* PCR). A primeira reação da PCR foi realizada em um volume final de 25 µL contendo 1× do tampão de PCR contendo MgCl₂ (Sinapse), 0,4 mM de dNTPs, 0,4 µM dos iniciadores senso e antissenso (Tabela 1), 1 U da enzima *Taq* DNA polimerase (Sinapse) e 1 µL da amostra de cDNA (diluição 1:10) (amplificação das *Cg-Defs*) ou 100 ng de gDNA (amplificação das *Cg-BigDefs*). As reações aconteceram a partir de uma desnaturação inicial a 95°C por 10 min, seguido de 35 ciclos de desnaturação a 95°C por 30 s, hibridização a 50-60°C por 20 s e extensão a 72°C por 30 s. A segunda reação da PCR foi aplicada seguindo as mesmas condições descritas anteriormente, porém utilizando como amostra 1 µL do produto da reação anterior (diluído 1:10). Os produtos de amplificação foram analisados em gel de agarose (1,5%) através de uma coloração de brometo de etídio (0,5 µg/mL).

Tabela 1. Sequências nucleotídicas dos iniciadores utilizados neste trabalho.

| Gene | GenBank | Sequência 5'– 3' (Senso e antissenso) | Amplicon |
|--|--------------|--|----------|
| Clonagem molecular e sequenciamento nucleotídico. | | | |
| Cg-Actin | XM_011434666 | S: TAATCCACATCTGCTGGAAGGTGG AS: TCACCAACTGGGATGACATGG | 846 pb |
| Cg-Defh3 | GT054002 | S: AGGCATAGATGCCGCACCTC AS: TAATAACTTATCTACAGTGG | 167 pb |
| Cg-Defh4 | CU682678 | S: CATGTTTAAGATTTTACTGC AS: CTCAACTCATCTACAATGAC | 187 pb |
| Cg-BigDef4 | GHAT01130943 | S: CGAAATAAAAAGTCCACCATG AS: CTGCTTGCATGGATTACCCC | 640 pb |
| Cg-BigDef5 | GECI01014168 | S: AATCAAGTCAACATGAACAG AS: TTATCCTAGATTTCTAGGTC | 655 pb |
| Cg-BigDef6 | GECI01022844 | S: ATGGCGAGGAAATTGTTCCAG AS: CTAAGTCTGCAGCATT | 928 pb |
| Análises de expressão gênica (RT-qPCR) | | | |
| Cg-RPL40 | XM_011439956 | S: AATCTTGCACCGTCATGCAG AS: AATCAATCTCTGCTGATCTGG | 146 pb |
| Cg-Defh3 | GT054002 | S: GAAAAGACTAACGTGCGAGC AS: CGTGAATCCAAGGGTAATGC | 97 pb |
| Cg-Defh4 | CU682678 | S: TAGATGCAGCTCCTGAGAAG AS: CTCAACTCATCTACAATGAC | 134 pb |
| Cg-BigDef4 | GHAT01130943 | S: TTTGGCCGAGACTCTGGAAG AS: GCACAAGAATGACTGTCTGAG | 198 pb |
| Cg-BigDef5 | GECI01014168 | S: CCATGAAGGATCTGAGAGAG AS: GCTGTCATTTGAGTTCTTGC | 176 pb |
| Cg-BigDef6 | GECI01022844 | S: AACATCGCCCGGTCTTATTC AS: GAGCATGGCTTTGAACTCC | 200 pb |
| Cg-BPI1 | AY165040 | S: GATAGAAATAGGAATGGACGG AS: GTTATAGATCCACGCTGCTCC | 115 pb |
| Análises de presença/ausência de genes (PAV) | | | |
| Cg-Actin | XM_011434666 | S: TAATCCACATCTGCTGGAAGGTGG AS: TCACCAACTGGGATGACATGG | 846 pb |
| Cg-Defh3 | GT054002 | S: AGGCATAGATGCCGCACCTC AS: TAATAACTTATCTACAGTGG | 276 pb |
| Cg-Defh4 | CU682678 | S: CATGTTTAAGATTTTACTGC AS: CAAGAACCAAGATGTCCAGG | 638 pb |
| Cg-BigDef4 | GHAT01130943 | S: CGAAATAAAAAGTCCACCATG AS: CTGCTTGCATGGATTACCCC | 640 pb |
| Cg-BigDef5 | GECI01014168 | S: AATCAAGTCAACATGAACAG AS: TTATCCTAGATTTCTAGGTC | 655 pb |
| Cg-BigDef6 | GECI01022844 | S: ATGGCGAGGAAATTGTTCCAG AS: CTAAGTCTGCAGCATT | 928 pb |

3.6. Clonagem molecular e sequenciamento nucleotídico

Os produtos obtidos a partir da amplificação das sequências de interesse foram ligados ao vetor pGEM-T Easy (Promega), utilizando a enzima T4 DNA Ligase. Em um volume de reação de 10 μ L, foi adicionado 1 \times do tampão de ligação, 50 ng do plasmídeo pGEM-T Easy, 2 μ L do produto de PCR fresco e 3 U da enzima T4 DNA Ligase e a reação foi incubada *overnight* a 4°C. Em seguida, bactérias quimiocompetentes *Escherichia coli* DH5 α foram transformadas por choque térmico (42°C por 45 s, seguido imediatamente de uma incubação no gelo por 2 min) e incubadas em meio SOC (Super Optimal broth with Catabolite repression: 2% tripton, 0,5% extrato de levedura, 0,05% NaCl, 250 mM KCl, 2 M MgCl₂, 20 mM glicose, pH 7,2), sob agitação a 37°C por 90 min. As bactérias transformadas foram selecionadas em meio LB sólido (Lysogeny Broth: 1% tripton, 0,5% extrato de levedura, 1% NaCl, 1,5% ágar, pH 7,2) contendo o antibiótico ampicilina (100 μ g/mL) e X-GAL (20 μ g/mL). Para a confirmação da presença do inserto no plasmídeo, as colônias bacterianas de coloração branca foram submetidas a técnica de PCR-*Colony*, utilizando os iniciadores M13 (Promega) (S: 5'-TGTAACGACGGCCAGT-3'; As: 5'-CAGGAAACAGCTATGACC-3'). Em um volume de reação de 10 μ L contendo um esfregaço de uma colônia bacteriana selecionada, foi adicionado 0,4 mM de dNTPs, 0,2 μ M dos iniciadores M13 senso e antissenso e 1 U da enzima *Taq* DNA polimerase (Sinapse). A reação se deu através de uma desnaturação inicial a 95°C por 15 min, seguido de 30 ciclos de desnaturação a 95°C por 1 min, hibridização a 56°C por 1 min, extensão a 72°C por 1 min. As colônias positivas para o inserto foram submetidas a extração e purificação do DNA plasmidial utilizando o kit Wizard Plus SV Minipreps DNA Purification System (Promega). O plasmídeo purificado foi suspenso em água livre de DNases/RNases a 65°C e mantido a -20°C até seu uso posterior.

O sequenciamento nucleotídico foi realizado pelo método de Sanger, a partir de amostras de plasmídeos e de produtos de PCR purificados com o kit ExoSap-IT (Thermo Fisher Scientific), seguindo as especificações do fabricante. O preparo da reação de sequenciamento foi realizado em um volume final de 10 μ L, adicionando 1 μ L do reagente BigDye Terminator (Thermo Fisher Scientific), 1,5 μ L do tampão de sequenciamento 5 \times (Thermo Fisher Scientific), 5 μ M do iniciador M13 senso ou M13 antissenso e 700 ng das amostras de plasmídeo ou 60 ng das amostras de produto de PCR purificado. A reação de sequenciamento aconteceu a partir de uma desnaturação inicial a 96°C por 10 s, seguido de 25 ciclos de desnaturação a 96°C por 10 s, hibridização a 50°C por 5 s e extensão a 60°C por 4 min. As amostras foram precipitadas em etanol e então foram sequenciadas no sequenciador

automatizado de DNA ABI3500 Genetic Analyzer (Applied Biosystems) do Laboratório Multiusuário de Estudos em Biologia (UFSC/CCB/LAMEB).

3.7. Análise de sequências

As sequências obtidas no sequenciamento nucleotídico foram analisadas utilizando o programa BioEdit Sequence Alignment Editor 5.0.9 e traduzidas através do programa ExPASy Translate Tool (<https://web.expasy.org/translate/>). A predição do peptídeo sinal e do pró-domínio foi realizada utilizando o programa Prop1 Server (<http://www.cbs.dtu.dk/services/Prop1/>). A massa molecular e o ponto isoelétrico teórico (*pI*) dos peptídeos maduros foram definidos através do programa ExPASy ProtParam Tool (<https://web.expasy.org/protparam/>). A estrutura gênica das sequências foi esquematizada utilizando o software Exon-Intron Graphic Maker (<http://wormweb.org/exonintron>). O programa MAFFT version 7 (<https://mafft.cbrc.jp/alignment/server/>) auxiliou o alinhamento e edição das sequências aminoacídicas realizadas no programa BioEdit Sequence Alignment Editor 5.0.9. O modelamento tridimensional das defensinas foi gerado utilizando como referência os peptídeos resolvidos e depositados no programa SWISS-MODEL (<https://swissmodel.expasy.org/>). As defensinas *Cg-Defm*, *Cg-Defh1*, *Cg-Defh2* e *Cg-Defdg* foram gerados a partir do peptídeo deduzido da *Cg-Defm* (PDB: 2B68), enquanto que a *Cg-Defh3* e *Cg-Defh4* foram resolvidos a partir do peptídeo Lucifensina (PDB: 2LLD). A estrutura tridimensional de todas as big defensinas foi resolvida a partir da estrutura da *Cg-BigDef1* (PDB: 6QBL). As estruturas tridimensionais foram editadas utilizando o programa UCSF Chimera 1.14 e a qualidade dos modelos tridimensionais foi medida utilizando a escala GMQE (*global model quality estimation*), que consiste em uma escala de 0 a 1, na qual quanto mais próximo de 1, mais fidedigna é a modelagem em relação ao modelo original.

3.8. Análises de expressão gênica

A expressão gênica das defensinas foi realizada através da técnica de transcrição reversa seguida de PCR quantitativa em tempo real (RT-qPCR), utilizando o equipamento StepOnePlus Real-Time PCR Systems (Thermo Fisher Scientific). Em um volume final de 15 μ L foi adicionado 7,5 μ L de SYBR Select Master Mix (Thermo Fisher Scientific), 0,4 μ M dos iniciadores senso e antissenso (Tabela 1) e 1 μ L da amostra de cDNA (diluição 1:10). Todas as reações foram realizadas em duplicatas técnicas. As reações de RT-qPCR aconteceram a partir de uma desnaturação inicial a 95°C por 10 min, seguido por 40 ciclos de desnaturação a 95°C por 15 s e hibridização/extensão a 60°C por 1 min. Para validar se a reação foi específica para

apenas um produto de amplificação, foi realizada uma curva de dissociação (*melting curve analysis*) entre 60 a 95°C com um aumento de temperatura de 0,3°C por segundo.

A expressão gênica foi calculada através da fórmula $2^{-\Delta\Delta Cq}$ (LIVAK; SCHMITTGEN, 2001). As amostras foram normalizadas utilizando o gene de referência Ubiquitina-60S proteína ribossomal L40 (GenBank: XM_011439956) e os resultados da análise da expressão gênica foram calibrados utilizando os valores de ΔCq das amostras do grupo “Salina” (Animais estimulados apenas para simular a injúria das injeções). A análise estatística foi realizada utilizando o teste de análise de variância (ANOVA), com separação de médias pelo teste de Tukey. Os grupos foram considerados significativamente diferentes quando as variações da expressão gênica foram superiores a 1,5 vezes e com o valor de $P < 0,05$. As análises estatísticas e os gráficos foram feitos utilizando o programa GraphPad Prism 8 (<https://www.graphpad.com/scientific-software/prism/>).

