

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
INSTITUTO DE BIOTECNOLOGIA
GRADUAÇÃO EM BIOTECNOLOGIA

THAILLYNE FERREIRA DE BRITO PASSOS

ORIGEM E EVOLUÇÃO DAS SNAKINAS NO REINO VEGETAL

PATOS DE MINAS
DEZEMBRO DE 2020

THAILLYNE FERREIRA DE BRITO PASSOS

ORIGEM E EVOLUÇÃO DAS SNAKINAS NO REINO VEGETAL

Monografia apresentada ao Instituto de Biotecnologia da Universidade Federal de Uberlândia como requisito final para a obtenção do título de Bacharel em Biotecnologia.

Orientador: Prof. Dr. Aulus Estevão Anjos de Deus Barbosa.

**PATOS DE MINAS
DEZEMBRO DE 2020**

THAILLYNE FERREIRA DE BRITO PASSOS

Origem e evolução das snakinas no reino vegetal

Monografia apresentada ao Instituto de Biotecnologia da Universidade Federal de Uberlândia como requisito final para a obtenção do título de Bacharel em Biotecnologia.

Banca examinadora:



Prof. Dr. Aulus Estevão Anjos de Deus Barbosa – IBTEC/UFU
Presidente



Me. Thaís Cunha de Sousa Cardoso - PPGGB/UFU
Membro



Prof. Dr. Gilvan Caetano Duarte - IBTEC/UFU
Membro

Patos de Minas, 15 de dezembro 2020.

AGRADECIMENTOS

Começo por agradecer a Deus por, ao longo desses anos de graduação, me guiar nos momentos de aflições e dificuldades e me ter feito ver o caminho, nos momentos que pensei em desistir. A ele eu devo minha gratidão.

Agradeço aos meus pais Hélio e Hércia, em especial a minha mãe Hércia por ser o meu pilar, estar ao meu lado e me fazer acreditar que tinha forças e capacidade para realizar o meu grande objetivo. Eu à agradeço do fundo do meu coração e com algumas lágrimas nos olhos.

Agradeço ao meu namorado Marcos por tanto amor e cuidado, principalmente quando a minha caminhada se tornou árdua e acima de tudo, pelo incentivo, pois muitas vezes foi o empurrão que eu precisava.

Agradeço ao meu orientador Prof. Aulus pela confiança, oportunidade e paciência que foi depositada em mim. Muito obrigada pelos diversos ensinamentos ao longo de todos esses anos.

A todos os professores eu agradeço a orientação repleta de conhecimento e sabedoria.

Agradeço de todo meu coração aos meus amigos de faculdade que “sofreram” tanto quanto eu, principalmente a Thais Oliveira, Beatriz Scatolin e Higor Donizete, que me transmitiram força e por todo o companheirismo ao longo desse percurso.

A quem eu não mencionei, mas estive junto eu quero deixar um agradecimento eterno: vocês foram imensamente importantes nessa jornada.

Por fim saibam que o meu coração será sempre grato a todos que estiveram comigo diretamente ou indiretamente ao longo da minha formação.

RESUMO

Existem diversos tipos de moléculas que são importantes no sistema imune inato dos organismos vivos, dentre elas os peptídeos antimicrobianos (AMPs). Os AMPs são moléculas de baixo peso molecular que atuam na defesa do hospedeiro contra patógenos. Devido às suas propriedades, eles têm sido apontados como uma possível solução para minimizar os efeitos negativos de agentes patogênicos, podendo ser aplicados no desenvolvimento de plantas mais resistentes a patógenos e de novos antibióticos. Dentre as famílias de peptídeos antimicrobianos encontradas, temos as Snakinas, que desempenham atividades antibacterianas e antifúngicas, além de estarem envolvidas no crescimento e desenvolvimento das plantas. No entanto, Snakinas nunca foram analisadas nas espécies de plantas mais antigas. Dessa forma, o presente estudo objetivou identificar e caracterizar *in silico* os genes codificadores de Snakinas nos seguintes grupos: algas verdes, briófitas (Hornworts, Marchantiales e Pelliales) e pteridófitas (Equisetales, Osmundales, Plagiogyriales e Salviniaceae). Através de ferramentas da bioinformática, como o banco de dados *1000 plants*, softwares como o Blast2GO, MEGA X e o SWISS-MODEL foi possível encontrar as sequências equivalentes à família das Snakinas. Foram identificados 50 genes de Snakinas, distribuídos em todas as linhagens vegetais analisadas. Com base nos resultados obtidos foi possível visualizar a evolução destes genes na linhagem vegetal, desde algas mais primitivas. No presente estudo também foi demonstrado a identificação e caracterização dos genes codificadores de AMPs da família das Snakinas, os produtos expressos, as estruturas tridimensionais, a análise dos domínios conservados e a ancestralidade dessas Snakinas comparadas com outras plantas por meio de árvore filogenética, todas elas possuindo duas α -hélices longas e seis ligações de dissulfeto entre as 12 cisteínas. A identificação de genes codificadores da família das Snakinas em plantas primitivas pode contribuir para o entendimento de seus respectivos aspectos evolutivos, possibilitando avanços no entendimento do funcionamento de peptídeos antimicrobianos e suas diversas aplicações biotecnológicas. E em virtude das considerações acima, tal estudo fornece uma nova perspectiva sobre as relações evolutivas das Snakinas no Reino Vegetal e sua relação com as circunstâncias de sobrevivência e adaptação das plantas.

Palavras-chave: Peptídeos Antimicrobianos. Snakinas. Bioinformática. Análise Filogenética.

ABSTRACT

There are several types of molecules that are important in the innate immune system of living organisms, they are antimicrobial peptides (AMPs). AMPs are low molecular weight molecules that act in the defense of the host against pathogens. They have been identified as a possible solution to minimize the negative effects of pathogens, which can be known in the development of plants more resistant to pathogens and new antibiotics. Among the families of antimicrobial peptides found, we have Snakin, which perform antibacterial and antifungal activities, in addition to being involved in plant growth and development. However, Snakin have never been analyzed in the oldest plant species. Thus, the present study aimed to identify and characterize *in silico* the genes encoding Snakinas in the following groups: green algae, bryophytes (Hornworts, Marchantiales and Pelliales) and pteridophytes (Equisetales, Osmundales, Plagiogyriales and Salviniiales). Through bioinformatics tools, such as the 1000 plants database, software such as Blast2GO, MEGA X and SWISS-MODEL, it was possible to find as sequences equivalent to the Snakin family. 50 Snakinas genes were identified, distributed in all the vegetal lines analyzed. Based on the results obtained, it was possible to visualize the evolution of these genes in the plant line, from more primitive algae. In the present study, it was also characterized the identification and characterization of the AMP encoding genes of the Snakin family, the expressed products, the three-dimensional structures, the analysis of the conserved domains and the ancestry of these Snakin in comparison with other plants through phylogenetic tree, all of them having two long α -propellers and six disulfide bonds between the 12 cysteines. Identification of the AMP encoding genes of the Snakin family genes in primitive plants may contribute to the understanding of their respective evolutionary aspects, enabling advances in the understanding of the functioning of antimicrobial peptides, and their diverse biotechnological applications. Our study provides a new perspective on the evolutionary relationships of Snakin in the Kingdom Vegetable and their relationship to the circumstances of plant survival and adaptation.

Keywords: *Antimicrobial Peptides. Snakinas. Bioinformatics. Phylogenetic Analysis.*

SUMÁRIO

1 INTRODUÇÃO.....	8
2 REFERENCIAL TEÓRICO	9
2.1 Peptídeos antimicrobianos.....	9
2.2 Snakinas em plantas.....	12
2.3 Bioinformática e genômica.....	13
3 OBJETIVOS.....	14
3.1 Objetivo geral	14
3.2 Objetivos específicos.....	14
4 MATERIAL E MÉTODOS.....	14
4.1 Busca pelas sequências de proteínas Snakinas em banco de dados	14
4.2 Alinhamento local das sequências de proteínas Snakinas	15
4.3 Análises dos domínios conservados	16
4.4 Obtenção das sequências de aminoácidos	16
4.5 Alinhamento global de Snakinas	16
4.6 Classificação dos genes por filogenia.....	16
4.7 Análises estruturais.....	17
5 RESULTADOS E DISCUSSÃO	17
5.1 Identificação de Snakinas expressas no Reino Vegetal.....	17
5.2 Análise das sequências de Snakinas	221
5.3 Análises filogenética das Snakinas nas linhagens vegetais.....	23
5.4 Análise estrutural.....	26
6 CONCLUSÃO.....	28
REFERÊNCIAS	29
APÊNDICES	33

1 INTRODUÇÃO

Há aproximadamente 450 milhões de anos, as plantas evoluíram de uma alga *Charophyceae* ancestral com características biológicas, celulares e bioquímicas (LIGRONE, 2019). Várias adequações fisiológicas e químicas das algas resultaram na passagem da vida aquática para a terrestre, e várias evidências propõem que elas desenvolveram um estilo de vida terrestre de forma independente (VRIES; ARCHIBALD, 2018).

