



BÚSQUEDA Y CARACTERIZACIÓN DE PÉPTIDOS ANTIMICROBIANOS EN *Acca sellowiana*

María Florencia Rossi Novoa

Trabajo Final de Grado de la Licenciatura en Ciencias Biológicas

Tutora: Susana Rodríguez Decuadro

Cotutora: Clara Pritsch

ÍNDICE

ÍNDICE	
RESUMEN.....	
INTRODUCCIÓN.....	
OBJETIVOS	
METODOLOGÍA	
Construcción del transcriptoma.....	
Identificación de las secuencias.....	
Alineamientos múltiples y árboles de distancia genética.....	
Validación por PCR.....	
Extracción de ADN.....	
PCR.....	
Electroforesis de ADN.....	
Secuenciación.....	
Alineamientos ADN genómico vs transcriptoma.....	
RESULTADOS	
Identificación de las secuencias.....	
Caracterización de secuencias.....	
Alineamientos múltiples y árboles de distancia genética.....	
Validación de secuencias por PCR.....	
Análisis de secuencias obtenidas por PCR.....	
AsDf1.....	
AsDf5.....	
AsDf6.....	
AsSn1.....	
AsHv1.....	
DISCUSIÓN	
CONCLUSIONES Y PERSPECTIVAS	
BIBLIOGRAFÍA	
MATERIAL SUPLEMENTARIO	
Anexo 1: <i>Querys</i> utilizados para la búsqueda mediante BLAST.....	
Anexo 2: Secuencias identificadas en transcriptoma de <i>A. sellowiana</i>	
Anexo 3: <i>primers</i> y su ubicación en las secuencias del transcriptoma.....	

RESUMEN

Los péptidos antimicrobianos (AMPs) son polipéptidos pequeños, presentes en una gran diversidad de organismos muy lejanos filogenéticamente, incluyendo las plantas. Los AMPs forman parte de la respuesta inmune innata, permitiendo una acción rápida a menor costo energético que el sistema inmune adaptativo de los vertebrados superiores. Se han demostrado diversas aplicaciones de estos péptidos como agentes anti-infecciosos, en campos como la agronomía, la veterinaria y la medicina. Sin embargo, a pesar de su importante potencial, la bioprospección de AMPs en la flora nativa de Uruguay es aún incipiente. El objetivo de este trabajo fue identificar y caracterizar genes que codifican péptidos antimicrobianos en las 6 familias principales: defensinas, esnaquinas, tioninas, *hevein-like*, lipid transfer proteins (LTPs) y ciclótidos, en la especie frutal nativa *Acca sellowiana* (Myrtaceae), denominada comúnmente guayabo del país. La búsqueda se realizó mediante BLAST, a partir de un transcriptoma *de novo* de hoja y flor y utilizando secuencias de referencia de otras especies. Se caracterizó la relación entre las proteínas identificadas, y con las secuencias de referencia, mediante alineamientos múltiples y árboles de distancia genética de las familias. Se validó su existencia en ADN genómico mediante PCR, y los fragmentos amplificados se secuenciaron para analizar la estructura exón-intrón e identificar posibles isoformas. Se identificaron 23 defensinas, 49 esnaquinas, 7 tioninas, 12 *hevein-like*, y 87 LTPs. No se encontró ninguna secuencia con homología a ciclótidos. Se determinaron grupos dentro de las familias; en el caso de las defensinas, tioninas y *hevein-like* algunos grupos fueron novedosos, mientras que para las esnaquinas y LTPs ya se habían caracterizado anteriormente en otras especies. Se seleccionaron 9 secuencias correspondientes a 3 familias, para su amplificación a partir de ADN genómico y se observó que los tamaños obtenidos eran correspondientes con AMPs. Se obtuvo para cinco de ellos la estructura exón-intrón, que coinciden con las esperadas para cada una de las familias. Queda por analizar las secuencias restantes con el fin de obtener el panorama completo de las familias. Como perspectiva a futuro se propone producir las proteínas de manera heteróloga y realizar ensayos de actividad, así como determinar se estructura terciaria.

INTRODUCCIÓN

Los péptidos antimicrobianos (AMP) constituyen armas de defensa antiguas con un posible rol en el éxito evolutivo de los organismos multicelulares. Forman parte de la respuesta inmune innata, permitiendo una respuesta rápida a menor costo energético que el sistema inmune adaptativo de los vertebrados superiores. Están presentes tanto en plantas como en animales, teniendo un rol importante en la formación de barreras de defensa pre o post infección con una amplia gama de organismos blanco, incluyendo bacterias, hongos, virus y protozoos (Zasloff, 2002). En algunos casos también pueden poseer propiedades antitumorales e inmunomoduladoras (Larrañaga et al., 2012).

Los AMPs son polipéptidos de menos de 200 aminoácidos (muy frecuentemente menos de 50 aminoácidos), y a diferencia de los metabolitos secundarios, son productos de genes únicos (Lay & Anderson, 2005; Benko-Iseppon, 2010). A pesar de la gran variedad existente, hay ciertas características compartidas entre las familias de AMP: tamaño pequeño, carga general positiva, estereogeometría anfipática, tolerancia a solventes ácidos y orgánicos, estabilidad térmica y amplia actividad biológica (Yeaman & Yount, 2003; Brodgen, 2005). Comúnmente contienen un número par de cisteínas que participan en puentes de disulfuro intracelulares, brindándole a la molécula, estabilidad estructural y terminodinámica (Lay & Anderson, 2005).

En plantas, los AMP son frecuentemente secretados, y se encuentran de manera constitutiva en los órganos de almacenamiento (semillas), y en las capas de células periféricas de los tejidos generativos (órganos reproductores, frutos y flores), además de ser producidos en tejidos vegetativos inducidos por infecciones o heridas (Granz, 2004; Benko-Iseppon, 2010). En estos organismos, los AMP no sólo actúan de manera local, sino que también se acumulan en territorios no infectados, desde el punto de infección. Este fenómeno es conocido como resistencia sistémica adquirida (SAR) (Terras et al., 1995; Lay & Anderson, 2005; Benko-Iseppon, 2010).

Si bien existe una gran variabilidad de AMP en plantas, pueden identificarse distintas familias de acuerdo con arreglos conservados y específicos de los residuos de cisteínas, lo que conlleva a patrones de plegados y estructuras similares, incluso en ausencia de alto grado de similaridad o identidad de de secuencia. Los principales grupos son defensinas, tioninas, *lipid transfer proteins* (LTP), ciclótidos, *hevein-like* y esnaquinas (Benko-Iseppon, 2010; Nawrot et al. 2014).

Defensinas

Las defensinas son los únicos AMPs descritos hasta el momento que parecen estar conservados en grupos de organismos tan distintos como plantas, invertebrados y vertebrados, constituyendo una superfamilia de proteínas (Charlet et al., 1996; Thomma et al., 2002; Oard & Karki, 2006; Revisado por Rodriguez Decuadro, 2018). Las defensinas de plantas son péptidos básicos, ricos en cisteínas, de tamaño pequeño (aprox. 5 kDa, 45 a 54 aminoácidos). Los identificados hasta el momento poseen 8 cisteínas que forman 4 puentes de disulfuro estabilizadores de su estructura tridimensional (Benko-Iseppon, 2010), compuesta por una hélice alfa y tres hojas beta antiparalelas (Lay et al., 2012) (Figura 1a). Esta estructura brinda resistencia a altas temperaturas, y degradación mediada por pH o proteasas (Hegedus & Marx, 2012). Se ha demostrado que algunas defensinas forman dímeros, lo cual aumenta su actividad antifúngica (Lay, 2012) (Figura 1b). La mayoría exhibe una conservación de secuencia limitada, con excepción de los 8 residuos de cisteínas y una glicina en la posición 34 que se encuentran estrictamente conservadas (Lay & Anderson, 2005), conformando los motivos CSq β (CX_nCXXXCX_nCX_nCXC) γ -core (GXCX₃₋₉C), donde C=Cisteína, X=cualquier aminoácido, y n un número no conservado de aminoácidos). La estructura primaria generalmente consiste en un péptido señal en el extremo amino terminal y un péptido maduro básico (De Coninck et al. 2013) mientras que la estructura de sus genes está integrada por dos exones y un intrón (Carvalho & Gómez, 2009). Sin embargo, existe una segunda clase de defensinas que posee un pro-dominio C-terminal ácido, que es removido durante el proceso de maduración. Este pro-dominio se ha identificado en las defensinas PhD1 y PhD2 de *Petunia hybrida*, NaD1 de *Nicotiana glauca*, y la defensina del desarrollo del grano de maíz ZmESR6. La función de este pro-dominio es aún desconocida (Azmi & Hussain, 2021).

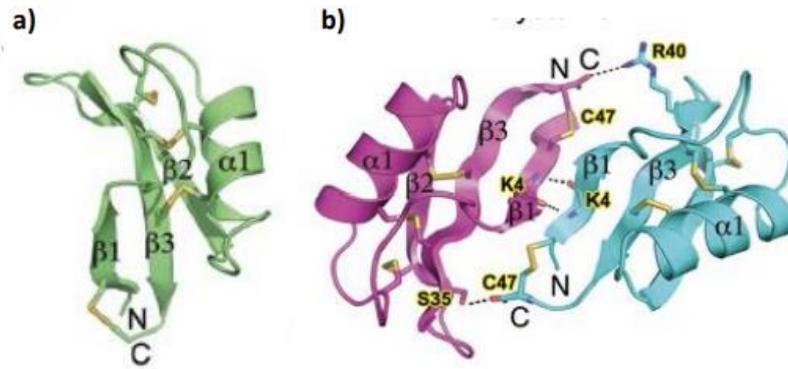


Figura 1. Estructura tridimensional de NaD1, defensina de *Nicotiana alata*. a) representación de la forma monomérica. Los enlaces disulfuro se muestran como barras. b) representación de la forma dimérica. Los residuos involucrados en la interfaz del dímero se muestran como barras y los enlaces de hidrógeno y los puentes salinos se muestran en negro como líneas de puntos. Extraído de Rodríguez Decuadro (2018). Adaptado de Lay et al. (2012).

Esnaquinas

Las esnaquinas son un grupo de péptidos encontrado solamente en plantas, que poseen un amplio espectro de actividades antimicrobianas *in vitro* (Oliveira-Lima et al., 2017). Además, tienen un rol en la respuesta de tolerancia al estrés biótico y abiótico (Bouteraa et al., 2023). Es una familia codificada por los genes snakin/GASA, y ha sido definida a partir de los péptidos Esnaquina-1 (StSN1) (Segura et al., 1999) y Esnaquina-2 (StSN2) (Berrocal-Lobo et al., 2002), aislados de tubérculos de papa (*Solanum tuberosum*). La estructura génica de StSN1 está conformada por dos exones y un intrón, mientras que StSN2 posee tres exones y dos intrones (Meiyalaghan et al. 2014). Posteriormente, Nahirñak et al (2016) describieron una tercer esnaquina de papa que difiere de las dos anteriores (StSN3). Los análisis filogenéticos permiten agrupar las esnaquinas de papa en tres subfamilias (I, II, III) representadas por StSN1, StSN2 y StSN3, respectivamente (Nahirñak et al., 2016). En relación a la estructura primaria, las esnaquinas se caracterizan por poseer tres dominios definidos: un péptido señal putativo en el extremo amino terminal, una región variable tanto en tamaño como en composición de secuencia, y una región C-terminal o dominio GASA con 12 residuos de cisteína en posiciones altamente conservadas y 60 aminoácidos de largo (Nahirñak et al., 2012). La caracterización de su estructura terciaria hasta la fecha es escasa. Porto & Franco (2013) realizaron una predicción la estructura de la StSN1 a partir de un modelo *ab initio* y comparativo, y un predictor de puentes disulfuro, resultando en una estructura compuesta de dos α -hélices largas y un patrón de 6 puentes disulfuro (Figura 2a). Yeung et al. (2016) presentaron una estructura determinada por una combinación de cristalización proteica racémica y fase inducida

por daños por radiación (RIP), compuesta por dos α -hélices cortas formando un dominio *hélix-turn-hélix*, con un patrón de 6 puentes disulfuro. Cabe agregar que, si bien ambos modelos presentan 6 puentes disulfuro, sólo coinciden en dos de ellos (Rodríguez Decuadro, 2021).

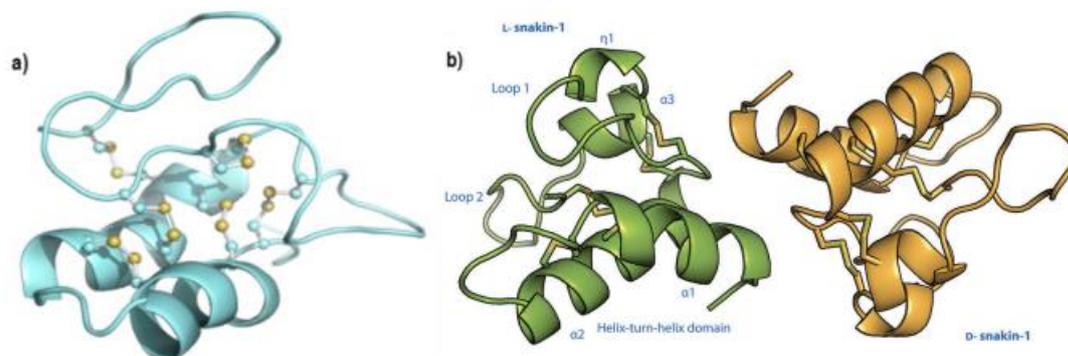


Figura 2. Estructura tridimensional de la StSN1 a) predicción tridimensional a partir de modelo *ab initio* y comparativo, y un predictor de puentes disulfuro, extraída de Porto & Franco (2013). b) modelo de la estructura L-esnaquina-1 y D-esnaquina-1, obtenida por cristalografía de rayos X, extraída de Yeung et al. (2016).

Lipid Transfer Proteins (LTP)

Los LTP vegetales son pequeños péptidos catiónicos, cuya función se ha asociado con su capacidad para unirse y transportar lípidos y otras moléculas hidrofóbicas (Amador et al. 2021). Además de la función de transferencia de lípidos, otras funciones se han descrito en varias especies vegetales, con énfasis en la defensa de la planta contra el ataque de fitopatógenos (Kader, 1996). Clásicamente se han subdividido en dos familias (LTP1 y LTP2), ambas con patrones conservados de residuos de cisteínas y una estructura tridimensional con una cavidad interior hidrofóbica que corresponde al sitio de unión de lípidos (Carvalho & Gomes, 2007). Los péptidos de la familia LTP1 tienen un tamaño de 10 KDa aprox, de 90 a 95 aminoácidos y un patrón de cisteína de 8 residuos conservados y puntos isoeléctricos de entre 9 y 10 (Carvalho y Gomes, 2007). Los péptidos de la familia LTP2 tienen un tamaño aproximado de 7 KDa, con 70 residuos en promedio, 4 puentes de disulfuro conservados y alto punto isoeléctrico (Castro et al., 2003). Si bien las dos familias poseen un péptido señal, este varía en el tamaño: para el caso de la LTP1 posee 21-27 aminoácidos, mientras que para LTP2 posee 27-35 aminoácidos (Carvalho & Gomes, 2007). Con el avance de la secuenciación masiva y las herramientas bioinformáticas, otras clasificaciones basadas en la conservación de secuencias, las modificaciones postraduccionales en los sitios de anclaje del glicosilfosfatidilinositol, las posiciones de los intrones y el

espaciado en las regiones entre residuos de cisteína, han sido propuestas (revisado por Santos-Silva et al. 2023).

Tioninas

Las tioninas fueron los primeros péptidos para los cuales se describió una actividad antipatogénica (Padovan et al., 2010). Son péptidos básicos (ricos en arginina, lisina y cisteína), y de bajo peso molecular (~5 kDa). Han sido reportados en varios tejidos vegetales como semillas, tallos y raíces, y se localizan principalmente en el espacio intracelular (aunque también pueden encontrarse extracelularmente) (Selitrennikoff, 2001). La pre-pro-proteína está constituida por un péptido señal o líder de 25 aminoácidos y un pro-péptido de 60 aminoácidos (que incluye el péptido maduro y una cola C-terminal acídica, que neutraliza la tionina madura). Las tioninas se pueden clasificar en 4 tipos según su secuencia y el patrón de puentes de disulfuro (Bohlman et al 1994) pero recientemente se ha propuesto una nueva clasificación más práctica para todas las variantes en clase 1 (8 cisteínas) y clase 2 (6 cisteínas) (Höng et al., 2021). La expresión transgénica de tioninas ha demostrado actividad antibacteriana en varias especies de plantas (Benko-Iseppon, 2010), y presentan toxicidad frente a bacterias, hongos, animales y células vegetales (Castro & Fontes, 2005).

Hevein-like

Las *hevein-like* son un grupo de proteínas que se caracterizan por su homología con la proteína heveína del árbol del caucho (*Hevea brasiliensis*) (Van Parijs et al., 1991). Los miembros de esta familia se caracterizan por presentar proteínas precursoras largas, y poseer un motivo conservado de unión a quitina, involucrado en la interacción con diversos patógenos. Generalmente poseen ocho cisteínas formando cuatro puentes disulfuro, pero también pueden encontrarse variantes de seis y diez cisteínas (Lee et al., 1991; Irving & Gehring, 2012). Son pequeños péptidos básicos de 29 a 45 aminoácidos, con actividad *in vitro* contra diversos hongos fitopatógenos con contenido de quitina. Dicha actividad se mantiene hasta después de la exposición a altas temperaturas (90°C por 10 minutos) (Slavokhotova et al. 2017a). El precursor de las *hevein-like* es una proteína de 210 aminoácidos compuesta por un péptido señal (17 aminoácidos), el dominio heveín-like (43 aminoácidos), un *linker* (6 aminoácidos), y un dominio C-terminal (144 aminoácidos) (Slavokhotova et al. 2017a).

Ciclótidos

Los ciclótidos son péptidos encontrados en plantas de 28-37 aminoácidos, cuya función principal es la protección de las plantas contra patógenos y pestes,

especialmente insectos (Craik, 2012). Se caracterizan por un esqueleto peptídico cíclico y una disposición anudada de tres enlaces disulfuro conservados, unidos en un motivo denominado "nudo de cisteína", y por la presencia de un péptido señal (Craik et al., 1999). Su estabilidad es excepcional, debido al nudo de cisteína, y tienen un gran rango de bioactividad. Se expresan en un amplio rango de tejidos vegetales, incluidos hojas, tallos, flores, semillas latentes y raíces (Craik, 2012; Trabi & Craik, 2004). Su alta estabilidad, así como su toxicidad selectiva para algunas líneas celulares de cáncer han posicionado a esta familia como destacados blancos para el diseño de nuevos fármacos (Craik et al., 2002). Estas potenciales aplicaciones llevaron al desarrollo de métodos para su producción, como la síntesis de péptidos de fase sólida y métodos recombinantes (Craik et al., 2012).

Aplicación y prospección de AMPs

Se han demostrado diversas aplicaciones de los AMPs como agentes anti-infecciosos (Zasloff, 2002). En el campo de la agronomía, se consideran compuestos prometedores para la prevención y el tratamiento de enfermedades de cultivos de importancia comercial y de su conservación post-cosecha (Kaur et al., 2011; Larrañaga et al., 2012; Shah & Read, 2013), así como también para aumentar la tolerancia a estrés abiótico (Carvalho & Gomes, 2011; Nahirñak et al. 2012). Esto es de vital importancia en un contexto en el que se busca lograr el aumento de la producción agrícola, tanto para alimentar a una población en constante crecimiento, como para contribuir al desarrollo de países y adoptar métodos de producción más eficientes y sostenibles (FAO, 2009). Por otro lado, se plantea la necesidad de búsqueda de nuevos compuestos con actividad antimicrobiana, tanto por el aumento de la resistencia a los pesticidas utilizadas actualmente, como por el cuestionamiento al uso de agroquímicos, perjudiciales para el ambiente y la salud humana (FAO 2016). Además, el interés reciente por el desarrollo de nuevos agentes terapéuticos humanos, responde al creciente problema de resistencia a los antibióticos convencionales y la necesidad de nuevos antibióticos (Zasloff, 2002).