3.9. Análises de presença/ausência de genes (PAV)

A validação do fenômeno de PAV (*presence/absence variation*) foi realizada através da técnica de PCR convencional. Em um volume final de reação de 15 μ L, adicionou-se 1 \times do tampão de PCR contendo $MgCl_2$ (Sinapse), 0,4 mM de dNTPs, 0,4 μ M dos iniciadores senso e antissenso (Tabela 1) e 1 U da enzima *Taq* DNA polimerase (Sinapse). Como amostra, foi utilizado 100 ng de gDNA de 16 animais provenientes do Brasil e de seis animais oriundos da França. Inicialmente foi realizada uma reação de PCR utilizando os iniciadores *Cg-Actin* (Tabela 1) para assegurar a qualidade das amostras. Posteriormente, utilizando os mesmos iniciadores usados na clonagem molecular (Tabela 1), as amostras foram testadas quanto à presença dos genes das novas *Cg-defs* e *Cg-bigdefs*. Como controle positivo da reação de PCR foi utilizado no lugar da amostra 1 ng do plasmídeo pGEM-T Easy contendo o gene de interesse. O produto de amplificação foi analisado em gel de agarose 1,5% e corado com brometo de etídio (0,5 μ g/mL). A presença do gene foi considerada como positiva, a partir da presença de uma banda detectável no gel de agarose, de tamanho igual ao apresentado pelo controle positivo.

4. RESULTADOS

4.1. Identificação de novas defensinas na ostra do Pacífico *Crassostrea gigas*

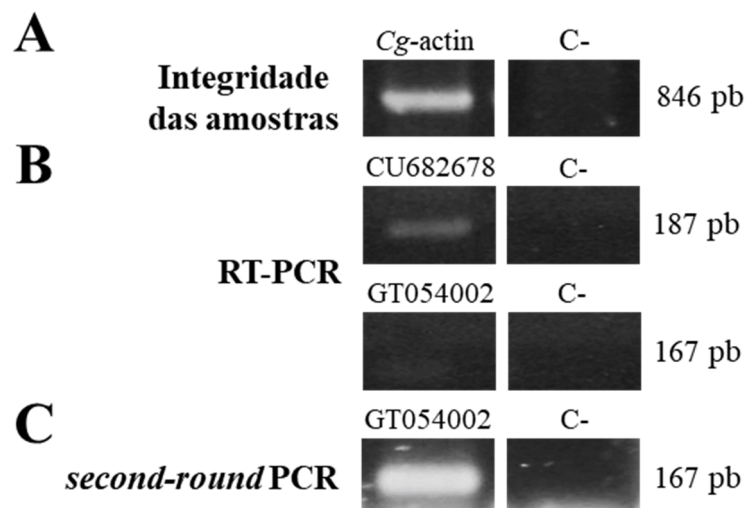
As buscas *in silico* de sequências codificantes para novas defensinas em *C. gigas* foram realizadas em bancos de dados públicos não anotados (EST e TSA), utilizando como sequências de entrada defensinas previamente descritas em espécies de moluscos. As sequências prospectadas foram traduzidas a partir da CDS das sequências nucleotídicas e analisadas quanto à presença de resíduos conservados de cisteínas, que correspondessem às assinaturas clássicas das defensinas.

Em bancos de dados de EST foram resgatadas duas sequências de defensinas do tipo CS $\alpha\beta$: a primeira (GenBank: GT054002) possui uma CDS parcial (sequência incompleta) de 207 pb, enquanto que a segunda sequência (GenBank: CU682678) possui uma CDS de 254 pb. Ambas sequências apresentam uma assinatura aminoacídica conservada contendo seis resíduos de cisteínas Cys-X₇₋₁₆-Cys-X₃-Cys-X₉-Cys-X₅-Cys-X₁-Cys, semelhante àquelas apresentadas pelas defensinas do tipo CS $\alpha\beta$ de insetos. Em relação à prospecção nos bancos TSA, foram encontradas três sequências com alta identidade com as big defensinas previamente descritas em *C. gigas* (*Cg-BigDefs*). As sequências resgatadas desse banco transcritômico (GenBank: GHAT01130943; GECI01014168; GECI01022844) consistem em três transcritos de 708, 423 e 369 pb, respectivamente. As três novas sequências possuem uma assinatura aminoacídica conservada contendo seis resíduos de cisteínas Cys-X₆-Cys-X₃-Cys-X₁₃-Cys-X₄-Cys-Cys, semelhante àquelas apresentadas pelas big defensinas de *C. gigas*.

4.2. Caracterização de novas defensinas do tipo CS $\alpha\beta$ (*Cg-Defs*)

A fim de atestar a veracidade das novas *Cg-Defs* encontradas nas buscas *in silico*, foram desenhados iniciadores específicos para cada nova sequência. As sequências das candidatas a novas *Cg-Defs* foram amplificadas a partir de amostras de cDNA de um *pool* de hemócitos de ostras provenientes da França. No entanto, apenas uma sequência, referente a *Cg-Def* (GenBank: CU682678), foi amplificada através da técnica de PCR, apresentando uma banda de intensidade fraca no gel de agarose corado com brometo de etídio. Optou-se então por realizar uma reação de *second-round* PCR, para obter a amplificação da outra *Cg-Def* que não havia sido amplificada anteriormente (GenBank: GT054002), utilizando como amostra o produto de amplificação da reação anterior, o que resultou em uma amplificação expressiva da sequência (Figura 6). Os produtos de amplificação foram então ligados a um vetor de clonagem e submetidos à técnica de sequenciamento nucleotídico pelo método de Sanger. O resultado do sequenciamento corroborou com as sequências encontradas *in silico*.

Figura 6. Eletroforese em gel de agarose (1,5%) corado com brometo de etídio representando a amplificação dos transcritos para as novas *Cg*-Defs em amostras de hemócitos de ostras *C. gigas*.



(A) Validação da integridade das amostras utilizando iniciadores para o gene da *Cg*-actin. (B) RT-PCR das amostras de hemócitos utilizando iniciadores específicos para cada nova *Cg*-Def. (C) *Second-round* PCR da *Cg*-Def (GenBank: GT054002), utilizando como amostra, 1 μ L do produto de amplificação da reação anterior, diluído 1:10. A primeira coluna de géis representa o resultado da amplificação dos transcritos, para cada um dos genes, enquanto que a segunda coluna corresponde ao controle negativo das reações (C-). Ao lado direito encontra-se o tamanho esperado para cada produto de amplificação em número de pares de bases.

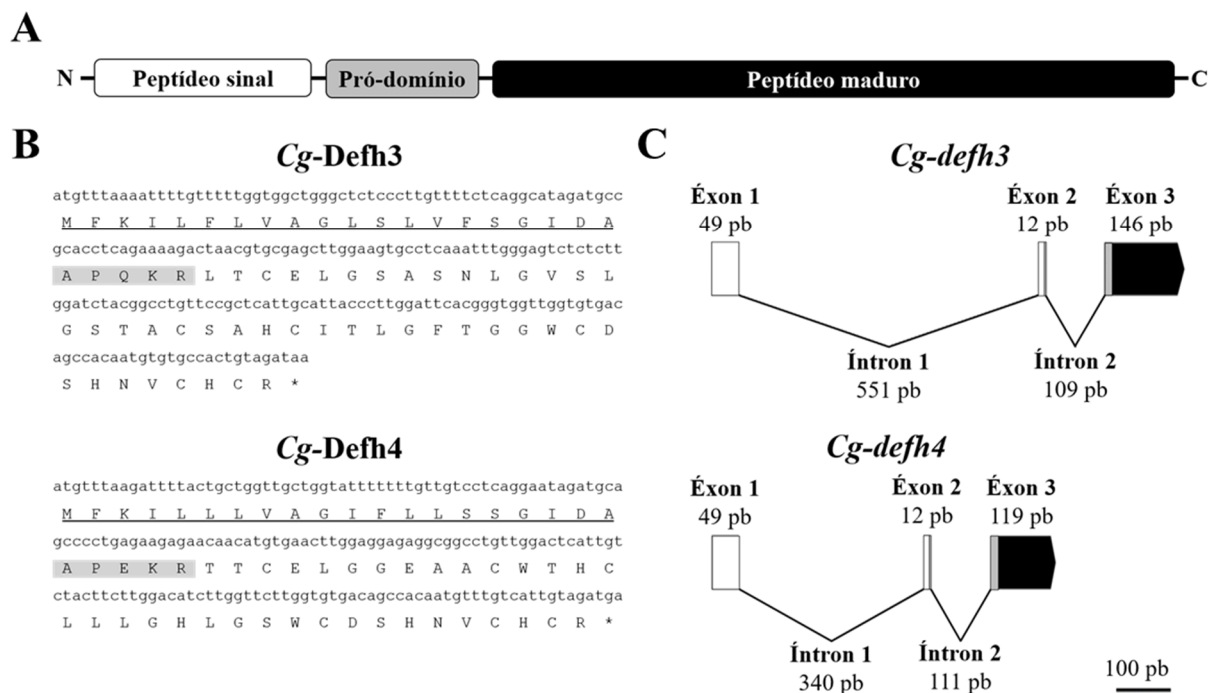
Como as novas *Cg*-Defs foram amplificadas a partir de amostras de hemócitos, elas receberam o nome de *Cg*-Defh3 (GenBank: GT054002) e *Cg*-Defh4 (GenBank: CU682678), sucedendo as defensas *Cg*-Defh1 e *Cg*-Defh2, anteriormente caracterizadas em *C. gigas* (GONZALEZ et al., 2007b). As novas *Cg*-Defs agora compõe um grupo de seis defensas do tipo CS α β em *C. gigas*, que além das quatro defensas hemocitárias (*Cg*-Defh1 a *Cg*-Defh4), também contém a *Cg*-Defm e *Cg*-Defdg (GUEGUEN et al., 2006; ZHANG et al., 2018).

A tradução da *Cg*-Defh3 evidenciou um precursor de 53 resíduos aminoacídicos, consistindo de um peptídeo sinal parcial (sequência incompleta) de cinco aminoácidos, seguido de um pró-domínio também de cinco aminoácidos contendo um sítio de clivagem dibásico (Lys-Arg) e um peptídeo maduro de 43 aminoácidos. No entanto, foi apenas a partir da publicação de um novo genoma que incluía o gene completo da *Cg*-Defh3, por Wang e colaboradores (2019), que foi possível caracterizar a porção faltante do peptídeo sinal, que consiste originalmente em uma sequência de 20 aminoácidos, totalizando um precursor de 68 resíduos aminoacídicos (Figura 7). Por sua vez, a sequência da *Cg*-Defh4 apresenta um precursor de 59 resíduos aminoacídicos, no qual também compreende um peptídeo sinal de 20 aminoácidos, seguido de um pró-domínio de cinco aminoácidos contendo um sítio de clivagem

dibásico (Lys-Arg), porém seguido de um peptídeo maduro de 34 resíduos aminoácídicos (Figura 7).

