



UNIVERSIDAD REGIONAL AMAZÓNICA IKIAM

FACULTAD DE CIENCIAS DE LA VIDA

CARRERA DE BIOTECNOLOGÍA

**REVELANDO PÉPTIDOS ANTIMICROBIANOS EN LA SECRECIÓN DE LA PIEL
DE LA RANA *Agalychnis spurrelli* MEDIANTE SECUENCIAMIENTO DE ARN
DIRIGIDO.**

Proyecto de investigación previo a la obtención del Título de:

INGENIERA EN BIOTECNOLOGÍA

AUTORA: JESSICA JACQUELINE VERDEZOTO PRADO

TUTORA: PhD. CAROLINA DEL CARMEN PROAÑO BOLAÑOS

Tena – Ecuador

2024

Carrera de Biotecnología

Declaración de derecho de autor, autenticidad y responsabilidad

Tena, 04 de septiembre de 2024

Yo, Jessica Jacqueline Verdezoto Prado, con documento de identidad N° 0106654049, declaro que los resultados obtenidos en la investigación que presento en este documento final, previo a la obtención del título de Ingeniería en biotecnología, son absolutamente inéditos, originales, auténticos y personales.

En virtud de lo cual, el contenido, criterios, opiniones, resultados, análisis, interpretaciones, conclusiones, recomendaciones y todos los demás aspectos vertidos en la presente investigación son de mi autoría y de mi absoluta responsabilidad.


.....

Jessica Jacqueline Verdezoto Prado

C.I: 0106654049

Carrera de Biotecnología

Autorización de publicación en el repositorio institucional

Tena, 04 de septiembre de 2024

Yo, Jessica Jacqueline Verdezoto Prado, con documento de identidad N° 0106654049, en calidad de autor/a y titular de los derechos morales y patrimoniales del trabajo de titulación “Revelando péptidos antimicrobianos en la secreción de la piel de la rana *Agalychnis spurrelli* mediante secuenciamiento de ARN dirigido”, de conformidad con el Art. 114 del CÓDIGO ORGÁNICO DE LA ECONOMÍA SOCIAL DE LOS CONOCIMIENTOS, CREATIVIDAD E INNOVACIÓN, reconozco a favor de la Universidad Regional Amazónica Ikiám una licencia gratuita, intransferible y no exclusiva para el uso no comercial de la obra, con fines estrictamente académicos.

Así mismo autorizo a la Universidad Regional Amazónica Ikiám para que realice la publicación de este trabajo de titulación en el Repositorio Institucional de conformidad a lo dispuesto en el Art. 144 de la Ley Orgánica de Educación Superior.



Jessica Jacqueline Verdezoto Prado

C.I: 0106654049

Carrera de Biotecnología

Certificado de dirección de trabajo de titulación

Tena, 04 de septiembre de 2024

Certifico que el trabajo de titulación “Revelando péptidos antimicrobianos en la secreción de la piel de la rana *Agalychnis spurrelli* mediante secuenciamiento de ARN dirigido”, aprobado bajo el mecanismo de titulación de proyecto de investigación en formato tesis, fue realizado por Jessica Jacqueline Verdezoto Prado, bajo mi dirección.

El mismo ha sido revisado en su totalidad y analizado por la herramienta de verificación de similitud de contenido; por lo tanto, cumple con los requisitos teóricos, científicos, técnicos, metodológicos y legales establecidos por la Universidad Regional Amazónica Ikiam, para su entrega y defensa.



PhD. Carolina del Carmen Proaño Bolaños

C.I: 1714818760

AGRADECIMIENTO

Agradezco profundamente a mis profesores que han constituido parte esencial de mi formación académica. Mis profes, Manuel Siavichai, Ruth Castro, Mónica Vintimilla, Washington Noroña, Norman Alvear, Dr. Cesar Guamán, Christian Chicaiza, Jaime Marti, Caroline Bacquet, Sara Álvarez, Rodrigo Espinosa. Especialmente a mi tutora Dra. Carolina Proaño Bolaños, por su invaluable orientación, apoyo, conocimientos y guía para poder realizar el presente trabajo.

Al Laboratorio de Biología Molecular y Bioquímica de la Universidad Regional Amazónica Ikiám por abrirme las puertas y brindarme las capacidades necesarias para iniciar el bello camino de la investigación. Agradezco con mucho cariño a los técnicos de laboratorio, Giovanna Morán, Walter Quilumbaquin, Andrea Salgado y Nina Espinosa de los Monteros que siempre supieron apoyarme y solventar mis dudas en cada desafío que superé.

Adicionalmente quiero agradecer a los amigos que la universidad me regaló a cambio de las noches de desvelo y las horas de estudio. Agradezco a Ale, Henry, Pao, Vicente, Mei, Majito, Leo por apoyarme en cada materia, examen y exposición. Por ser más que compañeros, mi segunda familia con la que he crecido profesional y sentimentalmente, gracias por no rendirse. A Thomy por todo el cariño y apoyo brindado, gracias por ser.

Finalmente, con broche de oro quiero agradecer a mis padres y hermanos por su comprensión, paciencia y amor incondicional. Sus llamadas de aliento siempre fueron claves para poder conseguir la meta. Todo este trabajo es posible gracias a ellos.

DEDICATORIA

Dedico este trabajo en su totalidad a mis padres, un reflejo vivo
de lo que soy gracias a ellos.
Representa la dedicación, el empeño, el amor y los valores que me inculcaron.
Es un trabajo extenso y meticuloso sinónimo fuerza y refugio.

TABLA DE CONTENIDO

PORTADA

DECLARACIÓN DE DERECHO DE AUTOR, AUTENTICIDAD Y RESPONSABILIDADii

AUTORIZACIÓN DE PUBLICACIÓN EN EL REPOSITORIO INSTITUCIONALiii

CERTIFICADO DE DIRECCIÓN DE TRABAJO DE TITULACIÓNiv

AGRADECIMIENTOSv

DEDICATORIAvi

TABLA DE CONTENIDOvii

ÍNDICE DE TABLASx

ÍNDICE DE FIGURASxii

ABREVIACIONESxiii

RESUMENxiv

ABSTRACTxv

1. CAPÍTULO I: PLANTEAMIENTO DEL PROBLEMA DE INVESTIGACIÓN 1

1.1. Antecedentes 1

1.1.1 Secreción cutánea de anfibios 1

1.1.2 Péptidos antimicrobianos 2

1.1.3 Secuenciación de ARN dirigido 3

1.1.4 Secuenciación por nanoporos..... 5

1.2. Planteamiento del problema 6

1.3. Justificación de la investigación 7

1.4. Objetivos de la investigación 9

1.4.1. General 9

1.4.2. Específicos 9

2. CAPÍTULO II: MARCO METODOLÓGICO 10

2.1. Preparación y secuenciación de bibliotecas de ADNc 10

2.1.1. Colección de muestras biológicas 10

2.1.2. Extracción del ARNm 10

2.1.3. Obtención de la librería de ADNc 11

2.1.4.	Amplificación del ADNc por RACE PCR	11
2.1.5.	Visualización de productos PCR por electroforesis	12
2.1.6.	Purificación de productos de la RACE PCR	13
2.1.7.	Ligación de productos PCR.....	13
2.1.8.	Secuenciación por nanoporos del ADNc	14
2.1.9.	Adquisición de datos y basecalling	15
2.2.	Diseño de pipeline para limpieza y tratamiento	15
2.2.1.	Análisis de datos obtenidos	15
2.2.2.	Análisis de secuencias putativas	17
2.2.3.	Análisis de espectrometría de masas	17
2.3.	Análisis de propiedades fisicoquímicas	18
2.3.1.	Caracterización fisicoquímica	18
2.3.2.	Predicción y modelamiento de estructuras 3D	18
3.	CAPÍTULO III: PRESENTACIÓN DE DATOS Y RESULTADOS	20
3.1.	Preparación y secuenciación de bibliotecas de ADNc	20
3.1.1.	Tratamiento de la muestra	20
3.1.2.	Reporte de la calidad de datos de secuenciación	20
3.2.	Diseño de pipeline para limpieza y tratamiento	22
3.2.1.	Secuencias precursoras obtenidas	22
3.2.2.	Clasificación de PAMs encontrados	27
3.2.3.	Clasificación de proteínas encontradas	53
3.2.4.	Análisis de PAMs por espectrometría de masas	61
3.3.	Análisis de propiedades fisicoquímicas	66
3.3.1.	Caracterización fisicoquímica de los péptidos nuevos	66
3.3.2.	Predicción de estructura secundaria de los nuevos PAMs	69
3.3.3.	Predicción de estructura terciaria de las nuevas proteínas	71
4.	CAPÍTULO IV: INTERPRETACIÓN Y DISCUSIÓN	73
4.1.	Preparación y secuenciación de bibliotecas de ADNc	73
4.1.1.	Tratamiento de la muestra	73
4.1.2.	Reporte de la calidad de datos de secuenciación	74
4.2.	Diseño de pipeline para limpieza y tratamiento	75

4.2.1. Secuencias precursoras obtenidas	75
4.2.2. Clasificación de PAMs encontrados	76
4.2.3. Clasificación de proteínas encontradas	79
4.2.4. Análisis de PAMs por espectrometría de masas	80
4.3. Análisis de propiedades fisicoquímicas	81
4.3.1. Caracterización fisicoquímica de los péptidos nuevos	81
4.3.2. Predicción de estructura secundaria de los nuevos PAMs	82
4.3.3. Predicción de estructura terciaria de las nuevas proteínas	83
5. CAPÍTULO V: CONCLUSIONES Y RECOMENDACIONES	85
REFERENCIAS BIBLIOGRÁFICAS	

ÍNDICE DE TABLAS

Tabla 1. Primers Forward y Reverse empleados para amplificar por RACE PCR.	11
Tabla 2. Resumen de las secuenciación y análisis bioinformático	20
Tabla 3. Lista de péptidos antimicrobianos encontrados por secuenciación de ARN dirigido	24
Tabla 4. Lista de proteínas de interés encontradas por secuenciación de ARN dirigido a.....	26
Tabla 5. Secuencias aminoacídicas que codifican a péptidos de la familia de dermaseptinas.....	28
Tabla 6. Alineamiento múltiple de las secuencias correspondientes a dermaseptinas	31
Tabla 7. Secuencias aminoacídicas que codifican a péptidos de la familia de triptofilinas.....	34
Tabla 8. Alineamiento múltiple de las secuencias correspondientes a triptofilinas	35
Tabla 9. Secuencias aminoacídicas que codifican a péptidos de la familia de filoseptinas.....	36
Tabla 10. Alineamiento múltiple de las secuencias correspondientes a filoseptinas ..	38
Tabla 11. Secuencias aminoacídicas que codifican a péptidos de la familia de filoquininas	39
Tabla 12. Alineamiento múltiple de las secuencias correspondientes a filoquininas ..	40
Tabla 13. Secuencias aminoacídicas que codifican a péptidos de la familia de medusinas.....	41
Tabla 14. Alineamiento múltiple de las secuencias correspondientes a medusinas ...	42
Tabla 15. Secuencias aminoacídicas que codifican a péptidos de la familia de plasticinas	43
Tabla 16. Alineamiento múltiple de las secuencias correspondientes a plasticinas.....	43
Tabla 17. Secuencias aminoacídicas que codifican a péptidos de la familia de caerinas	44
Tabla 18. Alineamiento múltiple de las secuencias correspondientes a caerinas.....	45
Tabla 19. Secuencias aminoacídicas que codifican a péptidos huérfanos	46

Tabla 20. Alineamiento múltiple de las secuencias aminoacídicas de péptidos huérfanos	49
Tabla 21. Secuencias aminoacídicas que codifican a proteínas de tipo Kazal.....	54
Tabla 22. Alineamiento múltiple de las secuencias correspondientes a proteínas de tipo Kazal	57
Tabla 23. Péptidos con secuencias confirmadas por LC-MS/MS	62
Tabla 24. Características fisicoquímicas de los péptidos correspondientes a dermaseptinas.....	66
Tabla 25. Características fisicoquímicas de los péptidos correspondientes a filoseptinas.....	67
Tabla 26. Características fisicoquímicas de los péptidos correspondientes a plasticinas	68
Tabla 27. Características fisicoquímicas de los péptidos correspondientes a caerinas	68
Tabla 28. Características fisicoquímicas de los péptidos correspondientes a péptidos huérfanos.....	68
Tabla 29. Predicción de estructuras para los nuevos péptidos	69
Tabla 30. Predicción de estructura 3D de proteínas inhibitorias de tipo Kazal.....	71

INDICE DE FIGURAS

Figura 1. Diagrama de flujo de trabajo para el análisis de datos de secuenciación de <i>A. spurrelli</i>	16
Figura 2. Depuración de datos de la secuenciación por nanoporos	21
Figura 3. Control de calidad de la secuenciación por nanoporos de <i>A. spurrelli</i>	22
Figura 4. Diseño del flujo de trabajo (pipeline) para la limpieza y tratamiento	23

ABREVIACIONES

PAMs: Péptidos antimicrobianos.

ARN: ácido ribonucleico que contiene información transcrita del ADN e incluye el ARN ribosomal, ARN transferencia, ARN mensajero, entre otros.

ARNm: ácido ribonucleico mensajero que contiene la información codificante proveniente del ADN, es esencial para la traducción de proteínas.

ADN: ácido desoxirribonucleico que contiene información genética para el desarrollo y funcionamiento de los organismos.

ADNc: ADN complementario, es una molécula de ADN sintetizada a partir de una molécula de ARNm mediante transcripción inversa.

RACE PCR: Rapid amplification of cDNA ends, amplificación rápida de los extremos del ADNc, técnica molecular que permite la amplificación del ADNc de longitud completa.

Ciencias Ómicas: son disciplinas de la biología que involucran el estudio de conjuntos completos de moléculas biológicas. Entre ellas la genómica, transcriptómica, proteómica, lipidómica, metabolómica, epigenómica, e interactómica.

dRNA-seq: direct RNA sequencing, método de secuenciación directa a partir de ARN

dcDNA-seq: direct cDNA sequencing, método de secuenciación directa a partir de ADNc

Base-calling: proceso por el cual se codifica las señales eléctricas registradas durante el paso de la molécula en la secuenciación.

ORF (Open Reading Frame): marco de lectura abierto, es una sección del ADN delimitada por codones de inicio y codones de parada (STOP), potencialmente codifica una proteína.

KB: Kilobyte unidad de medida de datos informáticos, no es lo mismo que kb/Kbp (Kilobase).

MB: Megabyte unidad de medida de datos informáticos superior a KB, es diferente a Mb (Megabase).

RESUMEN

El estudio y caracterización de péptidos antimicrobianos en la piel de anfibios tradicionalmente utiliza técnicas moleculares como el clonaje de ARNm y el análisis proteómico mediante espectrometría de masas. Sin embargo, una limitación del clonaje molecular es que cada plásmido clonado representa solo una secuencia de interés, lo que requiere la clonación y secuenciación de un gran número de clones para identificar ciertos péptidos. Para superar esta limitación, se puede amplificar el ARN utilizando primers específicos y secuenciar los amplicones con tecnología de tercera generación, como la secuenciación por nanoporos. Este estudio tiene como objetivo identificar péptidos antimicrobianos en la secreción cutánea de la rana mono planeadora *Agalychnis spurrelli* utilizando el dispositivo MinION de Oxford Nanopore Technology. La metodología se dividió en tres fases: (a) creación y secuenciación de ADNc amplificado, (b) diseño de un pipeline para la obtención de secuencias precursoras, y (c) análisis predictivo de potenciales péptidos antimicrobianos. El flujo de trabajo diseñado permitió identificar 163 secuencias precursoras dentro del marco de lectura abierto (ORF), codificando 39 péptidos agrupados en 7 familias. Entre estos, se identificaron: Dermaseptinas, Tryptofilinas, Filoseptinas, Filoquininas, Medusinas, Plasticinas, Caerinas y otros, además de inhibidores de proteínas de tipo Kazal. La caracterización fisicoquímica sugiere que 19 péptidos podrían tener actividad antimicrobiana. Estos resultados demuestran que la secuenciación por nanoporos al ARN dirigido es eficaz para dilucidar la diversidad de péptidos y proteínas en secreciones cutáneas de anfibios con potencial terapéutico.

Palabras clave: ADNc, *Agalychnis spurrelli*, secreción cutánea, secuenciación por nanoporos, secuenciación de ARN dirigido, péptidos, proteínas.

ABSTRACT

The study and characterization of antimicrobial peptides in amphibian skin traditionally uses molecular techniques such as mRNA cloning and proteomic analysis by mass spectrometry. However, a limitation of molecular cloning is that each cloned plasmid represents only one sequence of interest, requiring the cloning and sequencing of a large number of clones to identify certain peptides. To overcome this limitation, RNA can be amplified using specific primers and the amplicons can be sequenced using third-generation technology, such as nanopore sequencing. This study aims to identify antimicrobial peptides in the skin secretion of the monkey glider frog *Agalychnis spurrelli* using Oxford Nanopore Technology's MinION device. The methodology was divided into three phases: (a) creation and sequencing of amplified cDNA, (b) design of a pipeline for obtaining precursor sequences, and (c) predictive analysis of potential antimicrobial peptides. The designed workflow allowed the identification of 163 precursor sequences within the open reading frame (ORF), encoding 39 peptides grouped in 7 families. Among these, the following were identified: Dermaseptins, Tryptophilins, Phyloseptins, Phylloquinins, Medusins, Plasticins, Caerins and others, in addition to Kazal-type protein inhibitors. Physicochemical characterization suggests that 19 peptides may have antimicrobial activity. These results demonstrate that nanopore sequencing to targeted RNA is effective in elucidating the diversity of peptides and proteins in amphibian skin secretions with therapeutic potential.

Keywords: cDNA, *Agalychnis spurrelli*, skin secretion, nanopore sequencing, RNA-seq, peptides, proteins.

1. CAPÍTULO I: PLANTEAMIENTO DEL PROBLEMA DE INVESTIGACIÓN

1.1. Antecedentes

1.1.1. Secreción cutánea de anfibios

La secreción cutánea de los anfibios es producida por las glándulas granulares de la piel como respuesta defensiva frente a estímulos externos. Contiene una variedad de compuestos bioactivos, incluidas aminos biogénicas, alcaloides, ácidos nucleicos, péptidos y proteínas. Estos componentes se liberan como parte de una estrategia defensiva contra microorganismos patógenos potenciales y depredadores potenciales ocasionando irritaciones, dolor y entumecimiento. En el caso de los microorganismos, estos compuestos pueden interactuar con las membranas celulares de los patógenos, comprometiendo y destruyendo sus membranas celulares [1,2].

Hasta la fecha, se han registrado 1079 péptidos bioactivos provenientes de anfibios en la base de datos de péptidos antimicrobianos (APD3) de la Universidad del centro Médico de Nebraska [3,4]. Las principales familias de anfibios que producen péptidos bioactivos son Ranidae, Pipidae, Alytidae, Bombinatoridae, Hylidae, entre otras [5,6].

Los péptidos bioactivos presentan diversas actividades biológicas, entre ellas: inmunomoduladora, vasoconstrictora, inflamatoria, antioxidante, inhibitoria, anticoagulante, neuroestimulante, anticancerígena, antimicrobiana [7–9]. La actividad antimicrobiana destaca por el amplio espectro de acción que poseen los péptidos contra bacterias, hongos, virus y parásitos [10]. Por ejemplo, el estudio realizado por Thomson y colaboradores en 2007, logró identificar nueve péptidos antimicrobianos de las secreciones cutáneas de *Phyllomedusa hypochondrialis*. De los cuales, cinco nuevas dermaseptinas y cuatro filoseptinas presentaron actividad inhibitoria contra bacterias Gram positivas y Gram negativas (*Escherichia coli*, *Staphylococcus aureus* y *Micrococcus luteus*) [11].

Zannella y colaboradores en 2022, analizaron el péptido Temporin L (TL) en búsqueda de actividad antiviral y develaron que posee actividad significativa frente a herpes virus, paramyxovirus, virus de la influenza y coronavirus, dentro de los cuales incluye a SARS-CoV-2 [12]. Por otro lado, en 2012 Abbassi y colaboradores evaluaron la actividad antibacteriana y antiparasitaria del péptido temporin-SHd, demostrando que expresó amplio espectro de acción contra cepas multirresistentes de bacterias como *S. aureus* y actividad antiparasitaria frente al promastigote y amastigote de *Leishmania infantum* causando fuertes perturbaciones en la membrana fosolípida [13].

1.1.2. Péptidos antimicrobianos.

Los péptidos antimicrobianos (PAMs) son moléculas relativamente pequeñas conformadas de 12 a 60 aminoácidos. Su estructura puede variar en hélices α , láminas β y bucles, son catiónicos, con carga positiva a pH fisiológico, característica que les permite interactuar con las membranas de los microorganismos. Además, pueden adoptar estructuras anfipáticas con extremos hidrofóbicos e hidrofílicos que les permiten mayor interacción con la membrana. En las últimas décadas, la mayoría de PAMs caracterizados provienen de la secreción cutánea de anfibios [14,15].

Dentro de los anfibios, una de las subfamilias que más registros de péptidos antimicrobianos posee corresponde a *Phyllomedusinae*. En esta subfamilia se han descrito péptidos antimicrobianos pertenecientes a varias familias que incluyen: dermaseptinas, filoseptinas, bradiquininas, medusinas, plasticinas, triptofilinas, entre otros. Estos péptidos se caracterizan por presentar actividad inhibitoria contra hongos, bacterias, virus, parásitos, además que, algunos poseen actividad vasoconstrictora, inflamatoria y estimulante. Un ejemplo de ello es el estudio realizado por Leite y colaboradores en 2005, quienes aislaron seis nuevos péptidos de la secreción cutánea de *Pithecopus hypochondrialis* y *P. oreades* a los que llamaron filoseptinas (PS-1, -2, -3, -4, -5 y -6). Estas filoseptinas presentaron actividad antibacteriana y antiprotozoarias, destacando que la filoseptina-1 demostró bioactividad contra bacterias Gram positivas

y Gram negativas y nula actividad hemolítica [16,17]. Otro ejemplo es el estudio de Proaño y colaboradores en 2019, quienes realizaron el análisis de la secreción cutánea de *Agalychnis spurrelli* mediante clonaje molecular y espectrometría de masas, descubriendo once nuevos péptidos pertenecientes a las familias de filoquininas, medusinas, dermaseptinas, filoseptinas y péptidos huérfanos. El péptido denominado Dermaseptin-SP2 fue señalado como el péptido con mayor actividad inhibitoria contra *E. coli*, *S. aureus* ORSA y *Candida albicans* [18].

La técnica de búsqueda y descubrimiento de péptidos se basa actualmente en el clonaje molecular, que consiste en aislar el ARNm que codifica los péptidos de interés y posteriormente se emplea para sintetizar las cadenas de ADN complementario (ADNc). Luego, empleando cebadores (primers) específicos para los péptidos se amplifica el ADNc mediante RACE PCR y se clonan aplicando el sistema pGEM-T vector. Las regiones clonadas de secuencian por método Sanger y se realizan los respectivos análisis bioinformáticos para confirmar similitud, estructura y predicción de actividad de las secuencias precursoras de péptidos antimicrobianos. Sin embargo, esta técnica permite identificar una cantidad limitada de secuencias de interés. Puesto que al insertar el ADN en el plásmido solo se puede clonar una secuencia a la vez reduciendo la eficacia de la técnica y limitando el descubrimiento de péptidos de la secreción de la piel de los anfibios [19,20]. Una alternativa para sobrellevar esta limitación es la secuenciación de ARN dirigido empleando primers específicos para obtener secuencias precursoras de péptidos.

1.1.3. Secuenciación de ARN dirigido

La secuenciación de ARN dirigido es una técnica que se enfoca en secuencias de particular interés como las secuencias codificantes que dan como resultado proteínas específicas, incluyendo péptidos antimicrobianos. Esta técnica se basa en la extracción de ARN mensajero (ARNm) que permite la síntesis del ADN complementario (ADNc) y este a su vez es amplificado por primers específicos que flanquean las secciones de interés, mismas que serán secuenciadas y analizadas. Este enfoque dirigido es altamente

eficiente porque enfoca los recursos de secuenciación en las secuencias más relevantes, reduciendo el ruido y aumentando la profundidad de lectura en las secuencias de interés [21,22].

En el contexto de la identificación de péptidos antimicrobianos, la utilización de primers específicos en la amplificación del ARN dirigido permite focalizar la secuenciación en las regiones de interés. Estos primers son diseñados para unirse a secuencias específicas de ARNm que codifican péptidos conocidos o potencialmente activos, dirigiendo la secuenciación hacia estas áreas. Rončević y colaboradores en 2018, establecieron el protocolo exploratorio a la región altamente conservada del péptido señal en secuencias precursoras de PAMs pertenecientes a familias de anfibios Ranidae, Hylidae y Bombinatoridae. Demostrando que el diseño de cebadores degenerados específicos en conjunto con la secuenciación de ARN dirigido resultan una alternativa eficaz para obtener secuencias precursoras de péptidos antimicrobianos[23,24].

Por otro lado, en otro estudio realizado por Rončević y colaboradores en 2019, aplicaron secuenciación de ARN dirigido a la secreción cutánea de cinco especies diferentes de ranas Ranidae y lograron identificar más de cien péptidos de longitud completa, de las cuales 29 eran secuencias nuevas. Adicionalmente sintetizaron seis péptidos y uno de ellos demostró ser un antimicrobiano potente y de amplio espectro, activo contra cepas bacterianas ATCC [25].

La secuenciación de ARN dirigido se lleva a cabo empleando la técnica de secuenciación masiva de Illumina. Esta plataforma permite la secuenciación simultánea y masiva de una gran cantidad de ADN, presentando diversas ventajas como una mayor sensibilidad en la detección de variantes, un tiempo de ejecución reducido y menor costo de secuenciación por base. En el proceso, la plataforma Illumina realiza la síntesis y amplificación de fragmentos aleatorios de ADN mediante la técnica de PCR en puente, generando múltiples copias idénticas de la molécula que se conocen como clústeres [26,27].

Utilizar Illumina, supone conseguir datos de alta precisión y calidad debido al potencial que tiene la plataforma en su secuenciación masiva. Sin embargo, una de las limitaciones que presenta esta técnica es la fragmentación de la molécula de ADN en secuencias cortas de entre 50-150 pares de bases (pb), que luego son amplificadas por PCR en puente y secuenciadas en ciclos repetidos. En la actualidad, solo se logran lecturas completas de ADNc para un máximo de 50% de las lecturas secuenciadas, lo que dificulta el análisis y la interpretación de los datos. Esto limita en gran medida el descubrimiento de secuencias precursoras de potenciales péptidos e impide capturar la totalidad de los mismos. Como una alternativa a esta limitación surge la secuenciación de molécula única de lectura larga por nanoporos que evita el corte y posterior ensamblaje de las moléculas secuenciadas [28,29].

1.1.4. Secuenciación por nanoporos.

La secuenciación por Nanoporos o secuenciación de tercera generación, es una técnica avanzada de secuenciación caracterizada por su capacidad para secuenciar moléculas completas de ADN o ARN de extremo a extremo. Se basa en el paso de las moléculas a través de nanoporos, que son canales pequeños presentes en una membrana sobre la cual se coloca un voltaje. El paso de la molécula a través del poro genera cambios en la corriente eléctrica, los cuales son detectados y registrados por el secuenciador. La secuenciación por nanoporos tiene capacidad para generar lecturas de secuencias largas en tiempo real, diferenciándose de otras técnicas de secuenciación. Además, esta técnica tiene la capacidad de identificar y detectar modificaciones químicas en las moléculas de ADN o ARN, como metilaciones o adiciones de grupos funcionales [30,31].