Acredita-se que uma alga ancestral antes da origem das plantas terrestres, pode ter contribuído com as mais variadas adaptações que estão relacionadas com a evolução das plantas para a terra, por exemplo, a obtenção de paredes celulares, resistência a patógenos e adaptação a altas radiações (BOWMAN; SAKAKIBARA; FURUMIZU; DIERSCHKE, 2016). Portanto, toda a flora terrestre é proveniente de um único clado dentro das plantas.

Ao longo do processo evolutivo, as plantas estiveram constantemente expostas a várias situações de estresse e de fatores químicos e físicos desfavoráveis no ambiente, envolvendo a presença um vasto número de organismos fitopatogênicos. Desse modo, para a sobrevivência, as plantas necessitam de mecanismos eficientes de defesa (CASTRO; FONTES, 2005). As mais variadas estratégias de defesa desenvolvidas pelo Reino Vegetal durante a sua evolução fizeram com que a suscetibilidade se tornasse uma exceção enquanto a resistência, uma regra (MACCARI; MARIAGRAZIA; NIFOSÌ, 2015). A evolução das plantas e os incessantes desafios expostos à sobrevivência foram responsáveis pelo desenvolvimento de um sistema de defesa eficiente baseado na síntese de metabólitos secundários, como fenóis, terpenóides, quininos e taninos (BENKO-ISEPPON et al., 2010). Adicionalmente, os vegetais sintetizam além dos metabólitos secundários, os peptídeos antimicrobianos (AMPs). Devido à capacidade de se relacionar com diversos alvos, desde bactérias Gram-positivas e negativas, leveduras, fungos, vírus e protozoários, os AMPs são classificados como proteínas promíscuas (PELEGRINI et al., 2011). E como uma alternativa para controlar tais agentes patogênicos citados anteriormente, os AMPs são uma possível solução para minimizar os efeitos negativos dos mesmos, para isso a ciência busca através da pesquisa, plantas mais resistentes via melhoramento genético e transgenia, além de investirem em uma formulação para novos antibióticos (KEYMANESH; SOLTANI; SARDARI, 2009; LI et al., 2017).

Os AMPs de plantas são classificados com base em sua estrutura, sendo ricos em cisteínas que formam múltiplas ligações dissulfeto (GOYAL; MATTOO, 2016).

As snakinas - classificadas como membros da família Snakin/GASA (SEGURA et al. 1999; BERROCAL-LOBO et al. 2002), devido a sua variedade funcional, vêm recebendo um crescente destaque, além de atuarem como efetores do crescimento e de divisão celular. Os peptídeos membros da família Snakin/GASA são expressos em diversos tecidos (desde as raízes até as flores). Sua estrutura primária é composta por três motivos principais: (1) um peptídeo sinal na região N-Terminal, que pode conter de 18-29 resíduos de aminoácidos, (2) uma região variável que possui entre 7-31 resíduos de aminoácidos e o domínio C-Terminal GASA, composto por 60 aminoácidos em C-terminal, e região conservada com 12 cisteínas (ROXRUD et al 2007; FAN et al. 2017).

Desse modo, devido à sua variedade funcional e às suas características, as Snakinas se destacam com alvos biotecnológicos relevantes para uso tanto no desenvolvimento de novos fármacos, como na transformação de plantas (Oliveira-Lima et al., 2017).

Atualmente, os estudos que utilizam as ferramentas computacionais, nesse caso a bioinformática são imprescindíveis para a manipulação dos dados biológicos, e abordagens *in silico* para análise de peptídeos antimicrobianos, representando um pré-requisito de destaque para posterior validação experimental dos resultados. Portanto, a identificação de genes de Snakinas em plantas primitivas poderá colaborar para o entendimento dos seus aspectos evolutivos, viabilizando progressos no uso de peptídeos antimicrobianos em inúmeras aplicações biotecnológicas. (DESTOUMIEUX-GARZÓN et al., 2016).

2 REFERENCIAL TEÓRICO

2.1 Peptídeos antimicrobianos

As plantas, durante o seu processo evolutivo, adquiriram uma vasta gama de mecanismos de defesas para resistir a fatores bióticos (agentes patogênicos, como fungos, bactérias e vírus) e abióticos (estresses químico e físico, como secas, frio, poluentes, metais pesados e estresse salino). Por exemplo, após a contaminação por algum agente patogênico, as plantas acionam um conjunto de genes associados à resistência sistêmica adquirida (STINTZI et al., 1993), onde uma cascata complexa de moléculas é sintetizada para reter o invasor. São conhecidas, pelo menos, 17 famílias de genes que expressam uma extensa gama de mecanismos

moleculares de defesa, incluindo a produção de AMPs (AERTS et al., 2008; LÄMKE; BÄURLE, 2017).

Os AMPs são moléculas onipresentes, as quais desempenham um papel importante como defesa do organismo contra pragas e patógenos. Embora a maioria desses AMPs sejam classificados de acordo com sua estrutura, nesse caso, os peptídeos lineares de animais, insetos e plantas, apresentam um tamanho pequeno e sequência variável de aminoácidos. Além disso eles são anfipáticos, carregados eletricamente, ricos em determinado tipo de aminoácidos e podem ser classificados em várias classes e subclasses (JENSSEN; HAMILL; HANCOCK, 2006). De acordo com Maria-Neto (2011), os peptídeos antimicrobianos são classificados conforme sua estrutura e composição, como peptídeos lineares, peptídeos ricos em glicina e prolina e peptídeos estabilizados com cisteínas.

No geral, os AMPs das plantas compartilham diversos atributos comuns com os de microrganismos, animais e insetos. Dentre eles podemos citar que a maioria dos AMPs apresenta de 10 a 50 aminoácidos de comprimento; peso molecular de aproximadamente 10 kDa; carga líquida positiva devido aos aminoácidos apolares serem hidrofóbicos; podem ser expressos de forma constitutiva ou induzida, e possuem resíduos de cisteína que estabilizam a estrutura da proteína por pontes dissulfeto (PELEGRINI et al., 2011). Todos esses atributos estão relacionados principalmente ao seu mecanismo de ação clássico, o qual exerce seu papel no efeito geral da atividade antimicrobiana. (BAHAR; REN, 2013).

A carga líquida e a natureza anfipática estão relacionadas aos mecanismos de ação dos AMPs, nas quais propiciam a perturbação da membrana celular bacteriana. Segundo Mahlapuu et al. (2016), tal perturbação acontece devido às forças eletrostáticas entre os resíduos de aminoácidos positivos com as cargas negativas dos lipídeos da membrana celular. Isso acarreta a formação de poros, a saída do conteúdo celular e, por conseguinte, a morte do agente patogênico.

Os AMPs podem exercer diversas funcionalidades nas mais variadas condições ambientais, propriedade conhecida como “promiscuidade peptídica” (FRANCO, 2011). Tais peptídeos foram isolados de diversas fontes, em todos os reinos da vida (SILVA et al., 2011; BROGDEN, 2005) e, os mesmos podem ser classificados em dois vastos grupos, isto é, de acordo com a presença ou ausência de pontes dissulfeto (BROGDEN, 2005).

No caso de AMPs livres de pontes de dissulfeto, eles são compostos sobretudo por estruturas α -helicoidal e peptídeos não estruturados, ao passo que os peptídeos estabilizados por cisteína são formados por várias classes, as quais são subdivididas de acordo com os seus padrões de pontes de dissulfeto. Porém, devido essa alta versatilidade, a classificação em

famílias desses AMPs vegetais é complexa. Geralmente os peptídeos são classificados por: posições das pontes de dissulfeto, o número e o espaçamento de resíduos de cisteína, estrutura secundária e tridimensional e os tipos de aminoácidos na sequência. De acordo com Nawrot et al. (2014), diversas famílias de peptídeos antimicrobianos vegetais foram caracterizadas e até mesmo identificadas ao longo desses últimos anos, por exemplo as defensinas, α -hairpinin, heveínas, tioninas, ciclotídeos, proteínas de transferência de lipídeos não-específicas e Snakinas.

As principais famílias de AMPs vegetais com esses critérios (Tabela 1) apresentam as seguintes características (TAM et al.; 2015):

- Possuem tamanho mediano;
- Apresentam entre duas a seis ligações de pontes dissulfeto intramoleculares;
- Similaridade de aminoácidos baseados no motivo Cys e são conservados na estrutura se secundária e terciária;
- Dispõem de estabilidade estrutural devido as ligações de pontes dissulfeto;
- Apresentam número de resíduos Cys variados;
- Possuem estrutura compacta, com alta estabilidade química, térmica e enzimática.

Tabela 1. Principais famílias de AMPs de plantas.