A organização gênica de ambas *Cg-Defs* é bastante similar, consistindo em três éxons separados por dois íntrons. O primeiro e segundo éxons são de 49 e 12 pb, respectivamente, para ambas as *Cg-Defs*, sendo que o primeiro éxon codifica a parte inicial do peptídeo sinal, enquanto que o segundo éxon codifica a parte final do peptídeo sinal e o início do pró-domínio. O terceiro éxon, de 146 pb para a *Cg-Defh3* e 119 pb para a *Cg-Defh4*, codifica a parte final do pró-domínio e todo o peptídeo maduro de ambas as sequências (Figura 7).

Figura 7. Estrutura do precursor, sequências nucleotídica e aminoácídica deduzidas e estruturas gênicas das novas *Cg-Defs*.



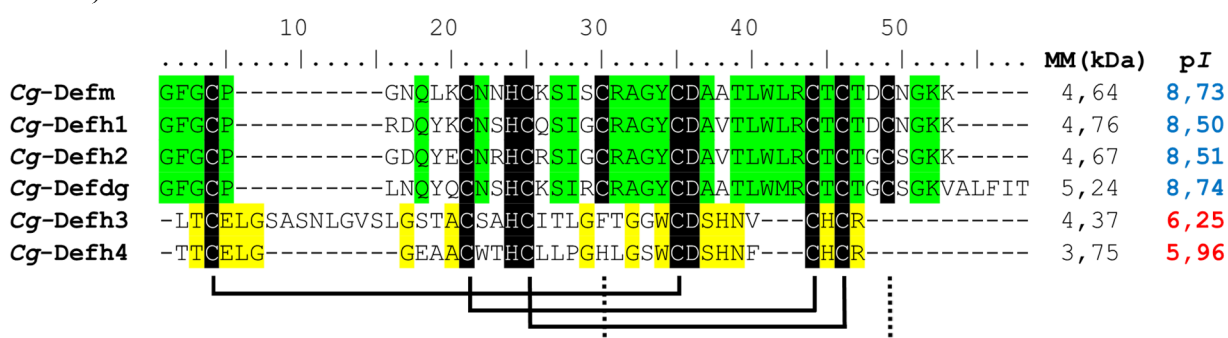
(A) Esquema ilustrativo do precursor das novas *Cg-Defs*. O retângulo branco representa a região do peptídeo sinal, em cinza o pró-domínio, enquanto que em preto é indicada a região codificante para o peptídeo maduro. As letras “N” e “C” indicam as regiões N- e C-terminal do peptídeo, respectivamente. (B) As sequências nucleotídicas e aminoácídicas destacam-se em letras minúsculas e maiúsculas respectivamente. O peptídeo sinal destaca-se sublinhado e o pró-domínio é correspondido pela região grifada em cinza. O códon de parada é representado por um asterisco. (C) Os retângulos e as linhas pretas representam os éxons e íntrons do gene, respectivamente. Os retângulos brancos representam a região codificante referente ao peptídeo sinal, em cinza referente ao pró-domínio, enquanto que em preto é indicada a região codificante para o peptídeo maduro. No canto inferior direito encontra-se a escala da representação, em um tamanho de 100 pares de bases.

As propriedades bioquímicas do peptídeo maduro das novas *Cg-Defs* foram caracterizadas. O peptídeo maduro da *Cg-Defh3* apresenta uma massa molecular de 4,37 kDa e um *pI* de 6,25, enquanto que a *Cg-Defh4* possui uma massa molecular de 3,75 kDa e um *pI*

de 5,96, demonstrando serem as primeiras defensinas do tipo CS α β com caráter aniônico identificadas em *C. gigas* (Figura 8).

As *Cg-Defh3* e *Cg-Defh4* apresentam uma identidade aminoacídica de 55% entre si, no entanto quando alinhadas com as outras *Cg-Defs* previamente caracterizadas, as novas defensinas apresentaram uma assinatura molecular diferente. Enquanto as defensinas *Cg-Defm*, *Cg-Defh1*, *Cg-Defh2* e *Cg-Defdg* são estabilizadas por quatro ligações dissulfeto (Cys₁₋₅Cys₂₋₆Cys₃₋₇Cys₄₋₈), as novas defensinas *Cg-Defh3* e *Cg-Defh4* compartilham apenas 12% de identidade com as defensinas previamente caracterizadas e são estabilizadas por apenas três ligações dissulfeto (Cys₁₋₄Cys₂₋₅Cys₃₋₆) (Figura 8).

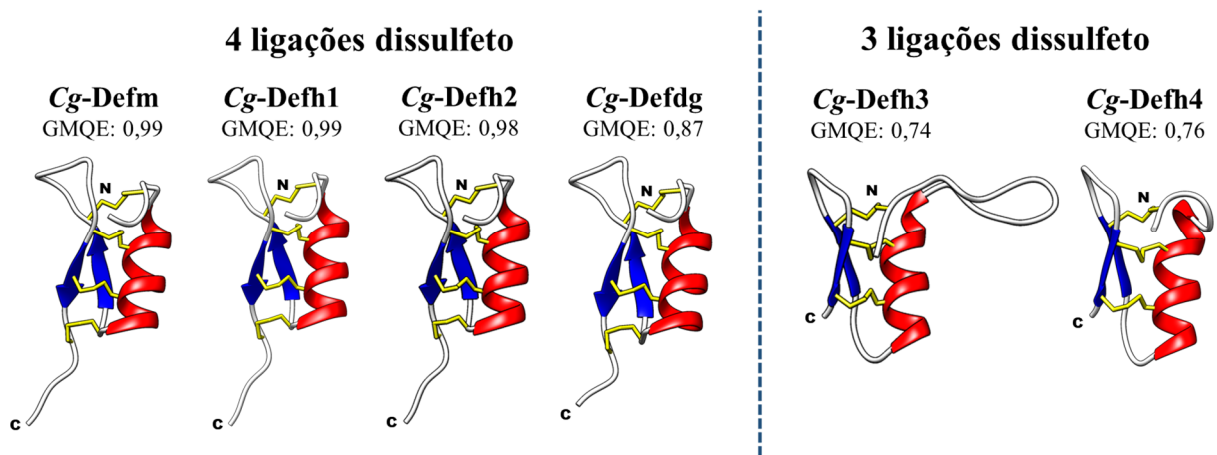
Figura 8. Alinhamento das defensinas do tipo CS α β de *Crassostrea gigas*: membros contendo oito resíduos de cisteínas (*Cg-Defm*; *Cg-Defh1*; *Cg-Defh2* e *Cg-Defdg*) e as novas defensinas identificadas neste estudo contendo seis resíduos conservados de cisteínas (*Cg-Defh3* e *Cg-Defh4*).



Os resíduos aminoacídicos de cisteínas, além daqueles conservados em todas as sequências estão destacados em preto enquanto que aqueles específicos das defensinas contendo oito e seis resíduos de cisteínas estão destacados em verde e amarelo, respectivamente. As barras abaixo das sequências correspondem à posição das ligações dissulfeto entre cisteínas e a barra serrilhada corresponde a ponte dissulfeto extra das defensinas contendo quatro pares de cisteínas. A massa molecular (MM) em quilodaltons (kDa) e o ponto isoelétrico teórico (pI) estão apresentados ao lado do alinhamento. Os valores de pI das defensinas com caráter catiônico e aniônico estão destacados em azul e vermelho, respectivamente.

Das defensinas do tipo CS α β previamente descritas em *C. gigas*, apenas a *Cg-Defm* teve sua conformação tridimensional resolvida, portanto as defensinas *Cg-Defh1*, *Cg-Defh2* e *Cg-Defdg* foram estruturadas a partir do modelo disponível da *Cg-Defm* (PDB: 2B68). No entanto, o modelamento tridimensional da *Cg-Defh3* e *Cg-Defh4* foi feito utilizando como modelo a estrutura resolvida do peptídeo Lucifensina (PDB: 2LLD), que consiste em uma defensina do tipo CS α β de inseto, presente na mosca *Lucilia sericata*. Os modelamentos resultaram em produtos com estruturas similares para todas as *Cg-Defs*, compostas de um α -hélice na porção N-terminal e duas folhas- β antiparalelas na porção C-terminal do peptídeo, estabilizadas por três ligações dissulfeto para as *Cg-Defh3* e *Cg-Defh4* e quatro ligações dissulfeto para as *Cg-Defh1*, *Cg-Defh2* e *Cg-Defdg* (Figura 9).

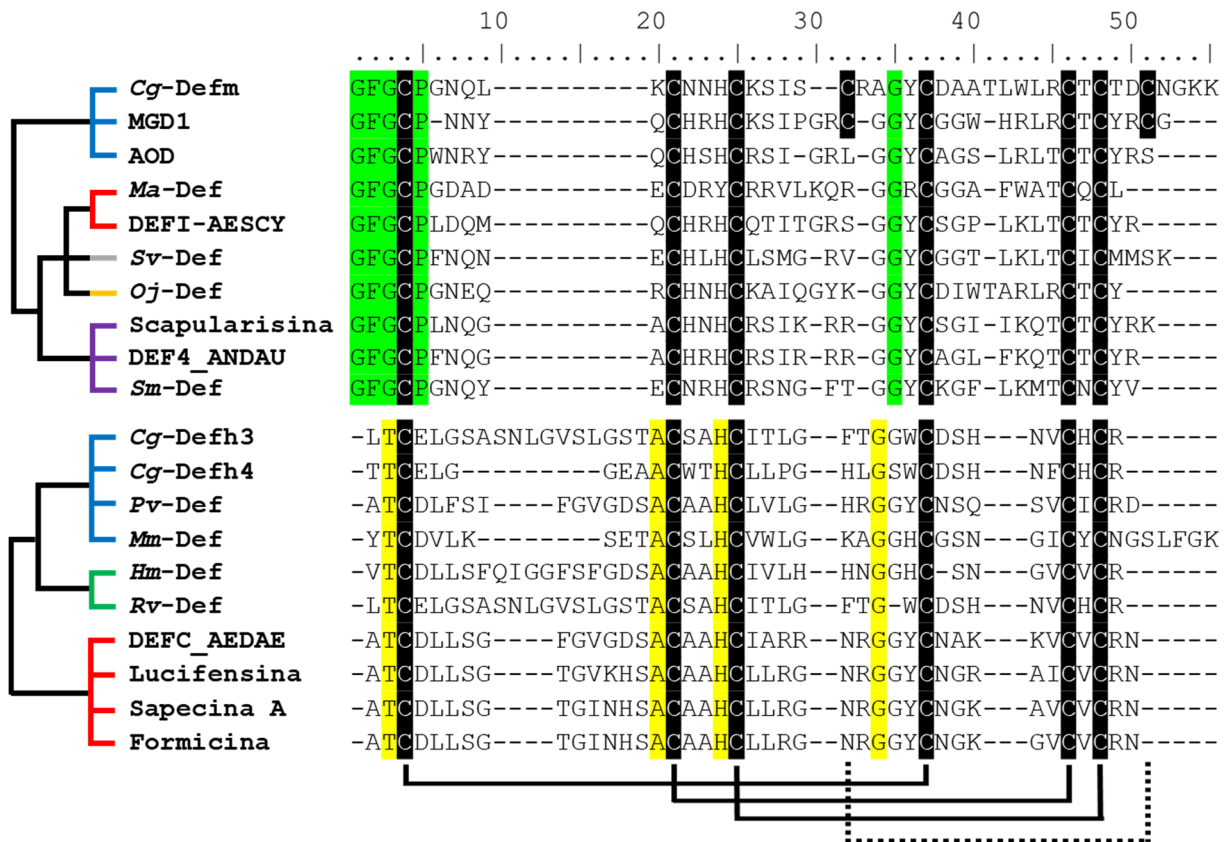
Figura 9. Modelamento tridimensional das Cg-Defs.