Existe un continuo avance en la secuenciación de molécula única o por nanoporos, siendo un ejemplo destacado los estudios realizados en organismos modelo implementado las tecnologías Oxford Nanopore (ONT). En un estudio efectuado en 2022 por Grünberger y su equipo, se evaluaron los diferentes protocolos de secuenciación de ARN basados en ONT, incluidos ARN directo, ADNc directo y PCR-ADNc. Describiendo que la secuenciación de PCR-ADNc brinda un mejor rendimiento y precisión en

comparación con la secuenciación directa de ARN en el organismo modelo procariótico *Escherichia coli* K-12 [32].

Del mismo modo, en 2022, Wongsurawat y colaboradores llevaron a cabo la secuenciación de ARN nativo y de ADNc en *Saccharomyces cerevisiae* para permitir comparaciones y evaluaciones relevantes de las técnicas. Los resultados obtenidos revelaron las ventajas y desventajas de usar ARN-seq directo o ADNc-seq directo para estudiar transcriptomas en levaduras, lo cual brindó información crucial a la hora de seleccionar un método de secuenciación de lectura larga apropiado para comprender las funciones de los genes, las vías y la caracterización funcional detallada. Por tal motivo, en el presente estudio se ha decidido implementar la secuenciación por ligación de ADN amplificado por RACE-PCR. Al emplear el kit Ligation Sequencing Kit V14 (SQK-LSK114) desarrollado por Oxford Nanopore Technology, en su dispositivo de secuenciación Minlon, con el propósito de obtener un gran número de datos y de buena calidad [33,34].

1.2. Planteamiento del problema

La resistencia antimicrobiana surge como una problemática a nivel mundial causada por el uso excesivo de los antimicrobianos. La Organización Mundial de la Salud (OMS) ha declarado en 2021, que la resistencia antimicrobiana se encuentra dentro de las 10 principales amenazas de salud pública debido a que causa 700 mil muertes cada año. Además, se estima que para el 2050 este problema puede ocasionar aproximadamente 10 millones de muertes y 100 billones de dólares en pérdidas económicas [35]. En Ecuador, se reportó por primera vez la bacteria *Klebsiella pneumoniae* productora de carbapenemasas en 2010. Adicionalmente el Instituto Nacional de Investigación en Salud Pública en el reporte de datos de resistencia a los antimicrobianos 2014- 2018, describió que los microorganismos: *Escherichia coli*, *K. pneumoniae*, *Staphylococcus aureus* y *Pseudomonas aeruginosa* son los microorganismos mayormente reportados anualmente y son sujetos a vigilancia de resistencia a los antimicrobianos [36].

Por lo tanto, existe la necesidad de buscar nuevos agentes antimicrobianos capaces de combatir a los microorganismos resistentes. Es así que los péptidos provenientes de la secreción cutánea de los anfibios surgen como una alternativa para combatir esta problemática.

A nivel de laboratorio, el flujo de trabajo para el descubrimiento de péptidos consiste en realizar clonaje molecular para crear bases de datos que se pueda contrastar con proteómica. Este proceso necesita la creación varios clones ya que cada plásmido solo puede contener una única secuencia de interés. De igual manera requiere que este gran número de clones sean secuenciados, lo que se traduce en que mayor cantidad de clones y secuencias van a proporcionar la misma información que se obtienen por medio de una secuenciación de molécula larga. Sin mencionar que al tener limitados clones se pierde gran diversidad de secuencias precursoras de péptidos en la muestra analizada [18].

Por otro lado, la secuenciación de ARN dirigido conlleva la aplicación de secuenciación Illumina, que a su vez implica la fragmentación de la molécula de ADN en secuencias cortas para completar la amplificación por puente. Lo que supone una gran limitación al momento de capturar el total de las secuencias precursoras del ADNc [28,29].

1.3. Justificación de la investigación

Dado el problema de la resistencia antimicrobiana es necesario encontrar alternativas que puedan combatir los microorganismos resistentes de manera eficaz. Por tal motivo, la búsqueda de péptidos antimicrobianos en la secreción cutánea de los anfibios resulta una línea de investigación prometedora para proporcionar nuevas alternativas terapéuticas antimicrobianas. Con el objetivo de llegar a una caracterización completa de los péptidos de la secreción de la piel se propone el uso de la técnica de secuenciación de ARN dirigido, donde se amplificará el ADNc con primers específicos para su posterior secuenciación mediante tecnología de nanoporos.

La tecnología desarrollada por Oxford Nanopore Technology (ONT) permite la secuenciación de molécula única mediante nanoporos obteniendo lecturas de ADN/ARN de alrededor de 2,2 millones de pares de bases de alta calidad en menor tiempo. Lo que significa que el dispositivo MinION (ONT), secuencia la molécula completa de ADN/ARN a una velocidad de mil millones de pares de bases por segundo, obteniendo de manera más eficiente los datos [31]. Esta ventaja facilita el descubrimiento de péptidos empleando secuenciación de ARN dirigido, de modo que se pueden obtener lecturas completas de ADNc que codifican una gran diversidad de moléculas, explorando la totalidad de los péptidos bioactivos presentes en la secreción cutánea de anfibios.

Por otro lado, se decidió trabajar con la especie *Agalychnis spurrelli* (Anura: Hylidae: Phyllomedusinae) como un modelo, debido a que en estudios previos donde se empleó clonaje molecular y proteómica se han identificado alrededor de 30 péptidos bioactivos presentes en su piel. Por lo cual, nos hemos apoyado en esta información para realizar la búsqueda de secuencias peptídicas y contrastarlas con datos ya reportados. Por lo tanto, en el presente estudio se plantea implementar la técnica de secuenciación de ARN dirigido, al amplificar con primers específicos el ADNc proveniente de la secreción de la piel de la rana arbórea *Agalychnis spurrelli* y secuenciarlo mediante secuenciación de molécula única desarrollada por las tecnologías Oxford Nanopore (ONT), con el propósito de descubrir potenciales péptidos que puedan ser utilizados en la lucha contra la resistencia antibacteriana.

1.4. Objetivos de la investigación

1.4.1. General

Identificar péptidos antimicrobianos presentes en la secreción cutánea de la rana *Agalychnis spurrelli* mediante secuenciación por nanoporos del ARN dirigido.

1.4.2. Específicos

- Secuenciar mediante nanoporos un set de amplicones generados empleando primers específicos a partir del ADNc de la secreción de la piel de la rana *Agalychnis spurrelli*.
- Establecer un flujo de trabajo (pipeline) para el tratamiento bioinformático de los datos obtenidos de la secuenciación de ARN-dirigido a la secreción de la piel de la rana *Agalychnis spurrelli*.
- Determinar la secuencia precursora de al menos dos péptidos de la secreción de la piel de la rana *Agalychnis spurrelli*.

2. CAPÍTULO II: MARCO METODOLÓGICO

2.1. Preparación y secuenciación de bibliotecas de ADNc

2.1.1. Colección de muestras biológicas

Los especímenes de *A. spurrelli* fueron proporcionados por el Centro Jambatu de Investigación y Conservación de Anfibios, bajo el contrato marco de acceso a recursos genéticos MAAE-DBI-CM-2021-0185, emitido por el Ministerio del Ambiente, Agua y Transición Ecológica. Se emplearon 2 especímenes adultos colectados en estado de cautiverio en la mañana y una vez completado el proceso de extracción de secreción, los individuos regresaron a sus terrarios sin daño aparente. Las secreciones cutáneas se extrajeron mediante estimulación por masaje dorsal, hasta observar una espuma blanquecina y percibir un olor pungente característico [18]. El exudado color blanco fue recolectado tras varios lavados del espécimen con agua desionizada en un vaso de precipitados de 1 litro. Inmediatamente, se trasvasó la muestra a tubos cónicos de 50 ml estériles y se conservaron en congelación a -80 °C por 24 horas. Posteriormente, las muestras fueron liofilizadas a -80 °C por 3 días en un VirTis BenchTop Pro Freeze Dryer (Sp Scientific, USA). A continuación, se pesaron y almacenaron a -20 °C hasta su uso posterior. Los especímenes

2.1.2. Extracción del ARNm

A partir de la secreción cutánea liofilizada se realizó la extracción del ARN mensajero, empleando el kit Dynabeads® ARNm DIRECT™ (DynaL Biotec, UK). El ARNm se aisló mediante hibridación con el oligo dT unido a las microesferas magnéticas de la siguiente manera: en primera instancia, se diluyeron 5 mg de secreción liofilizada en 1 ml de tampón de lisis/unión celular. Seguidamente, el ARNm se separó utilizando perlas magnéticas oligo-dT del kit. Se realizaron repetidos lavados a las perlas hibridadas con 600 µl de los tampones de lavado A y B del kit, siguiendo las instrucciones del kit.

Finalmente, se eluyó el ARNm de las perlas agregando 20 µl de tampón de elución (Tris HCl) y un shock térmico a 80 °C. La calidad del ARNm se midió con Nanodrop (Thermo Fisher Scientific, USA) y la concentración con Qubit empleando el kit Qubit® RNA HS Assay (Thermo Fisher Scientific, USA) [37–39].

2.1.3. Obtención de la librería de ADNc

La librería de ADN complementario (ADNc), se construyó con el sistema GoScript™ Reverse Transcription System para sintetizar la primera cadena de 3' ADNc por medio de transcripción reversa. Para favorecer la hibridación, se incubaron 4 µl del ARNm extraído y 1 µl del primer 3'CDS (20 µM) (5'- AAGCAGTGGTATCAACGCAGAGTACT30VN-3'; V=A, C, o G, N=A, T, C, o G) por 5 min a 70 °C seguidos de 5 min en hielo. A esta reacción se añadió 15 µl de máster mix compuesto por: 4 µl de Go Script 5x buffer, 3.2 µl MgCl₂, 1 µl dNTPs mix, 1 µl Go Script Retro transcriptasa y 5.8 µl de H₂O libre de nucleasas. Se obtuvieron 6 réplicas de esta librería de ADNc [18].

2.1.4. Amplificación del ADNc por RACE PCR

Una vez construida la librería de ADNc se realizó 3'-RACE PCR, para lo cual se prepararon 25 µl de Máster mix compuesto por: 12.5 µl de OneTaq® Hot Start 2X, 0.25 µl de Primer Forward (20 µM), 0.25 µl de Primer Reverse (20 µM), 5 µl de ADN muestra y 7 µl de agua libre de nucleasas. Para cada reacción de RACE PCR se empleó una combinación con uno de los diez Primers Forward y el Primer Reverse (Nested Universal Primer) como se observa en la Tabla 1. Se adicionó un control negativo en cada reacción empleando agua libre de nucleasas en lugar de ADNc. El primer SIC fue diseñado en el programa Geneious Prime v 2023.2, a partir de la región conservada (péptido señal) de los péptidos previamente reportados en *A. spurrelli* [40].

Tabla 1. Cebadores/Primers Forward y Reverse empleados para amplificar por RACE PCR.

N°	Primer Forward	Secuencia	Temperatura de annealing	Referencia
1	AC-S	5'-AGCAGCAAAAGAAGAAGCCATG-3'	60 °C	[18]
2	Aspur_F3	5'-TCCCAACATGGCTTTCCTGA-3'	51, 53, 55 y 57 °C	[18]
3	Aspur_F1	5'-GACCAACATGGCTTTCCTGA-3'	57 °C	[18]
4	Ccal1	5'-TAGACCAAACATGGCTTTCCTGA-3'	55 °C	[41]
5	Der-S1	5'-GACCTTCAGTACCCAGCACTTTC-3'	53, 55, 57 y 59 °C	[42]
6	SII	5'-GAGAAGAAGCCATGAAGACTCTGA-3'	59 °C	[18]
7	SIC	5'-TGGTCTCCCTTTCCHTCTGTG-3'	56 °C	Diseño propio
8	Brevinin-2G	5'-CCMRWCATGKCTTTCCHTDAAGAAATCT-3'	55 °C	[43]
9	PSA-S	5'-CAGCACTTCTGAATTACAAGACCAA-3'	56, 58, 60 y 62 °C	[41]
10	DERMO-PDS	5'-GACCAAAGATGTCWTTCTTGAAGAAAT-3'	55, 57, 59 y 61 °C	[18]
N°	Primer Reverse	Secuencia		
1	NUP	5'-AAGCAGTGGTATCAACGCAGAGT-3'	60 °C	[41]

H=A, C o T, M=A o C, R=A o G, W=A o T, K=G o T.

Realizado por: Jessica Verdezoto Prado, 2024

Posteriormente, se amplificó en el termociclador empleando el siguiente perfil térmico: denaturación inicial a 94 °C por 60 s, seguido por 40 ciclos con denaturación a 94 °C por 30 s, hibridación con un gradiente de temperatura entre 51-60 °C dependientes de la temperatura de annealing de los primers empleados (Tabla 1), por 30 s, extensión a 72 °C por 3 min y luego una extensión final a 72 °C por 10 min.

2.1.5. Visualización de productos de PCR por electroforesis

Se comprobó la presencia de secuencias amplificadas realizando electroforesis en gel de

agarosa al 2% en TBE 1X. Fue posible cargar directamente 5 µl de producto PCR en cada pocillo debido a que en la amplificación se empleó OneTaq® Hot Start 2X con buffer completo [44]. También se cargó el respectivo control negativo (C-) y el marcador de peso molecular 1000 bp DNA ladder (Promega, USA). La corrida electroforética se realizó con las siguientes especificaciones: 70 V por 90 minutos. Una vez transcurrida la corrida se tiñó el gel con una solución de 5 µl de Diamond™ Nucleic Acid Dye (Promega, USA) en 50 ml de TBE 1X por 30 minutos y finalmente se visualizó por medio del transiluminador Enduro GDS touch 1302 (Labnet, USA) con luz UV.

2.1.6. Purificación de productos de las RACE PCR

La purificación de productos amplificados se llevó a cabo siguiendo el protocolo establecido de PureLink™ PCR Purification Kit (Invitrogen™, USA). El kit elimina dímeros de dNTPs, enzimas, sales y restos de primers que no se acoplaron durante la reacción de PCR [45]. Este método de purificación se empleó para los productos PCR de los 10 cebadores utilizados por separado, de los cuales se escogieron los productos que contenían secuencias entre 300 y 500 pb y se determinó el volumen de los mismos. En resumen, en 1 volumen de productos PCR (50-100 µL) se agregaron 4 volúmenes de buffer de unión y se homogeneizó. Seguidamente, en una columna de centrifugación se colocó la mezcla con buffer de unión y se llevó a centrifugar a 10000 g por 1 minuto a temperatura ambiente. A continuación, se desechó el excedente y se lavó el ADN con 650 µL de buffer de lavado, se centrifugó a 10000 g durante 1 min y descartó el sobrante. Se volvió a centrifugar a 14000 g por 3 min y se descartó el residuo. En un tubo nuevo se colocó la columna y se agregaron 50 µL del buffer de elución, dejando incubar a temperatura ambiente por 1 minuto. Se centrifugó la columna a 14000 g por 2 min, se descartó la columna y se recuperó el líquido del fondo que contenía el producto de PCR purificado [45].

2.1.7. Ligación de productos PCR

Luego de la purificación de productos PCR se creó un pool de amplicones tomando

alícuotas de cada muestra de primer, con el propósito de tener una muestra diversa y con concentración de 100-200 fentomoles, requerida por el kit de secuenciamiento. A continuación, se empleó el kit Ligation Sequencing Kit V14 (SQK-LSK114) (Oxford Nanopore Technology, UK) para reparar extremos y se prepararon 30 µl de una solución detallada a continuación: 0.5 µl de ADN muestra control del kit (DCS), 24.5 µl de ADN, 3.5 µl de tampón de reacción de preparación final Ultra II y 1.5 µl de mezcla de enzimas de preparación final Ultra II. Posteriormente, la mezcla se incubó a 20 °C por 5 min y luego a 65 °C por 5 min en el termociclador, una vez terminado este paso la muestra se transfirió a un tubo 1.5 ml. Adicionalmente, se re-suspendieron las perlas AMPure XP (AXP) y se transfirieron 24 µl al tubo que contenía la mezcla anterior. Se homogeneizó por pipeteo suave y se incubó por 5 min a temperatura ambiente. Apoyados en una barra magnética, se lavaron las perlas dos veces con 200 µl de etanol fresco al 80%, que fue descartado y secado. El pellet fue suspendido en 31 µl de agua libre de nucleasas y nuevamente con la barra magnética se separaron las perlas y el sobrenadante que contenía el ADN reparado. Posteriormente se cuantificó mediante Qubit empleando el kit Qubit® dsDNA HS Assay (Thermo Fisher Scientific, USA) [46].

Seguidamente, se ligaron los adaptadores con 50 µl de la siguiente mezcla: 30 µl de ADN muestra, 12.5 µl de Tampón de ligadura (LNB), 5 µl de NEBNext Quick T4 ADN ligasa y 2.5 µl de Adaptador de ligadura (LA). Se mezcló e incubó a temperatura ambiente por al menos 10 min. Seguidamente, se separó la muestra usando 40 µl de perlas AMPure XP (AXP), una barra magnética y dos lavados de 200 µl de tampón de fragmentos cortos del kit (SFB). Se descartó el sobrenadante y conservó el pellet que fue re-suspendido en 15 µl de tampón de elución del kit (EB) e incubado a 37 °C por 10min. Nuevamente apoyados en la barra magnética, se separaron las perlas, recuperando el sobrenadante que contenía las librerías del ADN y se cuantificó por Qubit. Finalmente, se diluyó en ADN hasta una concentración final inferior a 30 fmol, tal como lo recomienda el protocolo del kit Ligation Sequencing Kit V14 (SQK-LSK114) (Oxford Nanopore Technology, UK) [34].

2.1.8. Secuenciación por nanoporos del ADNc

La secuenciación por nanoporos se llevó a cabo usando el dispositivo MinION Mk1B (Oxford Nanopore Technology, UK) acoplado con la Celda de Flujo Flongle FLO-FLG114 (Oxford Nanopore Technology, UK) y en conjunto con el kit de secuenciación por ligación Ligation Sequencing Kit V14 (SQK-LSK114) (Oxford Nanopore Technology, UK). Previo a la carga de la muestra se comprobó la disponibilidad de poros en la celda (55 poros) y se preparó 120.5 µl de la mezcla de cebador con: 117 µl de Lavado de celda de flujo (FCF), 0.5 µl de Albúmina sérica bovina a 50 mg/ml (BSA) y 3 µl de Anclaje de celda de flujo (FCT) [34].

Se cebó la celda introduciendo por el puerto de carga primero 80 µl de la mezcla previamente preparada, evitando formar burbujas en el interior. Seguidamente, se preparó 37.5 µl de mezcla de carga de muestra descrita a continuación: 15 µl de Buffer de secuenciación (SB), 12.75 µl de Perlas de librería previamente re-suspendidas (LIB) y 6 µl de Librería de ADN.

Enseguida, se completó el cebado al introducir 20 µl restantes de la mezcla por el puerto donde se introdujo la primera carga. A continuación, por el puerto SpotON se cargaron los 37.5 µl de mezcla de librería suavemente, asegurándose de que el líquido fluyera. Una vez completada la carga se cerraron los puertos de carga y SpotON, se colocó el protector de luz de la celda y se configuró el dispositivo para iniciar la secuenciación.

2.1.9. Adquisición de datos y basecalling

Previamente en el ordenador se instaló el protocolo MinKNOW v 23.11.2 como recomienda el proveedor (Oxford Nanopore Technology, UK). Dentro del programa se inició una nueva secuenciación y se completaron los datos correspondientes a: código de la celda de flujo (FLO-FLG114), identificación de la celda (ASJ157), dispositivo (MinION), guppy_version (7.2.11+0a03b4b93), identificación del protocolo (Aspurrelli), identificación de la muestra (ASRACE), kit de secuenciación (SQK-LSK114), tipo de

basecalling (super-precisa, en ingles *super-accurate*, 400pb), ubicación de salida de los datos (E:\Aspurrelli/ASRACE/20231218), tipo de reporte (Dorado v 7.2.11), software de reporte (MinKNOW Core v 5.8.0) [47].

2.2. Diseño de pipeline para limpieza y tratamiento

2.2.1. Análisis de datos obtenidos

Una vez completada la secuenciación se creó automáticamente una carpeta con los datos de la secuenciación: secuencias pasadas, secuencias erróneas, reportes de secuenciación, alineamiento de secuencias, actividad de los poros y un resumen total. Para la descarga se tomaron en consideración los reportes de secuenciación y la carpeta de secuencias pasadas. En un ordenador con sistema operativo Ubuntu de Linux se creó un script de análisis bioinformático en la terminal, siguiendo el flujo de trabajo detallado en la Figura 1. En primer lugar, se instaló y comprobó la versión de programas y librerías necesarias para el análisis de secuenciación por nanoporos como: NanoPlot, Nanofilt, Pychopper, Porechop, Python3, Amplicon_sorter y Orfipy.

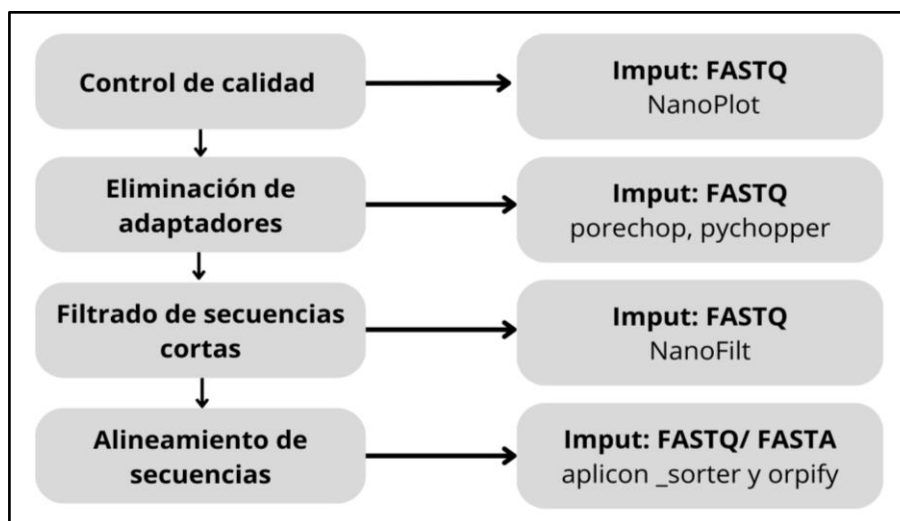


Figura 1. Diagrama de flujo de trabajo para el análisis de datos de secuenciación de *A. spurrelli*

Realizado por: Jessica Verdezoto Prado, 2024.

Se abrió la terminal y se ubicó en el directorio donde se encontraba la carpeta de secuencias pasadas, usando el comando *gunzip* y *cat* se descomprimió la carpeta y sus archivos fueron concatenados en un nuevo archivo. Seguidamente, utilizando NanoPlot v. 1.42.0 se efectuó un análisis estadístico para obtener la calidad y la longitud media de las secuencias reflejadas en histogramas. A continuación, se empleó porechop v. 0.2.4 para eliminar adaptadores, mientras que pychopper v. 1.0.0 para cortar primers y orientar las secuencias (poli A). Adicionalmente, se contabilizaron las lecturas usando la librería Awk v. 1.2.1 y se midió la calidad de secuencias con el comando NanoPlot.

Se realizó un filtrado con NanoFilt v. 2.8.0, eliminando las secuencias con longitud inferior a 200 pb y calidad menor a 12 ($>Q12$)[48]. Seguidamente, se clasificaron los amplicones teniendo en cuenta la similitud y longitud de las secuencias, utilizando amplicon_sorter.py para generar secuencias consenso [49]. Para obtener secuencias traducidas del ADNc que codifica proteínas y péptidos precursores se leyeron las secuencias con marcos de lectura abiertos (ORFs) empleando el comando orfipy v. 0.0.4[50].

2.2.2. Análisis de secuencias putativas

Las secuencias traducidas obtenidas se revisaron utilizando el programa Molecular Evolutionary Genetics Analysis, MEGA X (v 10). Donde se agruparon y compararon todas las secuencias por similitud, separándolas en dos archivos: péptidos y proteínas. Primero se analizaron las secuencias peptídicas con la herramienta BLAST Basic Local Alignment Search Tool del NCBI v 2.15.0, en busca de afinidad con otros péptidos antes descritos [51,52]. Para ello se empleó Protein BLAST (blastp) y en la caja de texto se insertó la secuencia aminoacídica completa, desde el codón de inicio hasta el codón STOP, pero sin el asterisco (*). A continuación, se modificó el parámetro de organismo, seleccionando amphibians (taxid:8292) mientras que los demás parámetros quedaron en las configuraciones por defecto (Standard databases, blastp (protein-protein BLAST))[53,54].

Por otro lado, también se utilizó BLAST con las mismas especificaciones para las secuencias proteínicas encontradas, con el fin de cotejarlas con proteínas ubicadas en la base de datos del National Center for Biotechnology Information (NCBI). Al final, si estas secuencias mostraban similitud con las familias de péptidos antimicrobianos, se las agrupó con ellas y se les asignó un nombre similar o nuevo correspondiente a la especie del anfibio, siguiendo la nomenclatura propuesta por Amiche y colaboradores, 2008 [6].

2.2.3. Análisis de espectrometría de masas

Se creó una base de datos con, un set de datos obtenidos mediante clonaje molecular y espectrometría de masas en tándem (espectrómetro de masas con trampa de iones LCQ Fleet ESI (Thermo Fisher San Jose, CA, EE.UU.) aplicados a la secreción cutánea de *A. spurrelli*, de acuerdo a la metodología propuesta por Proaño y colaboradores en 2019 [18]. Adicionalmente, se añadió otro set de datos obtenidos mediante clonaje molecular en el estudio de Proaño y colaboradores de los laboratorios de la Pontificia Universidad Católica del Ecuador y la Universidad Regional Amazónica Ikiam (datos no publicados). También, se agregó el set de datos obtenidos mediante secuenciación de ARN dirigido aplicada a la secreción cutánea de *A. spurrelli*, ampliando la base de datos.

Una vez actualizada la base de datos se empleó el software Mascot Distiller 2.8.5, que permitió la comparación de los espectros de las moléculas (péptidos fragmentados reales) presentes en la secreción cutánea, frente a las moléculas teóricas establecidas en la base de datos. Este análisis permite distinguir moléculas identificadas por clonaje molecular y secuenciación de ARN dirigido presentes en la muestra (secreción) mediante proteómica, aportando veracidad y robustés al estudio.

2.3. Análisis de propiedades fisicoquímicas

2.3.1. Caracterización fisicoquímica

Una vez establecida la base de datos de los potenciales nuevos péptidos, se analizaron las propiedades fisicoquímicas de los mismos empleando distintas herramientas bioinformáticas. Primero se calculó la masa teórica de cada péptido empleando el software de libre acceso PeptideSynthetics (Peptide Protein Research Ltd., Fareham, UK) donde en la caja de texto se colocó la secuencia correspondiente al péptido maduro y se corrió el programa. Luego, se predijo la estructura secundaria de los péptidos usando SOPMA (Self-Optimized Prediction Method with Alignment) de igual manera en la caja de texto se colocó la secuencia del péptido maduro, se seleccionó el tipo de conformación hélice alfa (α) y se corrió el programa. Finalmente, se calculó la carga neta y el punto isoeléctrico empleando Bachem Peptide Calculator, se seleccionó el formato de entrada de texto (aminoácido AA) seguidamente se colocó en la caja de texto la secuencia del péptido maduro y se corrió el programa [55–57].

2.3.2. Predicción y modelamiento de estructuras 3D

Se realizó el modelo de rueda helicoidal de los péptidos potencialmente antimicrobianos empleando el software heliQuest, se ingresó en la sección Analysis. Se modificaron los parámetros de selección: tipo de hélice se seleccionó alfa (α), window size (full), en la caja de texto se ingresó la secuencia del péptido maduro y se procedió con la predicción [58]. El modelamiento de la estructura tridimensional se obtuvo mediante la herramienta I-TASSER, para ello inicialmente se creó una cuenta en el software. En la caja de texto se colocó la secuencia peptídica, se llenó los campos de email y contraseña, se colocó el código de identificación del péptido finalmente se corrió (Run I-TASSER) el programa. Para la predicción de la estructura secundaria de las nuevas proteínas de interés se usó el software de libre acceso SWISS-MODEL. Se creó un proyecto denominado Proteins y en la caja de texto se insertó la secuencia aminoacídica correspondiente a la zona de la proteína madura. A continuación, se seleccionó construir

modelos y se eligió el mejor resultado respecto al mejor porcentaje de identidad y estado “monómero” [59,60].