Peptídeo	S-S Nº	Nome	Representação das famílias		Estrutura
			AA Nº	Motivo dissulfeto	
6C-Tionina	3	<u>Crambin</u>	46	2-C-0-C-11-C-8-C-5-C-7-C-6	<u>Gamma (Γ) fold</u>
8C-Tionina	4	<u>β-Purothionin</u>	45	2-C-0-C-7-C-3-C-8-C-3-C-1-C-7-C-6	<u>β1-α1-α2-β2-coil motif</u>
8C-Defensina	4	NaD1	47	2-C-10-C-5-C-3-C-9-C-6-C-1-C-3-C	<u>CSαβ motif</u>
10C-Defensina	5	PhD1	47	2-C-3-C-6-C-5-C-2-C-0-C-9-C-6-C-1-C-3-C	<u>β1-coil-α-β2-β3</u>
6C-Heveína	3	Ac-AMP1	29	3-C-4-C-4-C-0-C-5-C-6-C-1	<u>Gly & Cys rich</u>
8C-Heveína	4	<u>Hevein</u>	43	2-C-8-C-4-C-0-C-5-C-6-C-5-C-3-C-2	<u>Central β strands & (short</u>
10C-Heveína	5	EAFP1	41	2-C-3-C-3-C-4-C-0-C-5-C-6-C-5-C-1-C-1-C-2	<u>helical) side coils</u>
<u>Knottin</u>	3	PAFP-S	38	2-C-6-C-8-C-0-C-3-C-10-C-3	<u>Cystine knot</u>
<u>Cyclic Knottin</u>	3	<u>Kalata B1</u>	29	4-C-3-C-4-C-4-C-1-C-4-C-3	<u>Short β strand & coil</u>
<u>α-Hairpinin</u>	2	Ec-AMP1	37	6-C-3-C-13-C-3-C-8	<u>α1-turn-α2</u>
LTP	4	<u>Maize LTP1</u>	93	3-C-9-C-14-C-0-C-19-C-1-C-22-C-13-C-4	<u>Hydrophobic cavity</u>
LTP	4	<u>Wheat LTP2</u>	67	1-C-7-C-13-C-0-C-8-C-1-C-23-C-6-C	<u>α1-α2-α3-α4-coil</u>
<u>Snakina *</u>	6	Snakin-1	63	4-C-3-C-3-C-8-C-3-C-2-C-2-C-1-C-11-C-1-C-12-C-1	<u>α-helices</u>

*O motivo dissulfeto e a estrutura da Snakin-1 são previstos com base na modelagem de homologia.

Fonte: Adaptado de TAM et al.; (2015).

2.2 Snakinas em plantas

Os AMPs da família das Snakinas foram isolados, primeiramente, de tubérculos de batata (*Solanum tuberosum*), chamados de snakina-1 (63 aa) e snakina-2 (66 aa), com 12 cisteínas formando 6 pontes dissulfeto (PADOVAN; SCOCCHI; TOSSI, 2010). A estrutura das Snakinas tem duas α -hélices longas estabilizadas por ligações dissulfeto (PORTO; FRANCO, 2013).

Conforme Silverstein et al. (2007), as Snakinas foram classificadas em um amplo grupo de peptídeos ricos em cisteína (CRPs), e tanto o número quanto à disposição dos resíduos de cisteínas distingue as classes de CRPs das demais, como por exemplo tioninas, proteínas de transferência de lipídeos e defensinas.

Os genes das Snakinas codificam pequenas proteínas onde três domínios eminentes podem ser definidos: 1) um peptídeo de sinal putativo com 18-29 resíduos; 2) uma região variável exibindo elevada divergência entre os membros da família, tanto no comprimento de sequência e na composição de aminoácidos; e 3) uma região C-terminal de cerca de 63

aminoácidos com 12 resíduos de cisteína em posições conservadas nomeada domínio GASA (AUBERT et al., 1998).

O número de aminoácidos é uma característica que diferencia as subfamílias das Snakinas. Na Snakina-1, pode-se observar que é um peptídeo antimicrobiano rico em cisteínas, com um pH altamente básico, um pI de 8,98 e com 63 aminoácidos (AA). Em relação a estrutura tridimensional, a mesma é composta por duas longas hélices e com seis ligações dissulfeto. Já na subfamília da Snakina-2, as mesmas apresentam 38% de identidade com os aminoácidos da SN1 e na Snakina-3, também nomeada de domínio GASA, devido à sua similaridade com membros da família GASA (*Gibberellic Acid Stimulated, de Arabidopsis*) (BERROCAL-LOBO et al. 2002).

Um fato importante é que, a conservação das posições e número das cisteínas, ao longo da evolução sugere que esses resíduos desempenham um papel central na função das Snakinas (BEN-NISSAN et al., 2004), sendo fundamentais para as atividades antibacterianas, antifúngicas e antioxidantes (RUBINOVICH et al., 2014). Além disso, as Snakinas estão envolvidas em diversos aspectos do crescimento e desenvolvimento das plantas.

2.3 Bioinformática e genômica

A bioinformática é uma nova área que disponibiliza uma gama de dados genômicos obtidos por sequenciamento, além de ser uma ciência com aplicações de técnicas computacionais que permitem manipular, analisar, organizar e armazenar dados biológicos. Nesta nova área da ciência foram criados uma vasta quantidade de programas, banco de dados e algoritmos (NAVLAKHA; BAR-JOSEPH, 2011).

Atualmente, nesse ramo existem algumas plataformas digitais extremamente populares como por exemplo: O BLAST e o Genbank. O BLAST (*Basic Local Alignment Search Tool*), é uma ferramenta disponível em muitos sites, inclusive no banco de dados CNGB - *China National GeneBank*, o qual busca encontrar regiões de similaridade entre as sequências biológicas. Ou seja, compara a sequência de interesse (*query*) de nucleotídeos ou aminoácidos com um banco de dados, seja ele proteômico ou genômico e, posteriormente, calcula a significância estatística da mesma. Já o GenBank, disponível na plataforma do NCBI - *National Center for Biotechnology Information*, é um banco de dados biológicos onde armazena dados e informações de grande parte das sequências genéticas publicamente disponíveis.

As ferramentas da Bioinformática, além de terem um papel indispensável nas mais variadas análises dos dados de genomas, proteomas, metabolomas e transcriptomas, podem ajudar a diminuir o custo e até mesmo o tempo de diversas pesquisas científicas.

Além do mais, tais ferramentas são primordiais para a identificação e caracterização dos genes codificadores de AMPs, os produtos expressos, os aspectos funcionais, as estruturas tridimensionais e os domínios conservados (PESTANA-CALSA; RIBEIRO; CALSA, 2010).

3 OBJETIVOS

3.1 Objetivo geral

Identificar e analisar os genes codificadores dos AMPs da família das Snakinas na linhagem vegetal, desde algas verdes até pteridófitas.

3.2 Objetivos específicos

- Identificar, através de análises *in silico*, a presença de genes codificadores de AMPs da família Snakinas expressos nos grupos vegetais, de algas verdes até pteridófitas e seus domínios conservados;
- Classificar os genes encontrados de acordo com as suas subfamílias;
- Comparar as sequências de peptídeos encontradas com os peptídeos antimicrobianos da família das Snakinas descritos em Angiospermas;
- Analisar a estrutura 3D desses AMPs;
- Verificar através de um estudo filogenético a evolução das Snakinas identificadas e sua distribuição nos diferentes grupos vegetais.

4 MATERIAL E MÉTODOS

4.1 Busca pelas sequências de proteínas Snakinas em banco de dados

Inicialmente, foram realizadas revisões na literatura científica para identificar genes codificadores de peptídeos antimicrobianos da família das Snakinas já caracterizados em outras plantas. As sequências indicadas nas referências bibliográficas foram obtidas nos bancos de

dados *The Universal Protein Resource - UniProt* e *National Center for Biotechnology Information – NCBI*. Foram obtidas sequências da família das Snakinas do tipo I, II e III (Tabela 2). As sequências obtidas foram salvas no formato FASTA.

Tabela 2. Código de acesso das sequências de proteínas das Snakinas usados como “query”.

Plantas	Snakina-1	Snakina-2	Snakina-3
<i>Anthurium amnicola</i>		A0A1D1XY22	
<i>Arabidopsis thaliana</i>	NP_001077504.1		P46690; Q6GKX7
<i>Capsicum annuum</i>		XP_016539896.1	
<i>Medicago sativa</i>	JQ517286.1	U3GJY1	
<i>Petunia hybrida</i>			Q93WR4; CAA60677.1
<i>Solanum bulbocastanum</i>	D3G6B9		
<i>Solanum tuberosum</i>	NW_006239304.1	Q93X17	

Fonte: Própria autora.

4.2 Alinhamento local das sequências de proteínas Snakinas

As sequências de peptídeos das Snakinas foram submetidas ao algoritmo TBLASTN no banco de dados *1000 plants*, localizado no *China National GeneBank - CNGB*. O alinhamento local foi realizado contra cada genoma disponível nas seguintes linhagens vegetais: algas verdes, briófitas e pteridófitas (Tabela 3). Os resultados das análises foram organizados em arquivos FASTA e que foram utilizados em todas as análises seguintes.

Tabela 3. Número de genomas de cada grupo utilizado para construção dos bancos de dados.

	Nº de genomas
Algas verdes	168
Briófitas	82
Pteridófitas	98
Total	416

Fonte: Própria autora.

4.3 Análises dos domínios conservados

Usando o programa BLAST2GO (versão 5.1.2) (CONESA; GOTZ, 2005), as sequências de nucleotídeos dos genes de Snakinas foram submetidas ao alinhamento BLASTx e posteriormente analisadas pelo InterProScan (Tabela 6,7,8,9 e 10). Esta análise permitiu a identificação dos domínios conservados da família das Snakinas. As sequências que não apresentaram domínios conservados foram descartadas e não participaram das análises posteriores.