Em vermelho destaca-se a estrutura de uma α -hélice, em azul as duas folhas- β antiparalelas e em amarelo as ligações dissulfeto entre cisteínas. As letras “N” e “C” indicam as regiões N- e C-terminal do peptídeo, respectivamente.

As novas Cg-Defs foram alinhadas com outras defensinas do tipo CS $\alpha\beta$ encontradas em diferentes táxons (Bivalvia, Insecta, Collembola, Diplura, Arachnida e Gastropoda). O alinhamento permitiu a separação clara em dois grupos distintos de defensinas, que embora possuam assinaturas moleculares distintas, apresentam o posicionamento e pareamento entre cisteínas conservado entre todas as sequências. As defensinas presentes em moluscos bivalves e insetos se distribuem entre os dois diferentes grupos de defensinas. No entanto, alguns táxons, como os quelicerados e gastrópodes, apresentam apenas um tipo de assinatura molecular (Figura 10).

Figura 10. Alinhamento dos peptídeos maduros de defensinas do tipo CS β presentes em diferentes táxons.



Os resíduos aminoácidos de cisteínas estão destacados em preto, enquanto que os aminoácidos conservados apenas no agrupamento de defensinas superior e inferior estão destacados em verde e amarelo, respectivamente. As barras abaixo das sequências correspondem à posição das ligações dissulfeto e a barra serrilhada corresponde à ponte dissulfeto extra das defensinas contendo oito cisteínas. Ao lado esquerdo está representada uma filogenia simplificada da distribuição das defensinas do tipo CS β em diferentes táxons. Em azul estão representados indivíduos da classe Bivalvia (Genbank: Cg-Defm CAJ19280; MGD1 P80571; AOD P85008; Cg-Defh3 GT054002; Cg-Defh4 CU682678; Pv-Def MF362958; Mm-Def JI276475), em vermelho a classe Insecta (GenBank: Ma-Def GCMU01017328; DEFI-AESCY P80154; DEFC_AEDAE P81603; Lucifensinas HM243535; Sapecina A P18313; Formicina P10891), em cinza da classe Collembola (GenBank: Sv-Def GATZ02031221), em laranja da classe Diplura (GenBank: Oj-Def GAXJ02002559), em roxo da classe Arachnida (GenBank: Scapularisina EEC08934; DEF4_ANDAU P56686; Sm-Def KFM61041) e em verde da classe Gastropoda (GenBank: Hm-Def ALU63750; Rv-Def GDIA01212438).

4.3. Caracterização de novas big defensinas (Cg-BigDefs)

A autenticidade das novas Cg-BigDefs foi confirmada a partir da amplificação de suas sequências gênicas, pela técnica de PCR convencional, utilizando como amostras o gDNA de um *pool* de ostras provenientes do Brasil. Os produtos de amplificação foram clonados e submetidos à técnica de sequenciamento nucleotídico pelo método de Sanger. O resultado do sequenciamento validou a veracidade das sequências encontradas *in silico*.

As novas big defensinas identificadas a partir da prospecção *in silico* e sequenciadas receberam o nome de Cg-BigDef4 (GenBank: GHAT01130943), Cg-BigDef5 (GenBank: GECI01014168) e Cg-BigDef6 (GenBank: GECI01022844), sucedendo a lista de big

defensinas em *C. gigas* que previamente continha as *Cg-BigDef1*, *Cg-BigDef2* e *Cg-BigDef3* (ROSA et al., 2011).

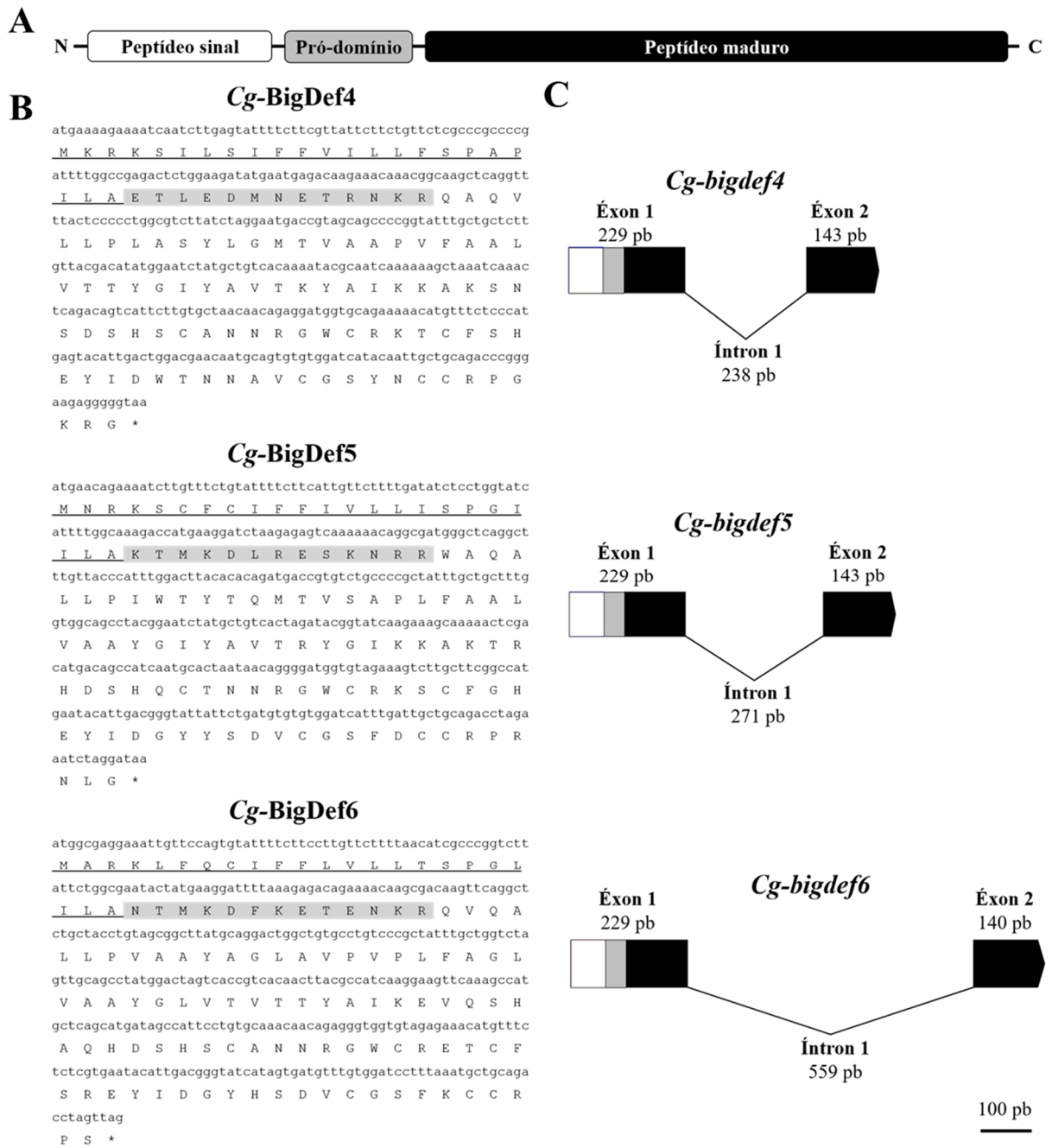
As sequências das três novas *Cg-BigDefs* foram traduzidas e apresentaram precursores com estruturas similares. Enquanto as *Cg-BigDef4* e *Cg-BigDef5* possuem um precursor de 123 resíduos aminoacídicos, a *Cg-BigDef6* contém um precursor de 122 aminoácidos. As três novas *Cg-BigDefs* são compostos por um peptídeo sinal de 23 aminoácidos, seguido de um pró-domínio de 13 aminoácidos, contendo um sítio de clivagem dibásico (Lys-Arg para *Cg-BigDef4* e *Cg-BigDef6*; Arg-Arg para *Cg-BigDef5*) e um peptídeo maduro de 87 aminoácidos para *Cg-BigDef4* e *Cg-BigDef5* e de 86 aminoácidos para *Cg-BigDef6* (Figura 11).

A organização gênica das três novas *Cg-bigdefs* é composta por dois éxons separados por um íntron. O primeiro éxon é de 229 pb para as três *Cg-bigdefs* e contém a sequência que codifica todo o peptídeo sinal, o pró-domínio e a parte inicial do peptídeo maduro. O segundo éxon é de 143 pb para *Cg-bigdef4* e *Cg-bigdef5* e de 140 pb para *Cg-bigdef6* e codifica a porção final do peptídeo maduro, responsável por codificar toda a região do domínio β -defensin-like (Figura 11).

As novas *Cg-BigDefs* foram caracterizadas quanto às propriedades bioquímicas de seus peptídeos maduros. A *Cg-BigDef4* apresenta um peptídeo maduro com uma massa molecular de 9,51 kDa e um *pI* catiônico de 9,30. Embora a *Cg-BigDef5* apresente o mesmo número de aminoácidos de sua antecessora, sua massa molecular é maior, devido a diferenciação dos resíduos aminoacídicos, apresentando uma massa de 9,87 kDa e um *pI*, também catiônico, de 9,00. Por sua vez a *Cg-BigDef6* contém uma massa molecular de 9,31 kDa, porém um ponto isoelétrico neutro de 7,04, sendo a única big defensina de *C. gigas* que não apresenta caráter catiônico (Figura 12).

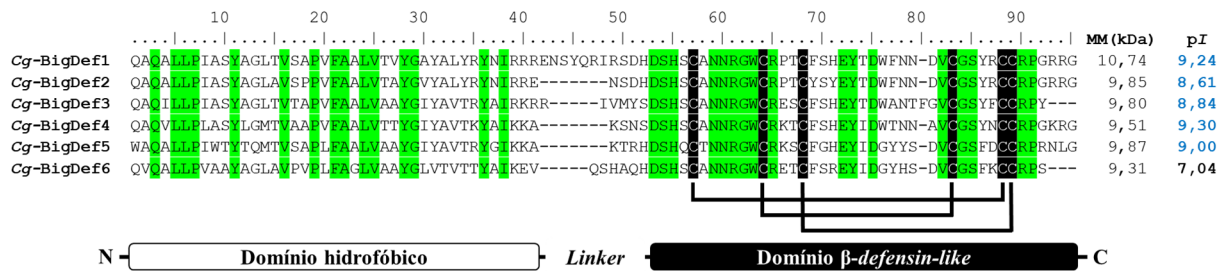
Assim como as demais *Cg-BigDefs*, as três novas sequências possuem uma região N-terminal hidrofóbica, conectada através de uma região *linker*, a uma domínio C-terminal β -defensin-like, que contém seis resíduos de cisteínas conservados e que estabilizam o peptídeo através de três ligações dissulfeto (Cys₁₋₅Cys₂₋₄Cys₃₋₆) (Figura 12).

Figura 11. Estrutura do precursor, seqüências nucleotídica e aminoacídica deduzidas e estruturas gênicas das novas *Cg-BigDefs*.