3. CAPÍTULO III: PRESENTACIÓN DE DATOS Y RESULTADOS

3.1. Preparación y secuenciación de bibliotecas de ADNc

3.1.1. Tratamiento de la muestra

Tras la extracción de secreción se obtuvieron dos tubos con 5 mg de secreción cutánea liofilizada de los que se extrajo el ARNm teniendo igualmente dos tubos con 8 μ l y 18 μ l de volumen, concentración de 66,4 ng/ μ l, calidad (Abs260/280=1,92; Abs260/230= 0,84) y concentración de 16,2 ng/ μ l, calidad (Abs260/280=1,36; Abs260/230= 1,07) respectivamente. La construcción de ADNc dio como resultado 6 réplicas de 20 μ l cada una y concentración entre 2,14 ng/ μ l y 2,58 ng/ μ l. En la amplificación por RACE PCR se obtuvieron 32 tubos con 20 μ l. Tras la purificación se creó un pool de muestra de 37 μ l y concentración 36,6 ng/ μ l (197,39 fmol). El producto de la ligación tenía un volumen de 15 μ l con concentración de 16,5 ng/ μ l (89 fmol), de los cuales se creó una alícuota de 6 μ l y concentración final 5.54 ng/ μ l (29.88 fmol).

3.1.2. Reporte de la calidad de datos de secuenciación

Una vez finalizado el basecalling se adquirió el reporte total de la secuenciación donde se observó datos sobre el número de lecturas crudas, longitud de las lecturas, número de PB, longitud N50. En base a esta información se creó la Tabla 2 que resume el contraste de los datos crudos y analizados. Adicionalmente la Figura 2, presenta el flujo de trabajo y la depuración de los datos conforme se realizó cada paso de limpieza.

Tabla 2. Resumen de la secuenciación y análisis bioinformático.

	Data cruda	Data analizada
Total de lecturas	271941,0 reads	163,0 reads
Longitud media de lectura	422,7 pb	66,4 aa
Longitud mediana de lectura	399,0 pb	67,0 aa
Calidad media de lectura	11,4 Q	-
Longitud N50	408,0 pb	69,0 aa
Bases totales	114961147,0 pb	10033,0 pb

- = no definido. **Realizado por:** Jessica Verdezoto Prado, 2024

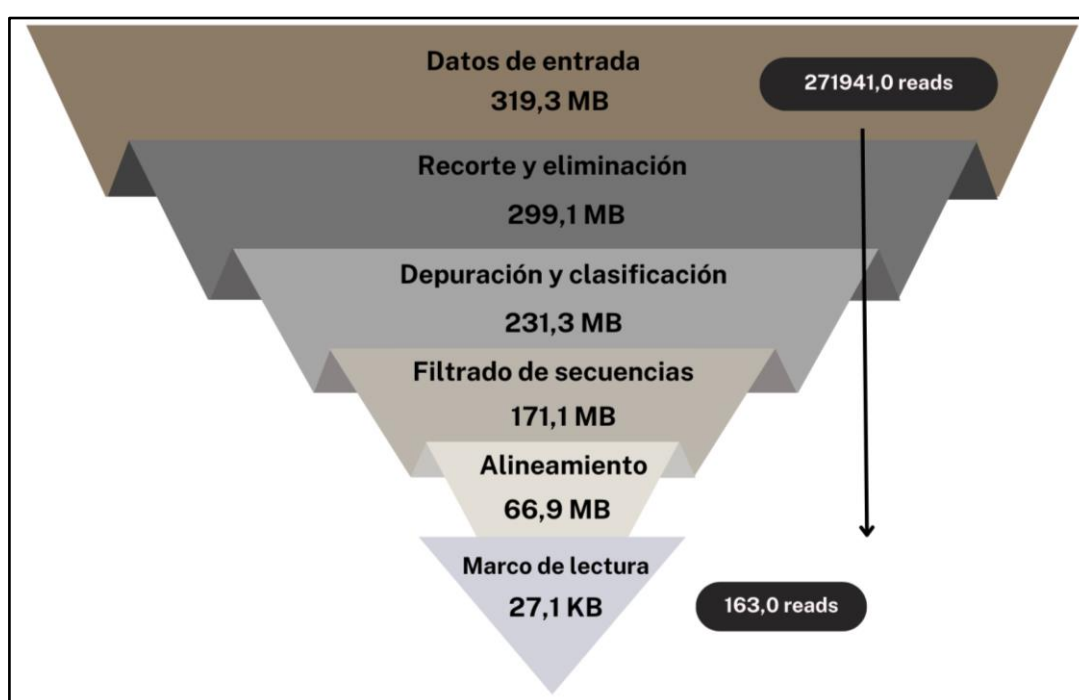


Figura 2. Depuración de datos de la secuenciación por nanoporos. El tamaño de los datos disminuye conforme cada paso de filtrado y limpieza.

Realizado por: Jessica Verdezoto Prado, 2024

Por otro lado, para iniciar el análisis bioinformático se realizó un primer histograma (plot) para evaluar la calidad de la secuenciación (Figura 3), aquí podemos observar la calidad respecto al tamaño de lecturas obtenidas. Se evidencia la distribución de las lecturas y su calidad que oscila entre Q9 y Q22. Sin embargo, el rango de calidad media de la secuenciación se encuentra entre Q11 y Q12.

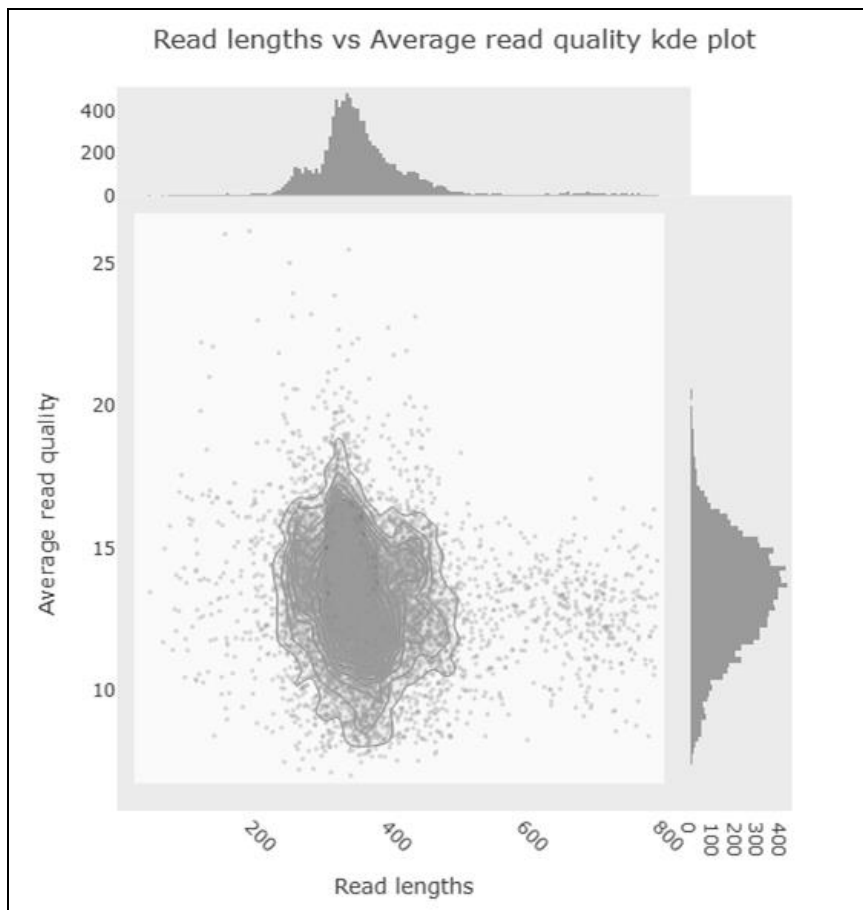


Figura 3. Control de calidad de la secuenciación por nanoporos de *A. spurrelli*. Eje “x” tamaño de lecturas (read lengths), eje “y” calidad promedio de las lecturas (Q score).

Realizado por: Jessica Verdezoto Prado, 2024

3.2. Diseño de pipeline para limpieza y tratamiento

3.2.1. Secuencias precursoras obtenidas

Inicialmente los comandos *gunzip* y *cat* concatenaron la información de las secuencias en un archivo denominado *Asp272.fastq* (319,3 MB). Al usar *porechop* se obtuvo el archivo *AspPore.fastq* (299,1 MB). Con *pychopper* se consiguió el archivo *UncAsp.fastq* (231,3 MB). Empleando *NanoFilt* se obtuvo el archivo *Unc2Fil.fastq* (171,1 MB). Tras aplicar *amplicon_sorter* se consiguió la carpeta *AmpUnc* donde reposaba el archivo *Unc2Fil_consens.fasta* (66,9MB). Utilizando *orfipy* se obtuvo el archivo *AspOrFin.fa* (27,1 KB) y finalmente en *MEGA X* se separaron las secuencias en dos archivos

AspOrFinPeptide.fa y AspOrFinProtein.fa. Esta información se resume en la Figura 4., diagrama de flujo de trabajo (pipeline) con el detalle de las herramientas bioinformáticas empleadas para la limpieza y tratamiento de los datos de secuenciación por nanoporos.

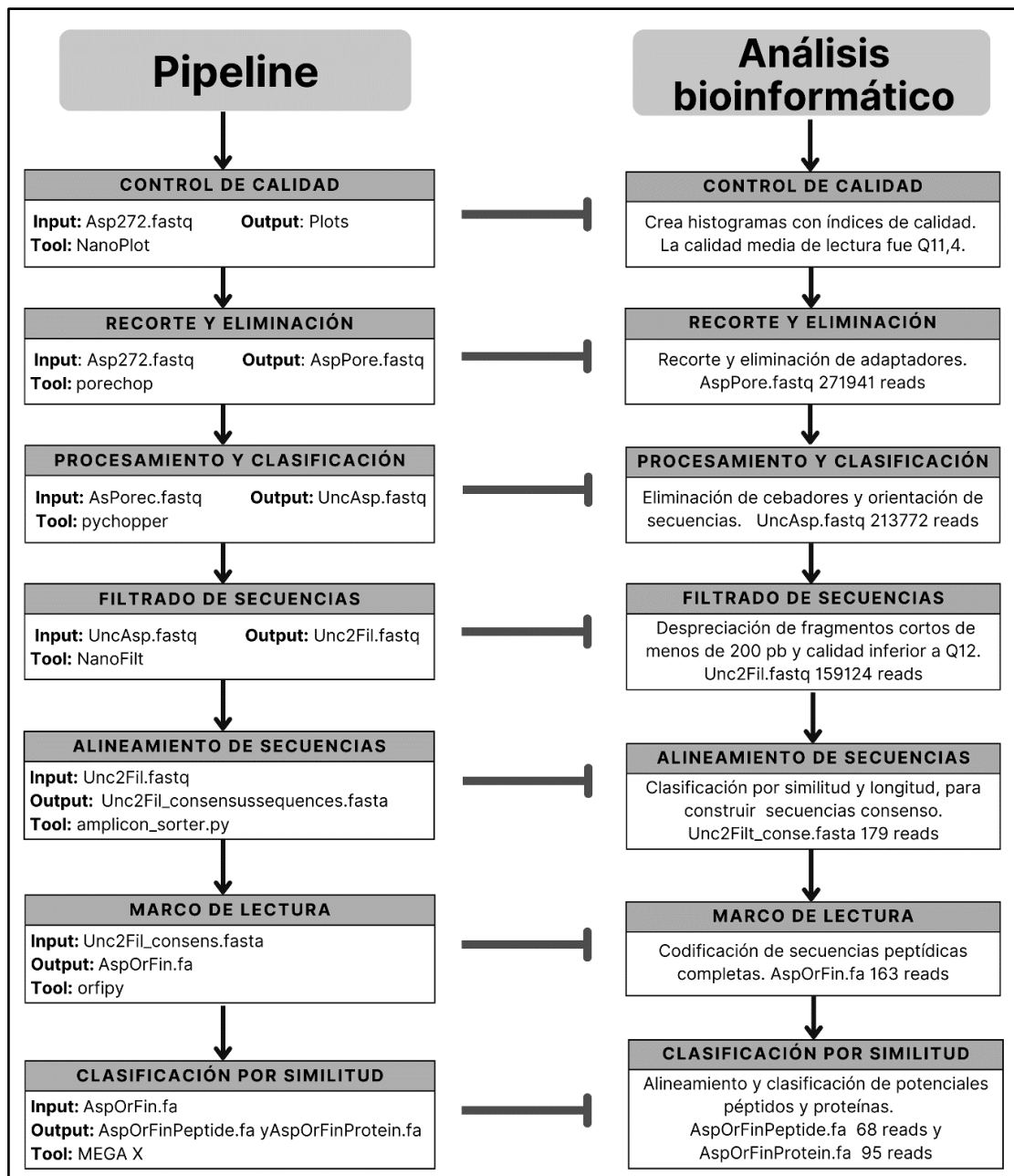


Figura 4. Diseño del flujo de trabajo (pipeline) para la limpieza y tratamiento de datos de la secuenciación por nanoporos.

Realizado por: Jessica Verdezoto Prado, 2024

Una vez realizada la limpieza y depuración de datos de secuenciación por nanoporos de la secreción de *Agalychnis spurrelli*, se obtuvieron ciento sesenta y tres secuencias precursoras. De las cuales tras aplicar BLAST y compararlas en la base de datos se obtuvieron sesenta y ocho secuencias precursoras de potenciales péptidos, veinte y cinco secuencias precursoras de potenciales proteínas y setenta secuencias precursoras que no fueron identificadas con moléculas de interés. Las secuencias precursoras de potenciales péptidos presentaron copias de la misma secuencia por lo que se redujeron a 39 secuencias, que se presentan en la Tabla 3. con sus contigs, número de copias y acompañadas del resultado del análisis Blast/p realizado a cada una en la base de datos del NCBI. Se puede observar que la familia de péptidos antimicrobianos (PAMs) con mayor registro de secuencias corresponde a las Dermaseptinas, seguida por las Tryptofilinas, Filoseptinas, Filoquininas, Plasticinas, finalmente las Medusinas y Caerinas con un único registro.

Tabla 3. Lista de potenciales péptidos encontrados por secuenciación de ARN dirigido a la secreción de *A. spurrelli*.

Identificación del transcrito	Contigs	Copias	Blast/p		Familia
			Cobertura	Identidad	
Cons_Unc2Fil 0 31	14892	5	100%	100.00%	Dermaseptinas
Cons_Unc2Fil 0 0	6329	1	98%	100.00%	Dermaseptinas
Cons_Unc2Fil 0 38	93	1	85%	98.28%	Dermaseptinas
Cons_Unc2Fil 0 75	1010	1	98%	97.37%	Dermaseptinas
Cons_Unc2Fil 0 19	890	1	98%	96.05%	Dermaseptinas
Cons_Unc2Fil 0 6	3625	2	98%	95.83%	Dermaseptinas
Cons_Unc2Fil 0 3	3007	3	94%	95.71%	Dermaseptinas
Cons_Unc2Fil 0 66	152	1	98%	93.67%	Dermaseptinas
Cons_FilAsUn 0 53	228	1	67%	90.57%	Dermaseptinas
Cons_Unc2Fil 0 5	1822	1	98%	89.98%	Dermaseptinas

Cons_Unc2Fil 0 28	352	1	100%	85.53%	Dermaseptinas
Cons_Unc2Fil 0 1	4664	1	52%	85.42%	Dermaseptinas
Cons_FilAspUnc 0 152	48	2	58%	78.95%	Dermaseptinas
Cons_Unc2Fil 0 12	599	1	68%	78.18%	Dermaseptinas
Cons_Unc2Fil 0 45	122	1	96%	91.94%	Triptofillinas
Cons_Unc2Fil 0 72	260	3	98%	87.10%	Triptofillinas
Cons_Unc2Fil 0 39	40	2	98%	86.67%	Triptofillinas
Cons_Unc2Fil 0 51	66	2	72%	81.82%	Triptofillinas
Cons_Unc2Fil 0 37	51	1	96%	74.19%	Triptofillinas
Cons_Unc2Fil 0 68	42	1	-	-	Triptofillinas
Cons_Unc2Fil 1 1	3722	3	98%	98.48%	Filoseptinas
Cons_Unc2Fil 1 7	1287	5	95%	98.46%	Filoseptinas
Cons_Unc2Fil 1 18	478	1	94%	97.06%	Filoseptinas
Cons_Unc2Fil 1 38	316	2	98%	94.12%	Filoseptinas
Cons_Unc2Fil 0 2	1524	1	100%	100.00%	Filoquininas
Cons_Unc2Fil 0 9	2426	1	98%	98.39%	Filoquininas
Cons_Unc2Fil 0 16	629	1	98%	91.94%	Filoquininas
Cons_Unc2Fil 0 8	686	4	98%	80.00%	Plasticinas
Cons_Unc2Fil 0 50	138	2	85%	78.26%	Plasticinas
Cons_Unc2Fil 2 0	1331	5	98%	98.53%	Medusinas
Cons_Unc2Fil 0 60	1718	2	98%	90.28%	Caerinas
Cons_Unc2Fil 1 2	3471	2	98%	100.00%	No definido
Cons_Unc2Fil 0 59	234	2	98%	100.00%	No definido

Cons_FilAsUn 0 25	3733	1	50%	97.30%	No definido
Cons_FilAspUnc 0 144	50	1	76%	88.24%	No definido
Cons_Unc2Fil 0 22	237	1	76%	87.50%	No definido
Cons_Unc2Fil 0 70	254	2	43%	83.33%	No definido
Cons_Unc2Fil 0 14	271	3	58%	80.56%	No definido
Cons_Unc2Fil 0 13	462	1	73%	68.89%	No definido

Copias= número de repeticiones que tuvo la misma secuencia en el archivo final.

Realizado por: Jessica Verdezoto Prado, 2024

Por otro lado, las principales secuencias precursoras de proteínas junto con sus contigs y análisis BLAST se agruparon en la Tabla 4. Se destacan 17 secuencias precursoras con alta similitud a proteínas de tipo Kazal descritas anteriormente en *Cruziophyla calcarifer* y *Pithecopus nordestinus*.

Tabla 4. Lista de potenciales proteínas de interés encontradas por por secuenciación de ARN dirigido a la secreción de *A. spurrelli*

Identificación del transcrito	Contigs	Copias	Blast/p		Familia
			Cobertura	Identidad	
Cons_Unc2Fil 1 56	130	2	76%	90.32%	Kazal
Cons_Unc2Fil 1 58	99	1	86%	89.77%	Kazal
Cons_Unc2Fil 1 12	3531	2	97%	89.16%	Kazal
Cons_Unc2Fil 1 92	132	1	97%	88.04%	Kazal
Cons_Unc2Fil 1 6	1983	1	98%	87.10%	Kazal
Cons_Unc2Fil 1 23	184	1	90%	85.19%	Kazal
Cons_Unc2Fil 1 4	1220	1	96%	78.67%	Kazal
Cons_Unc2Fil 1 30	508	1	95%	72.83%	Kazal

Cons_Unc2Fil 1 85	181	1	93%	69.23%	Kazal
Cons_Unc2Fil 1 9	1612	1	98%	68.67%	Kazal
Cons_Unc2Fil 1 8	1861	3	98%	67.86%	Kazal
Cons_Unc2Fil 1 20	544	2	98%	67.09%	Kazal
Cons_Unc2Fil 1 13	1438	1	98%	66.67%	Kazal
Cons_Unc2Fil 1 5	2202	1	76%	66.13%	Kazal
Cons_Unc2Fil 1 89	503	4	77%	65.08%	Kazal
Cons_Unc2Fil 1 40	315	1	98%	64.56%	Kazal
Cons_Unc2Fil 1 50	355	1	76%	63.49%	Kazal

Copias= número de repeticiones que tuvo la misma secuencia en el archivo final.

Realizado por: Jessica Verdezoto Prado, 2024.

3.2.2. Clasificación de PAMs encontrados

Las secuencias peptídicas se clasificaron en siete familias de péptidos con actividad antimicrobiana: dermaseptinas (35,90%), triptofilinas (15,38%), filoseptinas (10,26%), filoquininas (7,69%), plasticinas (5,13%), medusinas (2,56%) y caerinas (2,56%). Adicionalmente, las secuencias que no encajaron en las familias fueron agrupadas y denominadas como péptidos huérfanos (20,51%).

Dermaseptinas

Una vez formadas las secuencias consenso, 14 secuencias precursoras de dermaseptinas fueron identificadas, las cuales fueron nombradas: Dermaseptin-SP2, Dermaseptin-SP3, Dermaseptin-SP7, Dermaseptin-SP10, Dermaseptin-SP11, Dermaseptin-SP12, Dermaseptin-SP13, Dermaseptin-SP14, Dermaseptin-SP15, Dermaseptin-SP16, Dermaseptin-SP17, Dermaseptin-SP18, Dermaseptin-SP19, Dermaseptin-SP20. Dos de estas (DRS-SP2, DRS-SP3) ya descritas anteriormente en *A. spurrelli* [18].

Las secuencias precursoras codificaban un marco de lectura abierto (ORF) compuesto por: (1) un péptido-señal putativo de 22 aminoácidos; (2) un espaciador ácido de 23 aminoácidos que contienen dos sitios de procesamiento de propéptidos (KR). Y (3) la región del péptido maduro que variaba entre 24 a 53 aminoácidos, que se puede apreciar en la Tabla 5. Adicionalmente, algunas dermaseptinas contienen (4) una región de residuo de glicina donante de amida en el C-terminal.

Tabla 5. Secuencias aminoacídicas traducidas del marco de lectura abierto que codifican a péptidos de la familia de Dermaseptinas.

[illegible]

- = espacios añadidos para completar la alineación, *= codón de terminación (STOP). La región del péptido señal se encuentra subrayada (1-22 aa). En la zona ácida se visualiza los espaciadores ácidos marcados con negrita (23-45 aa). El péptido maduro se muestra resaltado en color gris (46-en adelante, aa) y la amidación se encuentra resaltada con color gris oscuro.

Realizado por: Jessica Verdezoto Prado, 2024.

Al comparar las secuencias aminoacídicas en la base de datos del NCBI se obtuvo la Tabla 6, donde se muestra que la dermaseptin-SP11 y dermaseptin-SP7 compartieron entre 97.37% y 96.05% de identidad con la secuencia aminoacídica de la dermaseptin-SP2 de *Agalychnis spurrelli* con una cobertura del 98% (número de acceso: A0A5P9K470.1). Mientras que sus péptidos maduros compartieron entre 100% y 93.55% de identidad con la dermaseptin-SP2 y una cobertura entre el 87% y 100% (número de acceso: A0A5P9K470.1).

Las dermaseptin-SP12, dermaseptin-SP13, dermaseptin-SP14 y dermaseptin-SP3 compartieron identidad de 85.42%, 93.67%, 78.95% y 90.57% respectivamente con la dermaseptin-B1 de *Phyllomedusa bicolor* con una cobertura de entre 52% y 98% (número de acceso: AFR78286.1). Al comparar los péptidos maduros se obtuvo que la región del péptido maduro de dermaseptin-SP12 y dermaseptin-SP13 compartieron entre 79.41% y 69.23% de identidad con dermaseptin-SS1 de *Phyllomedusa tarsius* y entre 97% y 100% de cobertura (número de acceso: WNH96388.1).

Para el caso de las dermaseptin-SP15, dermaseptin-SP16, dermaseptin-SP17 y dermaseptin-SP18 presentaron identidad de 78.18%, 89.98%, 90.50% y 98.28% con la dermaseptin-H3 de *Pithecopus hypochondrialis* y cobertura entre 68% y 94% (número de acceso: P84596.1). En cambio, sus péptidos maduros mostraron entre 77.78% y 92.86% de identidad con la dermaseptin-1 de *Pithecopus nordestinus* y 100% de cobertura (número de acceso: J7H8J4.1).

La dermaseptin-SP19 compartió identidad del 98.28% con la dermaseptin-H4 de *Pithecopus azureus* y cobertura del 98% (número de acceso: Q1EJP5.1). Mientras, que su péptido maduro mostró entre 92.86% de identidad con la dermaseptin-H4 de *Pithecopus azureus* y cobertura del 100% (número de acceso: Q1EJP5.1).

Para el caso de la dermaseptin-SP20, presentó 85.53% de identidad con la dermaseptin-H1 de *Pithecopus hypochondrialis* y 100% de cobertura (número de acceso: P81486.1).

Mientras que su péptido maduro obtuvo 88.98% de identidad con la dermaseptin-1 de *Phyllomedusa tomopterna* y 84% de cobertura (número de acceso: P85523.1).

La dermaseptin-SP10, presentó 95.83% de identidad con la dermaseptin-A4 de *Agalychnis annae* y 98% de cobertura (número de acceso: O93224.1). Y su péptido maduro obtuvo 93.10% de identidad con la dermaseptin-A4 de *Agalychnis annae* y 100% de cobertura (número de acceso: O93224.1).

Tabla 6. Alineamiento múltiple de las secuencias aminoacídicas correspondientes a la familia de Dermaseptinas, obtenidas por secuenciación de ARN dirigido a la secreción cutánea de *A. spurrelli*.

Nº de acceso	Péptido	Péptido señal	Espaciador ácido
		1	23. 45.
	Dermaseptin-SP7 .	M A F L K K S L F L V L F L G L V S L S I C E E E K R E N E D E E E Q E D D E Q S E E K R	
	Dermaseptin-SP11 .		
	Dermaseptin-SP2 .		E
A0A5P9K470.1	Dermaseptin-SP2 .		E
		Péptido maduro	
		46. 70.	
	Dermaseptin-SP7 .	A S W K D F F K N I G K A A G K A V L N S V T D M V N Q G E Q *	
	Dermaseptin-SP11 .	. . G . V . L	
	Dermaseptin-SP2 .	 V . L	
A0A5P9K470.1	Dermaseptin-SP2 .	 V . L	
Nº de acceso	Péptido	Péptido señal	Espaciador ácido
		1	23. 45.
	Dermaseptin-SP12 .	M A F L K K S L F L V L F L G L V S L S I C E E E K R E N E D E E E Q E D D E Q S E E K R	
	Dermaseptin-SP13 .		K M . .
	Dermaseptin-SP14 .		
	Dermaseptin-SP3 .	 M V L R .	
AFR78286.1	Dermaseptin-B1 .		K . - M . .
		Péptido maduro	
			98.
	Dermaseptin-SP12 .	A S W K D F F K K Y R K S S R K S C F K F S Y G Y G K S R R A I K Q K N V K S V I V P A L H *	
	Dermaseptin-SP13 .	. L . . . I L . . I G T V A L H A G K A A F G A A A D T I S Q G E . *	
	Dermaseptin-SP14 .	. A . T S I I . N M A . N A A . K V A . A A A N A A A K A A . N A I A K S V N N M A N Q E E Q E S K K M *	
	Dermaseptin-SP3 .	S L . S S I K D M A A A A G . A A L N A V N G I V N P G E Q *	
AFR78286.1	Dermaseptin-B1 .	. M . . . V L . . I G T V A L H A G K A A L G A V A D T I S Q G E .	

Tabla 6. Continuación de Tabla.[illegible]

Tabla 6. Continuación de Tabla.