La bioprospección consiste en la búsqueda sistemática de productos naturales y organismos de la biodiversidad para el desarrollo de productos de valor comercial. Hasta el momento, los cultivos comerciales han sido el principal blanco en la búsqueda de AMPs. En contraposición, la bioprospección de AMPs en las especies silvestres es aún incipiente. La mayor biodiversidad existente en plantas silvestres aumenta la pertinencia de aplicar la bioprospección por AMPs en esta categoría de especies (Pestana-Calsa et al., 2010). Hay antecedentes de trabajos de bioprospección de

AMPs en varias especies vegetales silvestres, nativas de Uruguay (*Peltophorum dubium* (ibirapita), *Erythrina crista-galli* (ceibo), y *Maytenus ilicifolia* (congorosa)) con datos de actividad microbiana confirmada mediante ensayos *in vitro* (Rodríguez Decuadro 2018; 2019 a-d), cuyos resultados son promisorios.

Acca sellowiana

El guayabo del país (*Acca sellowiana*) es una especie frutal leñosa, perteneciente a la tribu Myrteae (Myrtaceae), nativa de Uruguay y Brasil. Es una especie diploide, y su número cromosómico es $x=11$. Existen dos tipos botánicos de esta especie, el tipo Uruguay, que crece en el bioma de la Pampa, y el tipo Brasil, que crece en el bioma del bosque atlántico. Esta especie no sólo sirve como fuente de alimento para animales, sino que es utilizado por los humanos para consumo de su fruto, que tiene un sabor dulce-ácido que lo vuelve atractivo (Silva et al., 2024). Su genoma cloroplástico ha mostrado tener mucha similitud con el de otras especies de Myrtaceae, entre ellas *Eucalyptus grandis*, *Syzygium cumini*, *Eugenia uniflora* L, *Allosyncarpia ternata*, *Angophora costata*, *Corymbia maculata*, *Stockwellia quadrifida* (Oliveira de Machado et al., 2017).

Varios estudios han concluido que el guayabo del país es una buena fuente de vitaminas, polifenoles, fibra y minerales esenciales. En el caso de los polifenoles, se ha identificado que los ácidos fenólicos y los flavonoides componen la mayor parte de los compuestos fenólicos en la fruta (Thi Phan, 2019). A pesar de ser nativo de Uruguay y Brasil, su domesticación en estas zonas es aún incipiente (Vignale & Bisio, 2005; Donazzolo et al., 2020). Sin embargo, es plantado comercialmente en áreas de Nueva Zelanda, Colombia, Estados Unidos, Israel y España (Thorp y Bielecki, 2002).

Hay un interés creciente en la bioprospección de compuestos con actividad biológica en guayabo del país, siendo que se ha reportado actividad anti-oxidante, anti-inflamatoria, anti-microbiana, anti-cáncer y anti-diabetes, en extracto de hoja, flor y partes del fruto (Zhu, 2018). Sin embargo, los AMPs no han sido aún estudiados. Un bajo número de reportes indican la identificación de defensinas (incluyendo gamma-tioninas) en algunas especies de las tribus Eucalypteae y Syzygeae de Myrtaceae (Meyer et al., 2016; Tobias et al., 2018; Ozyigit et al., 2019), pero sin reportes en la tribu Myrteae, a la cual pertenece guayabo del país. Existe evidencia a partir de secuenciación y anotación automática de AMPs en mirtáceas, sin embargo, hasta el momento éstas no han sido revisadas y no existen publicaciones al respecto.

OBJETIVOS

Objetivo general

Identificar y caracterizar genes que codifican para péptidos antimicrobianos del tipo defensinas, esnaquinas, tioninas, *hevein-like*, LTPs y ciclótidos en la especie frutal nativa *Acca sellowiana*.

Objetivos específicos

1. Identificar transcritos que codifiquen defensinas, esnaquinas, tioninas, *hevein-like*, LTPs y ciclótidos en transcriptoma de *Acca sellowiana* mediante abordaje bioinformático
2. Conocer la relación de las proteínas identificadas, entre ellas y con AMP de otras especies, mediante alineamientos múltiples y árboles de distancia genética
3. Validar su existencia en ADN genómico mediante PCR
4. Caracterizar su estructura intrón-exón

METODOLOGÍA

Construcción del transcriptoma

Previo al comienzo de este trabajo, se construyó el transcriptoma *de novo* utilizando ARN de hoja y de tres estadios de flor. La preparación de las bibliotecas y la secuenciación fue llevada a cabo por Macrogen Inc. Para la construcción de las bibliotecas se utilizó TrueSeq stranded mRNA (PE 2x100). Los *reads* crudos fueron filtrados con Trimmomatic (Bolger et al. 2014). Los reads de buena calidad resultantes fueron ensamblados utilizando Trinity (Grabherr et al. 2011) con un tamaño de kmer de 25.

Identificación de las secuencias

Se partió de los resultados del ensamblado *de novo* del transcriptoma conformado por 100869 *contigs*, y se analizó los transcritos en búsqueda de secuencias putativas que codifiquen proteínas del tipo AMP, específicamente de las familias de las defensinas, esnaquinas, tioninas, hevein-like, ciclótidos y LTPs. Se utilizó el programa BioEdit (Hall, 1999) para la creación de una base de datos local con el pool de transcriptomas. Se utilizó la herramienta BLAST local, dentro del mismo programa, utilizando como *query* secuencias reportadas de proteínas AMP en diferentes especies, buscando en los seis marcos de lectura de los transcritos. Las secuencias de AMP reportadas fueron obtenidas de UniProt (<https://www.uniprot.org/>), Phytozome (<https://phytozome-next.jgi.doe.gov/>) y NCBI (<https://www.ncbi.nlm.nih.gov/>). (Anexo 1, Tabla S1)

Defensinas

Para la familia de las defensinas, se utilizaron como *query* secuencias de *Eucalyptus grandis* por su cercanía filogenética con la especie de estudio, debido a la variabilidad de estos péptidos en la secuencia aminoacídica no perteneciente al dominio conservado. También se utilizaron secuencias de *Nicotiana glauca* y de *Petunia hybrida* por ser defensinas con presencia de un pro-dominio particular. Una vez realizado este procedimiento, se utilizaron las secuencias identificadas en el transcriptoma de *A. sellowiana* como *query* con el fin de encontrar más secuencias.

Esnaquinas

Para la familia de las esnaquinas, dado su alto nivel de conservación de secuencia, se utilizaron las tres esnaquinas reportadas de *Solanum tuberosum* que definen las tres subfamilias en esta especie (subfamilia 1, 2 y 3).

Hevein-like

En el caso de las *hevein-like*, se utilizó el péptido heveína, de *Hevea brasiliensis*, y se agregaron secuencias de *Amaranthus caudatus*, *Stellaria media* y *Triticum kiharae*, para incluir secuencias *hevein-like* con anotación revisada.

Tioninas

Para la identificación de tioninas se utilizaron secuencias de *E.grandis*, por su cercanía filogenética con la especie de estudio, y se complementaron con secuencias de *Triticum aestivum* y *Hordeum vulgare* por poseer secuencias revisadas de alta calidad.

LTP y ciclótidos

Las LTP fueron identificadas utilizando secuencias de *Arabidopsis thaliana*, *Lens culinaris*, *Vigna radiata* y *Prunus armeniaca*. Para los ciclótidos se utilizaron secuencias de *Viola hederacea*, *Chassalia chartacea*, *Clitoria ternatea*, *Pombalia parviflor*, *Melicytus latifolius*, *Melicytus chathamicus*, y *Viola biflora*.

Las secuencias aminoacídicas utilizadas y su número de accesión se encuentran en material suplementario (Anexo1, Tabla S1). Se seleccionaron los hits con *E-value* menor de 0,01, y que coincidan al menos tres de las cisteínas con los motivos típicos de cada péptido. Los motivos de las familias se basaron en los descriptos por Slavokhotova (2015). De las secuencias resultantes, se analizó la presencia de péptido señal con SignalP (<https://services.healthtech.dtu.dk/services/SignalP-6.0/>) y la localización subcelular utilizando DeepLoc (<https://services.healthtech.dtu.dk/services/DeepLoc-2.0/>). En el caso de las defensinas, las secuencias además se agruparon según la presencia o ausencia de sus dominios y del péptido señal.

Alineamientos múltiples y árboles de distancia genética

Para cada familia de AMP se realizaron alineamientos múltiples de las proteínas identificadas en guayabo y secuencias de referencia utilizadas como *query*, mediante ClustalW utilizando BioEdit (Hall, 1999). Estos alineamientos se utilizaron para construir árboles de distancia genética mediante el programa MEGA, utilizando el método de Neighbor Joining y bootstrap de 5000 remuestreos. Se analizaron tanto los alineamientos como los árboles con el fin de identificar posibles subfamilias filogenéticas.

Validación por PCR

Extracción de ADN

Se realizó extracción de ADN genómico de hojas de *A. sellowiana* según el método de Doyle y Doyle (1991) con modificaciones. Se pulverizó una muestra de 0,5-1,0 g de tejido fresco en mortero estéril con nitrógeno líquido. El polvo se trasvasó a un eppendorf y se agregó 1 ml/0,2 g de buffer CTAB. Se mezcló y se incubó durante 30 minutos en agitación a 65°C. Se centrifugó 10 minutos a 12000 rpm y el sobrenadante se recuperó con un volumen de solución cloroformo isoamílico (24:1). Se centrifugó 2 minutos a 10000 rpm a 4°C y se precipitó utilizando 2/3 de volumen de isopropanol frío. Se repitió la centrifugación y el pellet se lavó con 500 µL de EtOH 70% frío y se centrifugó nuevamente. El pellet se dejó secar y se resuspendió en 385 µL de TE 1X, se agregaron 20 µL de ARNasa (10 µg/µL), y se incubó 30 minutos a 37°C. Posteriormente se agregaron 20 µL de NaCl 4M y precipitó con 2 volúmenes de EtOH abs 1 hora a -20°C. Se centrifugó 20 minutos a 10000g a 4°C. El pellet se secó y se resuspendió en 50 µL de agua mQ.

Buffer CTAB: 2% CTAB, 1,4 M NaCl, 0,2% B-mercaptoetanol, 20 mM EDTA, 100 mM Tris-HCL, pH 8,0.

PCR

Se seleccionaron algunas de las secuencias identificadas, candidatas de codificar AMPs, para su validación mediante PCR a partir de ADN genómico. Se diseñaron *primers* específicos para cada secuencia seleccionada, utilizando Primer3Plus (<https://www.primer3plus.com/>), ubicando los mismos en las regiones UTR para obtener la secuencia ORF completa de cada gen (en el caso de más de una isoforma se incluyó ambas posibilidades). La secuencia de los *primers* así como su ubicación en las secuencias de los transcritos se presenta en Material Suplementario (Anexo 3).

La enzima utilizada fue la DreamTaq DNA Polymerase (5 U/µL) de Thermo. La reacción utilizada se presenta a continuación:

DreamTaq Buffer 10X- 2 µl

Primer F 10 µM- 1 µl

Primer R 10 µM- 1 µl

dNTPs 10 µM c/u- 1 µl

Taq. Dream 5 u/µl- 0,2 µl

BSA 10 ng/µl- 0,1

ADN- 2 µl

Agua MQ- 12,7 µl

El programa utilizado fue el siguiente: 95°C por 30 segundos, seguido por 35 ciclos de 95°C por 30 segundos, 52-58°C por 40 segundos y 72°C por 40 segundos. Se realizó una puesta a punto del método variando la temperatura de *annealing* entre 50 y 60 hasta que se seleccionó una para cada uno de los *primers*, y se varió la concentración de ADN con diluciones entre 1 y 1/100.

Electroforesis de ADN

Para visualizar los productos de amplificación se realiza electroforesis en gel de agarosa 2%, en buffer TBE 0,5x con 0,5 µg/mL de GoodView®. Las muestras se sembraron con buffer de carga 5x. La corrida se realizó a 90 V durante 1 hora. Como referencia se sembró un marcador de peso molecular (6 µL).

Buffer Tris Borato (TBE) (10x): 108 g de Tris base, 55 g de ácido bórico, y 40 mL de EDTA 0,5 M (pH=8) en 1L de agua destilada.

Buffer de carga (5x): 0,25% de azul de bromofenol, y 30% de glicerol en agua.

Marcador de peso molecular GeneRuler™ 1 kb DNA Ladder (Thermo Scientific) (Figura 3).

GeneRuler™ 1 kb DNA Ladder O'GeneRuler™ 1 kb DNA Ladder, ready-to-use

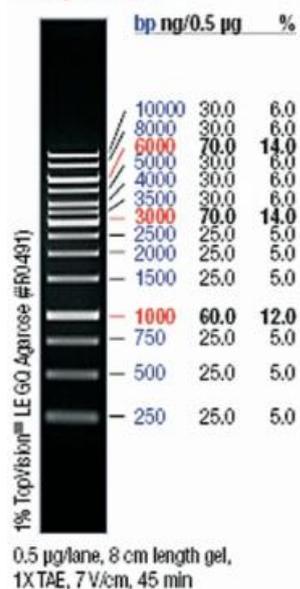


Figura 3. GeneRuler™ 1 kb DNA Ladder (Thermo Scientific).

Secuenciación

La identidad de los productos de PCR se verificó mediante secuenciación Sanger bidireccional, realizada por MacroGen Inc. (Seúl, Corea). Para ello, se enviaron a MacroGen. Los productos de PCR, junto con los *primers* específicos F y R, para cada secuencia

Alineamientos ADN genómico vs transcriptoma

Se realizaron alineamientos entre cada uno de los productos de PCR genómicos secuenciados y su transcripto respectivo, con el fin de validar los resultados, analizar la estructura exón-intrón, y observar posibles isoformas. Los alineamientos se realizaron utilizando las herramientas Clustal W y Pairwise Alignment de Bioedit, y Clustal Omega. Además de estas herramientas, en algunos casos se tuvo que terminar el alineamiento de forma manual utilizando Blast en Bioedit.

En los casos en que ambos productos (F y R) de secuenciación tuvieran buena calidad, se creó una secuencia consenso entre ambos en Bioedit y esta se utilizó para el alineamiento. Cuando esto no fuera posible, se utilizó solamente el producto de mejor calidad (F o R).

RESULTADOS

Identificación de las secuencias

Como resultado del proceso de BLAST y curado manual, se identificaron 178 secuencias con similitud a AMPs. Las secuencias identificadas comprenden 23 secuencias con similitud a defensinas (AsDf), 49 secuencias con similitud a esnaquinas (AsSn), siete secuencias con similitud a tioninas (AsTn), 12 secuencias con similitud a heveínas (AsHv), y 87 secuencias con similitud a LTP (AsLTP) (Tabla 1). No se identificaron secuencias con similitud a ciclótidos.

Para varias de las secuencias identificadas se encontraban anotadas al menos dos isoformas por gen (unigen) en el transcriptoma. Dentro de las isoformas, se identificaron algunas que diferían en su secuencia aminoacídica, y otras que eran idénticas. En las esnaquinas, la gran mayoría de las secuencias tuvieron al menos una isoforma (con la excepción de AsSn13, AsSn19 y AsSn42) (Tabla 1). Debido al gran número de secuencias encontradas, no se presentan las isoformas para los LTPs.

Las secuencias génicas fueron nombradas utilizando el prefijo As (*Acca sellowiana*), la familia (Df para defensinas, Sn para esnaquinas, Tn para tioninas, Hv para heveínas y

LTP para LTPs), y un número consecutivo. En el caso de las isoformas no idénticas (a nivel de proteína), se utilizó un número diferente para identificarlas. Para las isoformas idénticas, se utilizó el mismo número precedido de letras consecutivas.

Tabla 2. Número de secuencias putativas identificadas por familia e isoformas (según anotación del transcriptoma) por familias.

Familia	N°	N° genes	N° genes	Isoformas
	secue	sin	con	
	ncias	isoforma	isoforma	
Defensina	23	12	4	AsDf2a, AsDf2b, AsDf3, AsDf9a, AsDf9b
				AsDf6, AsDf7
				AsDf12, AsDf13
				AsDf19a, AsDf19b
Esnaquinas	49	4	14	AsSn1 AsSn2
				AsSn3, AsSn4
				AsSn5, AsSn6
				AsSn7, AsSn8, AsSn9, AsSn10
				AsSn11, AsSn12
				AsSn14a, AsSn14b, AsSn15
				AsSn16, AsSn17, AsSn18
				AsSn20a, AsSn20b
				AsSn21, AsSn22, AsSn23a, AsSn23b, AsSn24
				AsSn25, AsSn26a, AsSn26b, AsSn27, AsSn28, AsSn29
				AsSn30a, AsSn30b, AsSn30c, AsSn31, AsSn32, AsSn33
				AsSn34, AsSn35, AsSn36
				AsSn37, AsSn38
				AsSn39, AsSn40, AsSn41
Tioninas	7	5	1	AsTn3, AsTn4
Heveínas	12	10	1	AsHv5, AsHv6
Ciclótidos	0	-	-	
LTPs	87	-	-	

Caracterización de secuencias

Con el fin de caracterizar las secuencias proteicas identificadas, se analizó la presencia de péptido señal y su localización subcelular. En la familia de las tioninas, se encontró que todas las secuencias putativas poseían péptido señal. En relación a las defensinas, heveínas y esnaquinas, un 56,5%, 50,0% y 42,9% respectivamente de las secuencias identificadas poseían péptido señal (Figura 4). Para el caso de los LTPs, no se analizó la presencia de péptido señal debido al alto número de secuencias identificadas.

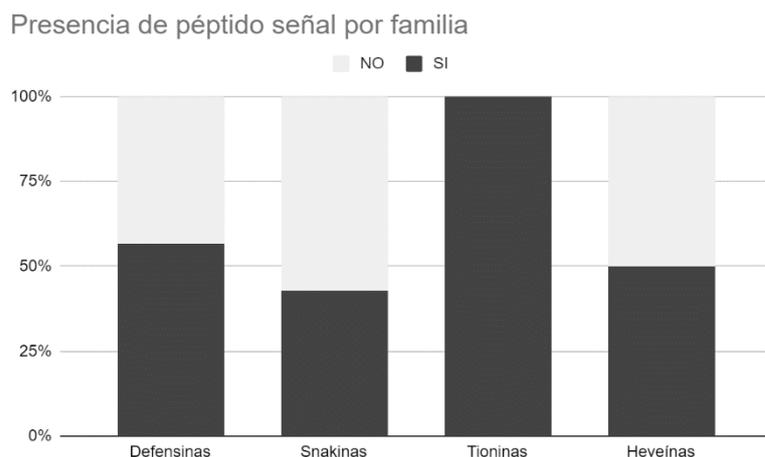


Figura 4. Porcentaje de las secuencias putativas con presencia de péptido señal en cada familia de AMPs.

En cuanto a la localización subcelular, todas las secuencias obtuvieron una alta probabilidad de localización extracelular (valores mayores a 60%). En algunos casos, también se detectó alta probabilidad de otras localizaciones como plástido (seis esnaquinas y una heveína) y citoplasma (una heveína).