(A) Esquema ilustrativo do precursor das novas *Cg-BigDefs*. O retângulo branco representa a região do peptídeo sinal, em cinza o pró-domínio, enquanto que em preto é indicada a região codificante para o peptídeo maduro. As letras “N” e “C” indicam as regiões N- e C-terminal do peptídeo, respectivamente. (B) As seqüências nucleotídicas e aminoacídicas destacam-se em letras minúsculas e maiúsculas, respectivamente. O peptídeo sinal destaca-se sublinhado e o pró-domínio é correspondido pela região grifada em cinza. O códon de parada é representado por um asterisco. (C) Os retângulos e as linhas pretas representam os éxons e íntrons do gene, respectivamente. Os retângulos brancos representam a região codificante referente ao peptídeo sinal, em cinza referente ao pró-domínio, enquanto que em preto é indicada a região codificante para o peptídeo maduro. No canto inferior direito encontra-se a escala da representação, em um tamanho de 100 pares de bases.

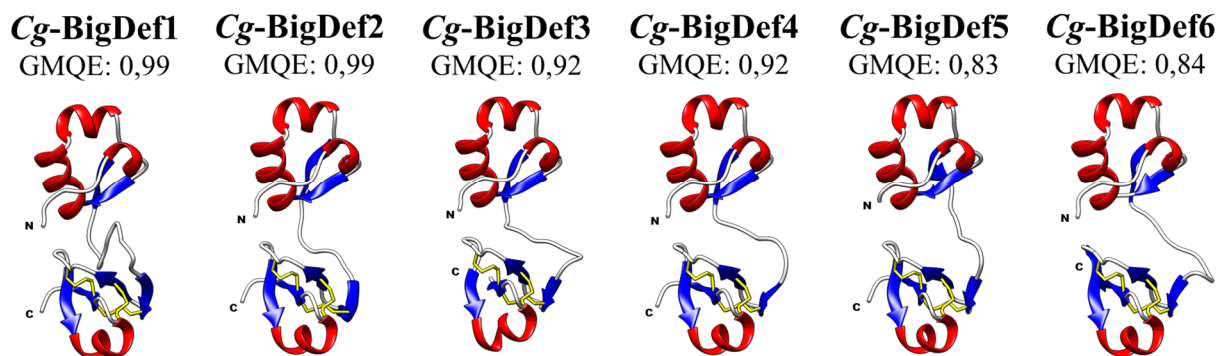
Figura 12. Alinhamento do peptídeo maduro e esquema ilustrativo dos domínios das big defensinas de *Crassostrea gigas* (*Cg-BigDef1*, *Cg-BigDef2*, *Cg-BigDef3*, *Cg-BigDef4*, *Cg-BigDef5* e *Cg-BigDef6*).



Os resíduos aminoacídicos conservados em todas as sequências estão destacados em verde enquanto que os resíduos de cisteínas estão destacados em preto. As barras abaixo das sequências correspondem à posição das ligações dissulfeto entre cisteínas (Cys₁₋₅Cys₂₋₄Cys₃₋₆). A massa molecular (MM) em kilodalton (kDa) e o ponto isoelétrico teórico (pI) estão apresentados ao lado do alinhamento. Os valores de pI das defensinas com caráter catiônico e neutro estão destacados em azul e preto, respectivamente. Abaixo do alinhamento, em escala, o retângulo branco destaca o domínio hidrofóbico, que se conecta ao retângulo preto, representando o domínio β -defensin-like, através de uma região *linker*. As letras “N” e “C” indicam as regiões N- e C-terminal do peptídeo, respectivamente.

O modelamento tridimensional de todas as *Cg-BigDefs* foi resolvido a partir da estrutura previamente elucidada da *Cg-BigDef1* (PDB: 6QBL). O resultado do modelamento demonstrou uma estrutura muito parecida entre todas as *Cg-BigDefs*, na qual elas possuem dois domínios globulares conectados por uma região *linker* (Figura 13).

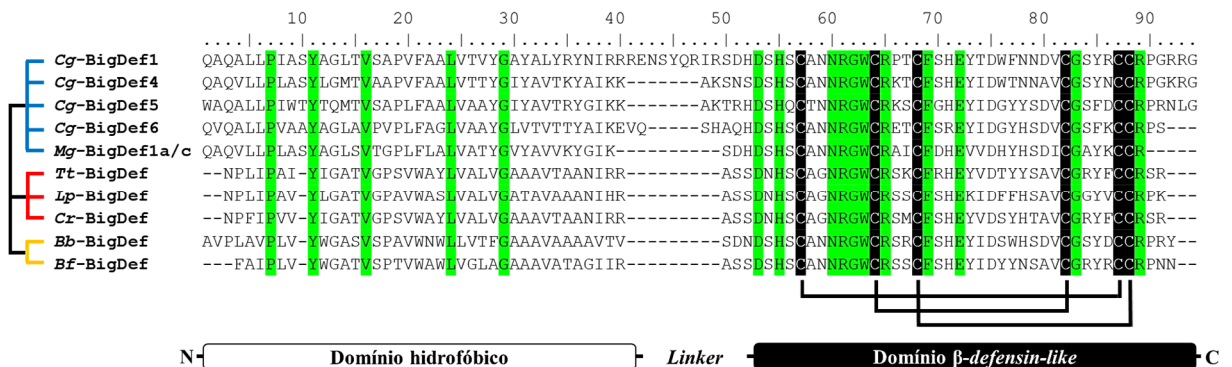
Figura 13. Modelamento tridimensional das *Cg-BigDefs*.



Em vermelho destacam-se as estruturas de α -hélice e em azul as folhas- β . O arranjo tridimensional do peptídeo é estabilizado através de três pontes dissulfeto (Cys₁₋₅Cys₂₋₄Cys₃₋₆) destacadas em amarelo. As letras “N” e “C” indicam as regiões N- e C-terminal do peptídeo respectivamente. As estruturas tridimensionais das big defensinas foram baseadas através da estrutura resolvida da *Cg-BigDef1* (PDB: 6QBL).

As novas *Cg-BigDefs* foram então alinhadas com outras big defensinas presentes em diferentes táxons (Mollusca, Arthropoda e Chordata). O resultado do alinhamento demonstrou que todas as big defensinas se agrupam em um mesmo grupo com assinatura molecular e pareamento entre cisteínas conservado entre todas as sequências (Figura 14).

Figura 14. Alinhamento dos peptídeos maduros de big defensinas presentes em diferentes táxons.



Os resíduos aminoacídicos conservados em todas as sequências estão destacados em verde enquanto que os resíduos de cisteínas estão destacados em preto. As barras abaixo das sequências correspondem à posição das ligações dissulfeto entre cisteínas (Cys₁₋₅Cys₂₋₄Cys₃₋₆). Abaixo do alinhamento, em escala, o retângulo branco destaca o domínio hidrofóbico, que se conecta ao retângulo preto, representando o domínio β-defensin-like, através de uma região linker. As letras “N” e “C” indicam as regiões N- e C-terminal do peptídeo, respectivamente. Ao lado esquerdo está representada uma filogenia simplificada da distribuição das big defensinas em diferentes táxons. Em azul estão representados indivíduos do filo Mollusca (Genbank: *Cg-BigDef1* JF703138; *Cg-BigDef4* GHAT01130943; *Cg-BigDef5* GECI01014168; *Cg-BigDef6* GECI01022844; *Mg-BigDef1a/c* CCC15007), em vermelho do filo Arthropoda (GenBank: *Tt-BigDef* P80957; *Lp-BigDef* XR_001291390; *Cr-BigDef* CK086629) e em laranja do filo Chordata (GenBank: *Bj-BigDef* AAO18674; *Bf-BigDef* ADH03420).























































4.4. Análise transcricional das novas defensinas de *C. gigas*

A fim de avaliar o perfil transcricional das novas *Cg-Defs* e *Cg-BigDefs* frente a infecções por patógenos de interesse ostreícola, foi delineada uma infecção experimental. De maneira geral, os resultados demonstraram que as novas defensinas apresentam uma expressão muito baixa, ou até mesmo ausente, ainda que após a infecção experimental. Enquanto que os transcritos para a *Cg-Defh3* foram encontrados em todas as amostras infectadas com o herpesvírus (OsHV-1 μVar), sua expressão foi indetectável em um pool de amostras controle e em um pool de animais infectados com *V. tasmaniensis* LGP32, por sua vez os transcritos referentes a *Cg-Defh4* foram encontrados somente em um dos três pools de ostras infectadas com *V. tasmaniensis* LGP32. Em relação às *Cg-BigDefs*, apenas a *Cg-BigDef4* foi detectada em dois pools de amostras de ostras infectadas com *V. tasmaniensis* LGP32 e em um pool de ostras infectadas com o herpesvírus (OsHV-1 μVar). Em todas as outras amostras e em relação as *Cg-BigDef5* e *Cg-BigDef6*, não houve qualquer sinal de detecção dos transcritos (Tabela 2).

Como não houve amplificação em todas as amostras, não foi possível quantificar a expressão gênica das defensinas, no entanto a detecção de um outro gene associado à imunidade atestou a qualidade das amostras. A expressão da proteína indutora da permeabilidade bacteriana (*Cg-BPI1*), que consiste em um gene de simples cópia, presente em todas as ostras

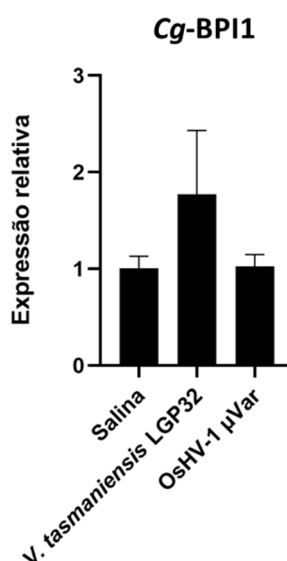
C. gigas e constitutivamente expresso nos hemócitos (SCHMITT et al., 2010b), demonstrou estar presente em todas as amostras (Figura 15), sendo detectada no ciclo de quantificação (Cq) próximo ao ciclo 23 da qPCR (Tabela 2).

Tabela 2. Análise da expressão das novas defensinas em ostras experimentalmente infectadas com a bactéria *Vibrio tasmaniensis* LGP32 e com o herpesvírus de ostra (OsHV-1 μ Var).

| Genes | | Salina | | | <i>V. tasmaniensis</i> LGP32 | | | OsHV-1 μ Var | | |
|-------------------|------------------|---|---|---|---|---|--|---|---|---|
| | | Pool 1 | Pool 2 | Pool 3 | Pool 1 | Pool 2 | Pool 3 | Pool 1 | Pool 2 | Pool 3 |
| Cg-Defh3 | Cq | 33,9 | - | 32,6 | 34,8 | - | 34,7 | 29,5 | 34,9 | 29,2 |
| | Tm (°C) | 81,1 | - | 81,1 | 81,1 | - | 81,0 | 81,0 | 81,0 | 81,1 |
| | Curva de melting |  |  |  |  |  |  |  |  |  |
| Cg-Defh4 | Cq | - | - | - | - | 33,9 | - | - | - | - |
| | Tm (°C) | - | - | - | - | 81,1 | - | - | - | - |
| | Curva de melting |  |  |  |  |  |  |  |  |  |
| Cg-BigDef4 | Cq | - | - | - | - | 34,7 | 34,3 | 34,8 | - | - |
| | Tm (°C) | - | - | - | - | 81,9 | 81,1 | 82,4 | - | - |
| | Curva de melting |  |  |  |  |  |  |  |  |  |
| Cg-BigDef5 | Cq | - | - | - | - | - | - | - | - | - |
| | Tm (°C) | - | - | - | - | - | - | - | - | - |
| | Curva de melting |  |  |  |  |  |  |  |  |  |
| Cg-BigDef6 | Cq | - | - | - | - | - | - | - | - | - |
| | Tm (°C) | - | - | - | - | - | - | - | - | - |
| | Curva de melting |  |  |  |  |  |  |  |  |  |
| Cg-BPI1 | Cq | 23,2 | 23,7 | 22,8 | 22,6 | 21,0 | 23,6 | 24,3 | 22,6 | 22,4 |
| | Tm (°C) | 82,9 | 83,2 | 82,7 | 83,0 | 82,9 | 82,9 | 83,0 | 82,9 | 83,3 |
| | Curva de melting |  |  |  |  |  |  |  |  |  |

Foram analisados os ciclos de quantificação (*cycle quantification* ou Cq) no qual os transcritos foram detectados, além da temperatura de hibridização (*melting temperature* ou Tm) dos produtos de amplificação, para garantir a amplificação específica apenas dos genes de interesse. Como amostras foram analisados três *pools* de cinco animais cada, de um grupo controle ou salina, que apenas simula a injúria do animal pela injeção do veículo (água do mar filtrada e autoclavada), do grupo infectado experimentalmente com bactérias *Vibrio tasmaniensis* LGP32 e do grupo infectado com o herpesvírus de ostra (OsHV-1 μ Var).