N° de acceso	Péptido	Péptido señal	Espaciador ácido
P81486.1	Dermaseptin-SP20 . Dermaseptin-H1 .	1	23. 45.
		<u>M A F L K K S L F L V L F L G L V S L S I C D E - K R E N E D E E E Q E D D E Q S E E K R</u>	
		. D I I E . E	
Péptido maduro			
P81486.1	Dermaseptin-SP20 . Dermaseptin-H1 .	77.	
		G L L K S L L K K V G I A A G K A V L N K V T D M V N Q G E Q *	
		. . W N . . V A . . A	
N° de acceso	Péptido	Péptido señal	Espaciador ácido
O93224.1	Dermaseptin-SP10 . Dermaseptin-A4 .	1	23. 45.
		<u>M A F L K K S L F L V L F L G M V S L S I C E E E K R E E E I E - - Q E D D E Q S E E K R</u>	
		. N . - -	
Péptido maduro			
O93224.1	Dermaseptin-SP10 . Dermaseptin-A4 .	75.	
		G M F T N M L K G I G K L A G Q A A L G A V K T L V G G Q *	
		 A . E .	

. = aminoácidos conservados, - = espacios añadidos para completar la alineación, *= codón de terminación (STOP). La región del péptido señal se encuentra subrayada, la zona ácida contiene espaciadores ácidos marcados con negrita. El péptido maduro esta resaltado en color gris y la amidación con color gris oscuro.

Realizado por: Jessica Verdezoto Prado, 2024.

Triptofilinas

Se identificaron 6 secuencias completas correspondientes a esta familia y fueron nombradas: Tryptophyllin-AST1, Tryptophyllin-AST1-2, Tryptophyllin-AST2-1, Tryptophyllin-AST2-3, TryptophyllinAST3 y Tryptophyllin-AST3-2. Las secuencias precursoras constan de una longitud entre 61 y 69 aminoácidos. La Tabla 7 muestra el marco de lectura abierto (ORF) que consistió en: (1) un péptido señal putativo de 22 aminoácidos; (2) un espaciador ácido de entre 25 y 30 aminoácidos que contienen dos sitios de procesamiento de propéptidos (KR/RR). Y (3) la región del péptido maduro que variaban entre 8 a 16 aminoácidos.

Tabla 7. Secuencias aminoacídicas traducidas del marco de lectura abierto que codifican a péptidos de la familia de triptofilinas.

Péptido	Péptido señal	Espaciador ácido
	1.	23. 35.
Tryptophyllin-AST1 .	<u>M D E L K K S L F L V L F L G L V S I S F C N E E K R Q D D D E G N E</u>	
Tryptophyllin-AST1-2 .	<u>M D E L K K S L F L I L F L G F V S I S F C D E E K R Q E D D E A N E</u>	
Tryptophyllin-AST2-1 .	<u>M D E L K K S L F L V L F L G F I S I S F C D E E K R E D D E Q G N E</u>	
Tryptophyllin-AST2-3 .	<u>M D E L K K S L L L V L F L G L V S I S F C D E E K R Q D D E E E N E</u>	
Tryptophyllin-AST3 .	<u>M A E L K K S L F F V L F L G L V S L S I C D K Q K R E T E K D E D N</u>	
Tryptophyllin-AST3-2 .	<u>M A E L K K S L F F V L F L G L V S L S I C D K Q K R E T E K D E D N</u>	
Péptido	Espaciador ácido	Péptido maduro
	36.	69.
Tryptophyllin-AST1 .	R E E K K E I H E E G N Q E E R R D K P P A W V P G K K * - - -	
Tryptophyllin-AST1-2 .	R E E K K E I H E M G N Q E A R R D K P P S W V P * - - - - - -	
Tryptophyllin-AST2-1 .	R E E K K E I H E E G N Q E E R R D M S P P W F G K K * - - - - -	
Tryptophyllin-AST2-3 .	R E E K K E I Q E E G N P E E T R D M D P P F N R * - - - - - -	
Tryptophyllin-AST3 .	I E E E E E N H E E - - - - - R R Q G K P Y W P P P F L P E K R Y *	
Tryptophyllin-AST3-2 .	I E E E E E N H E E - - - - - R R Q G K P Y W P P P L P E K R Y * -	

- = espacios añadidos para completar la alineación, *= codón de terminación (STOP). La región del péptido señal se encuentra subrayada, la zona ácida contiene espaciadores ácidos marcados con negrita. El péptido maduro esta resaltado en color gris.

Realizado por: Jessica Verdezoto Prado, 2024.

Tras el análisis de las secuencias aminoacídicas en la base de datos del NCBI se obtuvo la Tabla 8., donde se muestra el alineamiento de las tryptophyllin-AST1 y tryptophyllin-AST1-2 compartieron identidad de 91.94% y 86.67% respectivamente con la secuencia aminoacídica de la tryptophyllin-1 de *Agalychnis dacnicolor* con una cobertura del 98% y 96% (número de acceso: P83455.2). Mientras que sus péptidos maduros compartieron

entre 87.50% y 100% de identidad con la tryptophyllin-1 de *Agalychnis dacnicolor* y una cobertura entre el 88% y 83% (número de acceso: P83455.2).

Para el caso de la tryptophyllin-AST2-1 y tryptophyllin-AST2-3 se obtuvo entre 87.10% y 81.82% de identidad con la secuencia aminoacídica de tryptophyllin-2 de *Agalychnis dacnicolor* y entre 98% y 72% de cobertura (número de acceso: CAX48635.1). Su péptido maduro presentó entre 100.00% y 90.00% de identidad con tryptophyllin-2 de *Agalychnis callidryas* y una cobertura entre 90% y 75% (número de acceso: CDI71028.1).

Respecto a las tryptophyllinAST3 y tryptophyllin-AST3-2 compartieron identidad del 74.19% con la tryptophyllin-PhT3 de *Pithecopus hypochondrialis* y cobertura del 96% (número de acceso: AHF20803.1). Mientras, que sus péptidos maduros mostraron baja o nula identidad con péptidos o proteínas de anfibios antes descritos.

Tabla 8. Alineamiento múltiple de las secuencias aminoacídicas correspondientes a la familia de triptofilinas, obtenidas por secuenciación de ARN dirigido a la secreción cutánea de *A. spurrelli*.

N° de acceso	Péptido	Péptido señal	Espaciador ácido
P83455.2	Tryptophyllin-AST1	1. <u>M</u> <u>D</u> <u>E</u> <u>L</u> <u>K</u> <u>K</u> <u>S</u> <u>L</u> <u>E</u> <u>L</u> <u>V</u> <u>L</u> <u>F</u> <u>L</u> <u>G</u> <u>L</u> <u>V</u> <u>S</u> <u>I</u> <u>S</u> <u>E</u> <u>C</u> <u>N</u> <u>E</u> <u>E</u> <u>K</u> <u>R</u> <u>Q</u> <u>D</u> <u>D</u> <u>D</u> <u>E</u> <u>G</u> <u>N</u> <u>E</u>	23.
	Tryptophyllin-AST1-2	 I F D E A . .	
	Tryptophyllin-1	. N F D	
		Espaciador ácido	Péptido maduro
P83455.2	Tryptophyllin-AST1	36. R E E K K E I H E E G N Q E E R R D K P P A W V P G K K * 64.	
	Tryptophyllin-AST1-2	 M A S *	
	Tryptophyllin-1	 Q . D	
N° de acceso	Péptido	Péptido señal	Espaciador ácido
CAX48635.1	Tryptophyllin-AST2-1	1. <u>M</u> <u>D</u> <u>F</u> <u>L</u> <u>K</u> <u>K</u> <u>S</u> <u>L</u> <u>F</u> <u>L</u> <u>V</u> <u>L</u> <u>F</u> <u>L</u> <u>G</u> <u>F</u> <u>I</u> <u>S</u> <u>I</u> <u>S</u> <u>E</u> <u>C</u> <u>D</u> <u>E</u> <u>E</u> <u>K</u> <u>R</u> <u>E</u> <u>D</u> <u>D</u> <u>E</u> <u>Q</u> <u>G</u> <u>N</u> <u>E</u>	23.
	Tryptophyllin-AST2-3	 L L V Q E E . .	
	tryptophyllin-2	 R V D E N H G	
		Espaciador ácido	Péptido maduro
CAX48635.1	Tryptophyllin-AST2-1	36. R E E K K E I H E E G N Q E E R R D M S P P W F G K K * 63.	
	Tryptophyllin-AST2-3	 Q P . . T D . . F N R *	
	tryptophyllin-2	S R H	

N° de acceso	Péptido	Péptido señal	Espaciador ácido
		1.	23.
	Tryptophyllin-AST3 .	<u>M A F L K K S L F F V L F L G L V S L S I C D K Q K R E T E K D E D N</u>	
	Tryptophyllin-AST3-2 .		
AHF20803.1	tryptophyllin PhT-3 .	 I T . L T . . . E . . K A . E . . N K	
		Espaciador ácido	Péptido maduro
		36.	69.
	Tryptophyllin-AST3 .	I E E E E N H E E - - - - R R Q G K P Y W P P P F L P E K R Y *	
	Tryptophyllin-AST3-2 .	 - - - - - L P E K R Y *	
AHF20803.1	tryptophyllin PhT-3 .	N - - - - - D . . F I Y . M . . H	

. = aminoácidos conservados, - = espacios añadidos para completar la alineación, *= codón de terminación (STOP). La región del péptido señal se encuentra subrayada, la zona ácida contiene espaciadores ácidos marcados con negrita. El péptido maduro esta resaltado en color gris.

Realizado por: Jessica Verdezoto Prado, 2024.

Filoseptinas

Se obtuvieron 4 secuencias precursoras de péptidos alineadas a la familia de las filoseptinas y fueron nombrados de la siguiente manera: Phylloseptin-SP1, Phylloseptin-SP4, Phylloseptin-SP5 y Phylloseptin-SP2. De estos, PLS-SP1 y PLS-SP2 ya fueron reportados anteriormente en *A. spurrelli* [18].

Las secuencias precursoras contaron con una longitud entre 67 y 70 aminoácidos. La Tabla 9. muestra el marco de lectura abierto (ORF) que consistió en: (1) un péptido señal putativo de 22 aminoácidos; (2) un espaciador ácido de 24 aminoácidos que contienen dos sitios de procesamiento de propéptidos (KR/R). Y (3) la región del péptido maduro que variaban entre 19 a 22 aminoácidos. Adicionalmente, presentaron (4) un residuo de glicina donante de amida en el C-terminal.

Tabla 9. Secuencias aminoacídicas traducidas del marco de lectura abierto que codifican a péptidos de la familia de filloseptinas.

Péptido	Péptido señal	Espaciador ácido
	1	23. 46.
Phylloseptin-SP1 .	<u>MAFLKKSLELVLELGLVSLSI</u> C E E E E RETKEEENE Q E D D N R - E E K R	
Phylloseptin-SP4 .	<u>MAFLKKSLELVLELGLVSLSI</u> C E E E E RETKEEENE Q E D D N R - E E K R	
Phylloseptin-SP5 .	<u>MAFLKKSLELVLELGLVSLSI</u> C E E E E RETKEEENE Q E D D N R - E E K R	
Phylloseptin-SP2 .	<u>MAFLKKSLELVLELGLVSLSI</u> C E E E K R E T E E K E N E Q E D D D K S E E K R	

Péptido	Péptido maduro
	47. 66. 70.
Phylloseptin-SP1 .	F L S L I P H V I S A I P H V V N A L S N L G *
Phylloseptin-SP4 .	F L S L I P H V I S A I P H V V S A I S S L G *
Phylloseptin-SP5 .	F L S L I P H V I S A I P H V V N A L S G L G *
Phylloseptin-SP2 .	F L S L I P H A I N A V S A L A K H F G * - - -

- = espacios añadidos para completar la alineación, *= codón de terminación (STOP). La región del péptido señal se encuentra subrayada, la zona ácida contiene espaciadores ácidos marcados con negrita. El péptido maduro esta resaltado en color gris y la amidación con color gris oscuro.

Realizado por: Jessica Verdezoto Prado, 2024.

El alineamiento de las secuencias aminoacídicas en la base de datos del NCBI permitió crear la Tabla 10., donde se observa que la phylloseptin-SP1, phylloseptin-SP4 y phylloseptin-SP5 compartieron identidad 100.00%, 94.12% y 97.06% respectivamente con la secuencia aminoacídica de la phylloseptin-SP1 de *Agalychnis spurrelli* con una cobertura entre 94% y 98% (número de acceso: A0A5P9K461.1). Mientras que sus péptidos maduros compartieron entre 91.30% y 95.65% de identidad con la phylloseptin-SP1 de *Agalychnis spurrelli* con una cobertura del 100% (número de acceso: A0A5P9K461.1).

La phylloseptin-SP2 compartió identidad del 98.48% con la phylloseptin-Az2 de *Pithecopus azureus* con una cobertura de 98% (número de acceso: P85882.1). Al comparar su péptido maduro se obtuvo que compartió 95.00% de identidad con phylloseptin-Az2 de *Pithecopus azureus* con una cobertura de 100% (número de acceso: P85882.1).

Tabla 10. Alineamiento múltiple de las secuencias aminoacídicas correspondientes a la familia de filoseptinas, obtenidas por secuenciación de ARN dirigido a la secreción cutánea de *A. spurrelli*.

N° de acceso	Péptido	Péptido señal	Espaciador ácido
		123.46.	
	Phylloseptin-SP1 .	<u>M A F L K K S L F L V L F L G L V S L S I C E E E E R E T K E E E N E Q E D D N R - E E K R</u>	
	Phylloseptin-SP4 .		-
	Phylloseptin-SP5 .		-
A0A5P9K461.1	Phylloseptin-SP1 .	 K	-
		Péptido maduro	
		47.70.	
	Phylloseptin-SP1 .	F L S L I P H V I S A I P H V V N A L S N L G *	
	Phylloseptin-SP4 .	 S . I . S . .	
	Phylloseptin-SP5 .	 G . . .	
A0A5P9K461.1	Phylloseptin-SP1 .	 -	

N° de acceso	Péptido	Péptido señal	Espaciador ácido
		123.	
	Phylloseptin-SP2 .	<u>M A F L K K S L F L V L F L G L V S L S I C E E E K R E T E E K E N E Q E D D D K S E E K R</u>	
P85882.1	Phylloseptin-Az2 .		
		Péptido maduro	
		47.67.	
	Phylloseptin-SP2 .	F L S L I P H A I N A V S A L A K H F G *	
P85882.1	Phylloseptin-Az2 .	 I -	

. = aminoácidos conservados, - = espacios añadidos para completar la alineación, *= codón de terminación (STOP). La región del péptido señal se encuentra subrayada, la zona ácida contiene espaciadores ácidos marcados con negrita. El péptido maduro esta resaltado en color gris y la amidación con color gris oscuro.

Realizado por: Jessica Verdezoto Prado, 2024.

Filoquininas

Se identificaron 3 secuencias completas que codificaban a péptidos, que fueron nombrados: [Ser6, Val10, Asp11]-phyllokinin, Thr6, Val10, Asp11-phyllokinin y [Thr6, Ile10, Tyr11]-phyllokinin. [Ser6, Val10, Asp11]-phyllokinin y [Thr6, Val10, Asp11]-phyllokinin ya fueron reportadas anteriormente en *A. spurrelli* [18].

Las secuencias precursoras tenían una longitud de 63 aminoácidos. En la Tabla 11. se observa el marco de lectura abierto (ORF) que consistió en: (1) un péptido señal putativo de 22 aminoácidos; (2) un espaciador ácido de 25 aminoácidos que contienen dos sitios de procesamiento de propéptidos (KR), una secuencia de cuatro aminoácidos (ESPE/ESPD) en el extremo C-terminal; y (3) la región del péptido maduro de 11 aminoácidos.

Tabla 11. Secuencias aminoacídicas traducidas del marco de lectura abierto que codifican a péptidos de la familia de filoquininas.

Péptido	Péptido señal	
	1	22.
Ser6-Val10-Asp11-phyllokinin .	<u>M</u> <u>S</u> <u>F</u> <u>L</u> <u>K</u> <u>K</u> <u>S</u> <u>L</u> <u>L</u> <u>L</u> <u>V</u> <u>L</u> <u>F</u> <u>L</u> <u>G</u> <u>L</u> <u>V</u> <u>S</u> <u>F</u> <u>S</u> <u>I</u> <u>C</u>	
Thr6, Val10, Asp11-phyllokinin .	<u>M</u> <u>S</u> <u>F</u> <u>L</u> <u>K</u> <u>K</u> <u>S</u> <u>L</u> <u>F</u> <u>L</u> <u>V</u> <u>L</u> <u>F</u> <u>L</u> <u>G</u> <u>L</u> <u>V</u> <u>S</u> <u>F</u> <u>S</u> <u>I</u> <u>C</u>	
Thr6-Ile10-Tyr11-phyllokinin .	<u>M</u> <u>S</u> <u>F</u> <u>L</u> <u>K</u> <u>K</u> <u>S</u> <u>L</u> <u>F</u> <u>L</u> <u>V</u> <u>L</u> <u>F</u> <u>L</u> <u>G</u> <u>L</u> <u>I</u> <u>S</u> <u>F</u> <u>S</u> <u>I</u> <u>C</u>	
	Espaciador ácido	
	23.	47.
Ser6-Val10-Asp11-phyllokinin .	E E E K R E T E E E E N E D E M D E E S E E K K R E S P D	
Thr6, Val10, Asp11-phyllokinin .	E E E K R E T E E E E N E D E M N E E S E E - K R E S P E	
Thr6-Ile10-Tyr11-phyllokinin .	E E E K R E T E E E E N E D E I E E E S E E K K R E S P E	
	Péptido maduro	
	52.	63.
Ser6-Val10-Asp11-phyllokinin .	R P P G F S P F R V D *	
Thr6, Val10, Asp11-phyllokinin .	R P P G F T P F R V D *	
Thr6-Ile10-Tyr11-phyllokinin .	R P P G F T P F R I Y *	

- = espacios añadidos para completar la alineación, *= codón de terminación (STOP). La región del péptido señal se encuentra subrayada, la zona ácida contiene espaciadores ácidos marcados con negrita. El péptido maduro esta resaltado en color gris.

Realizado por: Jessica Verdezoto Prado, 2024.

Al comparar las secuencias aminoacídicas en la base de datos del NCBI se obtuvo la Tabla 12., donde se muestra que la [Ser6, Val10, Asp11]-phyllokinin y [Thr6, Val10, Asp11]-phyllokinin compartieron identidad de 98.39% y 100.00% con el Thr6-bradykinyl-ValAsp de *Agalychnis spurrelli* con una cobertura del 98% y 100% (número de acceso: A0A5Q0MUM5.1). Al comparar su péptido maduro se obtuvo que compartió 98.39% de identidad con Bradykinin de *Agalychnis callidryas* y 100% de cobertura (número de acceso: LOPJV8.1).

Para el caso de la [Thr6, Ile10, Tyr11]-phyllokinin, presentó 91.94% de identidad con el phyllokininogen-B1 de *Phyllomedusa bicolor* con una cobertura del 98% (número de acceso: AFR78288.1). Y su péptido maduro obtuvo 93.33% de identidad con el phyllokininogen-B1 con una cobertura del 98% (número de acceso: AFR78288.1).

Tabla 12. Alineamiento múltiple de las secuencias aminoacídicas correspondientes a la familia de filoquininas, obtenidas por secuenciación de ARN dirigido a la secreción cutánea de *A. spurrelli*.

N° de acceso	Péptido	Péptido señal	
		1	22.
	Ser6-Val10-Asp11-phyllokinin	<u>M S E L K K S L L L V L F L G L V S F S I C</u>	
	Thr6, Val10, Asp11-phyllokinin	 F	
A0A5Q0MUM5.1	Thr6-bradykinyl-ValAsp	 F	
		Espaciador ácido	
		23.	47.
	Ser6-Val10-Asp11-phyllokinin	E E E K R E T E E E E N E D E M D E E S E E K K R E S P D	
	Thr6, Val10, Asp11-phyllokinin	 N E	
A0A5Q0MUM5.1	Thr6-bradykinyl-ValAsp	 K E	
		Péptido maduro	
		63.	
	Ser6-Val10-Asp11-phyllokinin	R P P G F S P F R V D *	
	Thr6, Val10, Asp11-phyllokinin	 T	
A0A5Q0MUM5.1	Thr6-bradykinyl-ValAsp	 T	
		Espaciador ácido	
		23.	47.
	Thr6-Ile10-Tyr11- phyllokinin	E E E K R E T E E E E N E D E I E E E S E E K K R E S P E	
AFR78288.1	phyllokininogen-B1	 Q	
		Péptido maduro	
		63.	
	Thr6-Ile10-Tyr11- phyllokinin	R P P G F T P F R I Y *	
AFR78288.1	phyllokininogen-B1	 S	

. = aminoácidos conservados, - = espacios añadidos para completar la alineación, *= codón de terminación (STOP). La región del péptido señal se encuentra subrayada, la zona ácida contiene espaciadores ácidos marcados con negrita. El péptido maduro esta resaltado en color gris.

Realizado por: Jessica Verdezoto Prado, 2024

Medusinas

Se identificó una secuencia que codificaba un péptido de esta familia nombrado: Medusin-AS. Cabe destacar que este mismo péptido ya fue descrito en *A. spurrelli* anteriormente [18].

La secuencia precursora poseía una longitud de 69 aminoácidos. La Tabla 13. muestra el marco de lectura abierto (ORF) que consistió en: (1) un péptido señal putativo de 22 aminoácidos; (2) un espaciador ácido de 27 aminoácidos que contienen dos sitios de procesamiento de propéptidos (KR); (3) la región del péptido maduro de 18 aminoácidos y (4) una región de residuo de glicina donante de amida en el extremo C-terminal.

Tabla 13. Secuencias aminoacídicas traducidas del marco de lectura abierto que codifican a péptidos de la familia de medusinas.

Epítopos de la familia A (Medusinas)	
Péptido	Péptido señal
	122.
Medusin-AS .	<u>M</u> <u>A</u> <u>E</u> <u>L</u> <u>K</u> <u>K</u> <u>S</u> <u>L</u> <u>E</u> <u>L</u> <u>V</u> <u>L</u> <u>F</u> <u>L</u> <u>G</u> <u>L</u> <u>V</u> <u>S</u> <u>L</u> <u>S</u> <u>V</u> <u>C</u>
	Espaciador ácido
	2349.
Medusin-AS .	E E E K R E S E E E K N E Q E E D D R E E R S E E K R
	Péptido maduro
	5069.
Medusin-AS .	L L G M I P L A I S A I S A L S K L G * - - - -

- = espacios añadidos para completar la alineación, *= codón de terminación (STOP). La región del péptido señal se encuentra subrayada, la zona ácida contiene espaciadores ácidos marcados con negrita. El péptido maduro esta resaltado en color gris y la amidación con color gris oscuro.
Realizado por: Jessica Verdezoto Prado, 2024.

Tras el alineamiento de las secuencias aminoacídicas en la base de datos del NCBI se obtuvo la Tabla 14., donde se observa que la medusin-AS compartió 83.33% de identidad con la secuencia aminoacídica del Medusin-AS de *Agalychnis spurrelli* con una cobertura de 98% (número de acceso: A0A5Q0MU22.1). Mientras que su péptido maduro compartió 100.00% de identidad con Medusin-AS de *Agalychnis spurrelli* con una cobertura de 100% (número de acceso: A0A5Q0MU22.1)

Tabla 14. Alineamiento múltiple de las secuencias aminoacídicas correspondientes a la familia de medusinas, obtenidas por secuenciación de ARN dirigido a la secreción cutánea de *A. spurrelli*.

Nº de acceso	Péptido	Péptido señal	
		1	22.
	Medusin-AS .	<u>MAFLKKSLFLVLFLGLVSLSC</u>	
A0A5Q0MU22.1	Medusin-AS .		
		Espaciador ácido	
		23.	
	Medusin-AS .	EEEKRESEEEKNEQEEDDREERSEEEKR	
A0A5Q0MU22.1	Medusin-AS .	 D	
		Péptido maduro	
		50.	69.
	Medusin-AS .	LLGMIPLAISALS KL G*	
A0A5Q0MU22.1	Medusin-AS .	 -	

. = aminoácidos conservados, - = espacios añadidos para completar la alineación, *= codón de terminación (STOP). La región del péptido señal se encuentra subrayada, la zona ácida contiene espaciadores ácidos marcados con negrita. El péptido maduro esta resaltado en color gris y la amidación con color gris oscuro.

Realizado por: Jessica Verdezoto Prado, 2024.

Plasticinas

Se encontraron 2 secuencias completas correspondientes a esta familia y fueron nombradas: Plasticin-AS2, Plasticin-AS3. El tamaño de las secuencias precursoras variaba entre 70 y 81 aminoácidos. En la Tabla 15. se aprecia el marco de lectura abierto (ORF) que consistió en: (1) un péptido señal putativo de 22 aminoácidos; (2) un espaciador ácido de 22 aminoácidos que contienen dos sitios de procesamiento de propéptidos (KR). Y (3) la región del péptido maduro que variaban entre 25 a 36 aminoácidos.

Tabla 15. Secuencias aminoacídicas traducidas del marco de lectura abierto que codifican a péptidos de la familia de plasticinas.

Péptido	Péptido señal	
	1	22.
Plasticin-AS2 .	<u>MA</u> <u>F</u> <u>L</u> <u>K</u> <u>K</u> <u>S</u> <u>L</u> <u>F</u> <u>L</u> <u>V</u> <u>L</u> <u>F</u> <u>L</u> <u>A</u> <u>L</u> <u>V</u> <u>P</u> <u>L</u> <u>S</u> <u>I</u> <u>C</u>	
Plasticin-AS3 .	<u>MA</u> <u>F</u> <u>L</u> <u>K</u> <u>N</u> <u>S</u> <u>L</u> <u>F</u> <u>L</u> <u>V</u> <u>L</u> <u>F</u> <u>L</u> <u>T</u> <u>L</u> <u>V</u> <u>H</u> <u>L</u> <u>S</u> <u>I</u> <u>C</u>	
	Espaciador ácido	
	23.	44.
Plasticin-AS2 .	E E K R E G E I E K E Q E D D D Q S E A K R	
Plasticin-AS3 .	E E K R E G E N E K E Q E D D D Q S E A K R	
	Péptido maduro	
	45.	81.
Plasticin-AS2 .	G L L S G L L S T V T G L L G N L G G G G L K K I * - - - - -	
Plasticin-AS3 .	G L L S G L L G T V T G L I K S V L G S L S G G Q S S E S V Q S N D S K *	

- = espacios añadidos para completar la alineación, *= codón de terminación (STOP). La región del péptido señal se encuentra subrayada, la zona ácida contiene espaciadores ácidos marcados con negrita. El péptido maduro esta resaltado en color gris.

Realizado por: Jessica Verdezoto Prado, 2024.

Al comparar las secuencias aminoacídicas en la base de datos del NCBI se creó la Tabla 16., donde se observa que la plasticin-AS2 y plasticin-AS3 compartieron entre 80.00% y 78.26% de identidad con la secuencia aminoacídica del plasticin-S1 de *Phyllomedusa sauvagii* con una cobertura de entre 98% y 85% (número de acceso: Q1EN14.1). Mientras que sus péptidos maduros compartieron 92.00% de identidad con el plasticin-S1 con una cobertura del 96% (número de acceso: Q1EN14.1).

Tabla 16. Alineamiento múltiple de las secuencias aminoacídicas correspondientes a la familia de plasticinas, obtenidas por secuenciación de ARN dirigido a la secreción cutánea de *A. spurrelli*.

N° de acceso	Péptido	Péptido señal	
		1	22.
	Plasticin-AS2.	<u>M A F L K K S L E L V L F L A L V P L S I C</u>	
	Plasticin-AS3.	 N T . . H	
Q1EN14.1	Plasticin-S1 .		
		Espaciador ácido	
		23.	45.
	Plasticin-AS2.	E E - K R E G E I E K E Q E D D D Q S E A K R	
	Plasticin-AS3.	. . - N	
Q1EN14.1	Plasticin-S1 .	. . E N N . . . E .	
		Péptido maduro	
		46.	82.
	Plasticin-AS2.	G L L S G L L S T V T G L L G N L G G G G L K K I *	
	Plasticin-AS3.	 G I K S V L . S L S G G Q S S E S V Q S N D S K *	
Q1EN14.1	Plasticin-S1 .	. . V . D	

. = aminoácidos conservados, - = espacios añadidos para completar la alineación, *= codón de terminación (STOP). La región del péptido señal se encuentra subrayada, la zona ácida contiene espaciadores ácidos marcados con negrita. El péptido maduro esta resaltado en color gris.