En el caso de las defensinas, también se buscó la presencia o no de los motivos CSαβ y γ-core. En función de estos resultados, sumado a la presencia/ausencia de péptido señal y del número de cisteínas presentes, las 23 defensinas se dividieron en siete categorías. Se obtuvieron ocho secuencias en el grupo 1 (péptido señal, CSαβ completo, γ-core completo, 8 C conservadas): AsDf1 a AsDf7, tres secuencias en el grupo 2 (sin péptido señal, CSαβ completo, γ-core completo, 8 C conservadas), dos secuencias en el grupo 3 (CSαβ completo, γ-core completo, menos de 8 C), dos secuencias en el grupo 4 (8 C conservadas y CSαβ completo. γ-core con 10 residuos), una secuencia en el grupo 5 (menos de 8 C, CSαβ completo y γ-core ausente), una secuencia en el grupo 6 (menos de 8 C, CSαβ incompleto y γ-core presente) y seis secuencias en el grupo 7 (menos de 8 C, CSαβ incompleto y γ-core ausente). Del grupo 3 al 6 las secuencias pueden o no tener péptido señal.

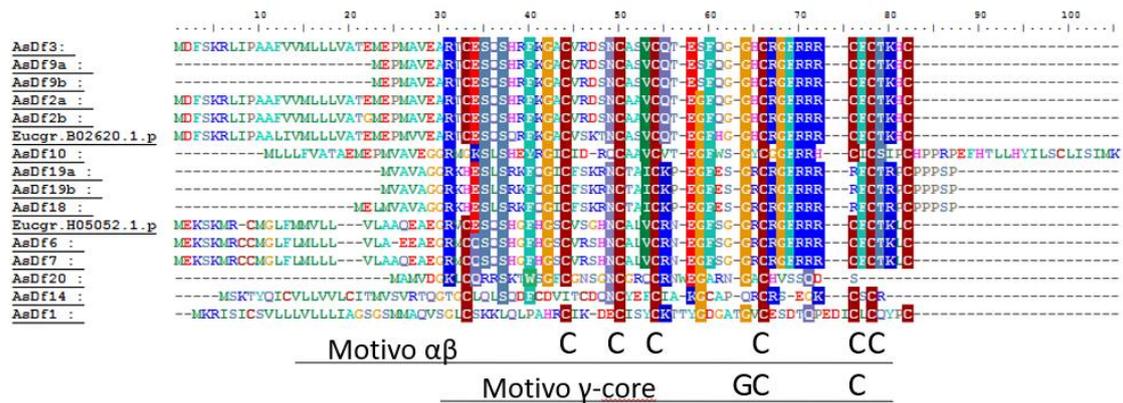
Alineamientos múltiples y árboles de distancia genética

Defensinas

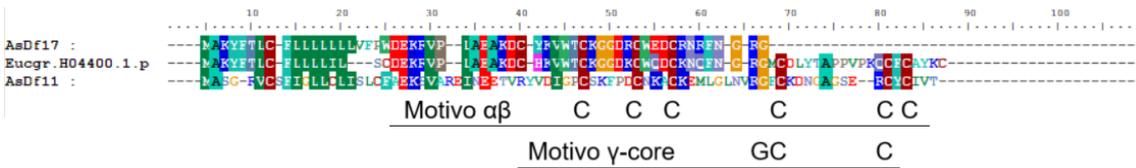
En el caso de las defensinas, se alinearon en conjunto con secuencias de *E. grandis* (Anexo 1 Tabla S1), identificando cuatro subgrupos (Figura 5 y Figura 6). El grupo 1 se caracteriza por la presencia de 10 aminoácidos (aa) entre la cisteína 1 y la cisteína 2, mientras que el grupo 2 tiene 5 aa y otra C en el extremo amino terminal, el grupo 3 tiene 6 aa y el grupo 4 8 aa. Por otro lado, cada uno de los grupos presenta ciertos

residuos conservados que se diferencian entre los grupos. Cabe agregar que los tres grupos tienen presencia de al menos una secuencia de *E. grandis* de referencia. En todos los grupos hay presencia de secuencias incompletas en las que no se detecta presencia de péptido señal o faltan alguno/s de los 8 residuos de cisteína. En los grupos 1 y 3 (alineamiento a y c) aparecen defensinas con un posible propéptido C-ter pero con baja similitud a otros pro-péptidos de defensinas de este tipo reportadas en la bibliografía.

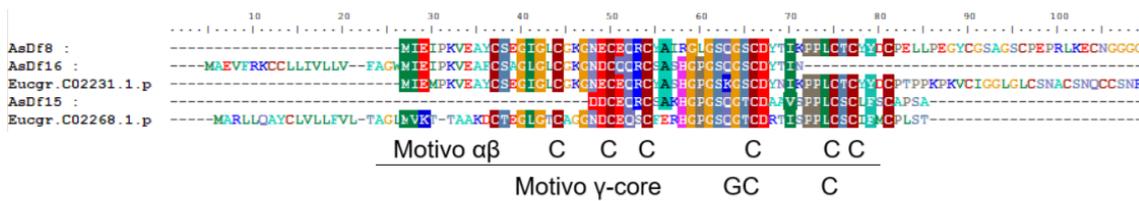
a)



b)



c)



d)

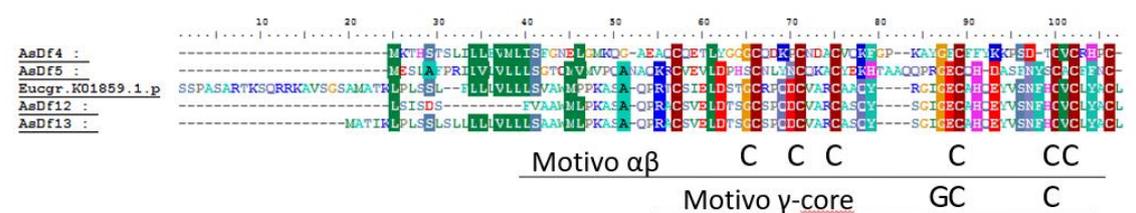


Figura 5. Alineamiento de defensinas identificadas con secuencias de *E. grandis* de referencia, a) grupo 1, b) grupo 2, c) grupo 3, y d) grupo 4. En los cuatro casos se presentan los motivos $\alpha\beta$ ($CX_nCXXXCX_nCX_nCXC$) y γ -core. ($GXCX_{3-9}C$)

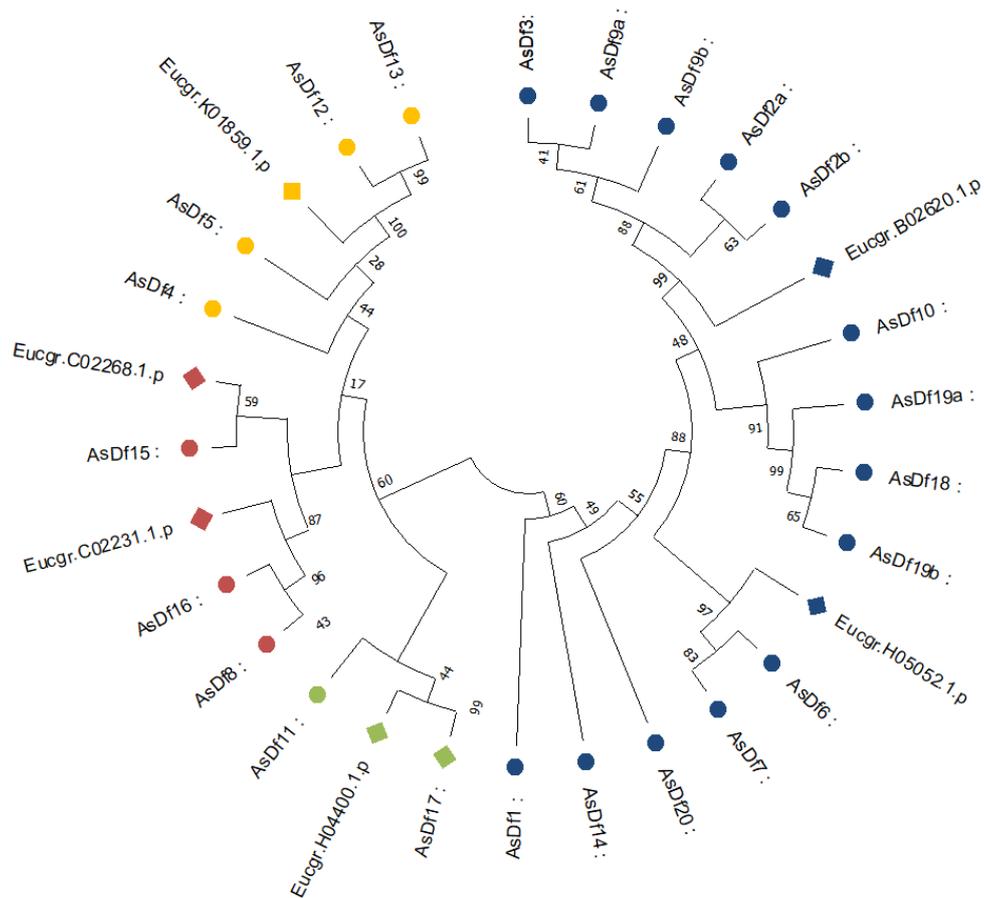


Figura 6. Árbol Neighbor Joining de distancia genética de péptidos defensinas putativos encontrados en los transcriptomas de guayabo, y defensinas de eucaliptus que se seleccionaron por similitud de secuencia aminoacídica. Los rombos representan a las defensinas de eucaliptus y los círculos a las proteínas de *A. sellowiana*. Los valores en los nodos muestran los valores de bootstrap (5000 réplicas). Se diferencian grupo 1 (azul), grupo 2 (verde), grupo 3 (rojo), y grupo 4 (amarillo).

Esnaquinas

En el caso de las esnaquinas, el alineamiento se realizó con las secuencias de referencia de las esnaquinas de papa (*S. tuberosum*) StSN1, StSN2, y StSN3, cada una perteneciente a una subfamilia distinta, la 1, 2 y 3 respectivamente. En el alineamiento también se identificaron tres grupos que corresponden con cada una de las subfamilias identificadas en papa (Figura 7 y Figura 8)

a)

```

      10      20      30      40      50      60
StSn1  SSFCDSRCKLRCSKAGLADRCIKYCSICCEEKRCVPSGTYGNKEEPCYRDAKNSKGGKPKCP
AsSn4  RCFCSSRCSARCSKAGLADRCIKYCSICCEKRCVPSGTYGNKEEPCYRDLKNSKGGKPKCP
AsSn3  SSFCSSRCSARCSKAGLADRCIKYCSICCEKRCVPSGTYGNKEEPCYRDLKNSKGGKPKCP
AsSn1  SSIKCSRCAVRCSEAGRMADRCIQYCSICCEKCCVPSGTYGHKEEPCYRDLKNSKGGKPKCP
AsSn2  -----KCAVRCSEAGRMADRCIQYCSICCEKCCVPSGTYGHKEEPCYRDLKNSKGGKPKCP
AsSn9  S-WCDTRCNVRCRSGADRCIKYCSICCEKCCVPSGTYGNKEEPCYRDLKNSKGGKPKCP
AsSn10 TGHCDTRCNVRCRSGADRCIKYCSICCEKCCVPSGTYGNKEEPCYRDLKNSKGGKPKCP
AsSn8  --SGLRCKDGRQNAQVADRCIKYCSICCEKCCVPSGTYGNKEEPCYRDLKNSKGGKPKCP
AsSn7  S-SGLRCKDGRQNAQVADRCIKYCSICCEKCCVPSGTYGNKEEPCYRDLKNSKGGKPKCP
AsSn10 -CYCEKRCGRQNAQVADRCIKYCSICCEKRCVPSGTYGNKEEPCYRDAKNSKGGKPKCP
AsSn11 VEGCEKRCGRQNAQVADRCIKYCSICCEKRCVPSGTYGNKEEPCYRDAKNSKGGKPKCP
AsSn6  PAFCEKRCRARCNAQVADRCIKYCSICCEKRCVPSGTYGNKEEPCYRDAKNSKGGKPKCP
AsSn5  TAFCEKRCRARCNAQVADRCIKYCSICCEKRCVPSGTYGNKEEPCYRDAKNSKGGKPKCP

```

b)

```

      10      20      30      40      50      60
StSn3  PRDCKKCFYRCSPTSHKKPCMFCCRCCARCLCVPPGTYGNKRCFCPCYNNWKTQGGPKCP
AsSn30a -LECCFRCKEERCSTAYKKPCMFCCRCCARCLCVPPGTYGNKRCFCPCYNNWKTTRGGPKCP
AsSn30b -LECCFRCKEERCSTAYKKPCMFCCRCCARCLCVPPGTYGNKRCFCPCYNNWKTTRGGPKCP
AsSn30c -LECCFRCKEERCSTAYKKPCMFCCRCCARCLCVPPGTYGNKRCFCPCYNNWKTTRGGPKCP
AsSn31 PCCECFRCKEERCSTAYKKPCMFCCRCCARCLCVPPGTYGNKRCFCPCYNNWKTTRGGPKCP
AsSn32 PCCECFRCKEERCSTAYKKPCMFCCRCCARCLCVPPGTYGNKRCFCPCYNNWKTTRGGPKCP
AsSn33 PCCECFRCKEERCSTAYKKPCMFCCRCCARCLCVPPGTYGNKRCFCPCYNNWKTTRGGPKCP
AsSn34 PCCECFRCKEERCSTAYKKPCMFCCRCCARCLCVPPGTYGNKRCFCPCYNNWKTTRGGPKCP
AsSn35 PCCECFRCKEERCSTAYKKPCMFCCRCCARCLCVPPGTYGNKRCFCPCYNNWKTTRGGPKCP
AsSn36 -----RCSCTSYKKPCMFCCRCCARCLCVPPGTYGNKRCFCPCYNNWKTTRGGPKCP
AsSn37 SYCCSCGSRRCCTSYKKPCLFCCRCCARCLCVPPGTYGNKRCFCPCYNNWKTTRGGPKCP
AsSn38 --ECSCGSRRCCTSYKKPCLFCCRCCARCLCVPPGTYGNKRCFCPCYNNWKTTRGGPKCP
AsSn39 PSDCPKCKEERCSTSHKKPCMFCCRCCARCLCVPPGTYGNKRCFCPCYNNWKTTRGGPKCP
AsSn40 PSDCPKCKEERCSTSHKKPCMFCCRCCARCLCVPPGTYGNKRCFCPCYNNWKTTRGGPKCP
AsSn41 VVDCPKCKEERCSTSHKKPCMFCCRCCARCLCVPPGTYGNKRCFCPCYNNWKTTRGGPKCP
AsSn42 PDCGAAADYRCSPTSHKKPCLFCCRCCARCLCVPPGTYGHKEEPCYNNWKTQGGPKCP

```

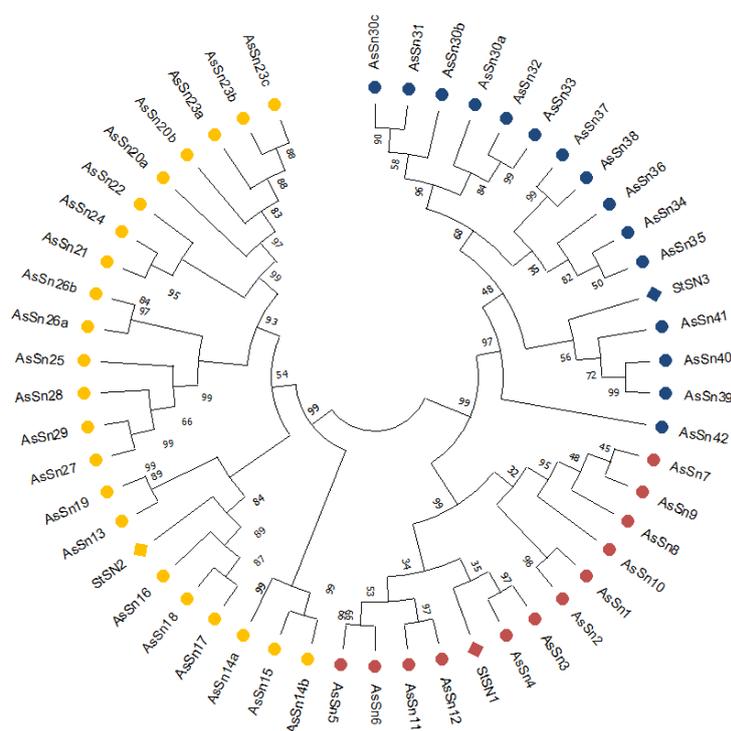
c)

```

      10      20      30      40      50      60
StSn2  KIDCGACARFCSLSSRIRLQFACSTCCRRCVPPGTSGNRLVC-FCYASLTTGHNKPKCP
AsSn13 KIDCGACARFCSLSSRIRLQFACSTCCRRCVPPGTSGNRLVC-FCYASLTTGHNKPKCP
AsSn14a SADCSSRCKLRCSKSRFMCDFACSTCCRRCVPPGTSGNRLVC-FCYANMTTGGSRPKCP
AsSn14b SADCSSRCKLRCSKSRFMCDFACSTCCRRCVPPGTSGNRLVC-FCYANMTTGGSRPKCP
AsSn15 KIDCSSRCKLRCSKSRFMCDFACSTCCRRCVPPGTSGNRLVC-FCYANMTTGGSRPKCP
AsSn16 CVDCGACARFCSLSSRIRLQFACSTCCRRCVPPGTSGNRLVC-FCYANMTTGGHNKPKCP
AsSn17 KIDCGACARFCSLSSRIRLQFACSTCCRRCVPPGTSGNRLVC-FCYANMTTGGHNKPKCP
AsSn18 -MDCGACARFCSLSSRIRLQFACSTCCRRCVPPGTSGNRLVC-FCYANMTTGGHNKPKCP
AsSn20a CADCGGLCKRCSLSSRIRLQFACSTCCRRCVPPGTSGNRELCRCYDLMTTGHNKPKCP
AsSn20b CADCGGLCKRCSLSSRIRLQFACSTCCRRCVPPGTSGNRELCRCYDLMTTGHNKPKCP
AsSn21 TLDCGGLCKRCSLSSRIRLQFACSTCCRRCVPPGTSGNRELCRCYDLMTTGHNKPKCP
AsSn22 TLDCGGLCKRCSLSSRIRLQFACSTCCRRCVPPGTSGNRELCRCYDLMTTGHNKPKCP
AsSn23a CADCGGLCKRCSLSSRIRLQFACSTCCRRCVPPGTSGNRELCRCYDLMTTGHNKPKCP
AsSn23b CADCGGLCKRCSLSSRIRLQFACSTCCRRCVPPGTSGNRELCRCYDLMTTGHNKPKCP
AsSn23c CADCGGLCKRCSLSSRIRLQFACSTCCRRCVPPGTSGNRELCRCYDLMTTGHNKPKCP
AsSn24 TLDCGGLCKRCSLSSRIRLQFACSTCCRRCVPPGTSGNRELCRCYDLMTTGHNKPKCP
AsSn25 KADDCIPLCDTRCSLSSRAFVCFACSTCCRRCVPPGTSGNRELCRCYDLMTTGHNKPKCP
AsSn26a AADDCIPLCDTRCSLSSRAFVCFACSTCCRRCVPPGTSGNRELCRCYDLMTTGHNKPKCP
AsSn26b AADDCIPLCDTRCSLSSRAFVCFACSTCCRRCVPPGTSGNRELCRCYDLMTTGHNKPKCP
AsSn27 KADDCIPLCDTRCSLSSRAFVCFACSTCCRRCVPPGTSGNRELCRCYDLMTTGHNKPKCP
AsSn28 KADDCIPLCDTRCSLSSRAFVCFACSTCCRRCVPPGTSGNRELCRCYDLMTTGHNKPKCP
AsSn29 KADDCIPLCDTRCSLSSRAFVCFACSTCCRRCVPPGTSGNRELCRCYDLMTTGHNKPKCP

```

Figura 7. Alineamiento del dominio GASA de 49 esnaquinas de guayabo con secuencias de referencia de *S. tuberosum*, a) subfamilia 1, b) subfamilia 2, y 3) subfamilia 3, representando los dominios snakin/GASA.