Figura 15. Análise da expressão de *Cg-BP1* nos hemócitos de ostras experimentalmente infectadas com a bactéria *Vibrio tasmaniensis* LGP32 e com o herpesvírus de ostra (OsHV-1 μ Var).



Os resultados consistem na média \pm desvio padrão de triplicatas biológicas (3 *pools* de 5 animais) por condição experimental: grupo controle (injetado com salina) e grupos desafiados com bactéria *Vibrio tasmaniensis* LGP32 e com o herpesvírus de ostra (OsHV-1 μ Var).

O resultado desse experimento levantou a hipótese de que as novas defensinas não são necessariamente pouco expressas, mas sim que estão sujeitas ao fenômeno genético de variação por presença/ausência gênica (PAV), estando completamente ausente no genoma de alguns indivíduos, gerando assim uma baixa ou nenhuma detecção dos transcritos em um *pool* de animais.

4.5. Investigação do fenômeno de presença/ausência de genes (PAV)

O fenômeno de PAV foi descrito pela primeira vez em invertebrados por Rosa e colaboradores (2015), sob a perspectiva da expressão das *Cg-BigDef1*, *Cg-BigDef2* e *Cg-BigDef3*. Tendo em vista a ocorrência desse fenômeno genético sob genes imunológicos de *C. gigas*, em específico de big defensinas, foi delineado um experimento para investigar o fenômeno de PAV relacionado às novas *Cg-Defs* e *Cg-BigDefs* identificadas neste trabalho. Para tal, foram coletadas amostras de gDNA de 16 animais provenientes do Brasil e de seis animais originários da França e avaliado a variação da presença/ausência dos genes da defensinas sob indivíduos de diferentes populações.

Os resultados demonstraram um massivo fenômeno de PAV sob os genes das defensinas, mesmo em indivíduos provenientes de um mesmo país. Além disso, foi notado uma enorme diversidade de combinações na distribuição desses genes por indivíduo, desde aqueles







que não apresentam nenhum dos cinco genes, até aqueles que apresentam apenas *Cg-Def3*, ou apenas *Cg-BigDefs* ou combinações de *Cg-Def3* e *Cg-BigDefs* (Figura 16).

O gene codificante para a *Cg-Defh3* encontrou-se presente em apenas dois dos 16 indivíduos provenientes do Brasil, porém manifestou-se em três dos seis animais franceses testados. Por outro lado, o gene *Cg-Defh4* não foi detectado em nenhum animal brasileiro, porém estava presente em um indivíduo oriundo da França.

Em relação às *Cg-BigDefs*, houve uma considerável diferença na presença dos genes entre os animais de diferentes países. Enquanto que o gene para a *Cg-BigDef4* encontrou-se presente em 12 dos 16 animais brasileiros, apenas um animal francês demonstrou conter esse gene. O mesmo ocorreu para a *Cg-BigDef6*, que foi encontrada em 4 indivíduos provenientes do Brasil, porém em nenhum animal francês. Por outro lado, a *Cg-BigDef5* demonstrou estar presente em 5 das 6 ostras francesas testadas, porém em apenas dois indivíduos brasileiros.

Figura 16. Padrão da variação por presença/ausência (PAV) dos genes relacionados as novas defensas em indivíduos de *C. gigas* provenientes do Brasil e da França.

A

| Ostras do Brasil | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | C+ | |
|-------------------|---|---|---|---|---|---|---|---|---|----|----|----|----|----|----|----|---|--------|
| <i>Cg-Defh3</i> | | | | | | | | | | | | | | | | |  | 276 pb |
| <i>Cg-Defh4</i> | | | | | | | | | | | | | | | | |  | 638 pb |
| <i>Cg-BigDef4</i> | | | | | | | | | | | | | | | | |  | 640 pb |
| <i>Cg-BigDef5</i> | | | | | | | | | | | | | | | | |  | 655 pb |
| <i>Cg-BigDef6</i> | | | | | | | | | | | | | | | | |  | 928 pb |
| <i>Cg-Actin</i> | | | | | | | | | | | | | | | | |  | 846 pb |

B

| Ostras da França | 1 | 2 | 3 | 4 | 5 | 6 | C+ | |
|-------------------|---|---|---|---|---|---|---|--------|
| <i>Cg-Defh3</i> | | | | | | |  | 276 pb |
| <i>Cg-Defh4</i> | | | | | | |  | 638 pb |
| <i>Cg-BigDef4</i> | | | | | | |  | 640 pb |
| <i>Cg-BigDef5</i> | | | | | | |  | 655 pb |
| <i>Cg-BigDef6</i> | | | | | | |  | 928 pb |
| <i>Cg-Actin</i> | | | | | | |  | 846 pb |

(A) Amostras de gDNA de 16 animais coletados no Brasil e (B) seis animais provenientes da França. Em vermelho destaca-se quando o gene não estava presente e em verde quando houve a detecção do gene pela técnica de PCR. Ao lado direito está destacado o controle positivo (C+) da técnica de PCR, utilizando no lugar da amostra de gDNA, 1 ng de plasmídeos contendo o gene de interesse. O gene da *Cg-actin* foi utilizado como controle endógeno, para confirmar a qualidade das amostras testadas. Os tamanhos correspondentes a cada produto de amplificação encontram-se a direita, em pares de bases.

5. DISCUSSÃO

As defensas são importantes efetores de defesa antimicrobiana, desempenhando um importante papel na regulação da microbiota e atuando como uma primeira barreira de defesa contra microrganismos invasores. Defensas do tipo CS $\alpha\beta$ e big defensas haviam sido anteriormente caracterizadas na ostra *Crassostrea gigas* (GONZALEZ et al., 2007b; GUEGUEN et al., 2006; ROSA et al., 2011; ZHANG et al., 2018), entretanto não havia relatos, até o presente estudo, da existência de tamanha variedade de genes, estruturas e características bioquímicas dessas moléculas, que são, na verdade, muito mais diversas em ostras do que anteriormente se acreditava.

No presente estudo, dois novos membros da família de defensas do tipo CS $\alpha\beta$ e três novas big defensas foram identificadas e caracterizadas, nas quais cada nova sequência encontrada é codificada por seu próprio gene. Eventos de duplicação gênica, seguidos por uma rápida diversificação das sequências de defensas foram propostos para explicar a alta diversidade desses peptídeos em animais, como em *C. gigas* (GERDOL et al., 2020a). As defensas sofreram uma forte pressão evolutiva a fim de manter uma alta afinidade por componentes estruturais exclusivos de microrganismos, onde a modificação de alguns resíduos aminoacídicos permite pequenas alterações de carga ou polaridade, aumentando a diversificação de estruturas moleculares e por fim, resultando na seleção de transcritos com características bioquímicas mais vantajosas em relação à atividade antimicrobiana (SCHMITT et al., 2010a).

As duas novas sequências de defensas do tipo CS $\alpha\beta$ identificadas em *C. gigas* possuem uma assinatura molecular e características bioquímicas diferentes daquelas descritas para as *Cg*-Defs previamente relatadas. Enquanto as *Cg*-Defm, *Cg*-Defh1, *Cg*-Defh2 e *Cg*-Defdg possuem uma assinatura molecular clássica de defensas identificadas em moluscos (*ancient invertebrate-type defensins* ou AITDs), contendo uma região N-terminal com uma assinatura aminoacídica conservada “GFGC” (Gly-Phe-Gly-Cys), as novas *Cg*-Defs possuem uma assinatura homóloga às defensas classicamente descritas em insetos (*classical insect-type defensins* ou CITDs), as quais possuem uma região N-terminal com assinatura aminoacídica “ATCDL” (Ala-Thr-Cys-Asp-Leu) (ZHU, 2008). Análises filogenéticas sugerem que a divergência entre as duas assinaturas moleculares homólogas ocorreu através de uma duplicação gênica há um bilhão de anos, quando o ancestral comum dos fungos e animais ainda não havia divergido (ZHU, 2008).

As diferentes assinaturas moleculares das defensas tipo CS $\alpha\beta$, presentes em moluscos e artrópodes (AITDs e CITDs), se mantêm conservadas há milhões de anos, mantendo

ainda uma alta identidade entre as sequências, permitindo aferir a importância da manutenção desses genes para a sobrevivência dos organismos. Estudos recentes sugerem ainda que, assim como as *Cg*-Defs com assinatura “GFGC”, as defensinas com assinatura “ATCDL” também possuem a capacidade de se ligar ao lipídio II, precursor da peptidoglicana da parede celular das bactérias, conferindo-lhes assim uma atividade antimicrobiana (SCHMITT; ROSA; DESTOUMIEUX-GARZÓN, 2016).

De maneira geral, as defensinas são AMPs catiônicos, no qual devido a sua carga carregada positivamente, interagem com componentes eletronegativos presentes nas membranas bacterianas, como por exemplo o LPS das bactérias Gram-negativas, ou com componentes da peptidoglicana de paredes bacterianas (BROGDEN, 2005). No entanto, as novas defensinas do tipo $CS\alpha\beta$ *Cg*-Defh3 e *Cg*-Defh4, aqui identificadas, são as primeiras defensinas de caráter aniônico descritas em ostras *C. gigas*. Testes antimicrobianos demonstraram que mesmo as defensinas aniônicas possuem atividade contra bactérias Gram-positivas e Gram-negativas (WEI et al., 2015; XU et al., 2016), contudo ainda com relatos restritos na literatura, em relação às defensinas catiônicas, cujo mecanismo de ação é melhor compreendido (SCHMITT et al., 2010a). Portanto, se fazem necessários experimentos futuros, como os de sinergismo entre as novas *Cg*-Defs aniônicas e outros AMPs catiônicos, para melhor compreender a diversidade bioquímica das defensinas.

Outro fato curioso é que existem diferenças estruturais importantes entre as *Cg*-Defs anteriormente descritas (*Cg*-Defm, *Cg*-Defh1, *Cg*-Defh2 e *Cg*-Defdg), que são estabilizadas por quatro ligações dissulfeto, e as novas *Cg*-Defs (*Cg*-Defh3 e *Cg*-Def4), estabilizadas por apenas três ligações dissulfeto. É importante ressaltar que a presença adicional de uma ligação dissulfeto na estrutura de defensinas, como ocorre na defensina MGD-1 da espécie de molusco bivalve *Mytilus galloprovincialis*, lhe confere uma maior estabilidade em meios com alta osmolaridade, como por exemplo, na água do mar (YANG et al., 2000).