4.4 Obtenção das sequências de aminoácidos

As sequências de nucleotídeos oriundas do BLAST2GO foram submetidas ao ORF Finder - *Open Reading Frame Finder* na plataforma do NCBI para a obtenção das sequências de aminoácidos. Para cada ORF gerada foi realizado o BLASTp contra o banco de proteínas não redundantes (NR) e foram registradas somente aquelas que apresentaram o domínio conservado referente à família das snakinas.

4.5 Alinhamento global de Snakinas

O alinhamento múltiplo de todas as sequências selecionadas foi feito através da ferramenta Muscle (EDGAR, 2004), do software MEGA X- *Molecular Evolutionary Genetics Analysis* (KUMAR et al., 2018). Os arquivos de alinhamento gerados pelo Muscle foram exportados como FASTA e submetidos ao programa BioEdit (versão 7.2.6.1) para um melhor aperfeiçoamento dos alinhamentos manuais e melhor visualização entre as sequências.

4.6 Classificação das Snakinas

As sequências das Snakinas foram classificadas em suas respectivas subfamílias através de árvores filogenéticas. Para elaboração das árvores filogenéticas, os alinhamentos foram submetidos ao programa MEGA X - *Molecular Evolutionary Genetics Analysis* (KUMAR et al., 2018). A árvore foi construída com o método *Neighbor-joining* (NJ) (SAITOU; NEI, 1987), modelo *Jones-Taylor-Thornton* (JTT), com teste de filogenia pelo método *bootstrap* com 1000 repetições.

4.7 Análises estruturais

Para a modelagem e visualização das estruturas 3D dos peptídeos antimicrobianos da família das Snakinas, foi utilizado o programa SWISS-MODEL (SCHWEDE et al., 2003). Em tal programa as sequências foram modeladas e comparadas com outros modelos já existentes de snakinas no banco de dados do próprio site.

5 RESULTADOS E DISCUSSÃO

5.1 Identificação de Snakinas expressas no Reino Vegetal

Dados transcriptômicos do banco de dados *1000 Plants* foram usados para identificar Snakinas expressas em diferentes linhagens vegetais. No presente estudo, foram identificadas 50 possíveis sequências de Snakinas, salientando a descoberta inédita desses genes em espécies de algas verdes, briófitas, isoetales e pteridófitas. O número de genes de Snakinas identificados varia de 6 em algas verdes a 24 em pteridófitas (Tabela 4). Tais resultados evidenciam a presença de genes de Snakinas desde as plantas mais primitivas, como as algas verdes, além de uma diversidade de número de genes de Snakinas identificados nas diferentes linhagens vegetais. Os dados mostram que o número de Snakinas, desde organismos aquáticos até as plantas vasculares demonstram uma tendência de aumento do número de genes ao longo do processo evolutivo. Um número maior de sequências foi identificado, porém as sequências que não apresentaram os domínios conservados de Snakinas ou apresentavam estarem incompletas foram retiradas das análises subsequentes. Os resultados do BLASTx e análise de domínios conservados para todas as sequências completas estão nas tabelas 5 a 7.

Tabela 4. Número de genomas de cada grupo utilizado para construção dos bancos de dados

	Genes encontrados
Algas Verdes	6
Briófitas	20
Pteridófitas	24
TOTAL	50

Fonte: Própria autora.

Tabela 5. Sequências de proteínas de Snakinas encontradas em algas verdes.

SeqName	Description	Length	e-Value	Similaridade	InterPro IDs
onekp_WDWX_scaffold_2002336	gibberellin-regulated protein 1-like	367	2.87921E-49	82.16	IPR003854 (PFAM)
onekp_IHOI_scaffold_2024897	gibberellin-regulated protein 1-like	452	4.45948E-53	95.25	IPR003854 (PFAM)
onekp_ZDIZ_scaffold_2002827	gibberellin-regulated protein 1-like	451	6.82643E-46	81.96	IPR003854 (PFAM)
onekp_WDWX_scaffold_2002337	gibberellin-regulated protein 1-like	287	1.02629E-27	75.67	IPR003854 (PFAM)
onekp_UKUC_scaffold_2098914	snakin-2-like isoform X1	320	2.41523E-57	85.43	IPR003854 (PFAM)
onekp_JMTE_scaffold_2016006	gibberellin-regulated protein 1	692	1.66918E-27	75.03	IPR003854 (PFAM)

Tabela 6. Sequências de genes codificadores de Snakinas encontradas em Briófitas

SeqName	Description	Length	e-Value	Similaridade	InterPro IDs
onekp_RXRQ_scaffold_2132160	protein GAST1-like	786	3.31639E-22	77.07	IPR003854 (PFAM)
onekp_WCZB_scaffold_2111668	gibberellin-regulated protein 4	639	2.01468E-24	77.91	IPR003854 (PFAM)
onekp_TFDQ_scaffold_2114143	protein RSI-1	617	4.97898E-61	76.44	IPR003854 (PFAM)
onekp_TFDQ_scaffold_2106904	Gibberellin-regulated protein 10	398	2.44745E-14	67.13	IPR003854 (PFAM)
onekp_PYHZ_scaffold_2066312	peamaclein-like	545	7.53029E-22	77.74	IPR003854 (PFAM)
onekp_PKOX_scaffold_2092056	peamaclein-like	626	1.14861E-21	79.21	IPR003854 (PFAM)
onekp_PKOX_scaffold_2090506	gibberellin-regulated protein 5-like	541	9.71666E-25	71.08	IPR003854 (PFAM)
onekp_PKOX_scaffold_2090828	Gibberellin-regulated protein 10	556	4.04249E-30	82.55	IPR003854 (PFAM)
onekp_PYHZ_scaffold_2004815	gibberellin-regulated protein 12-like	648	1.10278E-25	69.83	IPR003854 (PFAM)
onekp_PKOX_scaffold_2092783	gibberellin-regulated protein 12	679	3.82045E-24	72.8	IPR003854 (PFAM)
onekp_PYHZ_scaffold_2007583	Gibberellin-regulated protein 10	568	2.5986E-28	82.96	IPR003854 (PFAM)
onekp_PKOX_scaffold_2092923	Gibberellin-regulated family protein	690	7.68169E-20	72.85	IPR003854 (PFAM)
onekp_PYHZ_scaffold_2068297	Gibberellin-regulated family protein	673	1.45312E-20	73.34	IPR003854 (PFAM)
onekp_PYHZ_scaffold_2066053	gibberellin-regulated protein 10	532	9.97514E-23	71.35	IPR003854 (PFAM)
onekp_PKOX_scaffold_2089164	peamaclein	487	7.99826E-23	69.9	IPR003854 (PFAM)
onekp_PKOX_scaffold_2008326	Gibberellin-regulated family protein	771	9.27848E-14	63.14	IPR003854 (PFAM)
onekp_PKOX_scaffold_2090172	Gibberellin-regulated protein 10	527	1.60719E-14	66.97	IPR003854 (PFAM)
onekp_PYHZ_scaffold_2062874	Gibberellin regulated protein 4	414	1.53634E-4	55.29	IPR003854 (PFAM)
onekp_PKOX_scaffold_2091244	gibberellin-regulated protein 7	578	1.6262E-16	60.76	IPR003854 (PFAM)
onekp_PYHZ_scaffold_2066212	Gibberellin regulated protein	540	4.88234E-17	60.13	IPR003854 (PFAM)

Tabela 7. Sequências de genes codificadores de Snakinas encontradas em Pteridófitas.