9

Figura 8. Árbol Neighbor Joining de distancia genética de péptidos 49 Snakin/GASA putativos encontrados en el transcriptoma de guayabo y esnaquinas 1, 2 y 3 de papa. La secuencia aminoacídica deducida de los transcritos fue alineada con StSN1, StSN2 y StSN3, agrupándose en las tres subfamilias propuestas por Berrocal Lobo et al. (2002). Los rombos representan a las esnaquinas de papa y los círculos a las proteínas de *A. sellowiana*. En rojo se muestra la subfamilia I, en azul la subfamilia III y en amarillo la subfamilia II. Los valores en los nodos muestran los valores de bootstrap (5000 réplicas).

Tioninas

Para las tioninas, se realizó un alineamiento de las secuencias candidatas de *A. sellowiana* con las secuencias de referencia de *E. grandis*, y secuencias revisadas de *Triticum aestivum* (trigo), *Hordeum vulgare* (cebada), *Arabidopsis thaliana* y *Viscum album* en base al dominio hevein. Se observa que dos de las secuencias (AsTn2 y AsTn3) poseen una sección adicional en el extremo C-terminal con repetidos LS y CV que no se encuentra en ninguna otra secuencia, incluyendo las de referencia. Por otro lado, se observa un patrón de C en todas las secuencias de *A. sellowiana* y de *E. grandis*, que no aparece en *T. aestivum*, *H. vulgare*, *A. thaliana* y *V. album*, como se muestra en la Figura 9.

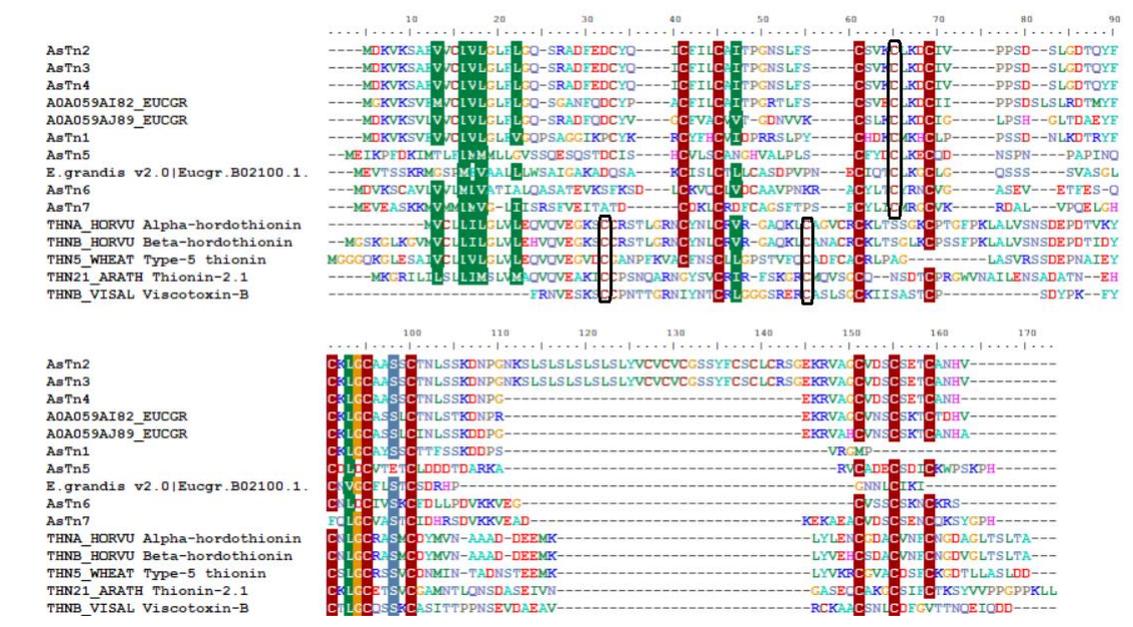


Figura 9. Alineamiento de siete tioninas de guayabo con secuencias de referencia de *E. grandis*, *T. aestivum*, *H. vulgare*, *A. thaliana* y *V. album*. En negro se marcan las diferencias en los patrones de cisteínas de las secuencias revisadas, en comparación con las putativas de *A. sellowiana* y *E. grandis*.

En cuanto al árbol, se observan tres grupos. El primero está conformado por las secuencias AsTn1, AsTn2, AsTn3, AsTn4 y dos secuencias de *E. grandis* anotadas como “thionin-like”, mientras que en el segundo se encuentran AsTn5, AsTn6, AsTn7 y una secuencia de *E. grandis*. Cabe agregar que las secuencias revisadas de *T. aestivum*, *H. vulgare*, *A. thaliana* y *V. álbum* conforman un tercer grupo que queda por fuera de los dos anteriores (Figura 10).

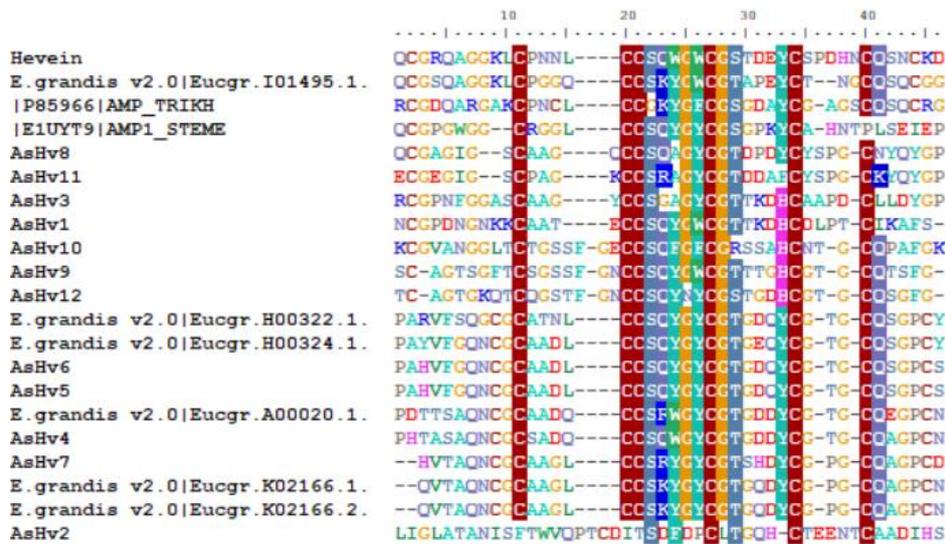


Figura 11. Alineamiento del motivo hevein de 12 hevein-like de guayabo con secuencias de referencia de *E. grandis*, *S. media*, *T. kiharae*, y la heveina de *H. brasiliensis*.

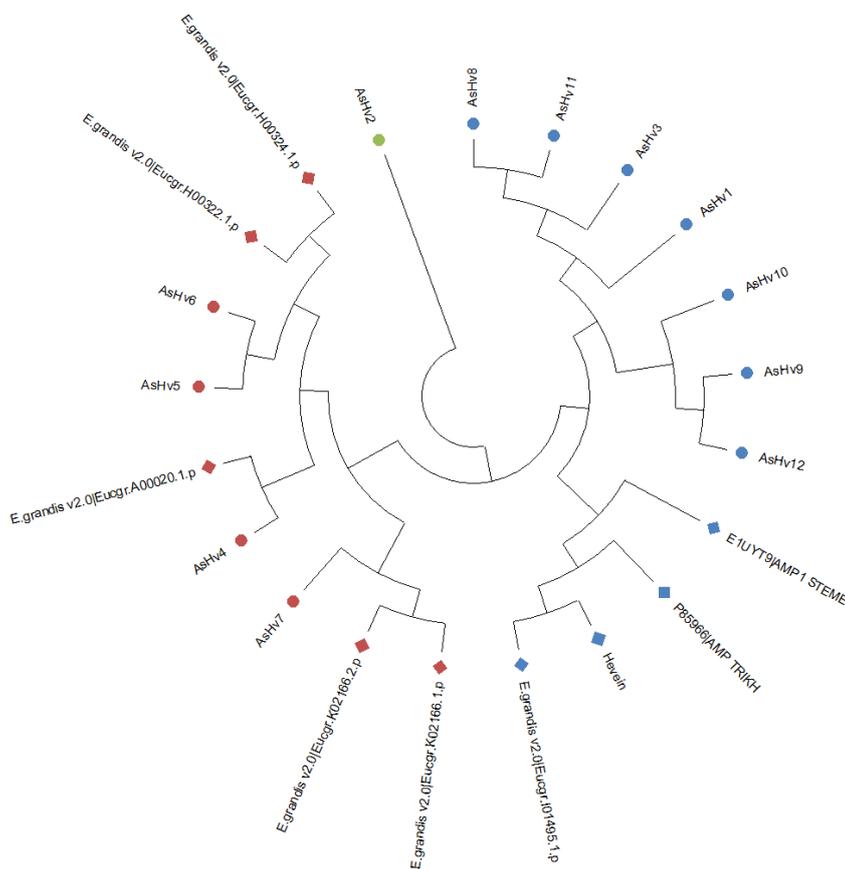


Figura 12. Árbol Neighbor Joining de distancia genética de 12 péptidos hevein-like putativos encontrados en el transcriptoma de guayabo y secuencias de referencia de *E. grandis*, *S. media*, *T. kiharae*, y la heveina de *H. brasiliensis*. Los rombos representan a las hevein-like de referencia y los círculos a las proteínas de *A. sellowiana*. En azul se muestra el grupo 1, en rojo el grupo 2 y en verde el grupo 3.

En el árbol de distancia genética, realizado a partir del alineamiento que se muestra en la figura 11, se detectaron tres grupos, (Figura 12). El grupo 1 incluye las proteínas de referencia de *H. brasiliensis*, *S. media*, *T. kiharae* y una proteína de *E. grandis*, el grupo 2 solo tiene referencias de *E. grandis*. Uno de estos grupos está conformado solamente por la AsHv2, que queda por fuera de los otros dos grupos (esta secuencia parece incompleta o mal ensamblada). Las proteínas AsHv se distribuyen de manera similar en ambos grupos.

LTPs

Para los LTP se realizó un alineamiento utilizando secuencias de referencia revisadas de *A. thaliana*, *Lens culinaris*, *Vigna radiata* y *Prunus armeniaca*. A partir de este alineamiento se construyó un árbol de distancia genética (Figura 13). El alineamiento no es presentado debido a la gran cantidad de secuencias putativas encontradas.

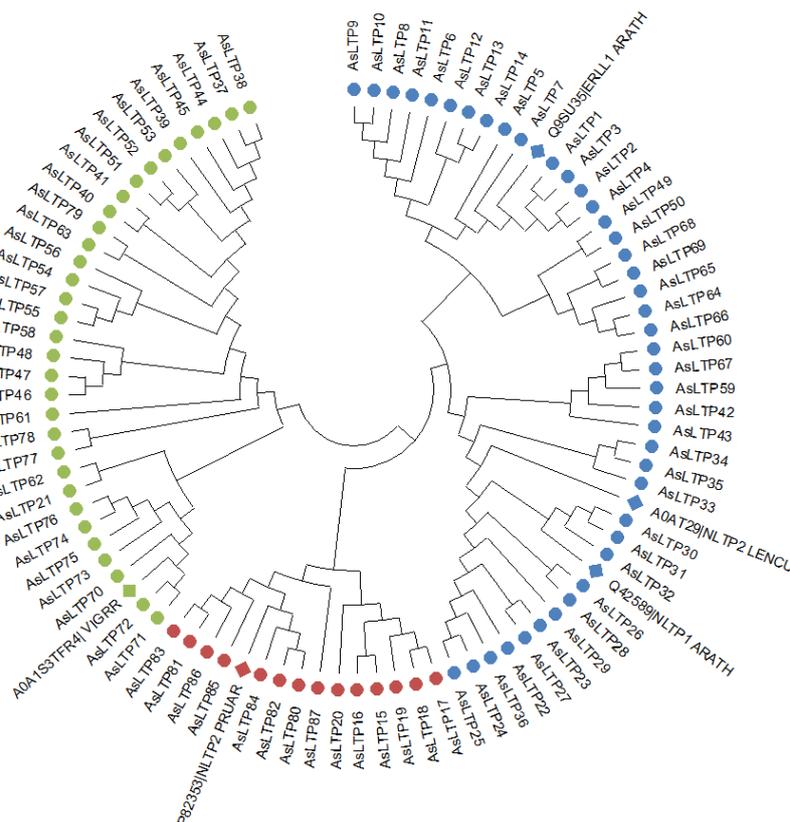


Figura 13. Árbol Neighbor Joining de distancia genética de 87 péptidos LTPs putativos encontrados en el transcriptoma de guayabo y secuencias de referencia de *A. thaliana*, *Lens culinaris*, *Vigna radiata* y *Prunus armeniaca*. Los rombos representan a los LTPs de referencia y los círculos a las proteínas de *A. sellowiana*. En verde se muestra el grupo 1, en azul el grupo 2 y en rojo el grupo 3.

A partir del árbol se identificaron tres grupos. De las cuatro secuencias de referencia que pertenecen a la subfamilia LTP1, tres de ellas quedaron en el grupo 2 (A0AT29, Q42589 y Q39176), y una quedó en el grupo 1 (A0A1S3TFR4). Una secuencia de referencia que pertenece a la subfamilia LTP2 (P82353), quedó en el grupo 3. Las proteínas AsLPT están mayoritariamente distribuidas (65 de 87) en los grupos 1 y 2 correspondientes a subfamilia LTP1, mientras que sólo 32 se localizaron en el grupo LTP2.

Validación de secuencias por PCR

Se seleccionaron 9 secuencias para su validación mediante PCR, cinco defensinas completas, una hevein-like completa y tres esnaquinas, y dentro de este último, se eligió una de cada una de las subfamilias. No se incluyeron representantes de las tioninas porque no presentan una secuencia similar a las tioninas típicas, ni de los LTP por la gran cantidad de secuencias putativas encontradas.

Tabla 3. secuencias seleccionadas para su validación en ADN genómico mediante PCR y posterior análisis de secuencia.

Defensinas	
AsDf1	MKRISICSVLLLVLIIIAGSGSMMAQVSGLCSSKLLQLPAHRCIKDECISYCKTTYGDGATGVCE SDTQPEDICLCQYPC
AsDf2	MDFSKRLIPAAFVVMLLLVAATEMEPMAVEARTCESQSHRFKGACVRDSNCAAVCQTEGFQG GHCRGFRRRCFCTKHC
AsDf4	MKTHSTSLILLFVMLISFGNELGMKQGAEAQCEQETLYGGGCQDKPCNDACVQKFGPKAYGF CFFYKPSDTCVCRHPC
AsDf5	MESLAFPRILVLLVLLSGTCMVMVPQANAQKRCVEVLDPHSCNLYNCQKACYEKHTAAQQPR GECQHDASFNYSCACFFNC
AsDf6	MEKSKMRCCMGLFLMLLLVLAEEAEGRMCQSQSHGFHFGSCVRSHNCALVCRNEGFSGGRC RGFRRRCFCTKLC
	MEKSKMRCCMGLFLMLLLVLAQAQAEGRMCQSQSHGFHFGSCVRSHNCALVCRNEGFSGGR CRGFRRRCFCTKLC
esnaquinas	
AsSn1	MMKKFALATFLLVALLSSFFDAATADSSLCDKCAVRCSKAGRMDRCLQYCGICCEKQC VPSGTYGHKDECPCYRDMKNSKGPKCP
	MRFVSVFVCFEGLCDKCAVRCSKAGRMDRCLQYCGICCEKQCVPSTGTYGHKDECPCYR DMKNSKGPKCP
AsSn13	MSLSKVLACLLSLLVLNQLVEAAEATELSSVESSEAAKKIDCGAACAAARCLASRQKICKRA CGTCCARCNCVPPGTAGNRDVCPCYASMTTHGGRLKCP
AsSn42	MASFLLRSSLFFLVASMCFVEVSLAGGEGSLRPDQCAAACDYRCSETSHRKPCLFFCNKC CAKCLCVPSGTYGHKEECPCYNNWKTQEGKPKCP
Hevein-like	
AsHv1	MRISTLAFLSLVSGALAGNCGPDNGNKKCAATECCSQYGWCGTTKDHCDLPTCIKAFSGASS PCEPSGSTPTTLKTSVVKQATSPATTFPSTVPNIDVCGSAQGGVTCPGVGADGYFYRCCSS AGHCGPKNDIQDQALYCGDGCQAGFGKCDSMAGPPVPTQTPGTAGDGETCGPIVNRKCGG SLCCSGSNFCGSGVDFCGAANWCQSKWGRCD

Se variaron las temperaturas de *annealing* y la concentración de ADN para determinar las mejores condiciones de amplificación (Figura 14). La temperatura de annealing óptima varió entre 54-58°C, dependiendo del gen a amplificar mientras que, respecto a la concentración de ADN, en ningún caso se observó amplificación con la muestra de ADN genómico sin diluir.