Em relação às três novas big defensinas identificadas e caracterizadas no presente trabalho, todas possuem uma assinatura molecular conservada e com alta identidade com as sequências previamente descritas de *Cg*-BigDefs. Todas as seis *Cg*-BigDefs descritas até então (*Cg*-BigDef1 a *Cg*-BigDef6), possuem uma região N-terminal hidrofóbica, conectada a um domínio C-terminal β -defensin-like através de uma região *linker* (GERDOL et al., 2020a).

Sabe-se que a atividade antimicrobiana da *Cg*-BigDef1 está associada à formação de *nanonets*, na qual as regiões N-terminais hidrofóbicas das big defensinas se agrupam para formar uma trama de fibras que permite o aprisionamento físico dos microrganismos, possibilitando assim que a região β -defensin-like desempenhe atividade antimicrobiana (LOTH

et al., 2019). Considerando a alta identidade encontrada entre as sequências das novas *Cg-BigDefs* com àquela da *Cg-BigDef1*, pode-se supor que o mecanismo de ação desses peptídeos seja similar, atuando na defesa antimicrobiana através da formação de *nanonets*. No entanto, essa hipótese deverá ser comprovada em ensaios futuros, juntamente com testes antimicrobianos envolvendo as demais *Cg-BigDefs*, pois ainda não é conhecido o espectro de atividade desses AMPs.

Estudos filogenéticos sugerem que o gene responsável por codificar as big defensas surgiu quando o ancestral de todos os moluscos, artrópodes e vertebrados ainda não havia divergido (GERDOL et al., 2020a). No entanto, ao longo da escala evolutiva, houve inúmeros casos de deleções gênicas em relação aos genes de big defensas, sugerindo que provavelmente houve uma forte pressão seletiva para a manutenção desses genes durante uma longa escala temporal, indicando a importância das big defensas para a sobrevivência dos organismos que as mantiveram (GERDOL et al., 2020a).

Curiosamente, a diversidade de genes codificantes para big defensas em moluscos bivalves muitas vezes é mais expressiva em nível intraespecífico do que entre espécies diferentes. Aparentemente, a diversificação das big defensas ocorre com bastante frequência em nível de espécie, por meio do surgimento de genes parálogos, originados a partir de diversas duplicações gênicas ao longo da evolução (GERDOL et al., 2020a). O mesmo foi anteriormente postulado para as defensas do tipo $CS\alpha\beta$ *Cg-Defm*, *Cg-Defh1* e *Cg-Defh2*, na qual a diversidade de *Cg-Defs* originou-se a partir de eventos de duplicação gênica, seguidos da especificação dessas defensas em diferentes tecidos, como no manto ou nos hemócitos (SCHMITT et al., 2010b).

O padrão de regulação da expressão das *Cg-Defs* e *Cg-BigDefs* é bastante variado e aparentemente cada diferente sequência se especializou em diferentes funções imunológicas. Em estudos com ostras *C. gigas* desafiadas por bactérias Gram-negativas do gênero *Vibrio*, o transcrito codificante para a defesa *Cg-Defm* não apresentou nenhuma modulação na expressão no manto (GUEGUEN et al., 2006), enquanto que a expressão das *Cg-Defh1* e *Cg-Defh2* sofreu uma forte repressão nos hemócitos, decorrente da infiltração hemocitária para outros tecidos (GONZALEZ et al., 2007b). De forma similar, os genes das big defensas *Cg-BigDef1* e *Cg-BigDef2* mostraram-se altamente induzidos frente a infecções bacterianas, enquanto que a *Cg-BigDef3* não foi modulada, mantendo-se constitutivamente expressa nos hemócitos (ROSA et al., 2011). Em relação a infecções virais, causadas pelo herpesvírus de ostra (OsHV-1 μ Var), Lorgeril e colaboradores (2018) conduziram um estudo em que as defensas *Cg-Defm* e *Cg-Defh1* não apresentaram modulação da expressão, enquanto que a

Cg-BigDef2 apresentou uma forte indução e a *Cg-BigDef3* apresentou forte repressão da expressão. Esses resultados indicam que a expressão das defensinas em *C. gigas* pode estar intimamente relacionada com seus papéis imunológicos de combate a diferentes fontes de infecções causadas por patógenos virais ou bacterianos.

No presente estudo não foi possível, contudo, identificar o perfil de expressão gênica das novas *Cg-Defs* e *Cg-BigDefs* em ostras desafiadas com bactérias *Vibrio tasmaniensis* LGP32 e com o herpesvírus de ostra (OsHV-1 μ Var), pois na maioria das amostras, a expressão das novas defensinas ficou abaixo do limiar de detecção pela técnica de RT-qPCR, não permitindo assim a amplificação das sequências.

Uma das possíveis hipóteses para a baixa detecção das sequências gênicas das novas *Cg-Defs* e *Cg-BigDefs* está relacionado ao fenômeno de PAV. O presente estudo demonstrou que todas as novas defensinas caracterizadas neste trabalho sofrem massivamente o fenômeno genético de PAV, no qual algumas ostras foram diagnosticadas positivas para os genes das novas defensinas, enquanto que esses genes estavam completamente ausentes em outras, mesmo em indivíduos de uma mesma população. Tais achados corroboram com os resultados obtidos por Rosa e colaboradores (2015), onde foi confirmado o fenômeno de PAV atuando sobre os genes das *Cg-BigDef1*, *Cg-BigDef2* e *Cg-BigDef3*.

De maneira interessante, os indivíduos analisados no presente estudo não apresentaram simultaneamente mais de dois genes codificantes para as defensinas, considerando os genes *Cg-Defh3*, *Cg-Defh4*, *Cg-BigDef4*, *Cg-BigDef5* e *Cg-BigDef6*. O mesmo resultado foi obtido por Rosa e colaboradores (2015), no qual foi avaliado a presença dos três genes das *Cg-BigDefs*, *Cg-BigDef1*, *Cg-BigDef2* e *Cg-BigDef3*, em 163 ostras *C. gigas*, e foi obtido que 60,7% das ostras apresentaram apenas um gene, 15,4% continham a combinação de duas big defensinas e nenhum animal apresentou simultaneamente os três genes. A ausência de genes para algumas *Cg-Defs* e *Cg-BigDefs* também é relatado nos genomas sequenciados e publicados da ostra *C. gigas* (WANG et al., 2012, 2019).

O fenômeno de PAV parece estar relacionado à sobrevivência e à adaptação evolutiva dos organismos (GERDOL et al., 2020b). Esse fenômeno permite com que uma mesma população mantenha uma grande diversidade de genes imunológicos, distribuídos entre diferentes indivíduos. Desta maneira se apenas uma pequena parcela de animais é selecionada por determinada condição ambiental, não há grande perda de variabilidade genética da população como um todo, pois os indivíduos sobreviventes possuem uma combinação de todos os genes (GERDOL et al., 2020b). Além disso, a manutenção de poucos genes imunológicos

em um único indivíduo permite que seu balanço energético se torne positivo, garantindo a sobrevivência das próximas gerações (GERDOL et al., 2020a).

Analisando conjuntamente nossos resultados, conclui-se que é necessário repetir o experimento de análise da expressão gênica das novas defensinas, porém utilizando indivíduos previamente conhecidos como portadores para cada um dos genes das novas *Cg-Defs* e *Cg-BigDefs*. Dessa maneira, será possível conhecer o perfil de distribuição da expressão das novas defensinas em diferentes tecidos, além da modulação da expressão após desafios com microrganismos patogênicos. Em relação à investigação do fenômeno de PAV envolvendo todas as *Cg-Defs* e *Cg-BigDefs*, novos experimentos deverão ser conduzidos com um maior número de animais, para que possa ser avaliado de maneira mais fidedigna se existe algum padrão diferenciado de distribuição dos genes para defensinas entre os indivíduos de uma população. Adicionalmente, a caracterização do espectro de atividade antimicrobiana das novas *Cg-Defs* e *Cg-BigDefs* deverá ser realizado através de testes *in vitro*, contra uma gama de microrganismos de interesse clínico e ambiental. Assim, será possível investigar se as diferentes assinaturas moleculares das defensinas estão relacionadas também a um maior espectro antimicrobiano.

A manutenção dos genes das *Cg-Defs* e *Cg-BigDefs* ao longo de milhões de anos na escala evolutiva sugere a importância dos mesmos no controle da microbiota e no combate a microrganismos patogênicos, o que torna esses peptídeos candidatos interessantes para o desenvolvimento de novos antibióticos naturais. Além disso, o expressivo fenômeno de PAV atuando sob os genes associado à imunidade poderia servir futuramente como marcadores moleculares para a seleção de ostras mais resistentes a doenças presentes nos cultivos desses animais.

6. CONCLUSÃO

Este trabalho demonstrou que ostras da espécie *Crassostrea gigas* apresentam uma grande diversidade de defensinas do tipo CS $\alpha\beta$ e big defensinas, ainda maior que o preconizado na literatura até o momento. A diversidade de defensinas ocorre tanto em nível de quantidade de sequências, como de variedade das assinaturas moleculares, padrões de expressão e características bioquímicas, revelando pela primeira vez, em ostras *C. gigas*, a presença de defensinas de caráter aniônico. O fenômeno de PAV também foi encontrado relacionado às novas sequências de defensinas identificadas, sugerindo que esse fenômeno genético seja comum em genes relacionados à imunidade de moluscos bivalves.

REFERÊNCIAS

- BACHÈRE, E. et al. The new insights into the oyster antimicrobial defense: Cellular, molecular and genetic view. **Fish and Shellfish Immunology**, v. 46, n. 1, p. 50–64, 2015.
- BROGDEN, K. A. **Antimicrobial peptides: Pore formers or metabolic inhibitors in bacteria?** *Nature Reviews Microbiology*, 2005.
- BULET, P.; STÖCKLIN, R.; MENIN, L. Anti-microbial peptides: From invertebrates to vertebrates. **Immunological Reviews**, v. 198, p. 169–184, 2004.
- DAS, S. et al. Comparative genomics and evolution of the alpha-defensin Multigene family in primates. **Molecular Biology and Evolution**, v. 27, n. 10, p. 2333–2343, 2010.
- DEFER, D. et al. Antimicrobial peptides in oyster hemolymph: The bacterial connection. **Fish and Shellfish Immunology**, v. 34, n. 6, p. 1439–1447, 2013.
- DESTOUMIEUX-GARZÓN, D. et al. Antimicrobial peptides in marine invertebrate health and disease. **Philosophical Transactions of the Royal Society B: Biological Sciences**, 2016.
- DUBOS, R. J. Studies on a bactericidal agent extracted from a soil bacillus: I. Preparation of the agent. Its activity in vitro. **Journal of Experimental Medicine**, v. 70, n. 1, p. 1–10, 1939.
- GABUR, I. et al. Gene presence-absence variation associates with quantitative *Verticillium longisporum* disease resistance in *Brassica napus*. **Scientific Reports**, v. 10, n. 1, p. 1–11, 2020.
- GANZ, T. et al. Defensins. Natural peptide antibiotics of human neutrophils. **Journal of Clinical Investigation**, v. 76, n. 4, p. 1427–1435, 1985.
- GANZ, T. The role of antimicrobial peptides in innate immunity. **Integrative and Comparative Biology**, v. 43, n. 2, p. 300–304, 2003.
- GERDOL, M. et al. Functional Insights From the Evolutionary Diversification of Big Defensins. **Frontiers in Immunology**, v. 11, n. April, p. 1–16, 2020a.
- GERDOL, M. et al. Massive gene presence/absence variation in the mussel genome as an adaptive strategy: first evidence of a pan-genome in Metazoa. 2020b.
- GONZALEZ, M. et al. Evidence of a bactericidal permeability increasing protein in an invertebrate, the *Crassostrea gigas* Cg-BPI. **Proceedings of the National Academy of Sciences of the United States of America**, v. 104, n. 45, p. 17759–17764, 2007a.
- GONZALEZ, M. et al. Molecular characterization of two isoforms of defensin from hemocytes of the oyster *Crassostrea gigas*. **Developmental and Comparative Immunology**, 2007b.
- GUANÍ-GUERRA, E. et al. **Antimicrobial peptides: General overview and clinical implications in human health and disease** *Clinical Immunology*, 2010.
- GUEGUEN, Y. et al. Immune gene discovery by expressed sequence tags generated from hemocytes of the bacteria-challenged oyster, *Crassostrea gigas*. **Gene**, v. 303, n. 1–2, p. 139–145, 2003.