SeqName	Description	Length	e-Value	Similaridade	InterPro IDs
onekp_JVSZ_scaffold_2001921	gibberellin-regulated protein 11-like	1176	4.0268E-20	63.68	IPR003854 (PFAM)
onekp_CAPN_scaffold_2038182	gibberellin-regulated protein 3-like	707	7.84043E-20	60.14	IPR003854 (PFAM)
onekp_JVSZ_scaffold_2126165	gibberellin-regulated protein 6	851	1.31183E-19	58.4	IPR003854 (PFAM)
onekp_JVSZ_scaffold_2123096	gibberellin-regulated protein	594	6.93197E-23	80.69	IPR003854 (PFAM)
onekp_CAPN_scaffold_2037324	gibberellin-regulated protein	556	2.29269E-22	73.9	IPR003854 (PFAM)
onekp_JVSZ_scaffold_2013082	gibberellin-regulated protein 6	738	2.63031E-28	78.94	IPR003854 (PFAM)
onekp_CAPN_scaffold_2007523	protein RSI-1	636	1.49082E-28	73.84	IPR003854 (PFAM)
onekp_CAPN_scaffold_2036054	Gibberellin-regulated protein 8	417	3.89879E-23	72.43	IPR003854 (PFAM)
onekp_JVSZ_scaffold_2123783	gibberellin-regulated protein 4 isoform X2	641	1.78463E-22	73.38	IPR003854 (PFAM)
onekp_JVSZ_scaffold_2008160	peamaclein-like	928	9.46905E-21	71.99	IPR003854 (PFAM)
onekp_UOMY_scaffold_2073303	protein GAST1-like	582	3.60605E-28	65.79	IPR003854 (PFAM)
onekp_UOMY_scaffold_2075302	protein RSI-1	824	1.7706E-30	76.2	IPR003854 (PFAM)
onekp_UOMY_scaffold_2076516	gibberellin-regulated protein 4-like	1046	4.02181E-17	66.01	IPR003854 (PFAM)
onekp_UOMY_scaffold_2008346	snakin-2-like isoform X1	1169	1.20551E-16	70.21	IPR003854 (PFAM)
onekp_UOMY_scaffold_2003973	gibberellin-regulated protein 10-like	980	4.22571E-22	74.55	IPR003854 (PFAM)
onekp_CVEG_scaffold_2017209	gibberellin-regulated protein 11-like	1008	1.99338E-18	67.42	IPR003854 (PFAM)
onekp_KIIX_scaffold_2078601	gibberellin-regulated protein 10-like	502	8.19906E-21	75.22	IPR003854 (PFAM)
onekp_CVEG_scaffold_2009601	gibberellin-regulated protein 11-like	1319	2.08022E-16	78.21	IPR003854 (PFAM)
onekp_CVEG_scaffold_2010681	protein GAST1-like	854	1.74944E-29	71.08	IPR003854 (PFAM)
onekp_CVEG_scaffold_2010682	protein GAST1-like	863	1.88359E-29	71.41	IPR003854 (PFAM)
onekp_KIIX_scaffold_2012325	protein GAST1-like	922	1.17527E-13	76.48	IPR003854 (PFAM)
onekp_KIIX_scaffold_2008305	gibberellin induced protein	1042	7.74512E-26	72.69	IPR003854 (PFAM)
onekp_CVEG_scaffold_2022320	gibberellin-regulated protein 6-like	698	2.80114E-14	80.27	IPR003854 (PFAM)
onekp_CVEG_scaffold_2022319	gibberellin-regulated protein 6-like	944	1.01599E-13	80.27	IPR003854 (PFAM)

Neste estudo, dentre as espécies analisadas, foram identificadas algas verdes pertencentes as seguintes espécies: *Chloromonas oogama*, *Pseudoscourfieldia marina*, e as demais são do gênero *Dunaliella*, ressaltando assim que foram encontradas em algas verdes 6 genes. Já as espécies de briófitas com 20 genes identificados foram *Phaeoceros carolinianus* sp, *Monoclea gottschei* sp, e algumas do gênero *Isoetes* sp. As espécies de pteridófitas com mais genes identificados foram *Azolla caroliniana* sp, que apresentou 6 genes, outras 10 do gênero *Equisetum*, 5 genes da espécie *Osmunda* sp e 3 da espécie *Pilularia globulifera* (Tabela 4).

A quantidade de genes de snakinas por espécie encontradas em grupos primitivos foram menores, em comparação com o relatado em *A. thaliana* (15), Maçã (26) (FAN et al., 2017) e (16) em batata (NAHIRNAK et al., 2016).

5.2 Análise das sequências de Snakinas

Por serem altamente conservadas em termos de composição de cisteínas, as Snakinas possuem 12 cisteínas conservadas e 6 ligações dissulfeto que são responsáveis pela estabilização e pela a manutenção da sua estrutura tridimensional, além das suas atividades biológicas (KUDDUS et al., 2016).

As sequências de Snakinas de algas verdes, briófitas e pteridófitas foram analisadas por meio de alinhamentos para identificar a presença do domínio conservado, bem como os motivos característicos de cada subclasse de Snakinas, (Figuras 1,2 e 3). Nesses alinhamentos é possível observar a grande conservação das posições das cisteínas (C), mesmo comparadas com as sequências referência das plantas angiospermas.

Figura 1. Alinhamento das Snakinas encontradas em algas verdes e seus ortólogos em plantas.

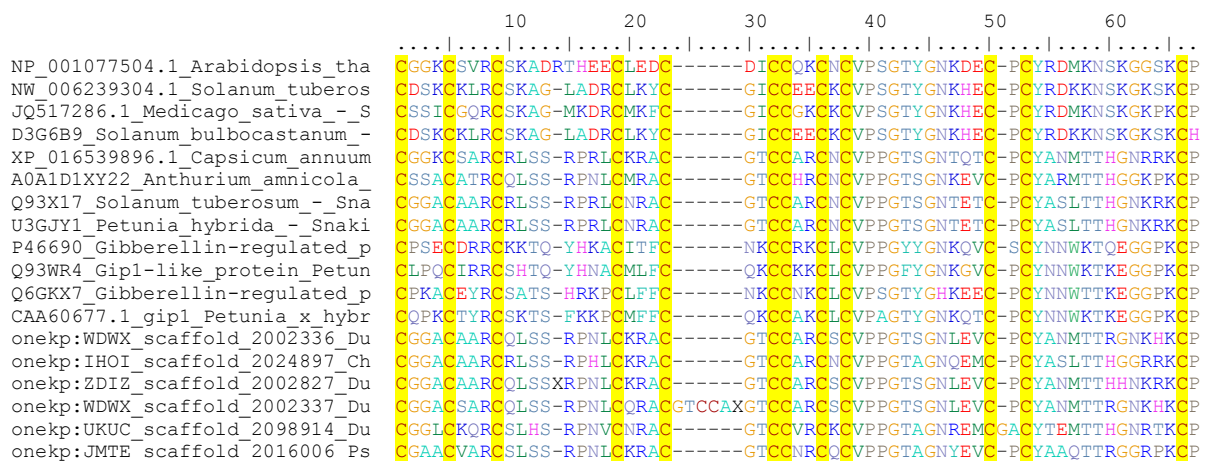


Figura 2. Alinhamento das Snakinas encontradas nas espécies de briófitas e seus ortólogos em plantas.

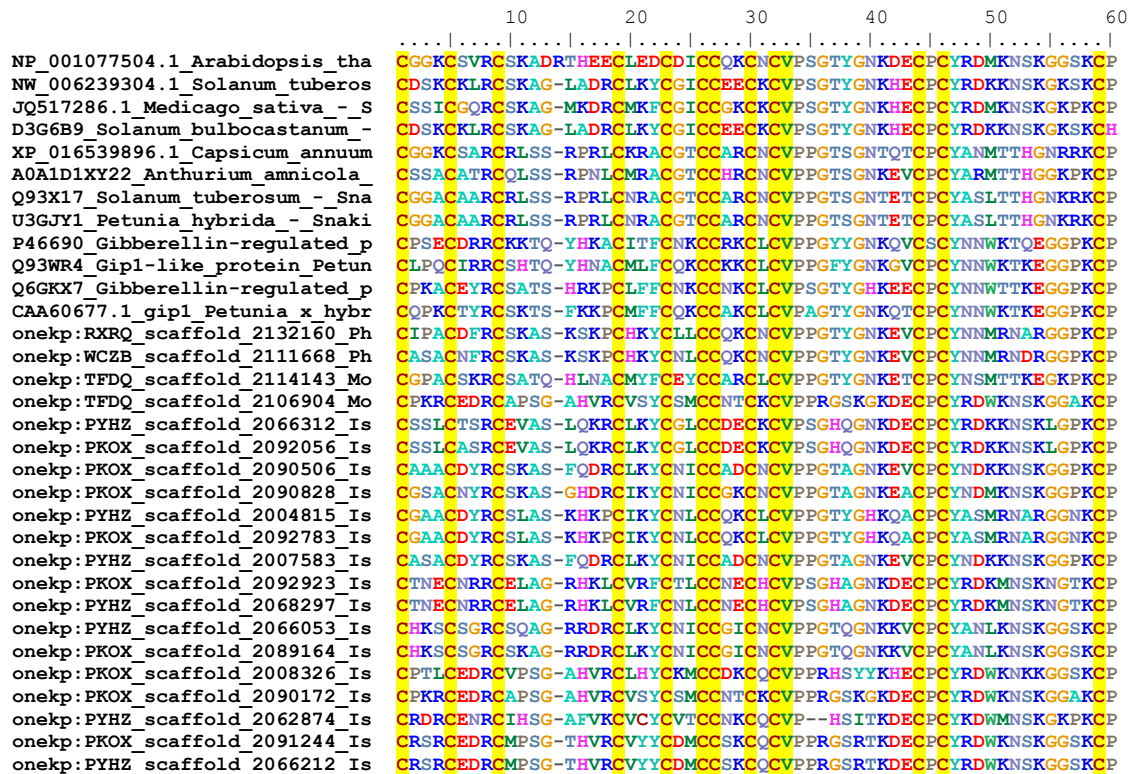


Figura 4. Alinhamento das Snakinas encontradas nas espécies de pteridófitas e seus ortólogos em plantas.



5.3 Análises filogenética das Snakinas nas linhagens vegetais

Para estudar a sua história evolutiva, as sequências de proteínas das Snakinas foram submetidas à análise filogenética. Foram construídas árvores filogenéticas utilizando como parâmetros o método *Neighbor-Joining*, que faz aproximações entre as sequências analisadas obtendo a relação de ancestralidade das mesmas (SAITOU; NEI, 1987; CALDART, 2016).