Realizado por: Jessica Verdezoto Prado, 2024

Caerinas

Se encontró 1 secuencia completa correspondiente a esta familia y fue nombrada: Caerin peptide-AS1. La secuencia precursora presentó una longitud de 73 aminoácidos. El marco de lectura abierto (ORF) se muestra en la Tabla 17. y consistió en: (1) un péptido señal putativo de 22 aminoácidos; (2) un espaciador ácido de 23 aminoácidos que contienen dos sitios de procesamiento de propéptidos (KR). Y (3) la región del péptido maduro de 27 aminoácidos.

Tabla 17. Secuencia aminoacídica traducida del marco de lectura abierto que codifica a péptidos de la familia de caerinas.

Péptido	Péptido señal	
	1	22.
Caerin Peptide-AS1 .	<u>M</u> <u>A</u> <u>F</u> <u>L</u> <u>K</u> <u>K</u> <u>S</u> <u>L</u> <u>F</u> <u>L</u> <u>V</u> <u>L</u> <u>F</u> <u>L</u> <u>G</u> <u>L</u> <u>V</u> <u>S</u> <u>L</u> <u>S</u> <u>I</u> <u>C</u>	
	Espaciador ácido	
	23.	45.
Caerin Peptide-AS1 .	D E E K R E N E D E E E Q E D D E Q S E E K R	
	Péptido maduro	
	46.	73.
Caerin Peptide-AS1 .	G M W G T L F K G A T A V A K H I L P H V F G S Q Q S *	

- = espacios añadidos para completar la alineación, *= codón de terminación (STOP). La región del péptido señal se encuentra subrayada, la zona ácida contiene espaciadores ácidos marcados con negrita. El péptido maduro esta resaltado en color gris.

Realizado por: Jessica Verdezoto Prado, 2024.

Al comparar la secuencia aminoacídica en la base de datos del NCBI se obtuvo la Tabla 18., donde se muestra que la caerin peptide-AS1 compartió 90.28% de identidad con la secuencia aminoacídica del caerin-regulated peptide de *Agalychnis callidryas* con una cobertura del 98% (número de acceso: B6HY14.1). Mientras que su péptido maduro compartió 85.29% de identidad con el caerin-regulated peptide con una cobertura del 100% (número de acceso: B6HY14.1).

Tabla 18. Alineamiento múltiple de la secuencia aminoacídica correspondiente a la familia de caerinas, obtenida por secuenciación de ARN dirigido a la secreción cutánea de *A. spurrelli*.

N° de acceso	Péptido	Péptido señal	
		1	22.
	Caerin Peptide-AS1.	<u>M A F L K K S L F L V L F L G L V S L S I C</u>	
B6HY14.1	Caerin-regulated	 L	
		Espaciador ácido	
		23.	45.
	Caerin Peptide-AS1.	D E E K R E N E D E E E Q E D D E Q S E E K R	
B6HY14.1	Caerin-regulated		
		Péptido maduro	
		46.	73.
	Caerin Peptide-AS1.	G M W G T L F K G A T A V A K H I L P H V F G S Q Q S *	
B6HY14.1	Caerin-regulated	 V . . . I K T L S -	

. = aminoácidos conservados, - = espacios añadidos para completar la alineación, *= codón de terminación (STOP). La región del péptido señal se encuentra subrayada, la zona ácida contiene espaciadores ácidos marcados con negrita. El péptido maduro esta resaltado en color gris.

Realizado por: Jessica Verdezoto Prado, 2024.

Péptidos huérfanos

Se identificaron 8 secuencias completas que codificaban a péptidos que coincidieron con más de una familia de PAMs, por tanto, fueron nombrados: Orphan Peptide-SP1, Orphan Peptide-SP2, Orphan Peptide-SP7, Orphan Peptide-SP8, Orphan Peptide-SP9, Orphan Peptide-SP10, Orphan Peptide-SP11, Orphan Peptide-SP12. Orphan Peptide-SP2 y Orphan Peptide-SP1 ya fueron reportadas anteriormente en *A. spurrelli* [18].

Las secuencias precursoras presentaron una longitud de entre 64 y 79 aminoácidos. El marco de lectura abierto (ORF) se muestra en la Tabla 19. y consistió en: (1) un péptido señal putativo de 22 aminoácidos; (2) un espaciador ácido de entre 22 a 30 aminoácidos que contienen dos sitios de procesamiento de propéptidos (KR). Y (3) la región del péptido maduro que variaba entre 11 a 26 aminoácidos. Adicionalmente, algunos péptidos contaban con (4) una región de residuo de glicina donante de amida en el extremo C-terminal.

Tabla 19. Secuencias aminoacídicas traducida del marco de lectura abierto que codifican a péptidos huérfanos.

Péptido	Péptido señal																							Espaciador ácido																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																									
	1																							23.										52.																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																															
OrphanPeptide-SP1 .	M	A	F	L	K	K	S	L	F	L	V	L	F	L	G	L	V	S	L	S	V	C	E	E	E	K	R	E	G	E	E	K	E	N	E	Q	E	E	D	D	R	E	E	R	S	E	E	-	-	-	K	R																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																													
OrphanPeptide-SP2 .	M	D	F	L	K	K	S	L	F	L	V	L	F	L	G	F	V	S	I	S	F	C	D	E	E	K	R	E	D	D	E	E	G	N	E	R	E	E	K	K	E	I	H	E	E	G	N	Q	E	E	K	R																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																													
OrphanPeptide-SP7 .	M	S	L	L	K	K	S	L	L	L	V	L	F	L	G	V	V	S	L	S	N	C	E	E	E	K	R	E	-	E	E	E	N	E	E	H	E	E	H	H	E	E	-	-	-	-	-	-	-	K	R																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																														
OrphanPeptide-SP8 .	M	A	F	L	K	K	S	L	F	L	V	L	F	L	G	V	V	S	L	S	I	C	E	E	E	K	R	E	G	E	N	E	G	Q	E	R	E	E	Q	K	E	E	-	-	-	-	-	-	-	-	K	R																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																													
OrphanPeptide-SP9 .	M	S	F	L	K	K	S	L	F	L	V	L	F	L	G	L	V	S	L	S	I	C	E	E	E	K	R	E	P	K	E	E	E	N	E	N	E	E	E	N	N	E	R	S	E	M	-	-	-	K	R																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																														
OrphanPeptide-SP10 .	M	D	F	L	K	K	S	L	F	L	V	L	F	L	G	F	V	S	I	S	F	C	D	E	E	K	R	Q	E	D	D	E	A	N	E	R	E	E	K	K	E	I	H	E	E	G	N	Q	E	A	K	-																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																													
OrphanPeptide-SP11 .	M	S	F	L	K	K	S	L	L	L	V	L	F	L	G	L	V	S	F	S	I	C	E	E	E	K	R	E	A	E	E	E	E	N	E	D	E	I	E	E	E	M	-	-	-	-	-	-	-	K	R																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																														
OrphanPeptide-SP12 .	M	S	F	L	K	K	S	L	L	L	V	L	F	L	G	L	V	S	F	S	I	C	E	E	E	K	R	E	T	E	E	E	E	N	E	D	E	W	T	K	K	V	K	R	-	-	-	-	-	-	K	R																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																													
Péptido maduro																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																	
	53.																							79.																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																									
OrphanPeptide-SP1 .	I	I	G	M	I	P	D	L	I	S	M	I	S	K	L	G	K	*	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

- = espacios añadidos para completar la alineación, *= codón de terminación (STOP). La región del péptido señal se encuentra subrayada, la zona ácida contiene espaciadores ácidos marcados con negrita. El péptido maduro está resaltado en color gris y la amidación con color gris oscuro.

Realizado por: Jessica Verdezoto Prado, 2024.

Tras el alineamiento de las secuencias aminoacídicas en la base de datos del NCBI se obtuvo la Tabla 20., donde se observa que el orphan peptide-SP1, presentó 100.00% de identidad con el Peptide SP-1 de *Agalychnis spurrelli* con una cobertura del 98% (número de acceso: QFZ95567.1). Y su péptido maduro obtuvo 100.00% de identidad con el Peptide SP-1 y una cobertura del 94% (número de acceso: QFZ95567.1).

Por otro lado, el orphan peptide-SP2 presentó 100.00% de identidad con el peptide SP-2 de *Agalychnis spurrelli* y 98% de cobertura (número de acceso: QFZ95568.1). Y su péptido maduro obtuvo 100.00% de identidad con el peptide SP-2 100% de cobertura (número de acceso: QFZ95568.1).

El orphan peptide-SP7 compartió 83.33% de identidad con la secuencia aminoacídica del Medusin-AS de *Agalychnis spurrelli* con una cobertura de 98% (número de acceso: A0A5Q0MU22.1). Adicionalmente compartió 76.92% de identidad con hylareleasin 11 de *Hyla simplex* con una cobertura de 76% (número de acceso: AEJ35109.1).

Para el caso del orphan peptide-SP8 mostró identidad del 83.33% con medusin-PT de *Phyllomedusa tarsius* y cobertura del 43% (número de acceso: A0A1M4BLT1.1). Mientras, que su péptido maduro obtuvo 60.87% de identidad con el ATP-binding cassette sub-family C member 3 de *Bufo gargarizans* y cobertura del 83% (número de acceso: XP_044152162.1).

El orphan peptide-SP9 presentó identidad de 80.56% con medusin-PT de *Phyllomedusa tarsius* y cobertura del 58% (número de acceso: A0A1M4BLT1.1). Su péptido maduro mostró 100.00% de identidad con el Peptide TRP-J1 de *Phasmahyla jandaia* y cobertura del 76% (número de acceso: P86709.1).

Respecto al orphan peptide-SP10 se obtuvo 86.27% de identidad con Tryptophyllin-T1 de *Agalychnis dacnicolor* y cobertura del 77% (número de acceso: P83455.2). Su péptido maduro mostró 75.00% de identidad con el T-cell immunoglobulin and mucin domain-

containing protein 4-like isoform X1 de *Geotrypetes seraphini* y cobertura del 100% (número de acceso: XP_033783407.1).

El orphan peptide-SP11 compartió 87.50% de identidad con la secuencia aminoacídica del phyllokininogen-B1 de *Phyllomedusa bicolor* con una cobertura de 76% (número de acceso: AFR78288.1). Mientras que su péptido maduro compartió 100.00% de identidad con la secuencia aminoacídica del Bradykinin-[Val1, Thr6]-bradykinyl-Ser, Pro, Ala de *Physalaemus nattereri* con una cobertura de 71% (número de acceso: A0A088MIT0.1). Además, de 100.00% de identidad con el T-kininogen 2-like de *Bufo gargarizans* con una cobertura del 64% (número de acceso: XP_044148410.1).

Para el caso del orphan peptide-SP12 mostró identidad del 93.70% con la secuencia aminoacídica del phyllokininogen-B1 de *Phyllomedusa bicolor* con una cobertura de entre 50% (número de acceso: AFR78288.1). Su péptido maduro presentó 71.43% de identidad con el hypothetical protein NDU88_002087 de *Pleurodeles waltl* y cobertura del 50% en el extremo N-terminal (número de acceso: KAJ1113846.1). Mientras, que en el extremo C-terminal obtuvo 50.00% de identidad con el uncharacterized protein LOC128502266 de *Spea bombifrons* y cobertura del 67% (número de acceso: XP_053328001.1).

Tabla 20. Alineamiento múltiple de las secuencias aminoacídicas correspondientes a péptidos huérfanos, obtenidas por secuenciación de ARN dirigido a la secreción cutánea de *A. spurrelli*.

N° de acceso	Péptido	Péptido señal	Espaciador ácido
QFZ95567.1	OrphanPeptide-SP1. peptide SP-1	1	49.
		MA F L K K S L F L V L F L G L V S L S V C E E E K R E G E E K E N E Q E E D D R E E R S E E K R	
		Péptido maduro	
QFZ95567.1	OrphanPeptide-SP1. peptide SP-1	50.	67.
		I I G M I P D L I S M I S K L G K *	
		-	
N° de acceso	Péptido	Péptido señal	Espaciador ácido
QFZ95568.1	OrphanPeptide-SP2. peptide SP-2	1	52.
		MD F L K K S L F L V L F L G F V S I S F C D E E K R E D D E E G N E R E E K K E I H E E G N Q E E K R	
		Péptido maduro	
QFZ95568.1	OrphanPeptideSP2 . peptide SP-2	53.	64.
		F L F L P F I G R R K *	
		-	

Tabla 20. Continuación de Tabla.

N° de acceso	Péptido	Péptido señal	Espaciador ácido
		1	52.
	OrphanPeptide-SP7	<u>M</u> S L L K K S L L L V L F L G V V S L S N C E E E K R E - E E E N E E H E E H H - E E - - - - -	
A0A5Q0MU22.1	Medusin-AS	. A F F L . . . V S . . . K N . Q . . D D R D . R S E E - - - - -	
AEJ35109.1	hylareleasin 11	. T F F L . . . I . . . K . . . E . . . E K R D G S E E E . . E K R D D S E E E	
		Espaciador ácido	
		53.	72.
	OrphanPeptide-SP7	- - - - - K R	
A0A5Q0MU22.1	Medusin-AS	- - - - -	
AEJ35109.1	hylareleasin 11	E E E K R D D E K E E E N E A D A E . .	
		Péptido maduro	
		73.	92.
	OrphanPeptide-SP7	- L L E G L L D T V T N L V G S I V * -	
A0A5Q0MU22.1	Medusin-AS	- . . G M I P L A I S A I S A L S K L G	
AEJ35109.1	hylareleasin 11	G . . D P V T N L L G G I L P G R K - -	
N° de acceso	Péptido	Péptido señal	Espaciador ácido
		1	49.
	OrphanPeptide-SP8	<u>M</u> A F L K K S L F L V L F L G V V S L S I C E E E K R E G E N E G Q E R E E Q K E E - - - - - K R	
A0A1M4BLT1.1	Medusin-PT	 F . . . F T D E - K E N E Q . D D R . E R S E E . .	
XP_044152162.1	ATP-binding cassette	- - - - -	. T
		Péptido maduro	
		50.	77.
	OrphanPeptide-SP8	L L M D - T L S K A T N L V D N - - V A N S A G G I I *	
A0A1M4BLT1.1	Medusin-PT Flags:	. . G M I P V A I T A I S A L S K L G - - - - -	
XP_044152162.1	ATP-binding cassette	. . A . E . . N H . D . A . . E P . . . E . R K - - -	

Tabla 20. Continuación de Tabla.

Nº de acceso	Péptido	Péptido señal	Espaciador ácido
	OrphanPeptide-SP9	1	49.
AOA1M4BLT1.1	Medusin-PT	<u>M</u> <u>S</u> <u>F</u> <u>L</u> <u>K</u> <u>K</u> <u>S</u> <u>L</u> <u>F</u> <u>L</u> <u>V</u> <u>L</u> <u>F</u> <u>L</u> <u>G</u> <u>L</u> <u>V</u> <u>S</u> <u>L</u> <u>S</u> <u>I</u> <u>C</u> <u>E</u> <u>E</u> <u>E</u> <u>K</u> <u>R</u> <u>E</u> <u>P</u> <u>K</u> <u>E</u> <u>E</u> <u>E</u> <u>N</u> <u>E</u> <u>N</u> <u>E</u> <u>E</u> <u>E</u> <u>N</u> <u>N</u> <u>E</u> <u>E</u> <u>R</u> <u>S</u> <u>E</u> <u>M</u> <u>K</u> <u>R</u>	
P86709.1	Peptide TRP-J1.	. A F . . . F T D . K . . . Q . D D R - E . .	
		- - - - -	
		Péptido maduro	
	OrphanPeptide-SP9	50.	62.
AOA1M4BLT1.1	Medusin-PT	L M Y Y T L P R P V G R *	
P86709.1	Peptide TRP-J1.	. L G M I P V A I T A I S A L S K L G	
		 - - -	
Nº de acceso	Péptido	Péptido señal	Espaciador ácido
	OrphanPeptide-SP10	1	22.
P83455.2	Tryptophyllin-1	<u>M</u> <u>D</u> <u>F</u> <u>L</u> <u>K</u> <u>K</u> <u>S</u> <u>L</u> <u>F</u> <u>L</u> <u>V</u> <u>L</u> <u>F</u> <u>L</u> <u>G</u> <u>F</u> <u>V</u> <u>S</u> <u>I</u> <u>S</u> <u>F</u> <u>C</u>	
XP_033783407.1	T-cell immunoglobuli	. N	
		- S H S A L L . W T L T L . - C I G Q T V L	
		Espaciador ácido	
	OrphanPeptide-SP10	23.	55.
P83455.2	Tryptophyllin-1	D E E K R Q E D D E A N E R E E K K E I H E E G N Q E A K - E T S	
XP_033783407.1	T-cell immunoglobuli	 D . . . G Q . D . . . E R R - - -	
	Tryptophyllin-1	T L T V S G T V G S P V T L P C R Y T V W G R . S C T F L G - - L	
		Péptido maduro	
	OrphanPeptide-SP10	56.	68.
P83455.2	Tryptophyllin-1	L L L G F L N K R N M L *	
XP_033783407.1	T-cell immunoglobuli	D K P P A W V P G . - - - -	
		. . . A . . A . . K . . K R	

Tabla 20. Continuación de Tabla.

N° de acceso	Péptido	Péptido señal	Espaciador ácido
		1	44.
	OrphanPeptide-SP11.	<u>M</u> S F L K K S L L L V L F L G L V S F S I C E E E K R E A E E E E N E D E I E E E M K R	
A0A088MIT0.1	Bradykinin .	. A F V . . L . F E H . . . K R . . E D A . S - -	
XP_044148410.1	T-kininogen 2 .	. R . . P V V . . C S C L . . S A N Q A S T . . - - - - - - - - - - - - - - - ,	
		Péptido maduro	
		45.	60.
	OrphanPeptide-SP11.	E E T V I K - G L S P L R G S *	
A0A088MIT0.1	Bradykinin .	- F V . . . - K R	
XP_044148410.1	T-kininogen 2 .	T R I . . . - S A M P	
N° de acceso	Péptido	Péptido señal	Espaciador ácido
		1	46.
	OrphanPeptide-SP12.	<u>M</u> S F L K K S L L L V L F L G L V S F S I C E E E K R E T E E E E N E D E W T K K V K R K R	
KAJ1113846.1	hypothetical protein .	- ,	
XP_053328001.1	uncharacterized prot .	- ,	
		Péptido maduro	
		47.	73.
	OrphanPeptide-SP12.	E S L Q I D L P V S V L F E L I N T L R K C N V I I *	
KAJ1113846.1	hypothetical protein .	. . . L V E . . . F . . C . N R V T . L S G D . F F A R	
XP_053328001.1	uncharacterized prot .	 V E . I L E N T . . I . S . D F . - - - V F K E E	

. = aminoácidos conservados, - = espacios añadidos para completar la alineación, *= codón de terminación (STOP). La región del péptido señal se encuentra subrayada, la zona ácida contiene espaciadores ácidos marcados con negrita. El péptido maduro está resaltado en color gris y la amidación con color gris oscuro.

Realizado por: Jessica Verdezoto Prado, 2024

3.2.3. Clasificación de las proteínas encontradas

Kazal

Las secuencias proteicas se agruparon en una familia de proteínas de tipo kazal, identificando 9 nuevas proteínas con actividad inhibidora de proteinasas. Tras el consenso de secuencias, se agruparon 17 secuencias precursoras que fueron cotejadas varias veces en una misma, y codificaban a proteínas de tipo kazal nombradas: ASKP-1a, ASKP-1b, ASKP-2, ASKP-3, ASKP-4a, ASKP-4b, ASKP-5a, ASKP-5b, ASKP-6a, ASKP-6b, ASKP-7a, ASKP-7b, ASKP-8, ASKP-9a, ASKP-9b, ASKP-9c, ASKP-19d. Se nombró a las proteínas de esta manera por proceder de *Agalychnis spurrelli*, pertenecer a una Kazal Protein y contener variantes a, b, c, d (ASKP-1a).

Las secuencias precursoras presentaron una longitud de entre 64 y 79 aminoácidos. El marco de lectura abierto (ORF) se muestra en la Tabla 21. y consistió en: (1) un péptido señal putativo de 22 a 33 residuos. Y (2) la región de la proteína inhibidora de tipo Kazal madura que variaba entre 53 a 69 residuos. Adicionalmente, las proteínas y sus variantes presentan el patrón Kazal CX (6,7) -CX (6,7) -CX (6,7) -YX (3) -CX (2) -CX (15-21) -C.

Tabla 21. Secuencias aminoacídicas correspondientes a proteínas de tipo Kazal, obtenidas mediante secuenciación de ARN dirigido a la secreción cutánea de *A. spurrelli*.

Proteína	Proteína señal																																										
	33.																																										
ASKP-1a .	M	K	T	L	I	S	S	A	L	L	F	C	I	L	A	A	A	L	W	P	V	P	E	A	E	G	-	-	-	-	-	-	-	-	-	-	-	-					
ASKP-1b .	M	K	T	L	I	S	S	A	L	L	F	C	I	L	A	A	A	L	W	P	V	P	E	A	E	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-			
ASKP-2 .	M	K	T	L	I	S	S	V	L	L	F	C	I	L	A	A	A	L	W	P	V	P	Q	A	E	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-			
ASKP-3 .	M	K	T	L	I	S	S	V	L	L	F	C	I	L	A	A	A	L	W	P	V	P	E	A	E	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		
ASKP-4a .	M	K	T	L	I	S	S	V	L	L	F	C	I	L	A	A	A	L	W	P	I	P	E	T	S	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		
ASKP-4b .	M	K	T	L	I	S	S	V	L	L	F	C	I	L	A	A	A	L	W	P	I	P	E	T	S	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		
ASKP-5a .	M	K	T	L	I	S	S	A	L	L	F	C	I	L	A	A	A	L	A	K	P	P	S	Q	A	F	E	K	L	V	F	K	H										
ASKP-5b .	M	K	T	L	I	S	S	A	L	L	F	C	I	L	A	A	A	L	A	K	P	P	S	Q	A	F	E	K	L	V	F	K	H										
ASKP-6a .	M	K	T	L	I	S	F	A	L	L	F	C	I	L	G	K	A	L	A	P	P	P	S	Q	P	F	E	K	L	V	F	K	H										
ASKP-6b .	M	K	T	L	I	S	F	A	L	L	F	C	I	L	G	K	A	L	A	P	P	P	S	Q	P	F	E	K	L	V	F	K	H										
ASKP-7a .	M	K	T	L	I	S	S	A	L	L	F	C	I	L	A	A	A	L	W	P	V	Q	E	A	E	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
ASKP-7b .	M	K	T	L	I	S	S	V	L	L	F	C	I	L	A	A	A	L	W	P	V	T	E	V	E	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
ASKP-8 .	M	K	T	L	I	S	S	V	V	L	L	C	I	L	A	S	A	L	A	P	L	P	S	Q	P	F	P	N	K	V	T	Q	L										
ASKP-9a .	M	K	T	L	I	S	S	A	L	L	F	C	I	L	A	A	A	L	W	P	I	P	E	A	E	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ASKP-9b .	M	K	T	L	I	S	S	V	L	L	F	C	I	L	A	A	A	L	W	P	I	H	E	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	E	A		
ASKP-9c .	M	K	T	L	I	S	S	V	L	L	F	C	I	L	A	A	A	L	W	P	V	H	E	A	E	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ASKP-9d .	M	K	T	L	I	S	S	V	L	L	F	C	I	L	A	A	A	L	W	P	I	Y	E	A	E	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	

Tabla 21. Continuación de tabla.

Proteína	Proteína madura																							
	I	II	III	IV	V	VI																		
			***	*	*	*																		
ASKP-1a .	A Y D P T	C R W Y H I F - G	C H S -	H N P V	C A T D G Q S	Y R S E	C H V	C K E	- - -	N R N G	-	K N L	Q I	E K E	G H C	S *	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -
ASKP-1b .	A Y A P T	C G W T Y N F - G	C H S -	H N P V	C A T D G Q S	Y R S E	C H V	C R E	- - -	N S N R E	Q K L L I	E K E	G H C	S *	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -
ASKP-2 .	A T A A K	C E N F R G F - -	C Y K	N Y V P V	C G T N G I	T Y S N Q	C M M	C F A	- - -	N R Q N K	N N V Q I	K K P	G P C	*	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -
ASKP-3 .	A T K P K	C P W F S S - - G	C P S	E Q N F V	C G T D G K S	Y M N E	C V M	C T W	- - -	N K N H G E	K V K I	E K K	G R C	*	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -
ASKP-4a .	A R K P D	C S K Y K P K E K	C P L	T M K V V	C G T D G R E	Y Y N E	C A L	C V S	- - -	N R D S K	K N V E I	K K D	S K C	*	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -
ASKP-4b .	A R K P D	C S K Y K P K E K	C P L	T M K V V	C G T D G R E	Y Y N E	C A L	C V S	- - -	N R D S K	K K K K K K K K K	S T L	R *	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -
ASKP-5a .	I T E P N	C K K Y E G K - K	C S V	D P N P V	C G T D E R T	Y Y N E	C A M	C L F	M R T S T	K K K K K K K K K	- - -	S T L	R *	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -
ASKP-5b .	I T E P N	C K K Y E G K - K	C S V	D P N P V	C G T D E R T	Y Y N E	C A M	C L F	M R T S T	K K E D K L I	K I	G K W G N	C R D *	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -
ASKP-6a .	I T E P D	C K R Y E G K - K	C P M	H Y S P V	C G T D E R T	Y Y N E	C A L	C V F	I R D S T	K K K D K L I	K I	G K W G K	C R G *	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -
ASKP-6b .	I T E P D	C K R Y E G K - K	C P M	H Y S P V	C G T D E R T	Y Y N E	C A L	C V F	I R D S T	K K K D K L I	K I	G K W G K	M Q R L T H Q C S A K *	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -
ASKP-7a .	A E D L Y	C P W Y Y A L - G	C Y D	- N N P V	C G K D G T T	Y G N K	C T L	C N L	- - -	N K L R E	K K V E F E K W G E	C *	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -
ASKP-7b .	A E D V T	C P W Y Y I F - G	C F D	- K N P L	C G T D G K T	Y P N K	C V L	C - - - -	- - - - -	K Q H G E N N I	Q I	E K W G N	C *	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -
ASKP-8 .	V F E P K	C Q K Y E G I - K	C D K	V L N P V	C G T N E R T	Y F N E	C A M	C V F	K R D S K	K S E D K L V K I	K K W G E	C S P K *	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -
ASKP-9a .	V I E P K	C S R Y P G I - -	C T Y	G P N A V	C G T D G K K	Y F N E	C T L	C V F	- - -	N R K N G	K D V K I	E K Y	G A C *	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -
ASKP-9b .	V I E P N	C K K Y E G K - K	A S L	D M N P V	C G T N G R T	Y Y N E	Y A L	C V F	I R D S T	K K A D K Q V K I	K K W G K	C *	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -
ASKP-9c .	V I E P N	C G K Y Q G I - -	C T F	G P K A I	C G T D G K R	Y F N E	C R L	C A S	- - -	N K E N G	Q D V K I	E K Y	G K C *	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -
ASKP-9d .	V I E P D	C V K Y H G S - -	C K K	S P N Y I	C G T D G T T	Y Y N E	C A L	C E A	- - -	N K G K K	K P A K I	K K Y	G K C *	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -
Patrón kazal	C	C	P2 P1	C	Y	C	C																	

- = espacios añadidos para completar la alineación, *= sitios conservados. El patrón Kazal está subrayado: (1) gris las 6 Cisteínas, (2) rosado la Tirosina y (3) rojo los sitios de activación P1 y P2.