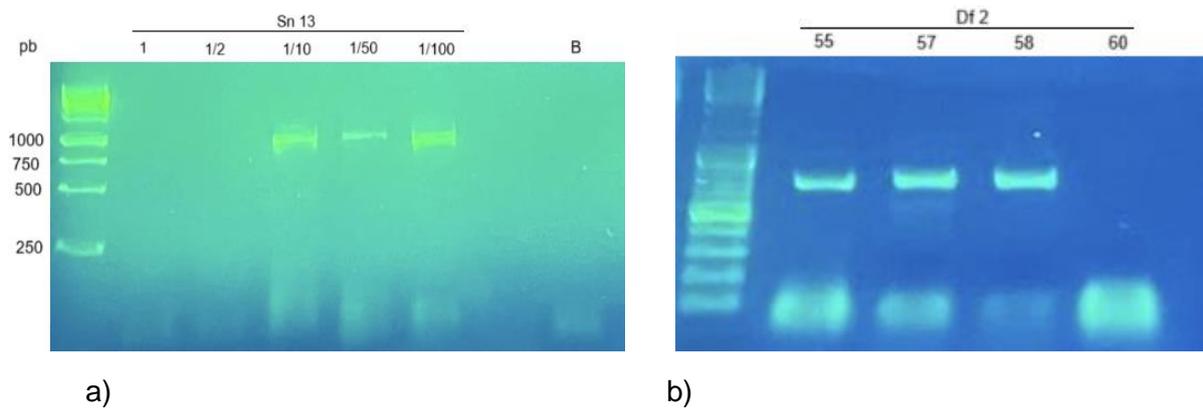
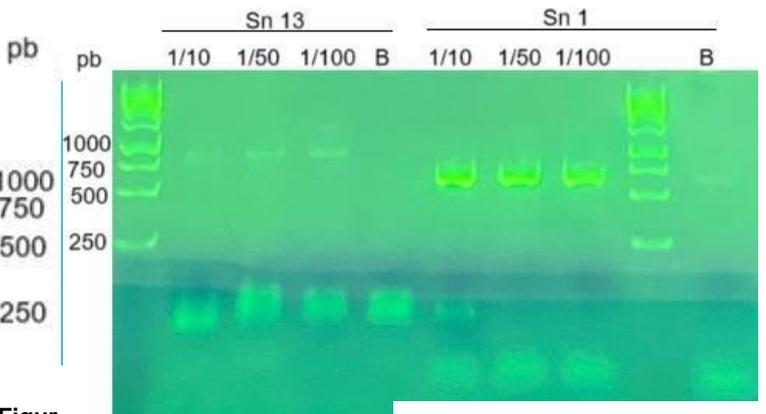
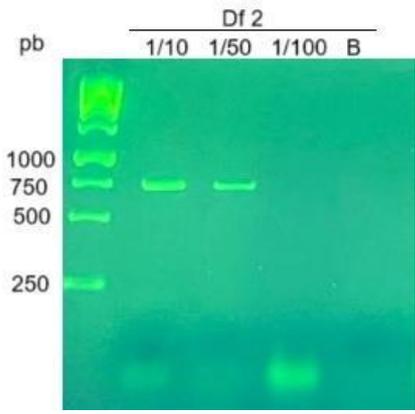
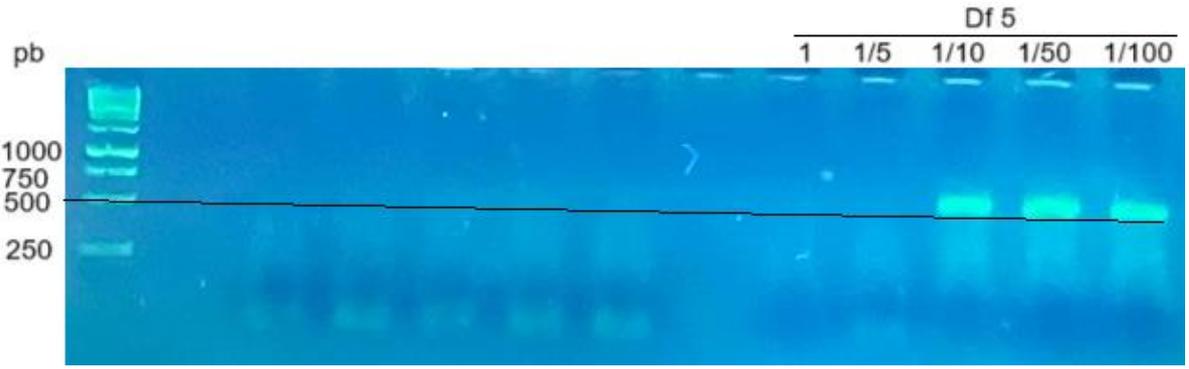
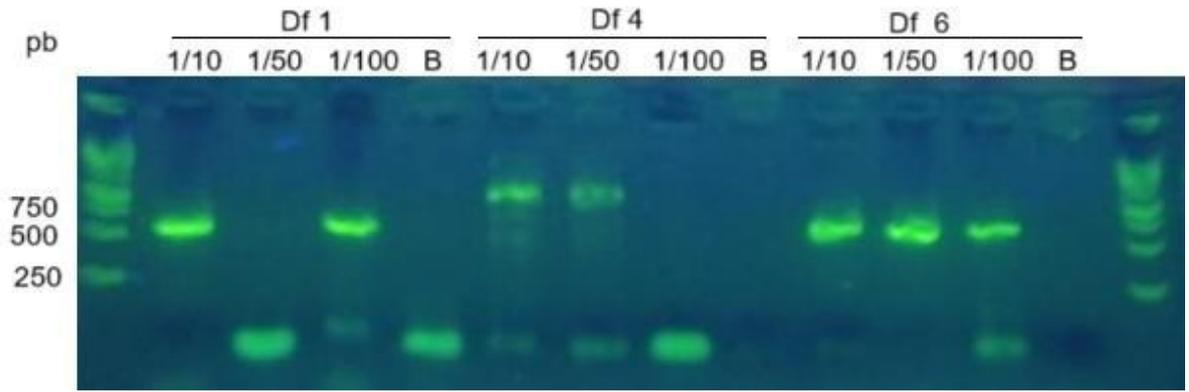


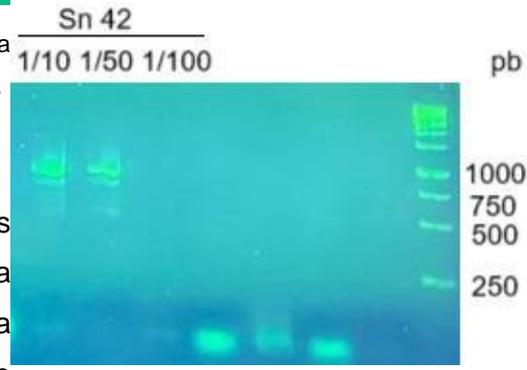
Figura 14. Pruebas de a) concentración de ADN, y b) temperatura de *annealing* (°C).

A concentraciones de ADN correspondientes a diluciones 1/10 y 1/100 todos los *primers* lograron amplificar de acuerdo con la visualización por electroforesis en gel de agarosa, si bien no se saben los tamaños del gen en el genoma, debido a los intrones, se espera que esté entre 500 y 1000 pares de bases (Figura 15).



Figur

a 15. visualización de la amplificación por PCR mediante geles de agarosa 2%. Para cada par de *primers* se muestran distintas concentraciones de ADN.



Análisis de secuencias obtenidas por PCR

A continuación, se presenta el análisis de los alineamientos entre las secuencias amplificadas por PCR y las obtenidas a través del transcriptoma. El producto de *AsSn42* no se envió a secuenciar debido a que se identificó más de una banda. Cabe resaltar que para *AsDf2*, *AsDf4*, *AsSn13* no se obtuvieron secuencias de buena calidad en la secuenciación posterior a la amplificación. Por esta razón no fue posible realizar alineamientos de estas secuencias y no son presentadas en esta sección. Es pertinente destacar que para *AsDf4* también se observa más de una banda en el gel.

AsDf1

Para la *Df1*, ambas secuencias resultantes (F y R) eran de buena calidad, por lo tanto, se realizó la secuencia consenso entre ellas. La misma fue usada posteriormente para el alineamiento con la secuencia de transcriptoma. Al alinear la secuencia consenso (*AsDf1* G) con la proveniente del transcriptoma (*AsDf* T) se observa la presencia de un gap en *AsDf* T que comienza en la posición 125 y culmina en la posición 267. La estructura coincide con la estructura exón-intrón-exón esperada para las defensinas (Figura 16). En la Figura 17 se presenta la estructura propuesta para esta defensina a partir del alineamiento. En el nucleótido 407 se agregó un nucleótido (A), debido a que la secuencia es de mala calidad, y en esta posición el marco de lectura cambiaba de +2 a +3.

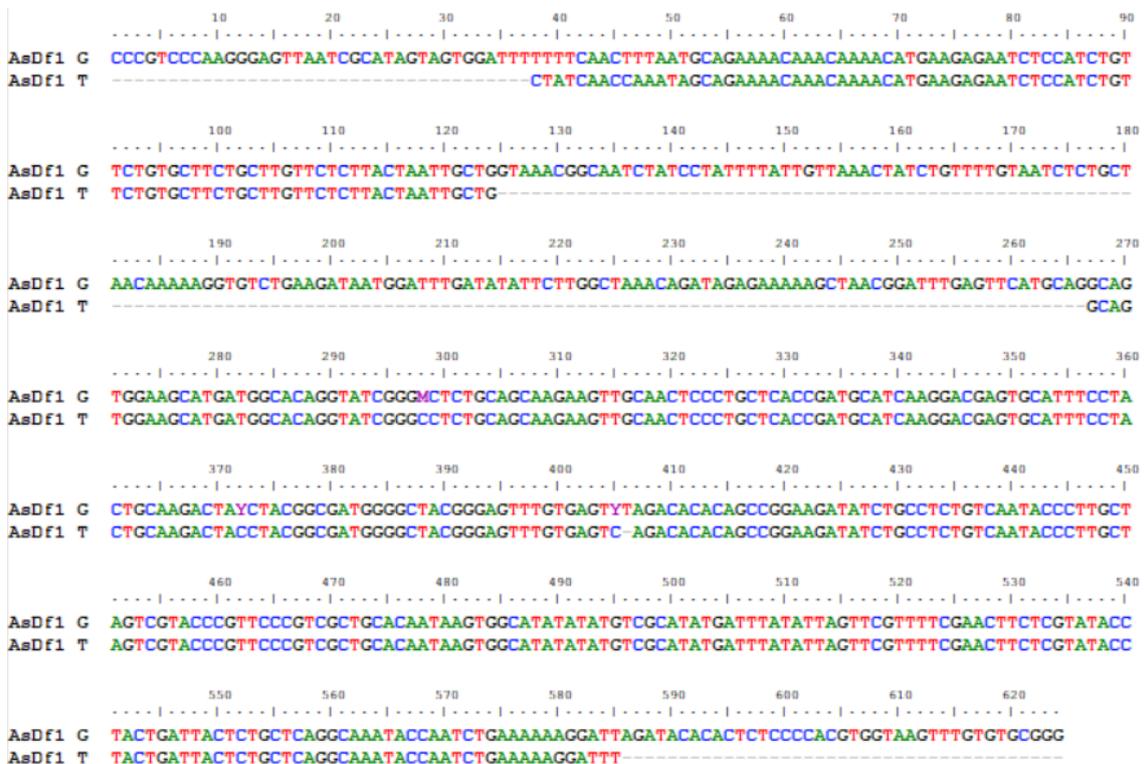


Figura 16. Alineamiento de la secuencia de *AsDf1* amplificada por PCR en ADN genómico (*AsDf1* G) y la obtenida a partir del transcriptoma de hoja y flor (*AsDf1* T). En el nucleótido 125 se observa el comienzo del intrón que finaliza en el nucleótido 267.



Figura 17. Estructura de la secuencia de *AsDf1* en ADN genómico. En gris se marcan las regiones 3' y 5' UTR, en verde claro el codón de comienzo, en rojo el codón de finalización, en celeste el intrón, y en amarillo la secuencia proteica. En fucsia se marca el nucleótido que fue agregado. Subrayados se muestran los nucleótidos o los aminoácidos que no tuvieron un consenso en la secuencia consenso entre la F y la R.

AsDf5

En el caso de la *AsDf5*, solamente se obtuvo la secuencia R de buena calidad, por lo tanto, se utilizó ésta (*AsDf* G) para realizar el alineamiento con la secuencia transcriptómica (*AsDf* T). En el alineamiento se observa un gap en *AsDf5* T desde la posición 123 a la posición 337 (Figura 18). Esta estructura también coincide con exón-intrón-exón, como es esperado para las defensinas. En la Figura 19 se presenta la estructura propuesta para esta defensina a partir del alineamiento. Se observa que la estructura propuesta no coincide exactamente con el alineamiento anterior, ya que el GT que da comienzo al intrón se encuentra en la posición 126 y el AG que finaliza se encuentra en la posición 327.

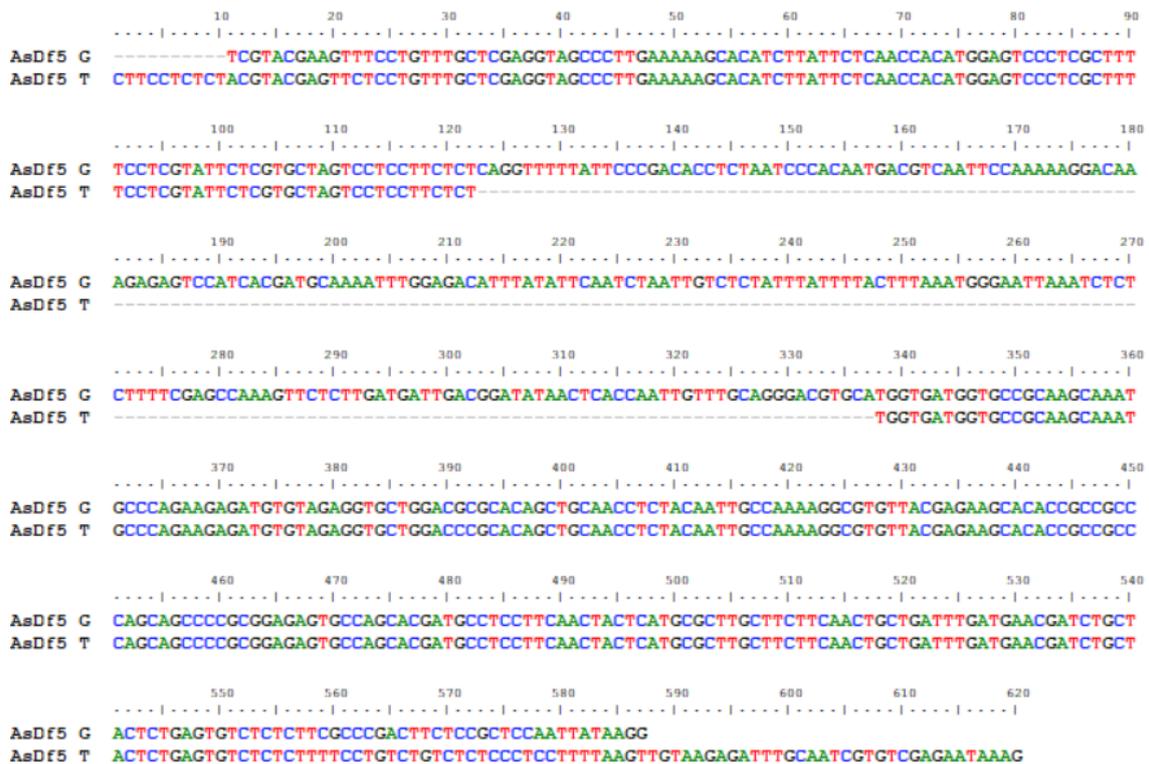


Figura 18. Alineamiento de la secuencia de *AsDf5* amplificada por PCR en ADN genómico (*AsDf5* G) y la obtenida a partir del transcriptoma de hoja y flor (*AsDf5* T). En el nucleótido 123 se observa el comienzo del intrón que finaliza en el nucleótido 337.



Figura 19. Estructura de la secuencia de *AsDf5* en ADN genómico. En gris se marcan las regiones 3' y 5' UTR, en verde el codón de comienzo, en rojo el codón de finalización, en celeste el intrón, y en amarillo la secuencia proteica.

AsDf6

Para la *AsDf6*, ambas secuencias eran de buena calidad (F y R), por lo cual se utilizó la secuencia consenso para el alineamiento (*AsDf6* G). En el alineamiento con la secuencia transcriptómica de *AsDf6* (AsDf6 T), se observa que, si bien alinean con 100% de identidad, no hay presencia de gap en *AsDf6* T (Figura 20). Esto indica la ausencia de intrón y no corresponde con la estructura esperada para una defensina. La *AsDf6* T presentaba una isoforma que también era amplificada por los *primers* utilizados (*AsDf7*). Se probó realizar el alineamiento de la *AsDf6* G con la secuencia transcriptómica de la *AsDf7* (*AsDf7* T), y en este caso sí se observó un gap, que inicia

en 154 y finaliza en 243 (Figura 21). La estructura observada para la *AsDf7* sí coincide con la esperada para las defensinas (exón-intrón-exón). En la Figura 23 se presenta la estructura propuesta para las defensina *AsDf7*.

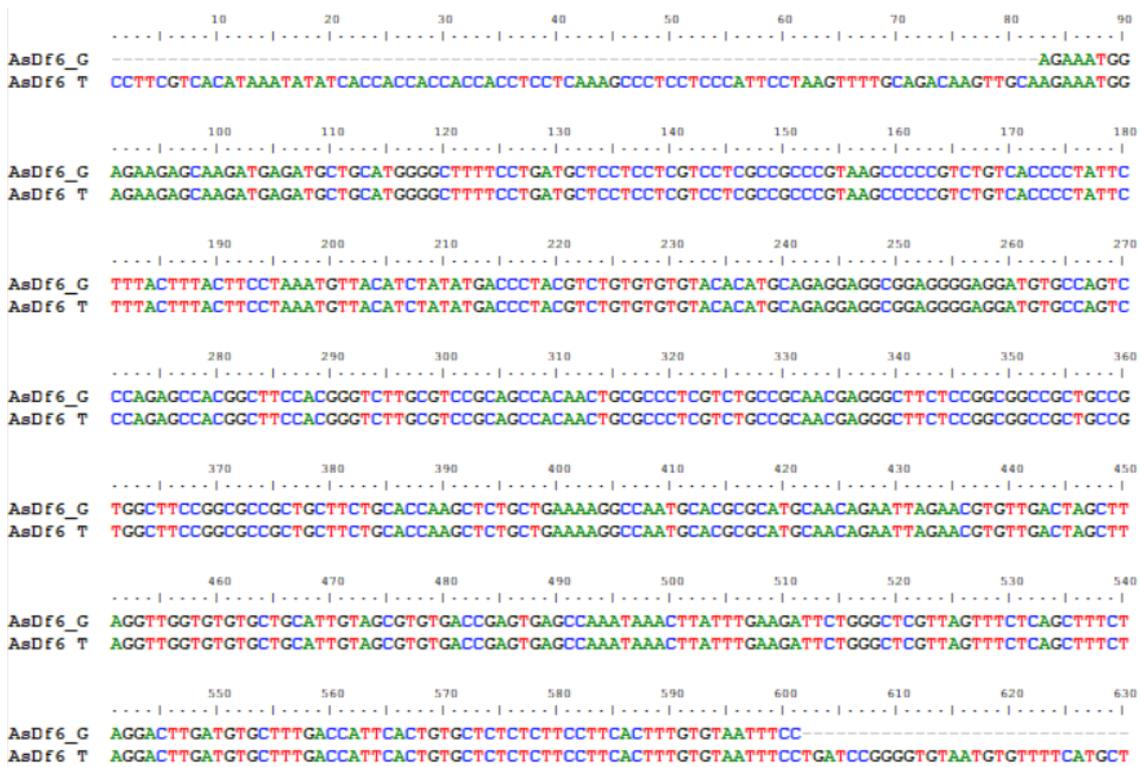


Figura 20. Alineamiento de la secuencia de *AsDf6* amplificada por PCR en ADN genómico (AsDf6 G) y la obtenida a partir del transcriptoma de hoja y flor (AsDf6 T).