A análise filogenética das possíveis Snakinas permitiu a visualização dessas proteínas em comparação com Snakinas ortólogas já descritas em plantas superiores. A árvore filogenética foi dividida em 4 clados diversos, sendo 3 grupos utilizados como referência e 1 novo grupo identificado neste trabalho (Figura 5). No novo grupo identificado no estudo (G4), foram observados o mesmo número cisteínas e nas mesmas posições em comparação com os grupos 1, 2 e 3.

Nos quatros grupos analisados, a árvore filogenética apresentou quais os clados e espécies de plantas possuem genes de Snakinas. Sendo assim, no grupo 2 incluiu 12 genes encontrados, sendo 6 referentes a espécies de algas verdes e 6 de espécies de pteridófitas. Já o grupo 3, temos a maioria dos genes correspondentes a espécies de pteridófitas e apenas um de briófitas. O novo grupo, denominado grupo 4 é composto por 4 genes referentes a espécies de briófitas, e 1 referente a espécies de pteridófitas. Além disso, o grupo 1 apresentou 23 genes, sendo 15 de espécies de briófitas, 8 referentes a espécies de pteridófitas.

No grupo 2, dentre as espécies de algas verdes analisadas, *Chloromonas oogama*, *Pseudoscurfieldia marina* e algumas do gênero *Dunaliella* foram as que se apresentaram mais próximas dos genes de referência de Snakina do tipo 2 (Figura 5). As demais sequências, são da espécie de pteridófitas, *Azolla caroliniana*, *Osmunda sp*, e outras três do gênero *Equisetum*, que apresentaram uma maior proximidade com gene de referência de *Anthurium amnicola*. As funções dos genes de Snakinas nesse grupo 2, podem estar associadas com componentes sintetizados no decorrer do desenvolvimento normal das plantas (fatores constitutivos para a resistência) e por mecanismos ativados após o contato com patógeno (fatores de resistência induzidos) (FREITAS et al., 2009). No entanto, conseguimos comprovar neste trabalho de que as snakinas existem pelo menos desde as algas verdes, atuando contra patógenos desde o início da evolução das plantas superiores.

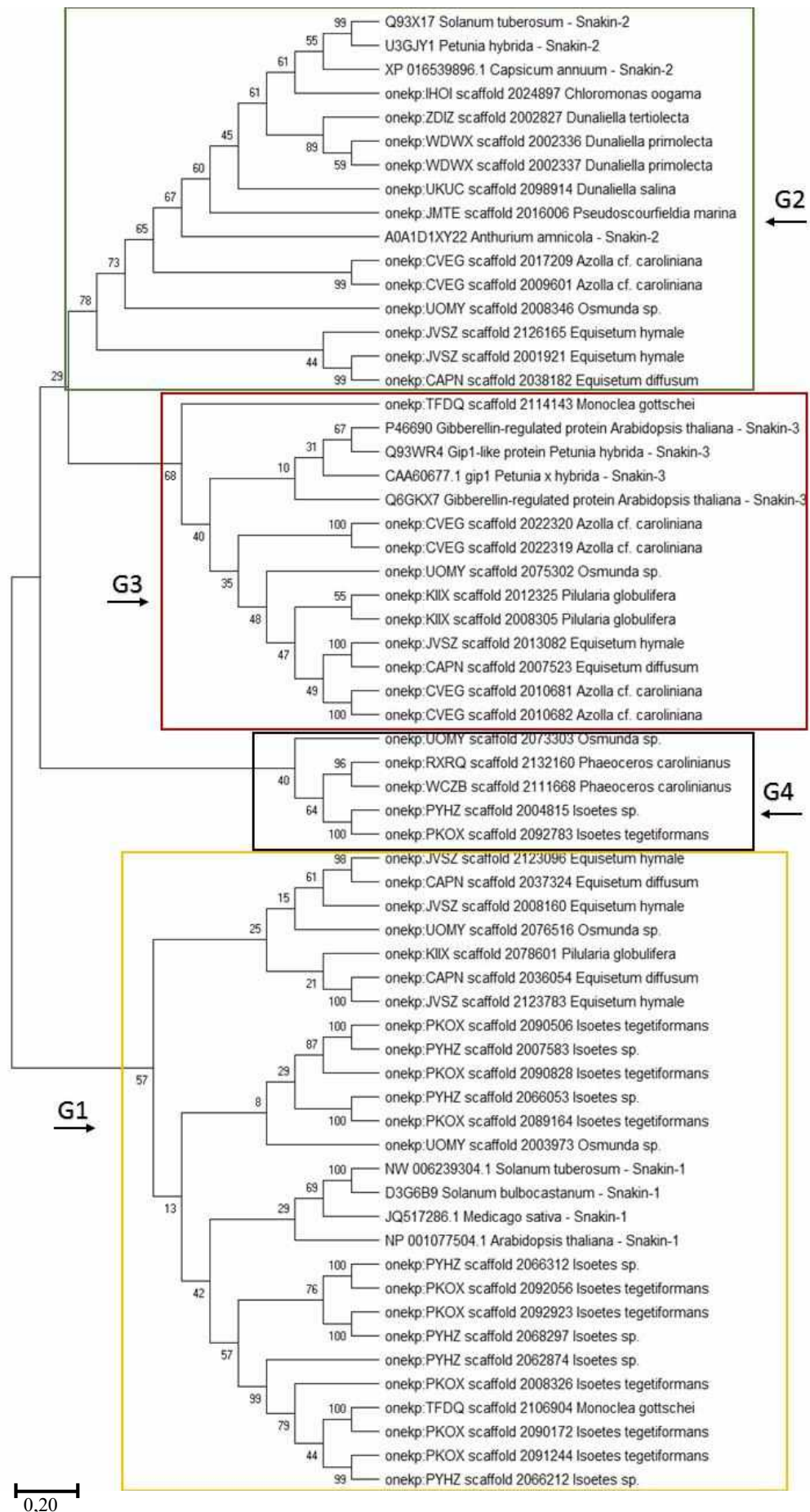
No grupo 3 (Figura 5), as pteridófitas *Azolla caroliniana*, *Osmunda sp*, *Pilularia globulifera*, *Esquisetum hymale* e *Esquisetum diffusum* apresentaram maior proximidade

genética entre as sequências de referência (PROCTOR; LIGRONE; DUCKETT; 2007; OLIVER; VELTEN; MISHLER, 2005).

No novo grupo, intitulado grupo 4, as snakinas descritas foram identificadas em espécies de briófitas e pteridófitas (Figura 5). Apesar de apresentar o domínio conservado de Snakinas, as mesmas não se agruparam com as com as sequências de referência dos 3 outros grupos, dando indícios de ser um novo grupo de Snakinas presentes apenas em plantas primitivas e por isso não foi descrito antes. Provavelmente, essas snakinas formaram um novo grupo devido a diferença em alguns aminoácidos. Por exemplo, a posição 60 do alinhamento, mostra que as demais sequências dos outros grupos possuem nessa posição os aminoácidos glutamina ou serina que fazem parte do grupo “polares não carregados”. No entanto, nas 5 sequências que pertence a esse novo grupo os aminoácidos dessa mesma posição são a alanina e o aspartato, que fazem parte dos grupos “apolar alifático” e “carregados negativamente”, respectivamente.

O grupo 1 apresentou 23 Snakinas, identificadas em briófitas e pteridófitas. É o que chama atenção nesse grupo, e que as Snakinas identificadas em espécies de briófitas do gênero *Isoetes* são mais próximas dos genes de referências do grupo 1 (Figura 5). Estas *Isoetes sp.* fazem parte de uma classe de licófitas, são herbáceas e vasculares.

Figura 5. Árvore filogenética das Snakinas identificadas em algas verdes, briófitas e pteridófitas.