Realizado por: Jessica Verdezoto Prado, 2024.

Una vez alineadas las secuencias aminoacídicas en la base de datos del NCBI se obtuvo la Tabla 22., donde se observa que la ASKP-1a y ASKP-1b compartieron 60.49% y 63.49% de identidad respectivamente con la secuencia de la CCKP-2 proteinase inhibitor de *Cruziophyla calcarifer* con una cobertura del 97% y 76% (número de acceso: ANN87741.1).

La ASKP-2 exhibió 66.13% de identidad con el trypsin inhibitor precursor de *Cruziophyla calcarifer* y 76% de cobertura (número de acceso: ANN87740.1). Por otra parte, la ASKP-3 obtuvo 90.32% de identidad con la CCKP-3 proteinase inhibitor de *Cruziophyla calcarifer* y 76% de cobertura (número de acceso: ANN87742.1).

Para el caso de la ASKP-4a y ASKP-4b presentaron 68.67% Y 69.33% de identidad con la la CCKP-2 proteinase inhibitor de *Cruziophyla calcarifer* con una cobertura ente 100% y 90% (número de acceso: ANN87741.1). Por otro lado, la ASKP-5a y ASKP-5b mostraron 85.19% y 87.10% de identidad con la PSKP-1 proteinase inhibitor de *Pithecopus nordestinus* y entre 90% y 98% de cobertura (número de acceso: AFY11406.1).

La ASKP-6a y ASKP-6b compartieron entre 88.04% y 89.77% de identidad con la secuencia de la PSKP-1 proteinase inhibitor de *Pithecopus nordestinus* con una cobertura de entre 97% y 86% (número de acceso: AFY11405.1). Por otro lado, la ASKP-7a y ASKP-7b exhibieron entre 64.56% y 78.67% de identidad con la CCKP-6a proteinase inhibitor de *Cruziophyla calcarifer* con una cobertura de entre 98% y 96% (número de acceso: ANN87750.1).

Para el caso de la ASKP-8 obtuvo 72.83% de identidad con la CCKP-7f proteinase inhibitor de *Cruziophyla calcarifer* y 96% de cobertura (número de acceso: ANN87757.1). Mientras que la ASKP-9a mostró 67.09% de identidad con la PSKP-1 proteinase inhibitor de *Pithecopus nordestinus* y 98% de cobertura (número de acceso: AFY11407.1).

La ASKP-9b, ASKP-9c y ASKP-9d compartieron 89.16%, 66.67% y 67.86 de identidad respectivamente con la secuencia de la CCKP-5d proteinase inhibitor de *Cruziohyla calcarifer* con una cobertura del 97%, 98% y 98% (número de acceso: ANN87749.1).

Tabla 22. Alineamiento múltiple de las secuencias correspondientes a proteínas de tipo Kazal, obtenidas por secuenciación de ARN dirigido.

N° de acceso	Proteína	Proteína señal
		1 26.
	ASKP-1a	MKTLISSALLFCILAAALWPVPEAEG
	ASKP-1b	
ANN87741.1	CCKP-2 prot.	
		Proteína madura
		27. * * * * *
	ASKP-1a	AYDPTCRWYHIFG-CHS-HNPVCATDGQSYRSECHVC--NRNGKNLQ-IEKEGH--CS*
	ASKP-1b	..A..G.TYN.....RE.S.REQKLL.....
ANN87741.1	CCKP-2 prot.	.TE.D.KK.P--K.PLAQ...G...RM.YN..AL.VFM.DS..KVK.QIKKMGK..
N° de acceso	Proteína	Proteína señal
		1 26.
	ASKP-2	MKTLISSVLLFCILAAALWPVPQAEG
ANN87740.1	trypsin inhi.	A.....E...
		Proteína madura
		27. * * * * *
	ASKP-2	ATAAKCENFRGF-CYKNYVPVCGTNGITYSNQCMCFANRQNKNNVQIKKPGPC*
ANN87740.1	trypsin inhi.	.VS.E.ARY-.LA.N.MLA...D.T.....L.YY..K..K.IE.RSR.R.-
N° de acceso	Proteína	Proteína señal
		1 26.
	ASKP-3	MKTLISSVLLFCILAAALWPVPEAEG
ANN87742.1	CCKP-3 prot.	K...
		Proteína madura
		27. * * * * *
	ASKP-3	ATKPKCPW-FSSGCPSEQNFVCGTDGKSYMNECVMCTWKNHGEKVKIEKKGRC*
ANN87742.1	CCKP-3 prot.	SL.....T.D.....N.....KM...N.G...VV...Y.-

Tabla 22. Continuación de tabla.

N° de acceso	Proteína	Proteína señal
		1 26.
	ASKP-4a	. MKT L I S S V L L F C I L A A A L W P I P E T S -
	ASKP-4b	 -
ANN87741.1	CCKP-2 proi.	 A V . . A E G
		Proteína madura
		27. * * * * *
	ASKP-4a	. A R K P D C S K Y K P K E K C P L T M K V V C G T D G R E Y Y N E C A L C V S N R D S K K N V E I K K D S - - K C *
	ASKP-4b	 K K K K . . K . T L R - *
ANN87741.1	CCKP-2 proi.	. T E . . . K . . - . G - . . A Q N P M F M N K . K . Q I K K M G K .
N° de acceso	Proteína	Proteína señal
		1 33.
	ASKP-5a	. MKT L I S S A L L F C I L A A A L A K P P S Q A F E K L V F K H
	ASKP-5b	
AFY11406.1	PSKP-1 prot.	 V . I T . . A Q
		Proteína madura
		34. * * * * *
	ASKP-5a	. I T E P N C K K Y E G K K C S V D P N P V C G T D E R T Y Y N E C A M C L F M R T S T K K K K K K K K - - - S T - L R * - - - -
	ASKP-5b	 E D . L I . I G K W G N C . D * - - - -
AFY11406.1	PSKP-1 prot.	 P . M L . V . R E D . L I . I G K W G K C . E T I N V K

N° de acceso	Proteína	Proteína señal
		1 33.
	ASKP-6a	MKTLISFALLFCILGKALAPPPSQPFELVFKH
	ASKP-6b	
AFY11405.1	PSKP-1 prot.	SM.AT.....TK
		Proteína madura
		34. * * * * *
	ASKP-6a	ITEPDCKRYEGKKCPMHYSPVCGTDERTYYNECALCVFIRDSTKKKDKLIIKIGKWGKCRG*-----
	ASKP-6b	-MQLRLTHQCSAK*
AFY11405.1	PSKP-1 prot.	L.K.....T.....EA..KTIPYIK---
N° de acceso	Proteína	Proteína señal
		1 26.
	ASKP-7a	MKTLISSALLFCILAAALWPVQEAE-
	ASKP-7b	V.....T.V.G
ANN87750.1	CCKP-6a pro.	IP.V.G
		Proteína madura
		27. * * * * *
	ASKP-7a	AEDLYCPWYYALGCYDNNPVCCKGDTTYGNKCTLCNLTNKLREKKVEFEKWGEC*
	ASKP-7b	...VT...IF..F.K..L..T..K..P...V.-KQHGNNIQI--...N..
ANN87750.1	CCKP-6a pro.	E..VA...VF..H.KYT...T..V..P...KI.GEDNI.IR--...N.-
N° de acceso	Proteína	Proteína señal
		1 34.
	ASKP-8	MKTLISSVLLCIIASALAPLP SQP-FPNKVTQL
ANN87757.1	CCKP-7f pro.	L.F.....T.....P.....Q.S...LKT
		Proteína madura
		35. * * * * *
	ASKP-8	VFEPKCQKYEGIKCDKVLNPVCGTNERITYFNECAMCVFKRDSKKSEDKLVKIKKWGECSPK*
ANN87757.1	CCKP-7f pro.	LA..N.K....K...LN.....G...Y...L...I...T.KA..M...H.....----

Tabla 22. Continuación de tabla.

N° de acceso	Proteína	Proteína señal
		1 26.
	ASKP-9a	MKTLISSALLFCILAAALWPIPEAEA
AFY11407.1	PSKP-1 prot.	 V V
		Proteína madura
		27. * * * * *
	ASKP-9a	V I E P K C S R Y P G I - C T Y G P N A V C G T D G K K Y F N E C T L C V F - - - - N R K N G K D V K I E K Y G A C *
AFY11407.1	PSKP-1 prot.	 D K N A D M - . . D D Y T P N . . S . . R . . L H N R E T - - - - . Q N . . . A H R . E . -
N° de acceso	Proteína	Proteína señal
		1 26.
	ASKP-9b	MKTLISSVLLFCILAAALWPIHEAEA
	ASKP-9c	 V
	ASKP-9d	 Y . . . G
ANN87749.1	CCKP-5d pro.	 T
		Proteína madura
		27. * * * * *
	ASKP-9b	V I E P N C K K Y E G K K A S L D M N P V C G T N G R T Y Y N E Y A L C V F I R D S T K K A D K Q V K I K K W G K C *
	ASKP-9c	 G . . Q . I - C T F G P K A I . . . D . K R . F . . C R . . A S - - - - N . E N G Q D . . . E . Y
	ASKP-9d	 D . V . . H . S - C K K S P . Y I . . . D . T C . . . E A - - - - N . G K K . P A Y
ANN87749.1	CCKP-5d pro.	 C D . N P E . F . C M -

- = espacios añadidos para completar la alineación, *= sitios conservados. El patrón Kazal está subrayado: (1) gris las 6 Cisteínas, (2) rosado la Tirosina y (3) rojo los sitios de activación P1 y P2.

Realizado por: Jessica Verdezoto Prado, 2024.

3.2.4. Análisis de PAMs por espectrometría de masas.

El análisis de cromatografía líquida acoplada a espectrometría de masas en tándem (LC-MS/MS) general de las secuencias permitió confirmar la presencia de al menos veinte y cinco, nuevos péptidos en la secreción cutánea de *A. spurrelli* como se muestra en la Tabla 23. Estos nuevos péptidos pertenecen a las siguientes familias de PAMs: Dermaseptinas, Triptofilinas, Filoseptinas, Filoquininas, Medusinas y Caerinas, por lo que cuentan con un rango de masa molecular entre 894,5260 y 3184,7540 Da.

Las masas moleculares monoisotópicas de los nuevos péptidos se determinaron con la herramienta de libre acceso Peptide Mass Calculator, PeptideSynthetics. Mientras que el software Mascot Distiller 2.8.5 permitió la identificación por probabilidad de coincidencia de cada uno de los péptidos respecto a sus masas moleculares, estableciendo el nivel de significación (tolerancia, 1). También se realizaron identificaciones con diferentes niveles de significación (tolerancia entre 0,5 y 1,5), que coincidieron con los péptidos ya encontrados en la selección predeterminada, tolerancia 1.

Tabla 23. Péptidos con secuencia confirmada por LC-MS/MS.

Péptido	Secuencia	Método de Identificación	% Cobertura	N° fragmentos	N° AAs	LCQ MW (Da)	Masa teórica (Da)	Score
Dermaseptin-SP2	ASWKVFLKNIGKAAGKAVLNSVTDMVNQ-NH ₂	cm, sec, ms ²	100	75/308	28	2987,163	2987,627	72
Dermaseptin-SP3	SLWSSIKDMAAAAGRAALNAVNGIVNP-NH ₂	cm, sec, ms ²	100	49/260	27	2695,207	2695,412	27
Dermaseptin-SP6	ASWKDFFKNIGKAAGKADLNSVTDMVNQ-NH ₂	cm	*	*	28	*	3053,528	*
Dermaseptin-SP7	ASWKDFFKNIGKAAGKAVLNSVTDMVNQ-NH ₂	cm, sec, ms ²	100	56/308	28	3037,323	3037,570	26
Dermaseptin-SP8	ASWKEFFKNVGKAAGKAVLNSVTNMVNQ-NH ₂	cm, ms ²	100	47/306	28	3036,101	3036,586	32
Dermaseptin-SP9	ASWKVFLKNIGKAAGKAVLNPVTDMVNQ-NH ₂	cm	*	*	28	*	2997,648	*
Dermaseptin-SP10	GMFTNMLKGIGKLAGQAALGAVKTLV-NH ₂	cm, sec, ms ²	100	68/276	26	2587,041	2587,460	29
Dermaseptin-SP11	ASGKVFLKNIGKAAGKAVLNSVTDMVNQ-NH ₂	sec	*	*	28	*	2858,569	*
Dermaseptin-SP12	ASWKDFFKKYRKSSRKSCFKFSYGYGKSRRRAIKQKNVK SVIVPALH	sec	*	*	46	*	5453,982	*
Dermaseptin-SP13	ALWKDILKKIGTVALHAGKAAFGAAADTISQ-NH ₂	sec	*	*	31	*	3161,776	*
Dermaseptin-SP14	AAWTSIIKNMAKNAAKKVAKAAANAAKAAANAIAK SVNNMANQEEQESKKM	sec	*	*	52	*	5369,801	*

Dermaseptin-SP15	GMWSTIKEIGIAAAKEAAKAAGKTISDVIKQ-NH ₂	sec, ms ²	100	75/334	31	3184,364	3184,754	48
Dermaseptin-SP16	GMWSKIKDVGKSAVKAAAKAAGKAALGALSEAV-NH ₂	cm, sec, ms ²	100	60/346	33	3182,424	3182,785	31
Dermaseptin-SP17	GLWSKIKQVGKEAAKAAGKAALNAVTDVVNQ-NH ₂	sec	*	*	31	*	3163,772	*
Dermaseptin-SP18	GLWSKIKQVGKEAAKAAEVRKRL	sec	*	*	20	*	2168,232	*
Dermaseptin-SP19	GLWSKIKQVGKEAAIAAGKAALGAL-NH ₂	sec	*	*	25	*	2449,443	*
Dermaseptin-SP20	GLLKSLKKVGIAAGKAVLNKVTDMVNQ-NH ₂	sec	28	*	28	*	2906,736	*
Dermatoxin-AS1	SFGSFMKGVGKGLATVGKFKVANKFGKLLEA-NH ₂	cm, ms ²	*	*	30	*	3086,700	*
Tryptophyllin- AST1	DKPPAWVP-NH ₂	cm, sec	*	*	8	*	908,481	*
Tryptophyllin- AST1-2	DKPPSWVP	cm, sec	*	*	8	*	924,460	*
Tryptophyllin AST2-1	DMSPPF-NH ₂	cm, sec	*	*	7	*	877,368	*
Tryptophyllin AST2-2	DTSPPF-NH ₂	cm, sec	*	*	7	*	847,375	*
Tryptophyllin-AT2- 3	DMDPPFNR	cm, sec	*	*	8	*	990,412	*

Tryptophyllin- AST3	QGKPYWPPPFLPE	cm, sec	*	*	13	*	1554,776	*
Tryptophyllin- AST3-2	QGKPYWPPPLPE	cm, sec	*	*	12	*	1407,708	*
Phylloseptin-SP1	FLSLIPHVISAIPHVVNALSNL-NH ₂	cm, sec, ms ²	100	67/210	22	2352,061	2352,357	52
Phylloseptin-SP2	FLSLIPHAINAVSALAKHF-NH ₂	cm, sec	*	*	19	*	2047,162	*
Phylloseptin-SP3	FLGIIPVAISALSRL-NH ₂	cm	*	*	15	*	1567,971	*
Phylloseptin-SP4	FLSLIPHVISAIPHVVAISSL-NH ₂	sec, ms ²	100	40/162	22	2298,053	2298,335	37
Phylloseptin-SP5	FLSLIPHVISAIPHVVNALSG-L-NH ₂	sec, ms ²	100	50/202	22	2295,075	2295,336	50
[Ser6, Val10, Asp11]- phyllokinin	RPPGFSPFRVD	cm, sec, ms ²	100	35/106	11	1273,498	1273,646	28
[Ser6, Val10, Asn11]-phyllokinin	RPPGFSPFRVN	cm	*	*	11	*	1272,662	*
[Thr6, Ile10, Tyr11]-phyllokinin	RPPGFTPFRIV	cm, sec	*	*	11	*	1349,714	*
[Thr6, Val10, Asp11]-phyllokinin	RPPGFTPFRVD	cm, sec, ms ²	100	41/106	11	1287,454	1287,661	18

Medusin-AS	LLGMIPLAISALS ₂ SKL-NH ₂	cm, sec, ms ²	100	60/148	18	1808,787	1809,105	63
Plasticin-AS1	GLVSGLLNTAGGLIGNLL-NH ₂	cm, ms ²	100	46/162	18	1679,698	1679,983	15
Plasticin-AS2	GLLSGLLSTVTGLLGNLGGGGL	sec	*	*	22	*	1968,115	*
Plasticin-AS3	GLLSGLLGTVTGLIKSVL-NH ₂	sec	*	*	18	*	1739,081	*
Caerin peptide-AS1	GMWGTLFKGATAVAKHILPHVF-NH ₂	cm, sec, ms ²	100	42/194	22	2378,740	2379,293	31
Sauvateide-AS1	FRPAILVRPK	cm, ms ²	85	32/86	14	1380,572	1379,877	*
Orphan Peptide-SP1	IIGMIPDLISMISKL-NH ₂	cm, sec, ms ²	100	51/124	15	1641,704	1641,945	68
Orphan Peptide-SP2	FLFLPFI-NH ₂	cm, sec, ms ²	100	12/24	7	894,192	894,526	28
Orphan Peptide-SP3	FLFPPFF-NH ₂	cm, ms ²	100	5/12	7	913,272	913,479	29
Orphan Peptide-SP4	IIGMIPDLISIFSKL-NH ₂	cm, ms ²	100	39/124	15	1657,601	1657,973	62
Orphan Peptide-SP6	FLFPPFLEEKSNML	cm	*	*	14	*	1710,858	*
Orphan Peptide-SP7	LLEGLLDVTNLVGSIV	sec, ms ²	100	34/152	17	1754,525	1754,992	97

Orphan Peptide- SP8	LLMDTLSKATNLVDNVANSA-NH ₂	sec	*	*	20	*	2088,077	*
Orphan Peptide- SP9	LMYYTLPRPV-NH ₂	sec	*	*	10	*	1250,674	*
Orphan Peptide- SP10	LLGFLNK	sec, ms ²	100	*	8	*	916,564	*
Orphan Peptide- SP11	IKGLSPLR-NH ₂	sec	*	*	12	*	1339,773	*
Orphan Peptide- SP12	IDLPVSVLFELINTL	sec, ms ²	42	*	26	*	2968,656	*

*= no encontrado, cm= clonaje molecular, sec= secuenciación de ARN dirigido, ms²= espectrometría de masas, -NH₂= amidación en el extremo C-terminal.

Realizado por: Jessica Verdezoto Prado, 2024.

3.3. Análisis de propiedades fisicoquímicas

3.3.1. Caracterización fisicoquímica de los péptidos nuevos.

Las características fisicoquímicas se obtuvieron tras el análisis *in silico* apoyado en herramientas bioinformáticas y se presentan a continuación por familias de PAMs. Para esta presentación se tomó en cuenta únicamente a los péptidos que fueron confirmados por los tres métodos de búsqueda, CM, TR y MS².

Dermaseptinas.

Tabla 24. Características fisicoquímicas de los péptidos correspondientes a la familia de dermaseptinas.

Péptido	Péptido maduro/Estructura secundaria	# AAs	Alpha helix (%)	Hidrofobicidad <H>	Momento hidrofóbico <μH>	Carga a pH 7	Masa teórica (Da)
	1 25. 34.						
	h t						
Dermaseptin-SP7	A S W K D F F K N I G K A A G K A V L N S V T D M V N Q a	28	96,43	0,290	0,349	+ 3	3037,570
Dermaseptin-SP8	 E V N	28	96,43	0,301	0,344	+ 4	3036,586
Dermaseptin-SP2	 V . L .	28	78,57	0,358	0,275	+ 4	3003,585
	h h h h h h h h h h h h h h h h h h t t c c e h h h h h h						
Dermaseptin-SP15	G M W S T I K E I G I A A A K E A A K A A G K T I S D V I K Q a	31	83,87	0,293	0,317	+ 3	3184,754
	h h						
Dermaseptin-SP16	G M W S K I K D V G K S A V K A A A K A A G K A A L G A L S E A V a	33	100	0,242	0,202	+ 5	3182,785
	h h						
Dermaseptin-H3	G L W S T I K N V G K E A A I A A G K A A L G A L	25	100	0,414	0,214	+ 3	2409,364

	h h							
Dermaseptin-SP10	GM F T N M L K G I G K L A G Q A A L G A V K T L V a	26	100	0,510	0,328	+ 4	2587,460	
Dermaseptin-A4	 A .	26	100	0,475	0,355	+ 4	2559,429	
	h h							
Dermaseptin-SP20	G L L K S L L K K V G I A A G K A V L N K V T D M V N Q a	28	100	0,372	0,159	+ 5	2906,736	
Dermaseptin-H1	. . W N . . V A . . A	28	96,43	0,399	0,189	+ 3	2866,574	

. = aminoácidos conservados, c = coli aleatorio, h = alpha hélix, e = cadena externa, t = giro beta, a= amidación en el extremo C-terminal.

Realizado por: Jessica Verdezoto Prado, 2024.

Filoseptinas.

Tabla 25. Características fisicoquímicas de los péptidos correspondientes a la familia de filoseptinas.

Péptido	Péptido maduro/Estructura secundaria	# AAs	Alpha helix (%)	Hidrofobicidad <H>	Momento hidrofóbico <μH>	Carga a pH 7	Masa teórica (Da)
	1 23. h						
Phylloseptin-SP4	F L S L I P H V I S A I P H V V S A I S S L a	22	100	0,903	0,474	+ 1	2298,335
Phylloseptin-SP5	 N . L . G . .	22	100	0,875	0,489	+ 1	2295,336
Phylloseptin-SP1	 N . L . N . .	22	100	0,848	0,516	+ 1	2352,357

. = aminoácidos conservados, h = alpha hélix, a= amidación en el extremo C-terminal.

Realizado por: Jessica Verdezoto Prado, 2024.

Plasticinas.

Tabla 26. Características fisicoquímicas de los péptidos correspondientes a la familia de plasticinas.

Péptido	Péptido maduro/ Estructura secundaria	# AAs	Alpha helix (%)	Hidrofobicidad <H>	Momento hidrofóbico <μH>	Carga a pH 7	Masa teórica (Da)
	1 h h h h h h h h h t t h h h h h h h 19.						
Plasticin-AS1	GLVSGLLNTAGGLIGNLLa	18	88,89	0,697	0,492	+ 1	1679,983
Plasticin-A1	 L . D . . .	18	100	0,682	0,493	0	1680,967

. = aminoácidos conservados, h = alpha hélix, t = giro beta, a= amidación en el extremo C-terminal.

Realizado por: Jessica Verdezoto Prado, 2024.

Caerinas.

Tabla 27. Características fisicoquímicas de los péptidos correspondientes a la familia de caerinas.

Péptido	Péptido maduro/Estructura secundaria	# AAs	Alpha helix (%)	Hidrofobicidad <H>	Momento hidrofóbico <μH>	Carga a pH 7	Masa teórica (Da)
	1 t c h h h h h t t h h h h h h h h c c t h c 23.						
Caerin Peptide-AS1	GMWGT L FKGATAVAKH I LPHV F a	22	63,64	0,689	0,380	+ 3	2379,293
Caerin-regulated	 V . . . I K T L S	23	73,91	0,640	0,419	+ 3	2552,398

. = aminoácidos conservados, c = coli aleatorio, h = alpha hélix, t = giro beta, a= amidación en el extremo C-terminal.

Realizado por: Jessica Verdezoto Prado, 2024.

Péptidos huérfanos.

Tabla 28. Características fisicoquímicas de los péptidos que no encajaron en las familias de PAMs.

Péptido	Péptido maduro/Estructura secundaria	# AAs	Alpha helix (%)	Hidrofobicidad <H>	Momento hidrofóbico <μH>	Carga a pH 7	Masa teórica (Da)
	17. c c c c c c c c c c c c c c c c						
OrphanPeptide-SP4	I I G M I P D L I S I F S K L a	15	0	0,953	0,645	+ 1	1657,973
OrphanPeptide-SP1	 M I	15	0	0,916	0,650	+ 1	1641,945
	h h h h h h h h h h h h h h h h						
OrphanPeptide-SP7	L L E G L L D T V T N L V G S I V	17	100	0,731	0,622	- 2	1754,992
	c c c c c c c c c c c c c c c c						
OrphanPeptide-SP12	I D L P V S V L F E L I N T L	26	0	0,904	0,534	- 2	2968,66

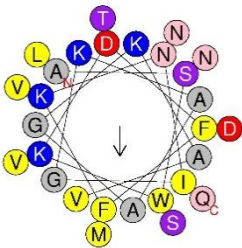
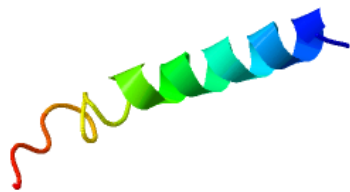
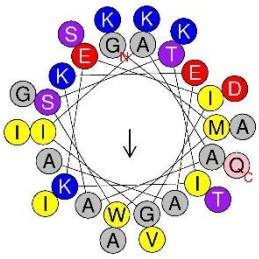
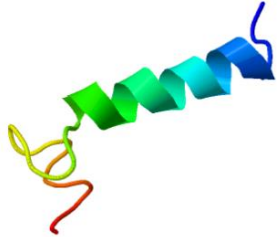
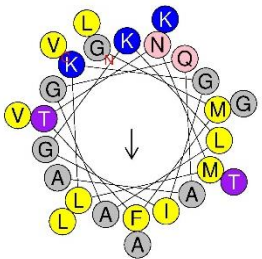
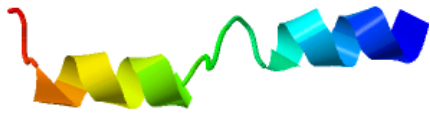
. = aminoácidos conservados, c= coli aleatorio, h= alpha hélix, a= amidación en el extremo C-terminal.

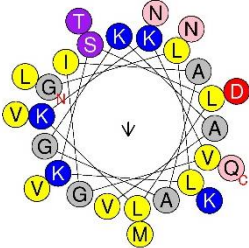
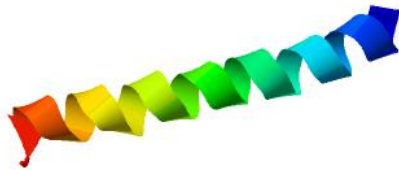
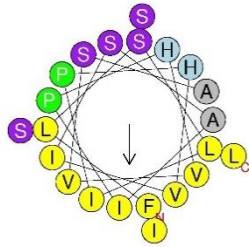
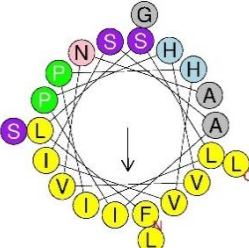
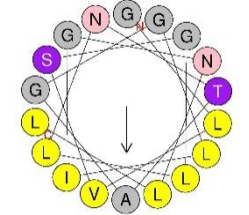
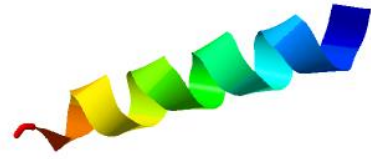
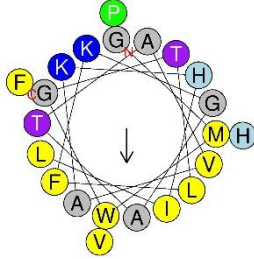
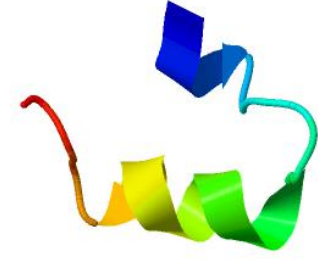
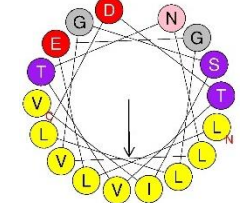
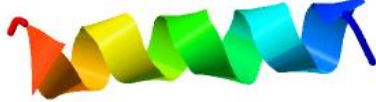
Realizado por: Jessica Verdezoto Prado, 2024.