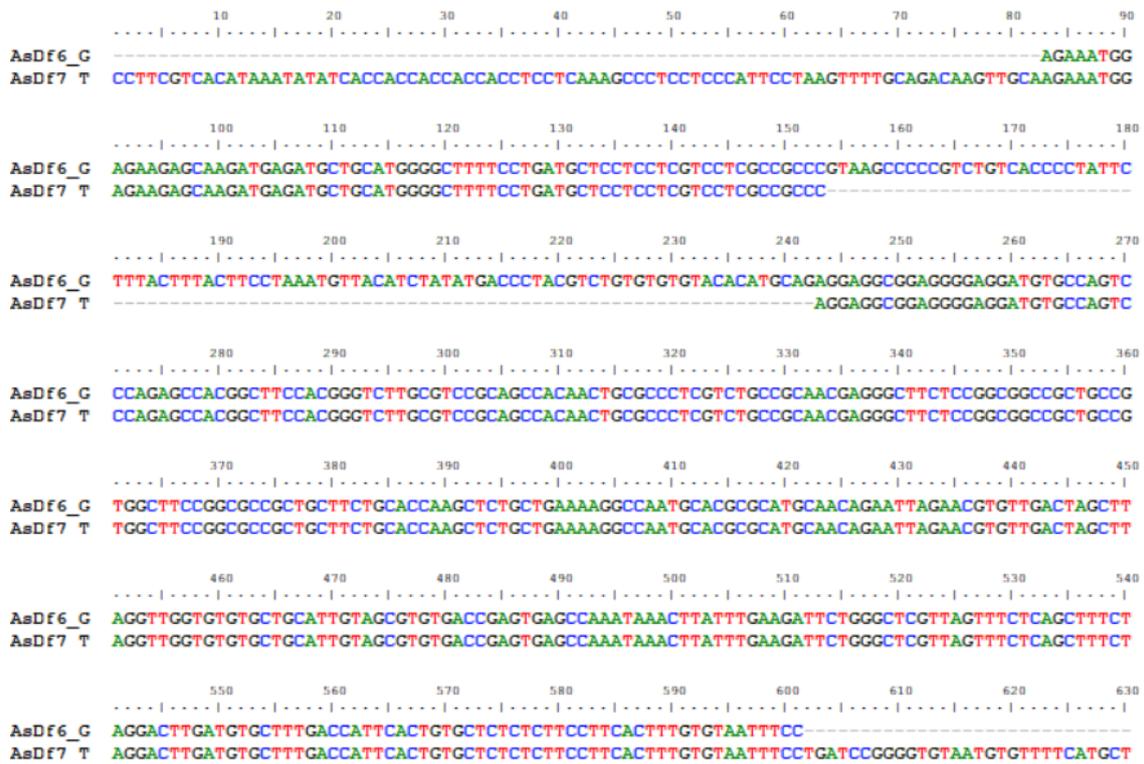


Figura 21. Alineamiento de la secuencia de *AsDf6* amplificada por PCR en ADN genómico (*AsDf6 G*) y la obtenida a partir del transcriptoma de hoja y flor (*AsDf7 T*). En el nucleótido 154 se observa el comienzo del intrón que finaliza en el nucleótido 243.



Figura 23. Estructura de la secuencia de *AsDf7* en ADN genómico. En gris se marcan las regiones 3' y 5' UTR, en verde el codón de comienzo, en rojo el codón de finalización, en celeste el intrón, y en amarillo la secuencia proteica

AsSn1

En el caso de la *AsSn1* se pudieron utilizar ambas secuencias (F y R) para la construcción de la secuencia consenso (*AsSn1 G*), y se alineó con la secuencia transcriptómica (*AsSn1 T*). Se presenta el alineamiento y se observa un gap que comienza en la posición 850 de la *AsSn1 T* y finaliza en la posición 957 (Figura 24). Esto coincide con una estructura exón-intrón-exón, como es esperado para las esnaquinas de la subfamilia 1. En la Figura 25 se presenta la estructura propuesta para esta defensina a partir del alineamiento.

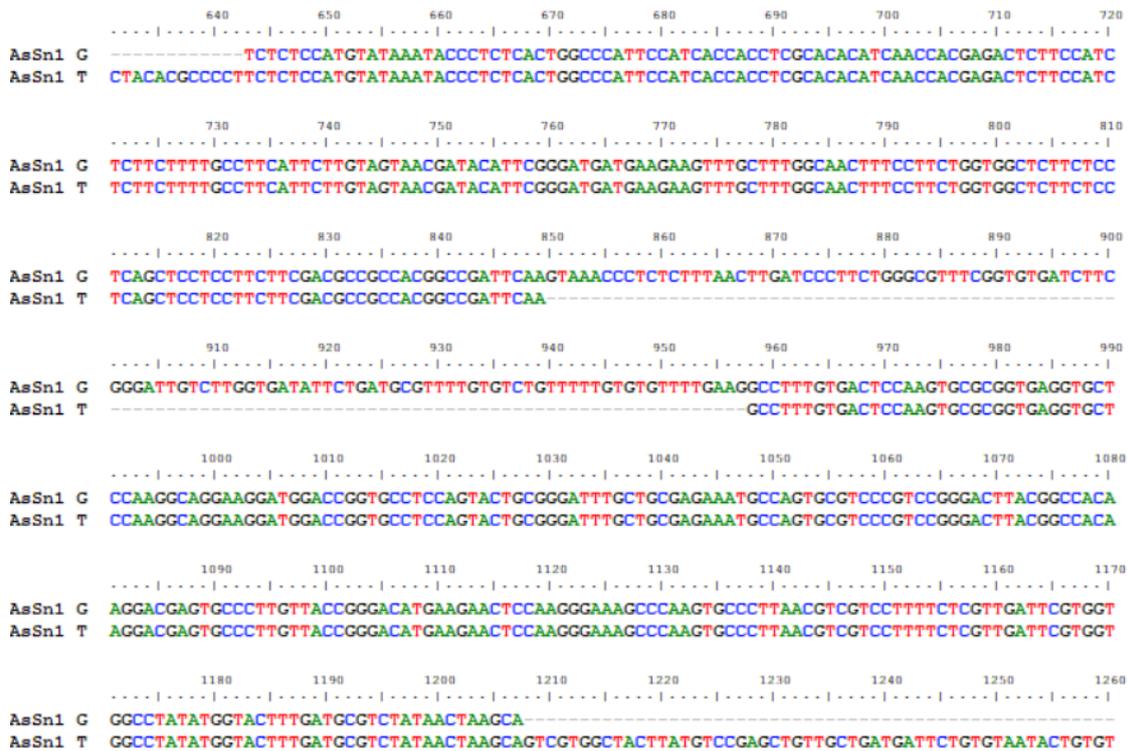


Figura 24. Alineamiento de la secuencia de *AsSn1* amplificada por PCR en ADN genómico (*AsSn1* G) y la obtenida a partir del transcriptoma de hoja y flor (*AsSn1* T). En el nucleótido 207 se observa el comienzo del intrón que finaliza en el nucleótido 315.



Figura 25. Estructura de la secuencia de *AsSn1* en ADN genómico. En gris se marcan las regiones 3' y 5' UTR, en verde el codón de comienzo, en rojo el codón de finalización, en azul el intrón, y en amarillo la secuencia proteica

AsHv1

Para la *AsHv1* se utilizó la secuencia F (*AsHv1* G), dado que la R era de mala calidad. Se realizó el alineamiento con la secuencia de transcriptoma (*AsHv1* T), y se observó un gap que inicia en la posición 221 de la *AsHv1* T, y finaliza en la posición 310 (Figura 26). Esto se corresponde con una estructura exón-intrón-exón. En la Figura 27 se presenta la estructura propuesta para esta defensina a partir del alineamiento. Cabe agregar que no se identificó el ATG del codón de inicio, sin embargo, se marcó donde debería estar el ATG, asumiendo que hacía 5' la secuencia de nucleótidos no es de buena calidad.

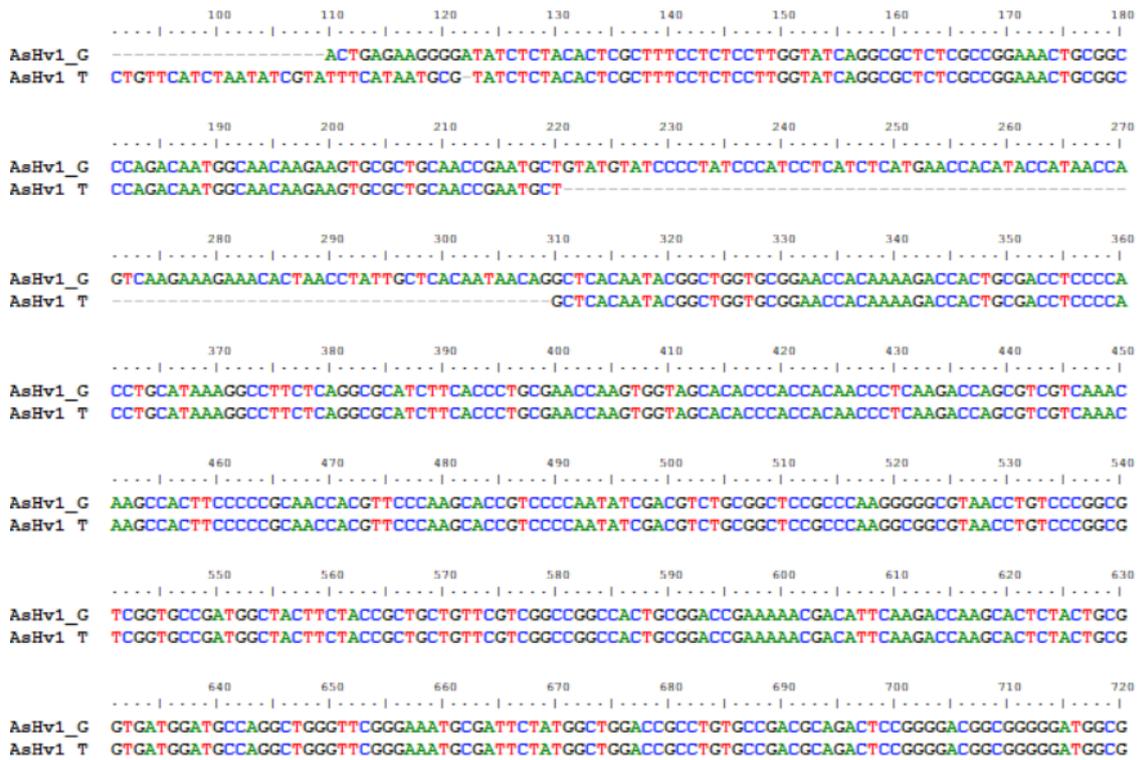


Figura 26. Alineamiento de la secuencia de *AsHv1* amplificada por PCR en ADN genómico (*AsHv1* G) y la obtenida a partir del transcriptoma de hoja y flor (*AsHv1* T). En el nucleótido 221 se observa el comienzo del intrón que finaliza en el nucleótido 310.

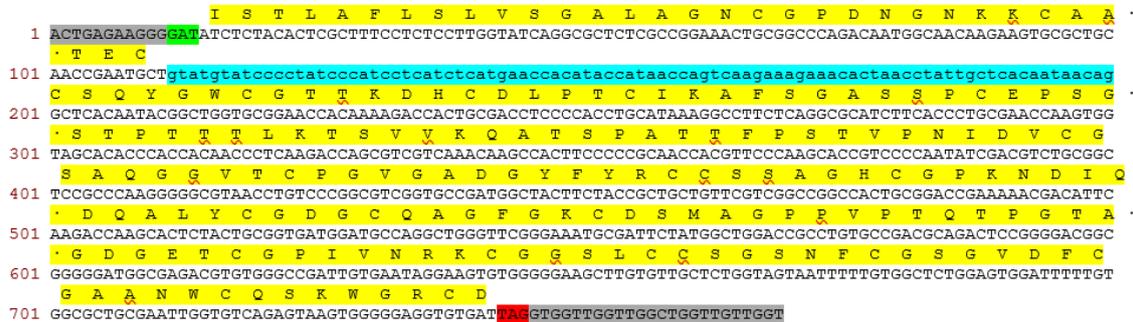


Figura 27. Estructura de la secuencia de *AsHv1* en ADN genómico. En gris se marcan las regiones 3' y 5' UTR, en verde el codón que se asume como el comienzo, en rojo el codón de finalización, en celeste el intrón, y en amarillo la secuencia proteica.

DISCUSIÓN

Los AMPs han demostrado ser moléculas sumamente promisorias para su uso anti-infeccioso en diversos ámbitos (Zasloff, 2002). Este trabajo tuvo como objetivo la búsqueda y caracterización de distintas familias de AMPs mediante la búsqueda de sus genes en el transcriptoma de *A. sellowiana*.

Búsqueda

La metodología de RNA-seq permitió detectar péptidos de todas las familias con excepción de los ciclótidos. En Phytozome tampoco se encuentra ninguna proteína anotada como “cyclotide” en el genoma de *E. grandis*. Hay antecedentes en nuestro país de búsqueda de AMPs en *Peltophorum dubium* (ibirapitá) y *Maytenus ilicifolia* (congorosa), donde tampoco se lograron detectar ciclótidos (Rodríguez Decuadro, 2018). Se plantea la discusión de si no hay ciclótidos en el genoma de mirtáceas o si la metodología utilizada no los permitió encontrar.

En cuanto a las defensinas, se detectaron 23 péptidos putativos, de los cuales 11 se presentaban en el transcriptoma como isoformas de otros. Rodríguez Decuadro (2018) utilizando la misma metodología obtuvo 14 defensinas para ibirapitá y 12 defensinas para congorosa. Slavokhotova et al. (2017b) identificaron 16 transcriptos que codifican para defensinas en semillas de *Stellaria media* sanas (no infectadas). En *E. grandis* se encontraron 15 defensinas anotadas (Phytozome, s.f.). La diferencia en la mayor cantidad obtenida en este trabajo puede deberse a isoformas resultantes de diferentes alelos por gen. Estas pueden ser reales o constructos del ensamblaje del transcriptoma, por lo que un refinamiento en la secuenciación o en el ensamblado podría reducir los resultados obtenidos. También puede deberse a que el transcriptoma es un pool de flor y hoja, donde se pueden estar expresando genes parálogos.

Con la esnaquinas ocurre una situación similar. Fueron detectados 49, de las cuales todas, con excepción de tres, se presentaban en el transcriptoma como isoformas de otras. En el trabajo de Rodríguez Decuadro (2018) se predijeron 18 transcriptos de esnaquinas putativas en ibirapitá, y 15 transcriptos en congorosa, mientras que en el trabajo de Li et al. (2022) en *Nicotiana tabacum* se predijeron 18 transcritos putativos. Shang et al. (2023) predijeron genes snakin-GASA en manglares, encontrando 27, 30 y 9 péptidos en *Avicennia marina*, *Kandelia obovata*, y *Aegiceras corniculatum* respectivamente. Sun et al. (2023) identificaron 15 genes snakin/GASA en el genoma de *Brassica rapa L. ssp. pekinensis*. Estos números son menores que los hallados en este trabajo por lo que el número mayor puede deberse a las isoformas. Sin embargo,

en el genoma de *E. grandis* hay tres proteínas anotadas como “snakin”, y 38 anotadas con el motivo GASA (Phytozome). El número similar al de *E. grandis* podría estar indicando que en la familia Myrtaceae el número de esnaquinas puede ser mayor.

Las hevein-like halladas fueron 12, con dos que se presentaban como isoformas. Estos números fueron similares a los reportados en otras especies con la misma metodología. En ibirapitá se encontraron ocho transcriptos hevein-like, mientras que en congrosa se encontraron 10 transcriptos (Rodríguez Decuadro, 2021). En *L. arenarius* (L.) se identificaron dos hevein-like (Slavokhotova et al., 2015), y en semillas de *S. media* sanas 14 hevein-like (Slavokhotova et al., 2017b). En eucaliptus se encuentran anotados en la base de datos Phytozome, cinco transcriptos para “hevein”.

En cuanto a las tioninas, se obtuvieron siete transcriptos putativos, de los cuales dos fueron presentados como isoformas. Los reportados para ibirapitá y congrosa son de 0 y 1, respectivamente (Rodríguez Decuadro, 2018). Para *L. arenarius* (L.) se reportaron 15 tioninas (Slavokhotova et al., 2015), y en transcriptoma de *S. media* se reportaron tres (Slavokhotova et al., 2017b). En *E. grandis* también hay siete tioninas anotadas (Phytozome), indicando que posiblemente se trate de un número asociado a la cercanía filogenética.

Con los LTPs sucede similar a las tioninas. Fueron encontrados 87, los cuales no fueron caracterizadas las isoformas debido a la gran cantidad. Rodríguez Decuadro (2018) reportó 28 LTPs putativos para ibirapitá, y 32 para congrosa, mientras que Slavokhotova et al. (2015) reportaron 51 LTPs, y 31 para semillas de *S. media* sanas (Slavokhotova et al., 2017b). Sin embargo, en *E. grandis* hay 91 proteínas anotadas como “lipid transfer protein” (Phytozome), por lo que, al igual que las tioninas, se puede asociar a la cercanía filogenética.

Caracterización

Se analizaron bioinformáticamente todos los transcriptos encontrados con el fin de predecir su localización subcelular e identificar la presencia de péptido señal, con la excepción de LTPs. Como era esperado, para todos los péptidos putativos se encontró que la localización extracelular fue la más probable ya que en plantas, los AMPs son en general secretados, por lo que su localización teórica es generalmente extracelular (Granz, 2004; Benko-Iseppon, 2010). El péptido señal, ha sido descrito en las cinco familias analizadas en este trabajo, y su presencia es de importancia para que la proteína madura sea secretada (De Coninck et al., 2013; Nahirňak et al., 2012; Slavokhotova et al., 2017a). El número de secuencias obtenidas con péptido señal varió según la familia. En el caso de las tioninas, el 100% presentó péptido señal, pero

en defensinas, heveínas y esnaquinas el número varió entre 56,5% y 42,9%. La ausencia de péptido señal en aproximadamente la mitad de estos transcritos puede deberse a errores en el ensamblado que genere transcritos donde falte el extremo 5'. Esto es apoyado por la gran cantidad de isoformas encontradas, que pueden ser constructos del ensamblado. Por otro lado, también podrían estar incompletos en el genoma como pseudogenes que aún se están expresando (Xie et al., 2019; Zou et al., 2009).

Las defensinas fueron el grupo donde se encontró mayor variabilidad de secuencia (Kovaleva et al., 2020; Lacerda et al., 2014; van der Weerden & Anderson, 2013) y estructura de cisteínas, por lo tanto, se dividieron en ocho grupos de acuerdo a la presencia los motivos CS $\alpha\beta$ y γ -core, la presencia de péptido señal y el número de cisteínas. Esto dio como resultado que de ocho de las 23 defensinas encontradas poseían todas las características de “defensinas verdaderas”, mientras que otras siete poseían menos de ocho C, CS $\alpha\beta$ incompleto y γ -core ausente. El resto se encontraban en algún punto intermedio. Esto mostró que efectivamente había una gran variabilidad en cuanto a la calidad de las secuencias y que una gran parte podían ser catalogadas como “defensinas raras”, que deberán confirmarse con estudios de ADN genómico.

Alineamientos y árboles de distancia génica

Se realizaron para todas las familias alineamientos múltiples y árboles de distancia génica.

En el caso de las defensinas, en el alineamiento se observó diferencias en los patrones de cisteínas que coinciden con los 4 grupos identificados en el árbol filogenético (Figuras 5 y 6). Si bien no se encontró que estos grupos se hayan reportado anteriormente en la literatura, cada uno de ellos presenta al menos una secuencia de *E. grandis* de referencia. Sin embargo, estos grupos no coinciden con los agrupados anteriormente según los motivos, cantidad de cisteínas y péptido señal. La mayoría de las defensinas “verdaderas” pertenecen al grupo 1, y dos de ellas pertenecen al grupo 4. Ninguna defensina “verdadera” pertenece al grupo 2 ni 3. Por otro lado, las defensinas “raras” también pertenecen en su mayoría al grupo 1, y una de ellas al grupo 2 y otra al grupo 3. Ninguna defensina “rara” pertenece al grupo 4.

En cuanto a las esnaquinas, se observa tanto en los alineamientos como en el árbol la existencia de 3 grupos, que coinciden con las subfamilias descritas en papa por Segura et al. (1999), Berrocal-Lobo et al. (2002), y Nahirñak et al. (2016). Tanto en los alineamientos como en el árbol están presentes la esnaquina de papa de referencia

perteneciente a cada una de las subfamilias (SNSt1, SNSt2 y SNSt3). Otros estudios también han dividido las proteínas snakin/GASA en tres subfamilias altamente conservadas en distintas especies (Li et al., 2022; Shang et al., 2023; Sun et al., 2023).