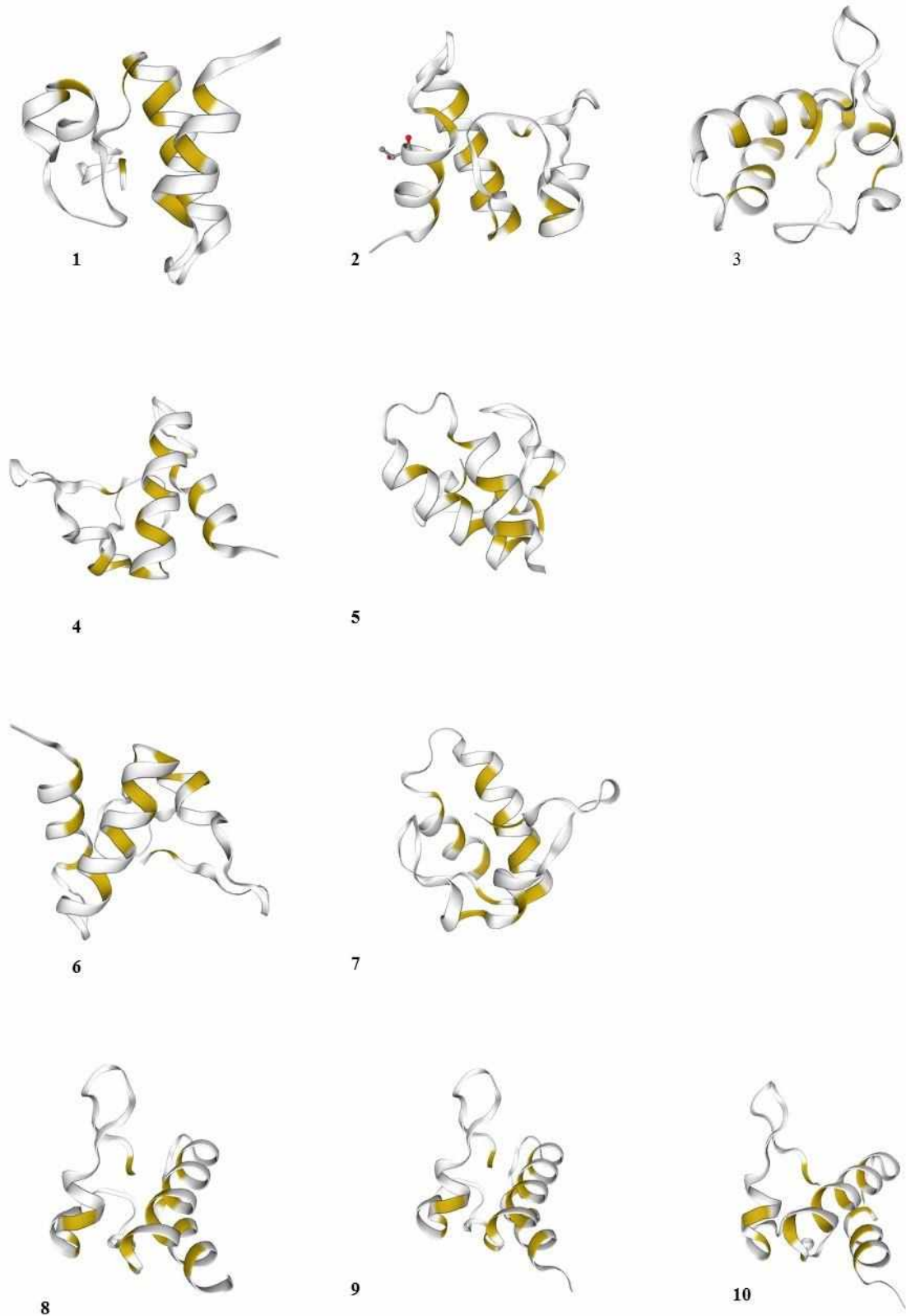
5.4 Análise estrutural

As estruturas de proteínas das Snakinas foram construídas por homologia com base no modelo – *5e5t.1.A*, através da ferramenta SWISS-MODEL e estão dispostas nas Figuras 6 e 7.

A modelagem da estrutura 3D permite a comparação de estruturas e subestruturas proteicas e reconhece as semelhanças espaciais (JAMBON; IMBERTY et al., 2003). Em algas verdes a snakina escolhida foi da espécie *Pseudoscourfieldia marina* do grupo 2, em briófitas a *Monoclea gottschei* representou o grupo 1 e o 3, isoetales a espécie *Isoetes sp* representou o grupo 1, já em pteridófitas as espécies *Pilularia globulifera*, *Azolla caroliniana* e *Osmunda sp* representou estruturalmente os grupos 1, 2 e 3 respectivamente. E para o novo grupo, o grupo 4 as espécies escolhidas foram *Osmunda sp*, *Phaeoceros carolinianus* e *Isoetes tegetiformans*.

As análises estruturais revelaram uma alta conservação entre os diferentes grupos de snakinas estudados, e tal conservação determina a funcionalidade desses peptídeos antimicrobianos que apresentam alfas hélices características.

Figura 6. Modelagem da estrutura 3D de Snakinas. (1) Briófitas – *Monoclea gottschei* (2) Briófitas – *Isoetes sp* (3) Pteridófitas – *Pilularia globulifera* – Snakin-1; (4) e (5) Algas verdes – *Pseudoscurfieldia marina* e Pteridófitas – *Azolla caroliniana*, respectivamente – Snakin-2; (6) e (7) Briófitas – *Monoclea gottschei* e Pteridófitas – *Osmunda sp*, respectivamente – Snakin-3; (8) Briófitas – *Isoetes tegetiformis* (9) Pteridófitas – *Osmunda sp* (10) Briófitas – *Phaeoceros caroliniana* – Snakin-4 (G4).



6 CONCLUSÃO

Foi possível identificar *in silico* 6 possíveis sequências completas classificadas como Snakinas em 5 espécies de Algas Verdes, 20 sequências em 4 espécies de Briófitas, e 24 sequências em 5 espécies de Pteridófitas. Quando comparadas as sequências presentes no genoma de plantas superiores, observamos que houve um aumento gradativo no número de Snakinas durante todo o processo evolutivo.

Com os resultados deste trabalho podemos inferir que os genes codificadores de Snakinas estão presentes desde espécies de algas verdes mais primitivas, porém a classificação pode ser aplicada apenas em algumas espécies dentro dos grupos analisados. No entanto, uma hipótese levantada é que a classificação existente para as Snakinas em plantas superiores não se aplica a algumas espécies primitivas, havendo assim, a necessidade de uma nova classificação incluindo o 4º grupo identificado neste trabalho.

Observa-se que as sequências analisadas possuem os domínios conservados e similares com os descritos na literatura, contudo alguns deles ficam distante filogeneticamente das sequências de referência. Experimentos adicionais devem ser feitos para a confirmação do modo de ação e funções do domínio conservado de snakinas nessas espécies primitivas de algas verdes, briófitas e pteridófitas.

REFERÊNCIAS

- AERTS, A. M. et al. The mode of antifungal action of plant, insect and human defensins. **Cellular and Molecular Life Sciences**, v. 65, n. 13, p. 2069–2079, jul. 2008.
- AUBERT, D.; CHEVILLARD, M.; DORNE, A.M.; ARLAUD, G.; and HERZOG M. Expression patterns of GASA genes in *Arabidopsis thaliana*: the GASA4 gene is up-regulated by gibberellins in meristematic regions. **Plant Molecular Biology**, v. 36, p. 871–883, 1998.
- BAHAR, A.; REN, D. Antimicrobial Peptides. **Pharmaceuticals**, v. 6, n. 12, p. 1543–1575, 2013.
- BEN-NISSAN G., J.Y. LEE, A. BOROHOF, AND D.WEISS. GIP, a *Petunia hybrida* GA-induced cysteine-rich protein: a possible role in shoot elongation and transition to flowering. **The Plant Journal**, v. 37, p. 229–238, 2004.
- BENKO-ISEPPON, A. M.; LINS GALDINO, S.; CALSA JR, T.; AKIO KIDO, E. et al. Overview on plant antimicrobial peptides. **Current Protein & Peptide Science**, v. 11, n. 3, p. 181-188, 2010.
- BERROCAL-LOBO, M.; SEGURA, A.; MORENO, M.; LOPEZ, G.; GARCIA-OLMEDO, F.; MOLINA, A. Snakin-2, an antimicrobial peptide from potato whose gene is locally induced by wounding and responds to pathogen infection. **Plant Physiology**, v.128, n.3, p.951-61, 2002
- BROGDEN, K. A. Antimicrobial peptides: pore formers or metabolic inhibitors in bacteria? **Nat Rev Micro**, v. 3, n. 3, p. 238–250, 2005.
- BOWMAN, J. L.; SAKAKIBARA, K.; FURUMIZU, C.; DIERSCHKE, T. J. A. r. o. g. **Evolution in the cycles of life**, v. 50, p. 133-154, 2016.
- CONESA, A. et al. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. **Bioinformatics**, v. 21, n. 18, p. 3674-6, 2005.
- CALDART, E. T. Análise filogenética: conceitos básicos e suas utilizações como ferramenta para virologia e epidemiologia molecular. **Acta Scientiae Veterinariae**, v. 44, p. 1392-1402, 2016.
- CASTRO, M.S.; FONTES, W. Plant defense and antimicrobial peptides. **Protein Pept Lett**, v. 12, p.13-18, 2005.
- DESTOUMIEUX-GARZÓN, D., ROSA, RD., SCHMITT. P., ET AL. Antimicrobial peptides in marine invertebrate. **Health And Disease**. 371, 2016
- EDGAR, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. **Nucleic Acids Research**, v. 32, n. 5, p.1792-1797. 2004.

FAN, S.; ZHANG, D.; XING, L.; QI, S.; DU, L.; WU, H.; SHAO.; H, LI.; Y., MA J.; HAN, M., Comprehensive analysis of GASA family members in the *Malus domestica* genome: identification, characterization, and their expressions in response to apple flower induction. **BMC Genomics** v. 18, p. 827

FRANCO, OL. Peptide promiscuity: an evolutionary concept for plant defense. **FEBS Lett**, v. 585 n. 7, p. 995–1000, 2011.

FREITAS, F. C.; VIEIRA JÚNIOR, J.R.; SILVA, D.S.G.; REIS, N.D; ANTUNES JÚNIOR, H; Mecanismos de defesa de plantas contra ataque de agentes fitopatogênicos, Embrapa Documentos, v. 133,2009.

GOYAL, R. K.; MATTOO, A. K. Plant antimicrobial peptides. **Host defense peptides and their potential as therapeutic agents**: Springer. p. 111-136, 2016

HALL, T. A. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. **Nucleic Acids Symposium Series**, v. 41, p. 95–98, 1999.