GUEGUEN, Y. et al. Characterization of a defensin from the oyster *Crassostrea gigas*: Recombinant production, folding, solution structure, antimicrobial activities, and gene expression. **Journal of Biological Chemistry**, v. 281, n. 1, p. 313–323, 2006.

GUEGUEN, Y. et al. Oyster hemocytes express a proline-rich peptide displaying synergistic antimicrobial activity with a defensin. **Molecular Immunology**, v. 46, n. 4, p. 516–522, 2009.

HANCOCK, R. E. W.; SAHL, H. G. **Antimicrobial and host-defense peptides as new anti-infective therapeutic strategies** *Nature Biotechnology*, 2006.

HE, X.; ZHANG, Y.; YU, Z. An Mpeg (macrophage expressed gene) from the Pacific oyster *Crassostrea gigas*: Molecular characterization and gene expression. **Fish and Shellfish Immunology**, v. 30, n. 3, p. 870–876, 2011.

KAISER, V.; DIAMOND, G. Expression of mammalian defensin genes. **Journal of leukocyte biology**, v. 68, n. 6, p. 779–84, 2000.

KIM, W. J.; JUNG, H.; GAFFNEY, P. M. Development of Type I Genetic Markers from Expressed Sequence Tags in Highly Polymorphic Species. **Marine Biotechnology**, v. 13, n. 2, p. 127–132, 2011.

KOO, H. B.; SEO, J. Antimicrobial peptides under clinical investigation. **Peptide Science**, v. 111, n. 5, 2019.

LEI, J. et al. The antimicrobial peptides and their potential clinical applications. **American Journal of Translational Research**, v. 11, n. 7, p. 3919–3931, 2019.

LI, Y. et al. CgRel involved in antibacterial immunity by regulating the production of CgIL17s and CgBigDef1 in the Pacific oyster *Crassostrea gigas*. **Fish and Shellfish Immunology**, v. 97, p. 474–482, 2020.

LIVAK, K. J.; SCHMITTGEN, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2- $\Delta\Delta$ CT method. **Methods**, 2001.

LORGERIL, J. et al. Immune-suppression by OsHV-1 viral infection causes fatal bacteraemia in Pacific oysters. **Nature communications**. 2018.

LOTH, K. et al. The ancestral N-terminal domain of big defensins drives bacterially triggered assembly into antimicrobial nanonets. **mBio**, v. 10, n. 5, p. 1–15, 2019.

MEADE, K. G.; O'FARRELLY, C. B-Defensins: Farming the microbiome for homeostasis and health. **Frontiers in Immunology**, v. 10, n. JAN, 2019.

MONTERO-ALEJO, V. et al. Panusin represents a new family of β -defensin-like peptides in invertebrates. **Developmental and Comparative Immunology**, v. 67, p. 310–321, 2017.

PHOENIX, D. A.; DENNISON, S. R.; HARRIS, F. Antimicrobial Peptides: Their History, Evolution, and Functional Promiscuity. **Antimicrobial Peptides**, p. 1–37, 2013.

RAMIREZ, J. L. et al. The *Aedes aegypti* IMD pathway is a critical component of the mosquito antifungal immune response. **Developmental and Comparative Immunology**, v. 95, p. 1–9, 2019.

ROSA, R. D. et al. Big defensins, a diverse family of antimicrobial peptides that follows different patterns of expression in hemocytes of the oyster *Crassostrea gigas*. **PLoS ONE**, v. 6, n. 9, 2011.

ROSA, R. D. et al. A hemocyte gene expression signature correlated with predictive capacity of oysters to survive *Vibrio* infections. **BMC Genomics**, 2012.

ROSA, R. D. et al. High polymorphism in big defensin gene expression reveals presence-absence gene variability (PAV) in the oyster *Crassostrea gigas*. **Developmental and Comparative Immunology**, v. 49, n. 2, p. 231–238, 2015.

SAHL, H.-G. et al. Mammalian defensins: structures and mechanism of antibiotic activity. **Journal of Leukocyte Biology**, v. 77, n. 4, p. 466–475, 2005.

SAITO, T. et al. A novel big defensin identified in horseshoe crab hemocytes: Isolation, amino acid sequence, and antibacterial activity. **Journal of Biochemistry**, v. 117, n. 5, p. 1131–1137, 1995.

SAUVAGE, C. et al. Single Nucleotide polymorphisms and their relationship to codon usage bias in the Pacific oyster *Crassostrea gigas*. **Gene**, v. 406, n. 1–2, p. 13–22, 2007.

SCHMITT, P. et al. Insight into invertebrate defensin mechanism of action: Oyster defensins inhibit peptidoglycan biosynthesis by binding to lipid II. **Journal of Biological Chemistry**, v. 285, n. 38, p. 29208–29216, 2010a.

SCHMITT, P. et al. Molecular diversity of antimicrobial effectors in the oyster *Crassostrea gigas*. **BMC Evolutionary Biology**, 2010b.

SCHMITT, P. et al. The antimicrobial defense of the Pacific oyster, *Crassostrea gigas*. How diversity may compensate for scarcity in the regulation of resident/pathogenic microflora. **Frontiers in Microbiology**, 2012.

SCHMITT, P.; ROSA, R. D.; DESTOUMIEUX-GARZÓN, D. An intimate link between antimicrobial peptide sequence diversity and binding to essential components of bacterial membranes. **Biochimica et Biophysica Acta - Biomembranes**, 2016.

SELSTED, M. E.; SZKLAREK, D.; LEHRER, R. I. Purification Activity Granulocytost. **Microbiology**, v. 45, n. 1, p. 150–154, 1984.

SEO, J. K. et al. Purification of a novel arthropod defensin from the American oyster, *Crassostrea virginica*. **Biochemical and Biophysical Research Communications**, v. 338, n. 4, p. 1998–2004, 2005.

SEO, J. K. et al. CgMolluscidin, a novel dibasic residue repeat rich antimicrobial peptide, purified from the gill of the Pacific oyster, *Crassostrea gigas*. **Fish and Shellfish Immunology**, v. 35, n. 2, p. 480–488, 2013.

SHAFEE, T. M. A. et al. The Defensins Consist of Two Independent, Convergent Protein Superfamilies. **Molecular Biology and Evolution**, v. 33, n. 9, p. 2345–2356, 2016.

SHAFEE, T. M. A. et al. Convergent evolution of defensin sequence, structure and function. **Cellular and Molecular Life Sciences**, v. 74, n. 4, p. 663–682, 2017.

SPRINGER, N. M. et al. Maize inbreds exhibit high levels of copy number variation (CNV) and presence/absence variation (PAV) in genome content. **PLoS Genetics**, v. 5, n. 11, 2009.

TARIS, N. et al. Transcriptome response of the Pacific oyster (*Crassostrea gigas*) to infection with *Vibrio tubiashii* using cDNA AFLP differential display. **Animal Genetics**, v. 40, n. 5, p. 663–677, 2009.

VIEIRA, C. S. et al. The NF- κ B inhibitor, IMD-0354, affects immune gene expression, bacterial microbiota and *Trypanosoma cruzi* infection in *Rhodnius prolixus* midgut. **Frontiers in Physiology**, v. 9, n. AUG, p. 1–11, 2018.

WANG, J. et al. The oyster genome reveals stress adaptation and complexity of shell formation. **Nature**, v. 490, n. 7418, p. 49–54, 2012.

WANG, X. et al. Nanopore Sequencing and De Novo Assembly of a Black-Shelled Pacific Oyster (*Crassostrea gigas*) Genome. **Frontiers in Genetics**, v. 10, 2019.

WEI, L. et al. The first anionic defensin from amphibians. **Amino Acids**, v. 47, n. 7, p. 1301–1308, 2015.

WILMES, M. et al. Antibiotic activities of host defense peptides: More to it than lipid bilayer perturbation. **Natural Product Reports**, 2011.

XU, X. X. et al. An anionic defensin from *Plutella xylostella* with potential activity against *Bacillus thuringiensis*. **Bulletin of Entomological Research**, v. 106, n. 6, p. 790–800, 2016.

YANG, Y. S. et al. Solution structure and activity of the synthetic four-disulfide bond Mediterranean mussel defensin (MGD-1). **Biochemistry**, v. 39, n. 47, p. 14436–14447, 2000.

ZHANG, Y. et al. Identification and bioactivity analysis of a newly identified defensin from the oyster *Magallana gigas*. **Developmental and Comparative Immunology**, v. 85, p. 177–187, 2018.

ZHU, S. Discovery of six families of fungal defensin-like peptides provides insights into origin and evolution of the CS α β defensins. **Molecular Immunology**, v. 45, n. 3, p. 828–838, 2008.

ZHU, S.; GAO, B.; TYTGAT, J. **Phylogenetic distribution, functional epitopes and evolution of the CS α β superfamily** *Cellular and Molecular Life Sciences*, 2005.

APÊNDICE 1 – Sequências de entrada utilizadas (*queries*)

| Família de Defensina | Espécie | GenBank | Nome da sequência |
|--|----------------------------------|----------------|----------------------------------|
| Defensinas do tipo CS α β | <i>Crassostrea gigas</i> | AJ565499 | <i>Cg-Defm</i> |
| | | DQ400101 | <i>Cg-Defh1</i> |
| | | DQ400102 | <i>Cg-Defh2</i> |
| | <i>Mytilus galloprovincialis</i> | KP125907 | <i>arthropod defensin-like 1</i> |
| Big Defensinas | <i>Crassostrea gigas</i> | JN251123 | <i>Cg-BigDef1</i> |
| | | JN251124 | <i>Cg-BigDef2</i> |
| | | JN251121 | <i>Cg-BigDef3</i> |

ANEXO 1 - Artigo científico em coautoria (relacionado à dissertação de Mestrado)



Functional Insights From the Evolutionary Diversification of Big Defensins

Marco Gerdol^{1*}, Paulina Schmitt², Paola Venier³, Gustavo Rocha⁴, Rafael Diego Rosa^{4*} and Delphine Destoumieux-Garzón^{5*}

¹ Department of Life Sciences, University of Trieste, Trieste, Italy, ² Laboratorio de Genética e Inmunología Molecular, Instituto de Biología, Pontificia Universidad Católica de Valparaíso, Valparaíso, Chile, ³ Department of Biology, University of Padova, Padova, Italy, ⁴ Laboratory of Immunology Applied to Aquaculture, Department of Cell Biology, Embryology and Genetics, Federal University of Santa Catarina, Florianópolis, Brazil, ⁵ IHPE, Université de Montpellier, CNRS, Ifremer, Université de Perpignan Via Domitia, Montpellier, France