Referente a las tioninas, las secuencias de referencia (con excepción de las de eucaliptus) conforman un grupo que queda separado del resto, y en los alineamientos se observan diferencias en el patrón de cisteínas entre estas y las encontradas en *Acca sellowiana*. Esto hace pensar que las tioninas encontradas son diferentes a las utilizadas como organismos de referencia. Sin embargo, las tioninas de eucaliptus sí presentan mayor homología con las de *A. sellowiana*, y en el árbol se observan agrupadas. Cabe agregar que las secuencias de *A. sellowiana* conforman dos grupos, que cada uno de ellos posee secuencias de *E. grandis* de referencia. En el trabajo de Almaghrabi et al. (2019) se encontraron thionin-like en *A. thaliana* con el mismo patrón de cisteínas que el de las tioninas putativas de *A. sellowiana*, particularmente en las familias CRP2310 y CRP2360. En este trabajo, estas familias quedaban agrupadas (junto con integrantes de otras familias) en un único grupo.

Las hevein-like forman dos claros grupos, con secuencias referentes cada uno de ellos, con excepción de AsHv2 que queda por fuera. Esta última, si bien posee 3 C del patrón CX_{3,8}CX_{4}CCX_{5}CX_{6}CX_{3,5}CX_{1,3}C (Shelenkov et al., 2020), que era condición para ser seleccionada, le faltan varias C, entre ellas la doble CC de la posición 20. Con excepción de esta secuencia, el resto de las 11 hevein-like halladas poseen todas las C del motivo hevein.

En cuanto a los LTPs, las proteínas identificadas en *A. sellowiana* se distribuyeron mayoritariamente en dos grupos (1 y 2) que contienen secuencias LTP1 de referencia. En el grupo 3 se encuentra la secuencia de referencia de LTP2 junto a varias secuencias de *A. sellowiana*. Llamativamente, el grupo 3 se encuentra más cercano al grupo 1 en el árbol NJ que el grupo 2 (Figura 13)., Rodríguez Decuadro et al. (2021) dividió los LTP1 de *P. dubium* en tres grupos, entre los cuales la LTP1 de *Vigna radiata* A0A1S3TFR4 quedó como secuencia de referencia en un grupo separado del resto de las secuencias de referencia de LTP1. De forma similar, en este trabajo, la secuencia de *Vigna radiata* se ubicó, junto con otras secuencias de *A. sellowiana*, en el grupo 2, separada del resto de las secuencias LTP1 de referencia. De acuerdo con Santos-Silva et al. (2023), las variables fisicoquímicas deberían asociarse con características fenéticas y funcionales para generar un sistema de clasificación

funcional para este grupo de péptidos ya que el sistema de clasificación en subfamilias 1 y 2 parece artificial.

Validación de secuencias por PCR

De las nueve secuencias seleccionadas para validación, en la amplificación de la *AsSn42* y *AsDf4* se observa más de una banda amplificada. Lo ideal hubiera sido realizar clonación y secuenciar colonias puras para poder analizar si existe más de una forma del gen, sin embargo, por cuestiones de tiempo esto no fue posible en esta instancia y queda pendiente para su realización posterior. Es pertinente mencionar que la *AsSn42* no presentaba isoformas según el ensamblado del transcriptoma, y de las que sí poseían isoformas (*AsSn1* y *AsDf6*), amplificó una sola banda.

Se realizaron alineamientos entre las secuencias de ADN genómico amplificadas por PCR y los *contigs* (transcriptos) obtenidos a través del transcriptoma. Esto permite analizar la estructura exón-intrón de los genes amplificados. De las nueve secuencias AMPs efectivamente amplificadas, para cinco AMPs (tres defensinas, una esnaquinas, una hevein-like) se obtuvieron secuencias de buena calidad que permitieron realizar análisis posteriores

Las defensinas poseen una estructura de exón-intrón-exón (Carvalho & Gómez, 2009). De las secuencias génicas de las defensinas analizadas, la *AsDf1* y la *AsDf5* presentan una clara estructura exón-intrón-exón que coincide con la esperada para este tipo de péptido. En cuanto a la *AsDf6*, el alineamiento entre la secuencia genómica amplificada y el *contig* (transcripto) correspondiente no reveló la presencia de ningún intrón. Sin embargo, al ser alineada con el transcripto *AsDf7* (isoforma de la *AsDf6*), se observó una estructura de exón-intrón-exón. Se puede pensar que *AsDf6* y *AsDf7* se traten de dos isoformas existentes, o que el transcripto *Df6*, erróneamente conservó un intrón sin procesar. La *AsDf1* presentó una parte de la secuencia que no estaba presente en la secuencia del transcriptoma y un cambio en el marco de lectura de +2 a +3. Esto puede ocurrir por la falta de un nucleótido por error en la secuenciación, que corra 1 nucleótido el marco de lectura.

La estructura génica de las esnaquinas varía según su subfamilia. La estructura de la subfamilia 1 está conformada por dos exones y un intrón, mientras que en la subfamilia 2 y la subfamilia 3 esta estructura es variable entre 2 y 3 intrones (Nahirñak et al., 2016). La secuencia génica *AsSn1* analizada presenta una estructura exón-intrón-exón, que corresponde con la estructura de la subfamilia 1. Este resultado coincide con los resultados obtenidos mediante alineamiento múltiple y árbol filogenético a partir de secuencia aminoacídica, según los cuales la esnaquina *AsSn1*

se encontraba dentro del grupo de la subfamilia 1. Queda pendiente el análisis de las secuencias génicas *AsSn13* y *AsSn42* para ver si sus estructuras coinciden con la de su subfamilia correspondiente.

Referente a la secuencia génica *AsHv1* (heveína), se logró identificar la presencia de un intrón. La literatura hallada referente a la estructura génica de la heveína es escasa. Van Damme et al., (1999) reportan dos intrones para un gen de hevein-like expresado en fruto de *Sambucus nigra*. Sin embargo, el gen de heveína de *H. brasiliensis* posee un intrón, conservado también en otros cuatro miembros de la familia génica hallados en esta especie (Pujade-Renaud et al., 2005).

CONCLUSIONES Y PERSPECTIVAS

Utilizando un abordaje bioinformático, se logró identificar por primera vez en transcriptoma *de novo* de hoja y flor de *Acca sellowiana*, AMPs de cinco familias distintas: 23 defensinas, 49 esnaquinas, 7 tioninas, 12 hevein-like, y 87 LTPs. No se encontró ninguna secuencia con homología a ciclotidos.

Mediante alineamientos con secuencias de referencia y construcción de árboles de distancia génica se caracterizaron los AMP codificados en el transcriptoma de *A. sellowiana* para cada una de las familias. Algunos AMPs de *A. sellowiana* se identificaron como miembros de grupos novedosos de las familias defensinas, tioninas y hevein-like, mientras que para esnaquinas y LTPs, las proteínas de *A. sellowiana* fueron similares a los ya caracterizados en otras especies. Se validó la existencia de nueve de estas secuencias en ADN genómico mediante PCR. Para todos los *primers* diseñados se obtuvieron secuencias del tamaño esperado para los AMPs. Sin embargo, para solo cinco de los fragmentos analizados, la calidad de secuencia fue buena, permitiendo identificar las estructuras exón-intrón. Además, para el gen *AsDf6* se logró identificar una posible isoforma.

Como perspectivas a este trabajo se continuará analizando las secuencias restantes de las cuales no se pudo obtener secuenciación de buena calidad. Para el caso de *AsSn42*, que se observaron dos bandas como resultado de la PCR, cada fragmento de amplificación será clonado mediante transformación de bacterias con plásmidos recombinantes, con el fin de comparar las secuencias clonadas en colonias aisladas.

A futuro se pretende expresar estas proteínas de forma heteróloga y realizar ensayos para evaluar su actividad antimicrobiana.

BIBLIOGRAFÍA

- Almaghrabi, B., Ali, M. A., Zahoor, A., Shah, K. H., & Bohlmann, H. (2019). Arabidopsis thionin-like genes are involved in resistance against the beet-cyst nematode (*Heterodera schachtii*). *Plant Physiology and Biochemistry*, *140*, 55–67. <https://doi.org/10.1016/j.plaphy.2019.05.005>
- Azmi, S., & Hussain, M. K. (2021). Analysis of structures, functions, and transgenicity of phytopeptides defensin and thionin: a review. *Beni-Suef University Journal of Basic and Applied Sciences*, *10*(1). <https://doi.org/10.1186/s43088-020-00093-5>
- Benko-Iseppon, A. M., Galdino, S. L., Calsa, T., Jr, Kido, E. A., Tossi, A., Belarmino, L. C., & Crovella, S. (2010). Overview on plant antimicrobial peptides. *Current Protein & Peptide Science*, *11*(3), 181–188. <https://doi.org/10.2174/138920310791112075>
- Berrocal-Lobo, M., Segura, A., Moreno, M., López, G., García-Olmedo, F., & Molina, A. (2002). Snakin-2, an antimicrobial peptide from potato whose gene is locally induced by wounding and responds to pathogen infection. *Plant Physiology*, *128*(3), 951–961. <https://doi.org/10.1104/pp.010685>
- Bohlmann, H., & Broekaert, W. (1994). The Role of Thionins in Plant Protection. *Critical Reviews in Plant Sciences*, *13*(1), 1–16. <https://doi.org/10.1080/0735268940970190>
- Boutera, M. T., Ben Romdhane, W., Baazaoui, N., Alfaifi, M. Y., Chouaibi, Y., Ben Akacha, B., Ben Hsouna, A., Kačániová, M., Čavar Zeljković, S., Garzoli, S., & Ben Saad, R. (2023). GASA proteins: Review of their functions in plant environmental stress tolerance. *Plants*, *12*(10), 2045. <https://doi.org/10.3390/plants12102045>
- Brogden, K. A. (2005). Antimicrobial peptides: pore formers or metabolic inhibitors in bacteria?. *Nature Reviews Microbiology*, *3*(3), 238–250. <https://doi.org/10.1038/nrmicro1098>
- Carvalho, A. de O., & Gomes, V. M. (2007). Role of plant lipid transfer proteins in plant cell physiology—a concise review. *Peptides*, *28*(5), 1144–1153. <https://doi.org/10.1016/j.peptides.2007.03.004>
- Carvalho, A. de O., & Gomes, V. M. (2009). Plant defensins—prospects for the biological functions and biotechnological properties. *Peptides*, *30*(5), 1007–1020. <https://doi.org/10.1016/j.peptides.2009.01.018>

- Castro, M S., Gerhardt, I.R., Orrù, S, Pucci, P, Bloch Jr, C. 2003. Purification and characterization of a small (7.3 kDa) putative lipid transfer protein from maize seeds. *Journal of Chromatography B*, 794: 109-114.
- Castro, M S, Fontes, W. 2005. Plant defense and antimicrobial peptides. *Protein Peptide Letters*, 12 11-16.
- Charlet, M., Chernysh, S., Philippe, H., Hetru, C., Hoffmann, J. A., & Bulet, P. (1996). Innate immunity: Isolation of several cysteine-rich antimicrobial peptides from the blood of a mollusc, *Mytilus edulis*. *The Journal of Biological Chemistry*, 271(36), 21808–21813. <https://doi.org/10.1074/jbc.271.36.21808>
- Craik, DJ. 2012. Host-defense activities of cyclotides. *Toxins*, 4: 139-156.
- Craik, DJ, Swedberg, JE, Mylne, JS, Emazar, M. 2012. Cyclotides as a basis for drug design. *Expert opinion on drug discovery*, 7: 179-194.
- Craik, D.J., Simonsen, S., & Daly, N. L. (2002). The cyclotides: novel macrocyclic peptides as scaffolds in drug design. *Current Opinion in Drug Discovery & Development*, 5(2), 251–260.
- Craik, DJ, Daly, NL, Bond, T, Waine, C. 1999. Plant cyclotides: a unique family of cyclic and knotted proteins that defines the cyclic cystine knot structural motif. *Journal of molecular biology*, 294: 1327-1336.
- De Coninck, B., Cammue, B. P., & Thevissen, K. (2013). Modes of antifungal action and in planta functions of plant defensins and defensin-like peptides. *Fungal Biology Reviews*, 26(4), 109-120.
- Donazzolo, J., Stefenon, V. M., Guerra, M. P., & Nodari, R. O. (2020). On farm management of *Acca sellowiana* (Myrtaceae) as a strategy for conservation of species genetic diversity. *Scientia Horticulturae*, 259(108826), 1-6, <https://doi.org/10.1016/j.scienta.2019.108826>
- Ganz, T. (2004). Defensins: antimicrobial peptides of vertebrates. *Comptes Rendus Biologies*, 327(6), 539-549, <https://doi.org/10.1016/j.crv.2003.12.007>
- Hall T (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl Acids Symp Ser* 41:95-98
- Hegedüs, N, Marx, F. 2012. Antifungal proteins: More than antimicrobials? *Fungal biology reviews*. 26:132-145

- Höng, K., Austerlitz, T., Bohlmann, T., & Bohlmann, H. (2021). The thionin family of antimicrobial peptides. *PloS One*, *16*(7), e0254549. <https://doi.org/10.1371/journal.pone.0254549>
- Irving, H. R., & Gehring, C. (Eds.). (2012). *Plant Signaling Peptides*. Springer Berlin Heidelberg.
- Kaur, J., Sagaram, U. S., & Shah, D. (2011). Can plant defensins be used to engineer durable commercially useful fungal resistance in crop plants? *Fungal Biology Reviews*, *25*(3), 128–135. <https://doi.org/10.1016/j.fbr.2011.07.004>
- Kovaleva, V., Bukhteeva, I., Kit, O. Y., & Nesmelova, I. V. (2020). Plant defensins from a structural perspective. *International Journal of Molecular Sciences*, *21*(15), 5307. <https://doi.org/10.3390/ijms21155307>
- Lacerda, A. F., Vasconcelos, E. A. R., Pelegrini, P. B., & Grossi de Sa, M. F. (2014). Antifungal defensins and their role in plant defense. *Frontiers in Microbiology*, *5*, 116. <https://doi.org/10.3389/fmicb.2014.00116>
- Larrañaga, P., Díaz-Dellavalle, P., Cabrera, A., Alem, D., Leoni, C., Almeida-Souza, A. L., Giovanni-De-Simone, S., & Dalla-Rizza, M. (2012). Activity of naturally derived antimicrobial peptides against filamentous fungi relevant for agriculture. *Sustainable agriculture research*, *1*(2), 211. <https://doi.org/10.5539/sar.v1n2p211>
- Lay, F. T., & Anderson, M. A. (2005). Defensins--components of the innate immune system in plants. *Current Protein & Peptide Science*, *6*(1), 85–101. <https://doi.org/10.2174/1389203053027575>
- Lay, F. T., Mills, G. D., Poon, I. K. H., Cowieson, N. P., Kirby, N., Baxter, A. A., van der Weerden, N. L., Dogovski, C., Perugini, M. A., Anderson, M. A., Kvensakul, M., & Hulett, M. D. (2012). Dimerization of plant defensin NaD1 enhances its antifungal activity. *The Journal of Biological Chemistry*, *287*(24), 19961–19972. <https://doi.org/10.1074/jbc.m111.331009>
- Lee, H. I., Broekaert, W. F., & Raikhel, N. V. (1991). Co- and post-translational processing of the hevein preproprotein of latex of the rubber tree (*Hevea brasiliensis*). *The Journal of Biological Chemistry*, *266*(24), 15944–15948. [https://doi.org/10.1016/s0021-9258\(18\)98499-1](https://doi.org/10.1016/s0021-9258(18)98499-1)
- Li, Z., Gao, J., Wang, G., Wang, S., Chen, K., Pu, W., Wang, Y., Xia, Q., & Fan, X. (2022). Genome-Wide Identification and Characterization of GASA Gene Family in

- Meiyalaghan, S., Thomson, S. J., Fiers, M. W., Barrell, P. J., Latimer, J. M., Mohan, S., Jones, E. E., Conner, A. J., & Jacobs, J. M. E. (2014). Structure and expression of GSL1 and GSL2 genes encoding gibberellin stimulated-like proteins in diploid and highly heterozygous tetraploid potato reveals their highly conserved and essential status. *BMC Genomics*, 15(1). <https://doi.org/10.1186/1471-2164-15-2>
- Nahirňak, V., Almasia, N. I., Hopp, H. E., & Vazquez-Rovere, C. (2012). Snakin/GASA proteins: involvement in hormone crosstalk and redox homeostasis: Involvement in hormone crosstalk and redox homeostasis. *Plant Signaling & Behavior*, 7(8), 1004–1008. <https://doi.org/10.4161/psb.20813>
- Nahirňak, V., Rivarola, M., Gonzalez de Urreta, M., Paniego, N., Hopp, H. E., Almasia, N. I., & Vazquez-Rovere, C. (2016). Genome-wide Analysis of the Snakin/GASA Gene Family in *Solanum tuberosum* cv. Kennebec. *American Journal of Potato Research: An Official Publication of the Potato Association of America*, 93(2), 172–188. <https://doi.org/10.1007/s12230-016-9494-8>
- Nawrot, R., Barylski, J., Nowicki, G., Broniarczyk, J., Buchwald, W., & Goździcka-Józefiak, A. (2014). Plant antimicrobial peptides. *Folia Microbiologica*, 59(3), 181–196. <https://doi.org/10.1007/s12223-013-0280-4>
- Oard, S., & Karki, B. (2006). Mechanism of beta-purothionin antimicrobial peptide inhibition by metal ions: molecular dynamics simulation study. *Biophysical Chemistry*, 121(1), 30–43. <https://doi.org/10.1016/j.bpc.2005.12.004>
- Oliveira-Lima, M., Benko-Iseppon, A. M., Neto, J. R. C. F., Rodriguez-Decuadro, S., Kido, E. A., Crovella, S., & Pandolfi, V. (2017). Snakin: Structure, roles and applications of a plant antimicrobial peptide. *Current Protein & Peptide Science*, 18(4), 368–374. <https://doi.org/10.2174/1389203717666160619183140>
- Oliveira de Machado, L., Vieira, L. do N., Stefenon, V. M., Oliveira Pedrosa, F. de, Souza, E. M. de, Guerra, M. P., & Nodari, R. O. (2017). Phylogenomic relationship of feijoa (*Acca sellowiana* (O.Berg) Burret) with other Myrtaceae based on complete chloroplast genome sequences. *Genetica*, 145(2), 163–174. <https://doi.org/10.1007/s10709-017-9954-1>
- Organización de las Naciones Unidas para la Alimentación y la Agricultura (FAO). (2009). *La agricultura mundial en la perspectiva del año 2050*.