3.3.2. Predicción de estructura secundaria de los nuevos PAMs.

La predicción de la estructura secundaria y la rueda helicoidal se llevó a cabo utilizando softwares de acceso libre, tomando en cuenta el análisis previo de la estructura secundaria y el porcentaje de hélice alfa. A continuación, en la tabla 29, se presentan las ruedas helicoidales con sus características anfipáticas y hélice alfa; en conjunto con las predicciones de estructura tridimensional para los péptidos con mejores características fisicoquímicas de actividad antimicrobiana previamente calculadas.

Tabla 29. Predicción de estructuras para los nuevos péptidos. La rueda helicoidal presenta diferentes colores para cada aminoácido mismos que indican la proyección de cada uno: amarillo = no polar hidrofóbico, en gris = Gly y Ala (no polar hidrofóbico), azul = básico, morado = polar sin carga, verde = Pro (apolar sin carga) y verde claro = His.

Péptido	Predicción de la rueda helicoidal	Predicción de la estructura 3D
Dermaseptin- SP7 DRS-SP7		
Dermaseptin- SP15 DRS-SP15		
Dermaseptin- SP10 DRS-SP10		

Dermaseptin- SP20 DRS-SP20		
Phylloseptin- SP4 PLS-SP4		
Phylloseptin- SP5 PLS-SP5		
Plasticin-AS1 PTC-AS1		
Caerin peptide- AS1 CRP-AS1		
OrphanPeptide- SP7 OP-SP7		

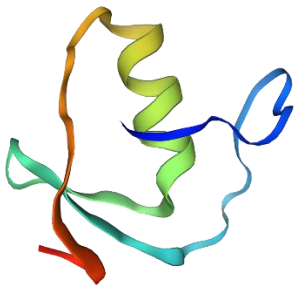
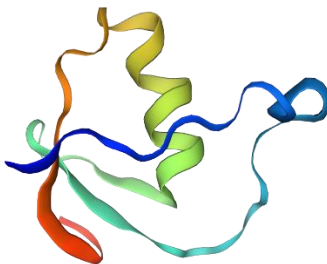
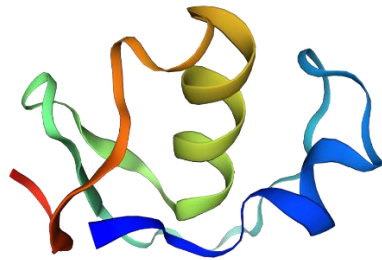
Realizado por: Jessica Verdezoto Prado, 2024.

La predicción de la estructura de cada péptido fue seleccionada en base al c-score proporcionado por el mismo programa. El c-score del modelado es el porcentaje de confianza para estimar la calidad que presenta cada modelo en base a las plantillas de ensamblaje de la secuencia. Se eligió el modelo con valor más positivo que representa el nivel de confianza de la predicción [61].

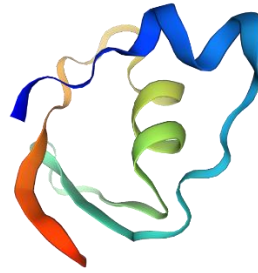
3.3.3. Predicción de estructura terciaria de las nuevas proteínas

La predicción de la estructura terciaria y actividad se llevó a cabo utilizando softwares de acceso libre. En la Tabla 30 se presentan las predicciones de estructura tridimensional para las proteínas de tipo Kazal más representativas del estudio.

Tabla 30. Predicción de estructura 3D de las proteínas de tipo kazal encontradas.

Péptido	Predicción de la estructura 3D
ASKP-2	
ASKP-3	
ASKP-4a	

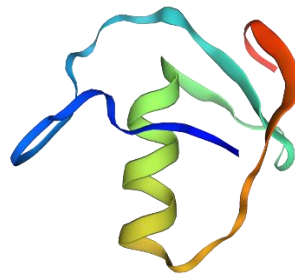
ASKP-7b



ASKP-8



ASKP-9d



Realizado por: Jessica Verdezoto Prado, 2024.

4. CAPÍTULO IV: INTERPRETACIÓN Y DISCUSIÓN

La aplicación de secuenciación dirigida de ARN es un método alternativo de alto rendimiento para la identificación de péptidos antimicrobianos que ofrece varias ventajas. Algunas de ellas son: la rapidez y eficiencia que presenta en comparación con otros métodos tradicionales, permitiendo la identificación de múltiples péptidos en un solo experimento, secuenciación. Además, se caracteriza por ser menos invasivo, lo que reduce el daño a las muestras biológicas, debido a que la cantidad de muestra empleada es mínima en relación a otras técnicas. Por otro lado, al utilizar la amplificación selectiva con primers específicos de PAMS a partir de bibliotecas de ADNc, es posible identificar una gran cantidad de péptidos de manera simultánea y con alto rendimiento [24,25].

Sin embargo, algunas de las limitaciones que se encontraron tras aplicar la tecnología de secuenciación de ARN dirigido en el presente estudio fueron. La dependencia de primers específicos para la amplificación de regiones de interés en el ADNc ya que se necesita diseñar algunas secuencias para encontrar las candidatas óptimas a primers específicos. Por otro lado, la creación de bibliotecas de ADNc y su posterior amplificación son pasos claves para obtener las secuencias precursoras de péptidos por lo que es inevitable contar con las habilidades, recursos y tiempo para llevarlas a cabo. Puesto que son técnicas avanzadas y de complejidad en estudios de biología molecular [62,63].

4.1. Preparación y secuenciación de bibliotecas de ADNc

4.1.1. Tratamiento de la muestra

La secreción cutánea de los anfibios posee moléculas bioactivas de gran interés, destacando especialmente los péptidos antimicrobianos. La mayoría de estos péptidos han sido identificados en anfibios de las familias Ranidae, Pipidae, Liopelmatidae, Alytidae, Bombinatoridae, Myobatrachidae, Hylidae, entre otras. En particular, la subfamilia *Phyllomedusinae* es ampliamente reconocida por su significativa

contribución como fuente de PAMs. Sin embargo, las técnicas de descubrimiento y la escasa información limitan en gran medida la identificación de los nuevos péptidos. La secuenciación por nanoporos promete reducir la brecha al proporcionar mayor diversidad de datos con alta calidad respecto a la poca cantidad de muestra analizada [31,64–66].

Después de realizar el tratamiento y limpieza de los datos de secuenciación, se obtuvieron 271.941 lecturas con una longitud total de 114.961.147 pb. Sin embargo, al intentar comparar estos datos para estimar la cantidad de secuencias obtenidas, no fue posible hacerlo con precisión, ya que el tamaño del genoma de *A. spurrelli* aún no ha sido descrito. Esto se debe a la falta de datos completos en los genomas referenciales de anfibios para la mayoría de especies [67]. A pesar de ello, es posible hacer una aproximación al comparar estos datos con el estudio de Andrade y colaboradores en 2022, en el que realizaron un borrador del genoma y un ensamblaje transcriptómico multi-tejido de la rana de hoja neotrópica *Phyllomedusa bahiana*, una especie perteneciente a la subfamilia *Phyllomedusidae*, cercana a *A. spurrelli*. Andrade reportó que *P. bahiana* tiene un tamaño total de genoma de 4,74 Gbp, mientras que para los datos de secuencias provenientes de la piel (dorso), se reportaron 55.109.630 pb. Esto sugiere que los datos obtenidos para *A. spurrelli* son relativamente pequeños. Esta diferencia podría deberse a que en el presente estudio se secuenció el ARNm de la secreción de la piel, y no de la piel como órgano completo, lo que podría explicar una discrepancia significativa en la cantidad de secuencias obtenidas [68].

4.1.2. Reporte de la calidad de datos de secuenciación

La secuenciación de molécula única mediada por nanoporos de Oxford Nanopore Technologies (ONT) permite detectar modificaciones en nucleótidos individuales lo que se traduce en la captura de diversidad de las moléculas y la fidelidad de las mismas. Al seleccionar un base-calling estricto (Dorado base-calling), que prioriza la calidad y precisión durante la secuenciación, se logra identificar variantes en las bases. Wescoe y colaboradores en 2014 mencionaron la capacidad de discriminación que puede lograr la

secuenciación por nanoporos tales como la C-5 de citosina (C, 5mC, 5hmC) y 6-metiladenina (6mA), obteniendo precisión de identificación entre 92% y 98% [47,69,70].

Respecto al índice de calidad inicial Q 11,4 obtenido usando NanoPlot (Figura 2), indica que el porcentaje de error es de 0.079, por ende, la interpretación o llamada de bases es escasamente incorrecta. Es decir que más del 90% de las bases fueron identificadas correctamente, lo que se traduce en que los datos de secuenciación son fieles a la muestra [71,72]. Adicionalmente se puede decir que es un buen índice de partida, puesto que posteriormente en el análisis bioinformático de limpieza se eliminará toda la información que genera ruido en los datos de secuenciación. Trivedi y colegas en 2014 describieron el conjunto de herramientas bioinformáticas y flujo de trabajo que puede aplicarse a los datos de secuenciación para mejorar el control de calidad. Mientras, que Linde y colaboradores en 2023 realizaron la comparación de la tecnología ONT frente a Illumina para analizar el genoma de tres bacterias, como resultados se obtuvo que la celda de flujo Flongle FLO-FLG114 presentaba mayor precisión en la secuenciación a diferencia de la celda 9.4.1, alcanzando datos similares a los datos de Illumina. Esto sugiere que el control de calidad apoyado en herramientas bioinformáticas y empleando una celda de última generación, una celda de flujo R10.4.1, puede alcanzar resultados de mayor precisión y confiabilidad [73–75].

4.2. Diseño de pipeline para limpieza y tratamiento

4.2.1. Secuencias precursoras obtenidas

Se realizó un análisis bioinformático de los datos de secuenciación, lo que resultó en el diseño de un flujo de trabajo para la limpieza y el tratamiento de los datos (Figura 4). Este flujo de trabajo (pipeline) permite recortar, eliminar y filtrar las secuencias precursoras, orientarlas, organizarlas en marcos de lectura abiertos (ORFs) y formar consensos de manera rápida, eficiente y accesible utilizando herramientas bioinformáticas de libre acceso. Sin embargo, se identificaron algunas limitaciones durante el desarrollo del pipeline. Una de ellas fue en el paso de formación de

consensos, debido al gran volumen de información de secuenciación. Otra limitación se observó en el paso de organización en ORFs, donde las secuencias se convierten de nucleótidos a aminoácidos, lo que genera “copias” de secuencias similares. Estas secuencias no pueden verificarse fácilmente para determinar si pertenecen a la misma secuencia o si la secuencia nucleotídica es diferente. Se recomienda implementar un paso de control antes y después de organizar las secuencias en ORFs para abordar esta problemática [50].

El análisis bioinformático, permitió distinguir secuencias precursoras de varios péptidos y proteínas de interés (Tabla 3, 4) entre los cuales se caracterizaron péptidos antimicrobianos correspondientes a 7 familias, y proteínas inhibidoras de proteinasa de tipo Kazal. Azevedo y colaboradores en 2011 mencionaron que los anuros de la subfamilia *Phyllomedusinae* resultan una fuente rica de moléculas peptídicas, pudiendo encontrarse con filoquininas o filocaerulinas, péptidos representativos de esta familia. Por otro lado, Proaño y colegas en 2019 lograron reportar péptidos pertenecientes a 4 familias de PAMs en la secreción de *A.spurrelli* empleando clonaje molecular, en este estudio destacaban los correspondientes a la familia dermaseptinas [5,18].

De igual manera, las proteínas encontradas exhibieron una gran diversidad de moléculas pertenecientes a proteínas inhibidoras de tipo Kazal. Wang y colaboradores en 2011, describen que la secreción cutánea de anfibios contiene inhibidores de proteasa con motivos Kunitz, Kazal y Bowman-Birk. Mientras que, Proaño y colegas en 2017 encontraron 18 nuevas proteínas de tipo Kazal en la secreción cutánea de *Cruziohyala calcarifer* demostrando que las especies del género *Phyllomedusa*, simbolizan una fuente prometedora de proteínas inhibitorias de proteasas. Sin embargo, para la especie *A.spurrelli* no se pudo localizar registros por lo que las proteínas descubiertas en este estudio se reportan por primera vez [76,77].

4.2.2. Clasificación de PAMs encontrados

En base a los resultados del BLAST se lograron clasificar las secuencias en familias de péptidos antimicrobianos que se describen a continuación, por el grado de similitud que presentaron las secuencias, además de poseer las regiones conservadas propias de las familias descritas. Por otro lado, las secuencias que no obtuvieron un buen porcentaje de similitud (60% +) con péptidos de la base de datos del NCBI fueron agrupados como péptidos huérfanos. Esto quiere decir que posiblemente pertenecen a familias de péptidos no reportadas, son similares a otras moléculas que no son péptidos o en el su defecto tuvieron errores en la conformación de los marcos de lectura de las secuencias durante el paso de limpieza y tratamiento de la muestra. Para resolver estas dudas se sugiere repetir el ensayo y verificar los procesos de limpieza descritos en el pipeline, adicionalmente comparar los datos obtenidos con otras secuencias precursoras que codifican a otras moléculas (proteínas de transporte, proteínas inhibidoras y otras) [54].

Dermaseptinas

Las secuencias peptídicas encontradas mediante secuenciación de ARN dirigido aplicada a la secreción cutánea de *A. spurrelli* fueron examinadas con la base de datos de NCBI empleando la herramienta BLAST (Basic Local Alignment Search Tool) que permitió la comparación por similitud de las secuencias. De manera que se obtuvieron 14 secuencias (Tabla 6) que compartieron entre 78% y 97% de similitud con las características correspondientes a dermaseptinas, cumplen con estos elementos característicos. Las dermaseptinas son una familia representativa de los péptidos antimicrobianos procedentes de ranas *Phyllomedusa*. Contienen un péptido maduro de entre 22 a 28 aminoácidos (aa) de longitud con regiones conservadas: en su 3er residuo Trp (W) y un motivo consenso AA(X)KAAL(X)A en la región media, donde X corresponde a cualquier aa. Además, frecuentemente poseen una modificación postraduccional Gly (G) donadora de amida en el extremo C-terminal [6,16]. Las nuevas dermaseptinas fueron nombradas DRS-SP7, DRS-SP10 hasta DRS-SP20 conforme a la nomenclatura propuesta por Amiche y colaboradores en 2008 [6].

Triptofilinas

Este estudio descubrió 6 secuencias peptídicas (Tabla 8) que compartieron entre 74% y 91% de similitud con péptidos correspondientes a triptofilinas, se separaron por grupos según sus concordancias con los motivos conservados. Las triptofilinas se caracterizan por ser secuencias cortas de entre 4 a 14 residuos y tener residuos de Trp y Pro. Se dividen en tres grupos identificados como: T-1, secuencias de entre 7 a 8 aa que en su N-terminal poseen Lys y Pro (KP) mientras que en su interior tienen variantes Pro-Trp (P-W); T-2, secuencias cortas de entre 4 a 7 aa pero contienen su interior conservado con Pro-Trp (P-W) y en algunos casos su extremo C-terminal amidado; T-3, secuencias de entre 13 y 14 aa con residuos conservados en la posición 5 Tyr/Phe (Y/F), en su interior residuos de Pro-Pro-Pro (PPP). Por tal motivo las nuevas triptofilinas se nombraron como triptofilinas T1, T2 y T3 según correspondieron a los grupos [78,79].

Filoseptinas

Para el caso de las filoseptina, el secuenciamiento de ARN dirigido de la secreción de *A. spurrelli* debeló 4 secuencias peptídicas (Tabla 10) con similitud entre 94% y 98% a secuencias peptídicas de filoseptinas, es decir que compartían las características de esta familia de péptidos. Presentaron secuencias de entre 18 a 20 aminoácidos cuyos residuos conservados son Phe-Leu-Ser-Leu-Ile (FLSLI) en el extremo N-terminal y una Gly (G) donante de un grupo amida en el extremo C-terminal. Sin embargo, de las cuatro secuencias obtenidas solo dos fueron denominadas nuevas, PLS-SP4 y PLS-SP5, debido a que no presentaron registros previos [16,80].

Filoquininas

Se identificaron 3 secuencias (Tabla 12) que compartieron similitud entre 91% y 100% con las características de la familia de filoquininas por loque fueron cotejadas en esta familia. Las filoquininas son consideradas como variantes de las bradiquininas (bradykinyl-Ile-Tyr(SO₃H) y provienen del no péptido (RPPGFSPFR), adicionalmente

contienen una secuencia precursora de cuatro aminoácidos (ESPE/ESPD) en el extremo C-terminal. Por tal motivo la nueva filoquinina identificada en el presente estudio fue denominada [Thr6, Ile10, Tyr11]-phyllokinin y compartió las características conservadas en los residuos 1-5, 7-9 de esta familia de péptidos. Adicionalmente presentó una modificación postraducciona de sulfatación en los residuos Tyr en el extremo N-terminal [16,81].

Plasticinas

Respecto a las 2 secuencias peptídicas nuevas (Tabla 16) que presentaron entre 85% y 98% de similitud con la plasticina-S1 de *Phyllomedusa sauvagii*, fueron denominadas como PTC-AS2 Y PCT-AS3 debido a que compartieron las características conservadas de esta familia de péptidos. Las plasticinas se caracteriza por contener al menos 23 aa y ser péptidos ricos en Gly, Leu (G, L), además de poseer un motivo consenso GXXXG, donde X corresponde a cualquier aa. Algunas secuencias contienen una modificación postraducciona Gly (G) donadora de amida en el extremo C-terminal [6,16].

Caerinas

El nuevo péptido denominado Caerin Peptide-AS1 presentó 98% de similitud con el caerin-regulated peptide de *Agalychnis callidryas* (Tabla 18), es decir que compartieron motivos conservados en el extremo N-terminal y la característica de contar con una longitud de entre 23 a 27 aa. Por tal motivo fue clasificado dentro de esta familia [82].

Péptidos huérfanos

Los péptidos que no encajaron en una de las familias de PAMs presentadas, fueron agrupados y denominados como péptidos huérfanos, nombrados así 8 péptidos como: OP-SP1, OP-SP2, OP-SP7 al OP-SP12 (Tabla 20). Estos péptidos no compartieron características significativas con una sola familia, sino por el contrario presentaron

características similares a péptidos de medusinas, péptidos SP, triptofilinas y bradikininas, lo que les impide ser clasificados en una familia determinada.

4.2.3. Clasificación de proteínas encontrados

En el presente estudio se identificaron 9 proteínas inhibidoras de tipo Kazal con sus respectivas variantes (Tabla 22), mediante secuenciación de ARN dirigido aplicada a la secreción cutánea de *A. spurrelli*. Las proteínas que se describen en este estudio comparten características conservadas propias de las proteínas inhibidoras de tipo Kazal reportadas en la subfamilia Phyllomedusinae teniendo a la PSKP-1 de *Phyllomedusa sauvagii* como una de sus principales representantes. Este tipo de proteínas se caracteriza por su longitud de entre 57 a 60 aa y contener 6 residuos de Cys (C) como motivo conservado CX₆-CX₁₅-CX₂₃-YX₃₀-CX₃₄-CX₃₇-C. Por otro lado, dependiendo de la actividad inhibitoria la proteína puede tener Pro, Asn, Asp, Ser (P, N, D, S) en su P2 y Lys, Arg, Leu, Ser, Asp (K, A, L, S, D) en su P1. Li y colaboradores en 2012, reportaron una proteína de tipo Kazal con actividad inhibitoria de tripsina (ACKTI) en *Agalychnis callidryas*, que a diferencia de la PSKP-1 presentaba Pro y Arg (P y A) en los motivos conservados P2 y P1 respectivamente [16,83].

El estudio más reciente realizado por Proaño y colaboradores en 2017, supone un referente para la comparación de proteínas inhibidoras de tipo Kazal presentes en la secreción cutánea de anfibios. Ya que aquí se reportaron 18 nuevas proteínas de tipo Kazal en la secreción de *Cruziohyla calcarifer*, nombradas CCKPs con sus respectivas variantes, correspondiendo a la nomenclatura previamente establecida y a su actividad inhibitoria. Destacando CCKP-1 por haber sido encontrada por espectrometría de masas y además de poseer actividad inhibitoria de tripsina [77].

4.2.4. Análisis de PAMs por espectrometría de masas

Tras la comparación de los datos obtenidos por clonaje molecular de estudios previos y secuenciación de ARN dirigido del presente trabajo con datos de espectrometría de

masas en tándem se logró confirmar la estructura de al menos 25 secuencias de péptidos presentes en la secreción cutánea de *A. spurrelli* (Tabla 23). De los cuales 21 de ellos lograron 100% de identificación, lo que se traduce a la fidelidad de las técnicas combinadas y robustez de las mismas. Por otro lado, las secuencias no reveladas o no descritas podrían estar sujetas a la baja expresión de las mismas o inclusive al tipo de muestra empleada, esto se debe principalmente a que las muestras pertenecen a distintos especímenes biológicos, lo que se traduce en diferencias significativas en cuanto a la madurez, desarrollo, género de los individuos y adicionalmente su alimentación [66,79].

Las muestras de secreción cutánea pertenecían a especímenes colectadas en cautiverio, lo que también aporta una variante en los resultados peptídicos obtenidos del material analizado. Puesto que un individuo colectado en cautiverio no presenta las mismas condiciones de estrés climático, alimentación, interacción y microbiota que un individuo colectado en campo. Esto se debe a que en cautiverio no enfrenta las mismas amenazas respecto a depredadores en su hábitat natural, por lo que no requieren activar los mismos mecanismos de defensa. Además, que la alimentación no autóctona y mantenimiento de los terrarios generan una reducción de secreción cutánea y alteración en los componentes que se encuentran en ella. El emplear individuos provenientes de cautiverio supone un limitante en la búsqueda de diversidad peptídica, por tal motivo se recomienda usar muestras de especie colectadas en campo y que estas pertenezcan a un número superior a dos individuos. [84,85]

4.3. Análisis de propiedades fisicoquímicas

4.3.1. Caracterización fisicoquímica de los péptidos nuevos

Las características fisicoquímicas de los péptidos confirmados mediante MS/MS son presentadas por familias donde la mayoría de PAMs han demostrado actividad antimicrobiana contra varias especies de hongos, bacterias, virus y parásitos. Este efecto se debe a que la mayoría de PAMs comparten cualidades respecto a su carga neta,

hidrofobicidad, afinidad con la membrana, entre otras. Respecto a la carga neta la mayoría de péptidos son catiónicos, es decir que poseen un rango de carga positiva entre +3 a +11, condición que les permite conectarse por afinidad a la superficie de la membrana bacteriana cargada negativamente [86].

Las dermaseptinas y sus subfamilias, filoseptinas, dermatoxinas presentan alta actividad antimicrobiana contra una gama amplia de microorganismos tanto en entornos silvestres como *in vitro*. Esta bioactividad es proveniente de la cationicidad de la molécula y de su hidrofobicidad, donde la cationicidad es directamente proporcional a la bioactividad antimicrobiana. Además, la amidación que poseen las dermaseptinas en su extremo C-terminal agrega una carga positiva a la molécula potenciando su actividad y estabilizando su estructura secundaria favoreciendo a la conformación hélice alfa del péptido [87].

Por otra parte, los péptidos pertenecientes a filoseptinas son altamente conservados, sin embargo, han demostrado que pese a tener pequeños cambios en los aminoácidos pueden lograr diferencias notables en la bioactividad. Al igual que las dermaseptinas, las filoseptinas tienen a menudo una amidación postraducciona y adoptan la estructura hélice alfa estable por la unión electrostática con la membrana lipídica. El formar esta estructura facilita el ingreso del péptido en la membrana celular, causando inestabilidad en la misma, incluso lisándola [5,16].

La cationicidad de los péptidos de la familia plasticinas es importante porque permite que estos se unan a la membrana celular y a los lípidos, por medio de la atracción electrostática de las cargas. Además, la interacción con los aminoácidos aromáticos, como el triptófano (Trp), pueden relacionarse a través de π -catión interacciones e influir en la distribución de los aminoácidos en la membrana y en la bioactividad [88]. Respecto a las caerinas, estas contienen residuos cargados positivamente que otorgan cationicidad y afinidad por la membrana bacteriana que posea carga negativa. Además, suelen adoptar conformación hélice alfa manteniendo una región flexible que les permite plegarse libremente, este plegamiento le otorga la característica anfipática que

resulta de suma importancia en el momento de la interacción del péptido con la membrana [89,90].

Finalmente, en esta sección se presentó únicamente los péptidos que fueron confirmados por MS/MS debido a que esta técnica de “confirmación” permite evidenciar con certeza las moléculas que se encontraron presentes en la muestra de secreción cutánea de *A. spurrelli*. Lo que nos indica que no se encontraron la mayoría de las secuencias por este método debido a que los ensayos fueron alejados, la muestra utilizada en secuenciación de ARN dirigido fue diferente a la muestra inyectada en el espectrómetro de masas. Para obtener 100% de identificación se recomienda emplear la misma muestra proveniente de los mismos individuos y en el mismo periodo de tiempo, además de mejorar el proceso de secuenciación y amplificación de ADNc [11].

4.3.2. Predicción de estructura secundaria de los nuevos PAMs

Las herramientas *in silico* permiten evaluar, modelar y predecir la estructura y posible actividad de las moléculas, en este caso secuencias peptídicas. Surgen como alternativa eficaz para reducir costos, acelerar procesos de investigación y mejorar la comprensión de la interacción molecular. HeliQuest y I-TASSER como herramientas de predicción presentadas en este estudio, brindan información valiosa para futuros ensayos con los péptidos encontrados. Son la base para aplicar procesos de evaluación antimicrobiana, diseño y síntesis de péptidos, en búsqueda de alternativas terapéuticas [91,92]. Adicionalmente se empleó I-TASSER porque es un software de confianza que ha ganado el primer lugar en predicciones de modelado de estructuras, debido a que usa como parámetros el c-score y el TM-score. El c-score representa el grado de confianza para estimarla calidad de las predicciones y el TM-score es la escala usada para medir la similitud estructural, es decir la precisión de las estructuras [61,93].

La estructura de los péptidos en medio acuoso puede ser indefinida, sin embargo, al conectarse con la membrana adoptan su estructura secundaria, que puede ser en forma de hélice alfa o lamina beta, potenciando su bioactividad. Las dermaseptinas pueden

adoptar una estructura hélice alfa en su extremo N-terminal al unirse a la membrana celular, no obstante, el poseer este tipo de estructura no determina únicamente la bioactividad, sino que en conjunto con la característica anfipática y la hidrofobicidad actúan en la interacción con la membrana [59].

A las plasticinas se las relaciona con las dermaseptinas debido a compartir la cualidad de ser péptidos ricos en G y L sin embargo tienen mayor diversidad estructural. Las plasticinas pueden adoptar conformaciones de hojas beta (β -sheets), horquilla beta (airpin- β) u otras formas en relación al entorno en el que se encuentren. Estas conformaciones les permiten adherirse a la membrana celular y alterar funciones específicas. Las plasticinas encontradas en *P. bicolor* y *P. sauvagei* (PTC-B1 y PTC-S1) presentan alta cationicidad por ende alto nivel de bioactividad incluyendo actividad antimicrobiana y hemolítica [16]. Por otro lado, Pukala y colaboradores en 2004 investigaron sobre los aminoácidos (residuos de Pro) que otorgan flexibilidad a las caerinas en su estructura y encontraron que alterar esta región cambiándola por otros aminoácidos produce la disminución de la actividad antimicrobiana en las caerinas [90].