JAMBON, M.; IMBERTY A.; DELEAGE, G.; and GEORJON, C. "A new bioinformatic approach to detect common 3D sites in protein structures." **Proteins** 52 v. 2, p. 137-145, 2003

JENSSEN, H.; HAMILL, P.; HANCOCK, R. E. W. Peptide antimicrobial agents. **Clinical Microbiology Reviews**, v. 19, n. 3, p. 491–511, 2006.

KEYMANESH, K., S. SOLTANI, S. SARDARI. Application of antimicrobial peptides in agriculture and food industry. **World J. Microbiol. Biotechnol.** v. 25, p. 933-944, 2009.

KUDDUS, M. R.; RUMI, F.; TSUTSUMI, M.; TAKAHASHI, R.; YAMANO, M.; KAMIYA, T., KIKUKAWA, M. DEMURA and AIZAWA, T. Expression, purification and characterization of the recombinant cysteine-rich antimicrobial peptide snakin-1 in *Pichia pastoris*. **Protein Expr Purif** v.122, p.15-22, 2016.

KUMAR, S. et al. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. **Molecular Biology and Evolution**, v. 35. n. 6. p. 1547–1549. 2018.

LÄMKE, J.; BÄURLE, I. Epigenetic and chromatin-based mechanisms in environmental stress adaptation and stress memory in plants. **Genome Biology**, v. 18, n. 1, p. 124, 2017.

LI, J. et al. Membrane Active Antimicrobial Peptides: Translating Mechanistic Insights to Design. **Frontiers in Neuroscience**, v. 11, p. 73, 2017.

LIGRONE, R. Biological innovations that built the world: **Four-billion-year Journey through Life and Earth History**, ed. 1, 2019.

MACCARI, G.; MARIAGRAZIA, D. L.; NIFOSÌ, R. In silico design of antimicrobial peptides, **Methods Mol. Biol.** v. 1268, p. 195–219, 2015.

MAHLAPUU, M. et al. Antimicrobial Peptides: An Emerging Category of Therapeutic Agents. **Frontiers in Cellular and Infection Microbiology**, v. 6, 2016.

MARIA-NETO, S., CANDIDO, EDE S., RODRIGUES, DR., DE SOUSA, DA., DA SILVA, EM., DE MORAES, LM., OTERO-GONZALEZ, ADE J., MAGALHAES, BS., DIAS SC., FRANCO, OL. Deciphering the magainin resistance process of *Escherichia coli* strains in light of the cytosolic proteome. **Antimicrob Agents Chemother** v.56, p. 1714–1724, 2011

National Center for Biotechnology Information – NCBI
(<https://www.ncbi.nlm.nih.gov/genbank/>)

NAHIRNAK, V.; RIVAROLA M., DE URRETA, MG.; PANIEGO N.; HOPP, HE.; ALMASIA, NI.; VAZQUEZ-ROVERE, C. Genome-wide analysis of the Snakin/GASA gene family in *Solanum tuberosum* cv. Kennebec. **Am J Potato Res**, v. 93, n. 2, p. 172–188, 2016

NAVLAKHA, S.; BAR-JOSEPH, Z. Algorithms in nature: the convergence of systems biology and computational thinking. **Molecular Systems Biology**, v. 7, p. 546, 8 nov. 2011.

NAWROT, R. et al. Plant antimicrobial peptides. **Folia Microbiologica**, v. 59, n. 3, p. 181–196, 2014.

OLIVEIRA-LIMA, M., A.M. BENKO-ISEPPON, J.R.C.F. NETO, S. RODRÍGUEZ-DECUADRO, E.A. KIDO, S. CROVELLA, et al. Snakin: Structure, roles and applications of a plant antimicrobial peptide. **Curr. Protein Pept. Sci.** v. 18, p. 368-74, 2017.

OLIVER, M. J.; VELTEN, J.; MISHLER, B. D. Desiccation tolerance in bryophytes: a reflection of the primitive strategy for plant survival in dehydrating habitats?. **Integrative and Comparative Biology**, v. 45, n. 5, p. 788-799, 2005.

ORF Finder - *Open Reading Frame Finder* - NCBI (<https://www.ncbi.nlm.nih.gov/orffinder/>).

PADOVAN, L.; SCOCCHI, M.; TOSSI, A. Structural aspects of plant antimicrobial peptides. **Current Protein & Peptide Science**, v. 11, n. 3, p. 210-219, 2010.

PELEGRINI, P. B. et al. Antibacterial peptides from plants: what they are and how they probably work. **Biochemistry Research International**, v. 2011, p. 9, 2011.

PESTANA-CALSA, M. C.; RIBEIRO, I. L. A. C.; CALSA, T. J. Bioinformatics-coupled molecular approaches for unravelling potential antimicrobial peptides coding genes in Brazilian native and crop plant species. **Current protein & peptide science**, v. 11, n. 3, p. 199–209, 2010.

PORTO, W.F., and FRANCO, O.L., Theoretical structural insights into the snakin/GASA family. **Peptides**, v. 44, p. 163–167, 2013.

PROCTOR, M. C. F.; LIGRONE, R.; DUCKETT, J. G. Desiccation tolerance in the moss *Polytrichum formosum*: physiological and fine-structural changes during desiccation and recovery. **Annals of Botany**, v. 99, n. 1, p. 75-93, 2007.

ROXRUD, I., LID SE., FLETCHER, JC., SCHMIDT, ED., OPSAHL-SORTEBERG, HG., GASA4, one of the 14-member Arabidopsis GASA family of small polypeptides, regulates flowering and seed development. **Plant Cell Physiol** v. 48, n. 3 p. 471–483, 2007

RUBINOVICH L., S. RUTHSTEIN, and WEISS, D. The Arabidopsis cysteine-rich GAS5 is a redox-active metalloprotein that suppresses gibberellin responses. **Molecular Plant**, v.7, p. 244–247, 2014.

SAITOU, N.; NEI, M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. **Mol Biol Evol**, v. 4, n. 4, p. 406-25, 1987.

SILVA, ON., MULDER KC., BARBOSA AE., OTERO-GONZALEZ AJ., LOPEZ-ABARRATEGUI C. et al. Exploring the pharmacological potential of promiscuous hostdefense peptides: from natural screenings to biotechnological applications. **Front Microbiol** 2 (232), 2011.

SILVERSTEIN K.A., W.A. MOSKAL JR., H.C. WU, B.A. UNDERWOOD, M.A. GRAHAM, C.D. TOWN, AND K.A. VANDENBOSCH. Small cysteine-rich peptides resembling antimicrobial peptides have been under-predicted in plants. **The Plant Journal**, v. 51, p. 262–280, 2007.

SEGURA, A., MORENO, M., MADUENO, F., MOLINA, A. E GARCÍA-OLMEDO, F. Snakin-1, a peptide from potato that is active against plant pathogens. **Mol Plant Microbe Interact.** v.12, p. 16-23, 1999

STINTZI, A.; HEITZ, T.; PRASAD, V.; WIEDEMANN-MERDINOGLU, S.; KAUFFMANN, S.; GEOFFROY, P.; LEGRAND, M.; FRITIG, B. Plant “pathogenesis-related” proteins and their role in defense against pathogens. **Biochimie**, v. 75, p. 687–706, 1993.

SWISS-MODEL (<https://swissmodel.expasy.org/>)

TAM, J. P. et al. Antimicrobial Peptides from Plants. **Pharmaceuticals (Basel)**, v. 8, n. 4, p. 711-57, 2015.

The Universal Protein Resource - UniProt (<http://www.uniprot.org/>)

VRIES J, ARCHIBALD JM. Plant evolution: landmarks on the path to terrestrial life. **New Phytologist** v.17: 1428–1434,2018.

1000 plants (<https://db.cngb.org/onekp/>)

APÊNDICES

Apêndice 1. Espécies onde foram identificados os genes de snakinas.

Cod.	Clado	Família	Espécie
WDWX	Green Algae	Dunaliellaceae	Dunaliella primolecta
IHOI	Green Algae	Chlamydomonadaceae	Chloromonas oogama
ZDIZ	Green Algae	Dunaliellaceae	Dunaliella tertiolecta
UKUC	Green Algae	Dunaliellaceae	Dunaliella salina
JMTE	Green Algae	Pycnococcaceae	Pseudoscourfieldia marina
RXRQ	Hornworts	Notothyladaceae	Phaeoceros carolinianus
WCZB	Hornworts	Notothyladaceae	Phaeoceros carolinianus
TFDQ	Liverworts	Marchantiaceae	Monoclea gottschei
PYHZ	Lycophytes	Isoetaceae	Isoetes sp.
PKOX	Lycophytes	Isoetaceae	Isoetes tegetiformans
JVSZ	Eusporangiate Monilophytes	Equisetaceae	Equisetum hymale
CAPN	Eusporangiate Monilophytes	Equisetaceae	Equisetum diffusum
UOMY	Leptosporangiate Monilophytes	Osmundaceae	Osmunda sp.
CVEG	Leptosporangiate Monilophytes	Salviniaceae	Azolla cf. caroliniana
KIIX	Leptosporangiate Monilophytes	Marsileaceae	Pilularia globulifera