4.3.3. Predicción de estructura terciaria de las nuevas proteínas

La estructura terciaria de las proteínas inhibidoras de tipo Kazal es altamente conservada, debido a que la estructura resulta crucial para la función inhibitoria de estas proteínas. Se caracterizan por poseer una serie de enlaces disulfuro que estabilizan su estructura tridimensional. Los dominios de tipo Kazal son pequeños (alrededor de 50 residuos), y están organizados en una disposición compacta que permite una interacción precisa con sus proteasas objetivos. La estructura incluye una región central con hélices alfa, entre dos laminas betas antiparalelas y una tercera hacia el extremo C-terminal, lo que contribuye a la rigidez y especificidad de la interacción inhibidor-proteasa [83,94].

La implementación de secuenciación de ARN dirigido a la secreción cutánea de anfibios para identificación y caracterización de estas biomoléculas provee información relevante para futuras investigaciones en áreas como la medicina, la biotecnología y la

farmacología. Además, podrá conducir al descubrimiento de nuevas terapias y aplicaciones bioactivas.

5. CAPÍTULO V: CONCLUSIONES Y RECOMENDACIONES

En el presente estudio se demuestra que la secreción cutánea de *A. spurrelli* es una fuente rica de péptidos antimicrobianos y proteínas. Se realizó la secuenciación por nanoporos al ADNc amplificado por primers específicos para obtener un gran número de datos de secuenciación que fueron tratados con el pipeline de limpieza y tratamiento apoyado en herramientas bioinformáticas. De esta manera se logrando identificar 39 potenciales péptidos antimicrobianos agrupados en 7 familias: 14 dermaseptinas, 6 triptofilinas, 4 filoseptinas, 3 filoquininas, 2 plasticinas, 1 caerinas, 1 medusinas y 8 péptidos huérfanos. Adicionalmente, se contrastó esta información por medio de espectrometría de masas confirmó la presencia de al menos 21 de estos péptidos en la secreción cutánea de *A. spurrelli*, otorgando robustez a la técnica implementada. Por otra parte, se realizó caracterización fisicoquímica de los potenciales péptidos y se obtuvo que al menos 9 péptidos son candidatos óptimos para presentar actividad antimicrobiana destacando las dermaseptinas (DRS-SP7, DRS-SP10, DRS-SP15, DRS16) y filoseptinas (PLS-SP4, PLS-SP5) debido a su cationicidad, hidrofobicidad, porcentaje alfa hélice y estructura tridimensional. Por lo que se recomienda realizar síntesis, ensayos inhibitorios y ensayos hemolíticos de cada uno de los péptidos para corroborar la respectiva actividad que poseen.

Es importante mencionar que también se identificaron 9 proteínas inhibidoras de tipo Kazal con sus variantes denominadas ASKP (1-9) de las cuales se predijo su estructura tridimensional. Sin embargo, al tener poca información de proteínas inhibitorias provenientes de *A. spurrelli*, la actividad y evaluación de las mismas es limitada, por lo que se sugiere continuar con las investigaciones. Finalmente, las péptidos y proteínas encontradas en este estudio evidencian la riqueza peptídica y proteica de la secreción cutánea de los anfibios, lo que se traduce en potenciales alternativas terapéuticas en la lucha contra la resistencia antimicrobiana y otras líneas biomédicas.

REFERENCIAS BIBLIOGRÁFICAS

1. Correia Santana CJ, Martins Magalhães AC, Dos Santos Júnior ACM, Ornelas Ricart CA, Lima BD, Morales Álvares A da C, et al. Figainin 1, a Novel Amphibian Skin Peptide with Antimicrobial and Antiproliferative Properties. *Antibiot Basel Switz.* 2020;9: 625. doi:10.3390/antibiotics9090625
2. Demori I, El Rashed Z, Corradino V, Catalano A, Rovegno L, Queirolo L, et al. Peptides for Skin Protection and Healing in Amphibians. *Molecules.* 2019;24: 347. doi:10.3390/molecules24020347
3. University of Nebraska Medical Center. Antimicrobial Peptide Database, APD3. 2024.
4. Wang G, Li X, Wang Z. APD3: the antimicrobial peptide database as a tool for research and education. *Nucleic Acids Res.* 2016;44: D1087–D1093. doi:10.1093/nar/gkv1278
5. Azevedo Calderon L de, Silva A de AE, Ciancaglini P, Stábeli RG. Antimicrobial peptides from Phyllomedusa frogs: from biomolecular diversity to potential nanotechnologic medical applications. *Amino Acids.* 2011;40: 29–49. doi:10.1007/s00726-010-0622-3
6. Amiche M, Ladram A, Nicolas P. A consistent nomenclature of antimicrobial peptides isolated from frogs of the subfamily Phyllomedusinae. *Peptides.* 2008;29: 2074–2082. doi:10.1016/j.peptides.2008.06.017
7. Chai J, Wu J, Li J, Liao H, Lu W, Guo R, et al. Novel Amphibian Bowman–Birk-Like Inhibitor with Antioxidant and Anticoagulant Effects Ameliorates Pancreatitis Symptoms in Mice. *J Med Chem.* 2023;66: 11869–11880. doi:10.1021/acs.jmedchem.3c00475
8. Dematei A, Costa SR, Moreira DC, Barbosa EA, Friaça Albuquerque LF, Vasconcelos AG, et al. Antioxidant and Neuroprotective Effects of the First Tryptophyllin Found in Snake Venom (*Bothrops moojeni*). *J Nat Prod.* 2022;85: 2695–2705. doi:10.1021/acs.jnatprod.2c00304
9. Han Z, Feng D, Wang W, Wang Y, Cheng M, Yang H, et al. Influence of Fatty Acid Modification on the Anticancer Activity of the Antimicrobial Peptide Figainin 1. *ACS Omega.* 2023;8: 41876–41884. doi:10.1021/acsomega.3c06806
10. Huan Y, Kong Q, Mou H, Yi H. Antimicrobial Peptides: Classification, Design, Application and Research Progress in Multiple Fields. *Front Microbiol.* 2020;11. doi:10.3389/fmicb.2020.582779
11. Thompson AH, Bjourson AJ, Orr DF, Shaw C, McClean S. Un enfoque combinado de espectrometría de masas y secuenciación de ADNc para el aislamiento y caracterización

- de nuevos péptidos antimicrobianos de las secreciones cutáneas de *Phyllomedusa hipocondrialis azurea*. *Peptides*. 2007;28: 1331–1343. doi:10.1016/j.peptides.2007.05.001
12. Zannella C, Chianese A, Palomba L, Marcocci ME, Bellavita R, Merlino F, et al. Broad-Spectrum Antiviral Activity of the Amphibian Antimicrobial Peptide Temporin L and Its Analogs. *Int J Mol Sci*. 2022;23: 2060. doi:10.3390/ijms23042060
 13. Abbassi F, Raja Z, Oury B, Gazanion E, Piesse C, Sereno D, et al. Antibacterial and leishmanicidal activities of temporin-SHd, a 17-residue long membrane-damaging peptide. *Biochimie*. 2013;95: 388–399. doi:10.1016/j.biochi.2012.10.015
 14. Rojas-Pirela M, Kemmerling U, Quiñones W, Michels PAM, Rojas V. Antimicrobial Peptides (AMPs): Potential Therapeutic Strategy against Trypanosomiasis? *Biomolecules*. 2023;13: 599. doi:10.3390/biom13040599
 15. Zhang L, Gallo RL. Antimicrobial peptides. *Curr Biol*. 2016;26: R14–R19. doi:10.1016/j.cub.2015.11.017
 16. Xu X, Lai R. The Chemistry and Biological Activities of Peptides from Amphibian Skin Secretions. *Chem Rev*. 2015;115: 1760–1846. doi:10.1021/cr4006704
 17. Leite JRSA, Silva LP, Rodrigues MIS, Prates MV, Brand GD, Lacava BM, et al. Phylloseptins: a novel class of anti-bacterial and anti-protozoan peptides from the *Phyllomedusa* genus. *Peptides*. 2005;26: 565–573. doi:10.1016/j.peptides.2004.11.002
 18. Proaño-Bolaños C, Blasco-Zúñiga A, Almeida JR, Wang L, Llumiquinga MA, Rivera M, et al. Unravelling the Skin Secretion Peptides of the Gliding Leaf Frog, *Agalychnis spurrelli* (Hylidae). *Biomolecules*. 2019;9: 667. doi:10.3390/biom9110667
 19. Gao Y, Wu D, Xi X, Wu Y, Ma C, Zhou M, et al. Identification and Characterisation of the Antimicrobial Peptide, Phylloseptin-PT, from the Skin Secretion of *Phyllomedusa tarsius*, and Comparison of Activity with Designed, Cationicity-Enhanced Analogues and Diastereomers. *Molecules*. 2016;21: 1667. doi:10.3390/molecules21121667
 20. He D, Cao Z, Zhang R, Li W. Molecular Cloning and Functional Identification of the Antimicrobial Peptide Gene Ctri9594 from the Venom of the Scorpion *Chaerilus tricosatus*. *Antibiotics*. 2021;10: 896. doi:10.3390/antibiotics10080896
 21. Borodina T, Adjaye J, Sultan M. A strand-specific library preparation protocol for RNA sequencing. *Methods Enzymol*. 2011;500: 79–98. doi:10.1016/B978-0-12-385118-5.00005-0
 22. Hrdlickova R, Toloue M, Tian B. RNA-Seq methods for transcriptome analysis. *Wiley Interdiscip Rev RNA*. 2017;8: 10.1002/wrna.1364. doi:10.1002/wrna.1364

23. Levin JZ, Yassour M, Adiconis X, Nusbaum C, Thompson DA, Friedman N, et al. Comprehensive comparative analysis of strand-specific RNA sequencing methods. *Nat Methods*. 2010;7: 709–715. doi:10.1038/nmeth.1491
24. Rončević T, Gerdol M, Spazzali F, Florian F, Mekinić S, Tossi A, et al. Parallel identification of novel antimicrobial peptide sequences from multiple anuran species by targeted DNA sequencing. *BMC Genomics*. 2018;19: 827. doi:10.1186/s12864-018-5225-5
25. Rončević T, Krce L, Gerdol M, Pacor S, Benincasa M, Guida F, et al. Membrane-active antimicrobial peptide identified in *Rana arvalis* by targeted DNA sequencing. *Biochim Biophys Acta Biomembr*. 2019;1861: 651–659. doi:10.1016/j.bbamem.2018.12.014
26. Rubio S, Pacheco-Orozco R, Gómez AM, Perdomo S, Gacía-Robles R. Secuenciación de nueva generación (NGS) de ADN: presente y futuro en la práctica clínica. *Univ Medica*. 2020;61. doi:https://doi.org/10.11144/Javeriana.umed61-2.sngs
27. Zhong Y, Xu F, Wu J, Schubert J, Li MM. Application of Next Generation Sequencing in Laboratory Medicine. *Ann Lab Med*. 2021;41: 25–43. doi:10.3343/alm.2021.41.1.25
28. Qiao L, Yang W, Fu J, Song Z. Transcriptome profile of the green odorous frog (*Odorrana margaretae*). *PloS One*. 2013;8: e75211. doi:10.1371/journal.pone.0075211
29. Bai Y, Meng Y, Luo J, Wang H, Li G, Li C. Full-length transcriptome assembly of *Andrias davidianus* (amphibia: caudata) skin via hybrid sequencing. *Biosci Rep*. 2021;41: BSR20210511. doi:10.1042/BSR20210511
30. Greninger AL, Naccache SN, Federman S, Yu G, Mbala P, Bres V, et al. Rapid metagenomic identification of viral pathogens in clinical samples by real-time nanopore sequencing analysis. *Genome Med*. 2015;7: 99. doi:10.1186/s13073-015-0220-9
31. Wang Y, Zhao Y, Bollas A, Wang Y, Au KF. Nanopore sequencing technology, bioinformatics and applications. *Nat Biotechnol*. 2021;39: 1348–1365. doi:10.1038/s41587-021-01108-x
32. Grünberger F, Ferreira-Cerca S, Grohmann D. Nanopore sequencing of RNA and cDNA molecules in *Escherichia coli*. *RNA N Y N*. 2022;28: 400–417. doi:10.1261/rna.078937.121
33. Wongsurawat T, Jenjaroenpun P, Wanchai V, Nookaew I. Native RNA or cDNA Sequencing for Transcriptomic Analysis: A Case Study on *Saccharomyces cerevisiae*. *Front Bioeng Biotechnol*. 2022;10. doi:10.3389/fbioe.2022.842299
34. Oxford Nanopore Technologies. Ligation sequencing kit V14 (SQK-LSK114). In: Oxford Nanopore Technologies [Internet]. 2023 [cited 10 Dec 2023]. Available: <http://community.nanoporetech.com/protocols/ligation-sequencing-amplicons-sqk-lsk114>

35. Organización Mundial de la Salud. Resistencia a los antimicrobianos. 2021 [cited 26 Jun 2024]. Available: <https://www.who.int/es/news-room/fact-sheets/detail/antimicrobial-resistance>
36. Ministerio de Salud Pública. Reporte de datos de resistencia a los antimicrobianos en Ecuador 2014-2018. 2018.
37. Life Technologies Corporation. Invitrogen® Dynabeads® mRNA DIRECT® Purification Kit -. 2012 [cited 29 May 2024]. Available: <https://www.fishersci.es/shop/products/dynabeads-mrna-direct-purification-kit-1/10563755>
38. Thermo Fisher Scientific. Qubit RNA HS Assay Kit User Guide. 2022: 8.
39. Torrejón D, Cárdenas J, Juárez D, Espinoza J, Proleón A, Agurto-Arteaga A, et al. Comparison of Four Methods of RNA Extraction and cDNA Synthesis from The Venom of Peruvian Snakes of the Genus *Bothrops* of Clinical Importance. *Int J Mol Sci.* 2023;24: 11161. doi:10.3390/ijms241311161
40. Geneious Dotmatics. Designing Taqman® and qPCR primers in Geneious Prime. In: Geneious [Internet]. 13 Dec 2023 [cited 11 Aug 2024]. Available: <https://help.geneious.com/hc/en-us/articles/360044627212-Designing-Taqman-and-qPCR-primers-in-Geneious-Prime>
41. Proaño-Bolaños C, Zhou M, Wang L, Coloma LA, Chen T, Shaw C. Peptidomic approach identifies cruzioseptins, a new family of potent antimicrobial peptides in the splendid leaf frog, *Cruziohyla calcarifer*. *J Proteomics.* 2016;146: 1–13. doi:10.1016/j.jprot.2016.06.017
42. Raja Z, André S, Piesse C, Sereno D, Nicolas P, Foulon T, et al. Structure, Antimicrobial Activities and Mode of Interaction with Membranes of Bovell Phylloseptins from the Painted-Belly Leaf Frog, *Phyllomedusa sauvagii*. *PLOS ONE.* 2013;8: e70782. doi:10.1371/journal.pone.0070782
43. Xiong W, Li J, Feng Y, Chai J, Wu J, Hu Y, et al. Brevinin-2GHk, a Peptide Derived from the Skin of *Fejervarya limnocharis*, Inhibits Zika Virus Infection by Disrupting Viral Integrity. *Viruses.* 2021;13: 2382. doi:10.3390/v13122382
44. New England Biolabs. Protocol for OneTaq Hot Start 2X Master Mix with Standard Buffer (M0484) | NEB. [cited 30 Jan 2024]. Available: <https://www.neb.com/en/protocols/2012/09/10/onetaq-hot-start-2x-master-mix-with-standard-buffer-m0484>
45. Life Technologies Corporation. purelink_pcr_man.pdf. 201125: 20.
46. Thermo Fisher Scientific. Qubit dsDNA HS Assay Kit. 8 Mar 2022: 8.

47. Oxford Nanopore Technologies. Data analysis - On-demand basecalling using the Dorado software. In: Oxford Nanopore Technologies [Internet]. 2024 [cited 28 May 2024]. Available: https://community.nanoporetech.com/technical_documents/data-analysis/v/datd_5000_v1_revr_22aug2016/on-demand-basecalling-using-the-dorado-software
48. Batut B, Doyle M, Cormier A, Bretaudeau A, Leroi L, Corre E, et al. Quality Control. In: Galaxy Training Network [Internet]. Galaxy Training Network; 56:50 [cited 11 Aug 2024]. Available: <https://training.galaxyproject.org/training-material/topics/sequence-analysis/tutorials/quality-control/tutorial.html>
49. Vierstraete AR, Braeckman BP. Amplicon_sorter: A tool for reference-free amplicon sorting based on sequence similarity and for building consensus sequences. *Ecol Evol.* 2022;12: e8603. doi:10.1002/ece3.8603
50. Singh U, Wurtele ES. orfipy: a fast and flexible tool for extracting ORFs. *Bioinformatics.* 2021;37: 3019–3020. doi:10.1093/bioinformatics/btab090
51. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol Biol Evol.* 2018;35: 1547–1549. doi:10.1093/molbev/msy096
52. Tamura K, Stecher G, Kumar S. MEGA11: Molecular Evolutionary Genetics Analysis Version 11. *Mol Biol Evol.* 2021;38: 3022–3027. doi:10.1093/molbev/msab120
53. National Library of Medicine NC for BI. Protein BLAST: search protein databases using a protein query. In: BLAST blastp suite [Internet]. [cited 12 Feb 2024]. Available: https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastp&PAGE_TYPE=BlastSearch&LIN K_LOC=blasthome
54. Tamang S. BLAST (Bioinformatics)- Definition, 5 Types, Steps, Uses. 3 Aug 2023 [cited 11 Aug 2024]. Available: <https://microbenotes.com/blast-bioinformatics/>
55. BACHEM peptide calculator. Peptide & Amino Acids Calculator | Bachem. 2024 [cited 15 May 2024]. Available: <https://www.bachem.com/knowledge-center/peptide-calculator/>
56. NPS@ : SOPMA secondary structure prediction. Available: https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_sopma.html
57. Geourjon C, Deléage G. SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. *Bioinformatics.* 1995;11: 681–684. doi:10.1093/bioinformatics/11.6.681
58. Roy A, Kucukural A, Zhang Y. I-TASSER: a unified platform for automated protein structure and function prediction. *Nat Protoc.* 2010;5: 725–738. doi:10.1038/nprot.2010.5

59. Keller RCA. Identification and in silico analysis of helical lipid binding regions in proteins belonging to the amphitropic protein family. *J Biosci.* 2014;39: 771–783. doi:10.1007/s12038-014-9479-z
60. Bienert S, Waterhouse A, de Beer TAP, Tauriello G, Studer G, Bordoli L, et al. The SWISS-MODEL Repository-new features and functionality. *Nucleic Acids Res.* 2017;45: D313–D319. doi:10.1093/nar/gkw1132
61. Zhang Y. I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics.* 2008;9: 40. doi:10.1186/1471-2105-9-40
62. Wan M, Faruq J, Rosenberg JN, Xia J, Oyler GA, Betenbaugh MJ. Achieving high throughput sequencing of a cDNA library utilizing an alternative protocol for the bench top next-generation sequencing system. *J Microbiol Methods.* 2013;92: 122–126. doi:10.1016/j.mimet.2012.10.010
63. Vvedenskaya IO, Goldman SR, Nickels BE. Preparation of cDNA libraries for high-throughput RNA sequencing analysis of RNA 5' ends. *Methods Mol Biol Clifton NJ.* 2015;1276: 211–228. doi:10.1007/978-1-4939-2392-2_12
64. Shibao PYT, Cologna CT, Morandi-Filho R, Wiesel GA, Fujimura PT, Ueira-Vieira C, et al. Deep sequencing analysis of toad *Rhinella schneideri* skin glands and partial biochemical characterization of its cutaneous secretion. *J Venom Anim Toxins Trop Dis.* 2018;24: 36. doi:10.1186/s40409-018-0173-8
65. Conlon JM. The contribution of skin antimicrobial peptides to the system of innate immunity in anurans. *Cell Tissue Res.* 2011;343: 201–212. doi:10.1007/s00441-010-1014-4
66. Brand GD, Santos RC, Arake LM, Silva VG, Veras LMC, Costa V, et al. The Skin Secretion of the Amphibian *Phyllomedusa nordestina*: A Source of Antimicrobial and Antiprotozoal Peptides. *Molecules.* 2013;18: 7058–7070. doi:10.3390/molecules18067058
67. Hong Y-H, Huang H-M, Wu L, Storey KB, Zhang J-Y, Zhang Y-P, et al. Characterization of Two Mitogenomes of *Hyla sanchiangensis* (Anura: Hylidae), with Phylogenetic Relationships and Selection Pressure Analyses of Hylidae. *Anim Open Access J MDPI.* 2023;13: 1593. doi:10.3390/ani13101593
68. Andrade P, Lyra ML, Zina J, Bastos DFO, Brunetti AE, Baêta D, et al. Draft genome and multi-tissue transcriptome assemblies of the Neotropical leaf-frog *Phyllomedusa bahiana*. *G3 GenesGenomesGenetics.* 2022;12: jkac270. doi:10.1093/g3journal/jkac270
69. Wescoe ZL, Schreiber J, Akeson M. Nanopores Discriminate among Five C5-Cytosine Variants in DNA. *J Am Chem Soc.* 2014;136: 16582–16587. doi:10.1021/ja508527b

70. Jain M, Olsen HE, Paten B, Akeson M. The Oxford Nanopore MinION: delivery of nanopore sequencing to the genomics community. *Genome Biol.* 2016;17. doi:10.1186/s13059-016-1103-0
71. Griffiths S. Quality Scores And Read Accuracy. In: EPI2ME Labs [Internet]. 1 Jul 2021 [cited 12 Aug 2024]. Available: <https://labs.epi2me.io/quality-scores/>
72. Edgar R. Quality (Phred) scores. Available: https://www.drive5.com/usearch/manual/quality_score.html
73. Trivedi UH, Cézard T, Bridgett S, Montazam A, Nichols J, Blaxter M, et al. Quality control of next-generation sequencing data without a reference. *Front Genet.* 2014;5. doi:10.3389/fgene.2014.00111
74. Linde J, Brangsch H, Hölzer M, Thomas C, Elschner MC, Melzer F, et al. Comparison of Illumina and Oxford Nanopore Technology for genome analysis of *Francisella tularensis*, *Bacillus anthracis*, and *Brucella suis*. *BMC Genomics.* 2023;24: 258. doi:10.1186/s12864-023-09343-z
75. Delahaye C, Nicolas J. Sequencing DNA with nanopores: Troubles and biases. *PloS One.* 2021;16: e0257521. doi:10.1371/journal.pone.0257521
76. Wang H, Wang L, Zhou M, Yang M, Ma C, Chen T, et al. Peptidomia funcional de la secreción cutánea de anfibios: un nuevo inhibidor de quimotripsina tipo Kunitz de la rana hiperólida africana, *Kassina senegalensis*. *Biochimie.* 2012;94: 891–899. doi:10.1016/j.biochi.2011.12.008
77. Proaño-Bolaños C, Li R, Zhou M, Wang L, Xi X, Tapia EE, et al. Novel Kazal-type proteinase inhibitors from the skin secretion of the Splendid leaf frog, *Cruziohyla calcarifer*. *EuPA Open Proteomics.* 2017;15: 1–13. doi:10.1016/j.euprot.2017.02.001
78. Chen C. Searching for intellectual turning points: Progressive knowledge domain visualization. *Proc Natl Acad Sci.* 2004;101: 5303–5310. doi:10.1073/pnas.0307513100
79. Rates B, Silva LP, Ireno IC, Leite FSF, Borges MH, Bloch C, et al. Peptidomic dissection of the skin secretion of *Phasmahyla jandaia* (Bokermann and Sazima, 1978) (Anura, Hylidae, Phyllomedusinae). *Toxicon.* 2011;57: 35–52. doi:10.1016/j.toxicon.2010.09.010
80. Neiva M, Vargas DC, Conceição K, Rádis-Baptista G, Assakura MT, Jared C, et al. Gene expression analysis by ESTs sequencing of the Brazilian frog *Phyllomedusa nordestina* skin glands. *Toxicon.* 2013;61: 139–150. doi:10.1016/j.toxicon.2012.10.016
81. Chen T, Shaw C. Cloning of the (Thr6)-phyllokinin precursor from *Phyllomedusa sauvagei* skin confirms a non-consensus tyrosine *O*-sulfation motif. *Peptides.* 2003;24: 1123–1130. doi:10.1016/j.peptides.2003.06.002

82. Vanhoye D, Bruston F, Nicolas P, Amiche M. Antimicrobial peptides from hylid and ranin frogs originated from a 150-million-year-old ancestral precursor with a conserved signal peptide but a hypermutable antimicrobial domain. *Eur J Biochem.* 2003;270: 2068–2081. doi:10.1046/j.1432-1033.2003.03584.x
83. Li R, Wang H, Jiang Y, Yu Y, Wang L, Zhou M, et al. A novel Kazal-type trypsin inhibitor from the skin secretion of the Central American red-eyed leaf frog, *Agalychnis callidryas*. *Biochimie.* 2012;94: 1376–1381. doi:10.1016/j.biochi.2012.03.009
84. Passos LF, Garcia G, Young RJ. Comparing the bacterial communities of wild and captive golden mantella frogs: Implications for amphibian conservation. *PLoS ONE.* 2018;13: e0205652. doi:10.1371/journal.pone.0205652
85. Kueneman JG, Bletz MC, Becker M, Gratwicke B, Garcés OA, Hertz A, et al. Effects of captivity and rewilding on amphibian skin microbiomes. *Biol Conserv.* 2022;271: 109576. doi:10.1016/j.biocon.2022.109576
86. Islam MdM, Asif F, Zaman SU, Arnab MdKH, Rahman MdM, Hasan M. Effect of charge on the antimicrobial activity of alpha-helical amphibian antimicrobial peptide. *Curr Res Microb Sci.* 2023;4: 100182. doi:10.1016/j.crmicr.2023.100182
87. Wang L, Zhou M, McClelland A, Reilly A, Chen T, Gagliardo R, et al. Novel dermaseptin, adenoregulin and caerin homologs from the Central American red-eyed leaf frog, *Agalychnis callidryas*, revealed by functional peptidomics of defensive skin secretion. *Biochimie.* 2008;90: 1435–1441. doi:10.1016/j.biochi.2008.04.016
88. Pirtskhalava M, Vishnepolsky B, Grigolava M, Managadze G. Physicochemical Features and Peculiarities of Interaction of AMP with the Membrane. *Pharm Basel Switz.* 2021;14: 471. doi:10.3390/ph14050471
89. Xiao L, Yang X, Li J, Zhang P, Tang S, Cao D, et al. Caerin 1 Peptides, the Potential Jack-of-All-Trades for the Multiple Antibiotic-Resistant Bacterial Infection Treatment and Cancer Immunotherapy. *BioMed Res Int.* 2022;2022: 7841219. doi:10.1155/2022/7841219
90. Pukala TL, Brinkworth CS, Carver JA, Bowie JH. Investigating the Importance of the Flexible Hinge in Caerin 1.1: Solution Structures and Activity of Two Synthetically Modified Caerin Peptides. *Biochemistry.* 2004;43: 937–944. doi:10.1021/bi035760b
91. Mourtada R, Herce HD, Yin DJ, Moroco JA, Wales TE, Engen JR, et al. Design of stapled antimicrobial peptides that are stable, nontoxic and kill antibiotic-resistant bacteria in mice. *Nat Biotechnol.* 2019;37: 1186–1197. doi:10.1038/s41587-019-0222-z
92. Sharma P, Sharma S, Joshi S, Barman P, Bhatt A, Maan M, et al. Design, characterization and structure–function analysis of novel antimicrobial peptides based on the N-terminal CATH-2 fragment. *Sci Rep.* 2022;12: 12058. doi:10.1038/s41598-022-16303-2

93. Yang J, Zhang Y. Protein Structure and Function Prediction Using I-TASSER. *Curr Protoc Bioinforma*. Ed Board Andreas Baxevanis Al. 2015;52: 5.8.1-5.815. doi:10.1002/0471250953.bi0508s52
94. Rimphanitchayakit V, Tassanakajon A. Structure and function of invertebrate Kazal-type serine proteinase inhibitors. *Dev Comp Immunol*. 2010;34: 377–386. doi:10.1016/j.dci.2009.12.004