- Organización de las Naciones Unidas para la Alimentación y la Agricultura (FAO). (2016). *El Plan de acción de la FAO sobre la resistencia a los antimicrobianos 2016-2020*.
- Padovan, L., Scocchi, M., & Tossi, A. (2010). Structural aspects of plant antimicrobial peptides. *Current Protein & Peptide Science*, 11(3), 210–219. <https://doi.org/10.2174/138920310791112093>
- Pestana-Calsa, M. C., Ribeiro, I. L. A. C., & Calsa, T., Jr. (2010). Bioinformatics-coupled molecular approaches for unravelling potential antimicrobial peptides coding genes in Brazilian native and crop plant species. *Current Protein & Peptide Science*, 11(3), 199–209. <https://doi.org/10.2174/138920310791112138>
- Porto, W. F., & Franco, O. L. (2013). Theoretical structural insights into the snakin/GASA family. *Peptides*, 44, 163–167. <https://doi.org/10.1016/j.peptides.2013.03.014>
- Pujade-Renaud, V., Sanier, C., Cambillau, L., Pappusamy, A., Jones, H., Ruengsri, N., Tharreau, D., Chrestin, H., Montoro, P., & Narangajavana, J. (2005). Molecular characterization of new members of the *Hevea brasiliensis* hevein multigene family and analysis of their promoter region in rice. *Biochimica et Biophysica Acta - Gene Structure and Expression*, 1727(3), 151–161. <https://doi.org/10.1016/j.bbaexp.2004.12.013>
- Rodríguez Decuadro, S. (2018). Prospección de defensinas y esnaquinas de plantas nativas para el desarrollo de nuevos agentes antimicrobianos. Universidad de la República.
- Rodríguez-Decuadro, S., Barraco-Vega, M., Dans, P. D., Pandolfi, V., Benko-Iseppon, A. M., & Cecchetto, G. (2018). Antimicrobial and structural insights of a new snakin-like peptide isolated from *Peltophorum dubium* (Fabaceae). *Amino Acids*, 50(9), 1245–1259. <https://doi.org/10.1007/s00726-018-2598-3>
- Rodríguez-Decuadro, S., Dans, P. D., Borba, M. A., Benko-Iseppon, A. M., & Cecchetto, G. (2019a). Gene isolation and structural characterization of a legume tree defensin with a broad spectrum of antimicrobial activity. *Planta*, 250(5), 1757–1772. <https://doi.org/10.1007/s00425-019-03260-w>
- Rodríguez Decuadro S. et al (2019b). Genes de defensinas en los taxa tetraploides del grupo Dilatata de *Paspalum* (Gramineae). XVII Congreso Latinoamericano de Genética. Mendoza, Argentina.

- Rodríguez Decuadro S. et al (2019c). Búsqueda de péptidos antimicrobianos en transcriptomas de novo de *Peltophorum dubium* y *Maytenus ilicifolia*, dos especies vegetales nativas de Sudamérica. XVII Congreso Latinoamericano de Genética. Mendoza, Argentina.
- Rodríguez Decuadro S., Ustra, MJ et al. (2019d). Evolución de genes Snakin/GASA en alotetraploides del grupo Dilatata de *Paspalum* (Poaceae) respecto a sus progenitores diploides. XII *Simposio de Recursos Genéticos para las Américas y el Caribe*. Rocha, Uruguay. <https://sirgeac.org/>
- Rodríguez-Decuadro, S., da Rosa, G., Radío, S., Barraco-Vega, M., Benko-Iseppon, A. M., Dans, P. D., Smircich, P., & Cecchetto, G. (2021). Antimicrobial peptides in the seedling transcriptome of the tree legume *Peltophorum dubium*. *Biochimie*, 180, 229–242. <https://doi.org/10.1016/j.biochi.2020.11.005>
- Santos-Silva, C.A.d.; Ferreira-Neto, J.R.C.; Amador, V.C.;Bezerra-Neto, J.P.; Vilela, L.M.B.;Binneck, E.; Rêgo, M.d.S.;da Silva, M.D.;Mangueira de Melo, A.L.T.;da Silva, R.H.; et al. From Gene to Transcript and Peptide: A Deep Overview on Non-Specific Lipid Transfer Proteins (nsLTPs). *Antibiotics* 2023, 12, 939. <https://doi.org/10.3390/antibiotics12050939>
- Segura, A., Moreno, M., Madueño, F., Molina, A., & García-Olmedo, F. (1999). Snakin-1, a peptide from potato that is active against plant pathogens. *Molecular Plant-Microbe Interactions: MPMI*, 12(1), 16–23. <https://doi.org/10.1094/MPMI.1999.12.1.16>
- Selitrennikoff, C. P. (2001). Antifungal proteins. *Applied and Environmental Microbiology*, 67(7), 2883–2894. <https://doi.org/10.1128/AEM.67.7.2883-2894.2001>
- Shah D.M., & Read N.D. (2013) Antifungal peptides come of age. *Fungal Biology Reviews*, 26, 107-108.
- Shang, C., Ye, T., Zhou, Q., Chen, P., Li, X., Li, W., Chen, S., Hu, Z., & Zhang, W. (2023). Genome-wide identification and bioinformatics analyses of host defense peptides Snakin/GASA in mangrove plants. *Genes*, 14(4), 923. <https://doi.org/10.3390/genes14040923>
- Shelenkov, A., Slavokhotova, A., & Odintsova, T. (2020). Predicting antimicrobial and other cysteine-rich peptides in 1267 plant transcriptomes. *Antibiotics (Basel, Switzerland)*, 9(2), 60. <https://doi.org/10.3390/antibiotics9020060>

- Silva, L. R., Moura, A. P. C., Gil, B. V., Rohr, A., Almeida, S. M. Z., Donazzolo, J., Perboni, A. T., Oliveira, F. L. R., Sant'Anna-Santos, B. F., Galon, L., & Danner, M. A. (2024). Morphophysiological changes of *Acca sellowiana* (Myrtaceae: Myrtoideae) saplings under shade gradient. *Brazilian Journal of Biology*, *84*. <https://doi.org/10.1590/1519-6984.252364>
- Slavokhotova, A. A., Shelenkov, A. A., & Odintsova, T. I. (2015). Prediction of *Leymus arenarius* (L.) antimicrobial peptides based on de novo transcriptome assembly. *Plant Molecular Biology*, *89*(3), 203–214. <https://doi.org/10.1007/s11103-015-0346-6>
- Slavokhotova, A. A., Shelenkov, A. A., Andreev, Y. A., & Odintsova, T. I. (2017a). Hevein-like antimicrobial peptides of plants. *Biochemistry. Biokhimiia*, *82*(13), 1659–1674. <https://doi.org/10.1134/s0006297917130065>
- Slavokhotova, A. A., Shelenkov, A. A., Korostyleva, T. V., Rogozhin, E. A., Melnikova, N. V., Kudryavtseva, A. V., & Odintsova, T. I. (2017b). Defense peptide repertoire of *Stellaria media* predicted by high throughput next generation sequencing. *Biochimie*, *135*, 15–27. <https://doi.org/10.1016/j.biochi.2016.12.017>
- Sun, B., Zhao, X., Gao, J., Li, J., Xin, Y., Zhao, Y., Liu, Z., Feng, H., & Tan, C. (2023). Genome-wide identification and expression analysis of the GASA gene family in Chinese cabbage (*Brassica rapa* L. ssp. *pekinensis*). *BMC Genomics*, *24*(1). <https://doi.org/10.1186/s12864-023-09773-9>
- Terras, F. R., Eggermont, K., Kovaleva, V., Raikhel, N. V., Osborn, R. W., Kester, A., Rees, S. B., Torrekens, S., Van Leuven, F., & Vanderleyden, J. (1995). Small cysteine-rich antifungal proteins from radish: their role in host defense. *The Plant Cell*, *7*(5), 573–588. <https://doi.org/10.1105/tpc.7.5.573>
- Thi Phan, A. D., Chaliha, M., Sultanbawa, Y., & Netzel, M. E. (2019). Nutritional Characteristics and Antimicrobial Activity of Australian Grown Feijoa (*Acca sellowiana*). *Foods (Basel, Switzerland)*, *8*(9), 376. <https://doi.org/10.3390/foods8090376>
- Thomma, B. P. H. J., Cammue, B. P. A., & Thevissen, K. (2002). Plant defensins. *Planta*, *216*(2), 193–202. <https://doi.org/10.1007/s00425-002-0902-6>
- Thorp G y Bielecki R. (2002). In: Thorp G, Bielecki R (Eds.). *Feijoas: Origins, cultivation and uses*. David Bateman Ltd, Auckland
- Trabi, M., & Craik, D. J. (2004). Tissue-specific expression of head-to-tail cyclized miniproteins in Violaceae and structure determination of the root cyclotide *Viola*

- hederacea* root cyclotide1[W]. *The Plant Cell*, 16(8), 2204–2216.
<https://doi.org/10.1105/tpc.104.021790>
- van der Weerden, N. L., & Anderson, M. A. (2013). Plant defensins: Common fold, multiple functions. *Fungal Biology Reviews*, 26(4), 121–131.
<https://doi.org/10.1016/j.fbr.2012.08.004>
- Van Parijs, J., Broekaert, W. F., Goldstein, I. J., & Peumans, W. J. (1991). Hevein: an antifungal protein from rubber-tree (*Hevea brasiliensis*) latex. *Planta*, 183(2), 258–264. <https://doi.org/10.1007/bf00197797>
- Van Damme, E. J. M., Charels, D., Roy, S., Tierens, K., Barre, A., Martins, J. C., Rougé, P., Van Leuven, F., Does, M., & Peumans, W. J. (1999). A gene encoding a hevein-like protein from elderberry fruits is homologous to PR-4 and class V chitinase Genes1. *Plant Physiology*, 119(4), 1547–1556.
<https://doi.org/10.1104/pp.119.4.1547>
- Vignale B & Bisio L. (2005). Selección de frutos nativos en Uruguay. *Agrociencia*, 9(1), 35-39
- Xie, J., Li, Y., Liu, X., Zhao, Y., Li, B., Ingvarsson, P. K., & Zhang, D. (2019). Evolutionary origins of pseudogenes and their association with regulatory sequences in plants. *The Plant Cell*, 31(3), 563–578.
<https://doi.org/10.1105/tpc.18.00601>
- Yeaman, M. R., & Yount, N. Y. (2003). Mechanisms of antimicrobial peptide action and resistance. *Pharmacological Reviews*, 55(1), 27–55.
<https://doi.org/10.1124/pr.55.1.2>
- Yeung, H., Squire, C. J., Yosaatmadja, Y., Panjikar, S., López, G., Molina, A., Baker, E. N., Harris, P. W. R., & Brimble, M. A. (2016). Radiation damage and racemic protein crystallography reveal the unique structure of the GASA/snakin protein superfamily. *Angewandte Chemie (International Ed. in English)*, 55(28), 7930–7933.
<https://doi.org/10.1002/anie.201602719>
- Zasloff, M. (2002). Antimicrobial peptides of multicellular organisms. *Nature*, 415(6870), 389–395. <https://doi.org/10.1038/415389a>
- Zhu, F. (2018). Chemical and biological properties of feijoa (*Acca sellowiana*). *Trends in Food Science & Technology*, 81, 121–131.
<https://doi.org/10.1016/j.tifs.2018.09.008>

Zou, C., Lehti-Shiu, M. D., Thibaud-Nissen, F., Prakash, T., Buell, C. R., & Shiu, S.-H. (2009). Evolutionary and expression signatures of Pseudogenes in Arabidopsis and rice. *Plant Physiology*, 151(1), 3–15. <https://doi.org/10.1104/pp.109.140632>

MATERIAL SUPLEMENTARIO

Anexo 1: Secuencias problema (*query*) utilizadas para la búsqueda de AMPs mediante BLAST.

Tabla S1: Secuencias *query* utilizados para la búsqueda mediante BLAST por familia de AMPs. El número de accesoión corresponde a las bases de datos Uniprot, Phytozome y NCBI, dependiendo de la proteína.

AMP	Especie	Número de accesoión	Secuencia <i>query</i>
Esnaquinas	<i>Solanum tuberosum</i>	Q948Z4 · SNAK1_SOLTU	MKLFLLTLLLVTLVITPSLIQTTMAGSSFCDKSK KLRCCKAGLADRCLKYCGICCEECKCVPSGTY GNKHECPCYRDKKNSKSKGSKCP
		Q93X17 · SNAK2_SOLTU	MAISKALFASLLLSLLLLLEQVQSIQTDQVTSNAI SEAAYSYKKIDCGGACAARCRLSSRPRLCNRA CGTCCARCNCVPPGTSGNTETPCPYASLTH GNKRKCP
		M1BA38 · M1BA38_SOLTU	MAKSGYNASFLLLLSMFLILLTFSNVVEGYNKL RPRDCKPKCTYRCSATSHKKPCMFFCQKCCA TCLCVPKGVYGNKQSCPCYNNWKTQEGKPK CP
Heveína	<i>Hevea brasiliensis</i>	P02877 HEVE_HEVB	EQCGRQAGGKLCPNLCCSQWGWGCGSTDE YCSPDHNCQSCKD
	<i>Amaranthus caudatus</i>	P27275 AMP_AMACA	MVNMKCVLIVIVMMAFMMVDPSMGVGEVCR GRCPGSMCCSQFGYCGKGPKYCGRASTTV HQADVAATKTAKNPTDAKLAGAGSP
	<i>Stellaria media</i>	E1UYT9 AMP1_STEME	MLNMKSFALVMLFATLVGVTIASGPNQCGPG WGGCRGGLCCSQYGYCGSGPKYCAHNTPLS EIEPTDAGRCSSRGTCSSGRCSSKYGYCGTG PAYCGLGMCQGSCLPDMNHPAQIARTEAA QAEAQAEAYNQANEAQVEAYYQATQAQTQA QPQVEPAVTKAP
	<i>Triticum kiharae</i>	P85966 AMP_TRIKH	AQRCDGQARGAKCPNCLCCGKYGFCGSGDA YCGAGSCQSQCRCR
LTP	<i>Arabidopsis thaliana</i>	Q9SU35 ERLL1_ARATH	MASKNSASLALFFALNILFFTLTVATNCNCKPS PKPKPVSPKPKPVQCPPPPRPSVPSNPRP VTPRTPGSSGNSCPIDALKLGVCANVLSLL NIQLGQPSSQCCSLIQGLVDVDAIICLCTALR ANVLGINLNPISLSVLLNVCNRKLPSTGQCA
		Q42589 NLTP1_ARATH	MAGVMKLACLLLACMIVAGPITSNAALSCGSV NSNLAACIGYVLQGGVIPPACCSGVKNLNSIAK TTPDRQQACNCIQGAARALGSLNAGRAAGIP KACGVNIPYKISTSTNCKTVR

AMP	Especie	Número de accesión	Secuencia query
	<i>Lens culinaris</i>	A0AT29 NLTP2_LENCU	MARGMKLACVVLVICMVIAPMAEGAISCGAV TSDLSPCLTYLTGGPGSPQCCGGVKKLLAAA NTTPDRQAACNCLKSAAGSITKLNNTNAAAALP GKCGVNIPYKISTTTNCNTVKF
	<i>Vigna radiata</i> <i>var. radiata</i>	A0A1S3TFR4 _VIGRR	MEGVVKFACLVGFFVVLVSVAKVDSAGECGKS TTPDNEAIKLAPCASAAQDENASVSQSCCAQV KKIGQNPSCLCAVLLSNTAKMAGVNPQIAVTIP KRCNLANRPVGYKCGPYTLP
	<i>Prunus armeniaca</i>	P82353 NLTP2_PRUAR	VTCSVPQLSPCLGPINS GAPSPPTCCQKLREQ RPCLCGYLKNPSLRQYVNSPNARKLASNCGV PVPQC
Tionina	<i>Hordeum vulgare</i>	P01545 · THNA_HORVU	MVCLLILGLVLEQVQVEGKSCCRSTLGRNCYN LCRVRGAQKLCAGVCRCKLTSSGKCTGFPK LALVSNSEPDPTVKYCNLGCRAEMCDYMVNA AADDEEMKLYLENCGDACVNFNGDAGLTSL TA
		P21742 · THNB_HORVU	MGSKGLKGMVCLLILGLVLEHVQVEGKSCCR STLGRNCYNLCRVRGAQKLCANACRCKLTSG LKCPSSFPKLALVSNSEPDPTIDYCNLGCRA MCDYMVNAAADDEEMKLYVEHCSDACVNF NGDVGLTSLTA
	<i>Triticum aestivum</i>	Q05806 · THN5_WHEAT	MGGGQKGLSAIVCLLVGLVLEQVQVEGVDC GANPFKVACFNSSLGPSTVFQCADFCACRLP AGLASVRSSDEPNIAIEYCSLGCRRSSVCDNMIN TADNSTEEMKLYVKRCGVACDSFCKGDTLLAS LDD
	<i>Eucalyptus grandis</i>	A0A059AJ89_EUCGR	MDKVKSVLVCLVGLFLGQSRADFQDCYVG CFVACVVTGDNVVKCSLKCLKDCIGLPSHGLT DAEYFCKLGCASSLCINLSSKDDPGEKRVVHC VNNSCKTCANHA
A0A059AI82_EUCGR		MGKVKSVFMVCLVGLFLGQSGANFQDCYPA CFILCAITPGRTLFSCSVECLKDCIIPSDSLSLR DTMYFCKLGCASSLCTNLSTKDNPREKRVAG CVNSCKTCTDHV	
Defensinas	<i>Eucalyptus grandis</i>	Eucgr.B02620.1.p	MDFSKRLIPAALIVMLLLVATEMEPMVVEARTC ESQSQRFGACVSKTNCASVCQTEGFHGGH CRGFRRRCFCTKHC
		Eucgr.C02231.1.p	MIEMPKVEAYCSEGIGLCGKNECEQRCYAS HGP GSKGSCDYNIKPLCTCYDYDCPTPPKPKV CIGGLGLCSNACSNQCCSNRCAAKFSQGRGY CDNSAGPWLCQCQYPC
		Eucgr.H04400.1.p X	MAKYFTLCFLLLLILSCDEKRVPLAEAKDCHKV WTCKGGDKCWQDCKNQFNGRGMCDLYTAP PVPKQCFCAYKC
		Eucgr.H05052.1.p	MEKSKMRCMGLFMMVLLVLAQEAEGRVCES QSHGFHGSVCVSGHNCALVCRNEGFSGGRCR GFRRRCFCTKLC

AMP	Especie	Número de accesión	Secuencia <i>query</i>
		Eucgr.C02234.1.p	MARLLQAYCLVLLFVLTAGLMVKTTAAKVCTE GLGTCAGGNDCEQSCFERHGPSQGTCDRTI SPPLCSCIFMCPLSA
		Eucgr.K01859.1.p	MQQSKSRPSIALFSEALSSSPASARTKSQRRK AVSGSAMATKLPSSLFLLLVLVLLSVAWMPPKA SAQPRTCSELDSTGCRPQDCVARCAAQYRGI GECAHQEYVSNFHCVCLYACL
		Eucgr.C02268.1.p	MARLLQAYCLVLLFVLTAGLMVKTTAAKDCTE GLGTCAGGNDCEQSCFERHGPSQGTCDRTI SPPLCSCIFMCPLST
	<i>Acca sellowiana</i>	TRINITY_DN119808_c0_g1 _i1 (este estudio)	MKRISICSVLLLVLVLLIAGSGSMMAQVSGLCSSK KLQLPAHRCIKDECISYCKTTYGDGATGVCES DTQPEDICLCQYPC
	<i>Nicotiana alata</i>	Q8GTM0 · DEF_NICAL	MARSLCFMAFAILAMMLFVAYEVQARECKTES NTFPGICITKPPCRKACISEKFTDGHCSKILRR LCTKPCVFDEKMTKTGAELAAEAKTLAAALLE EEIMDN
	<i>Petunia hybrida</i>	Q8H6Q1 · DEF1_PETHY	MARSICFFAVAILALMLFAAYDAEAATCKAACP TWDSVCINKKPCVACCKKAKFSDGHCSKILRR CLCTKECVFEKTEATQTETFTKDVNTLAEALLE ADMMV
Q8H6Q0 · DEF2_PETHY		MARSICFFAVAILALMLFAAYETEAGTCKAACP TWEGICINKAPCVKCCKAQPEKFTDGHCSKIL RRCLCTKPCATEEATATLANEVKTMAEALVEE DMME	

Anexo 2: Secuencias identificadas en transcriptoma de *A. sellwoiana*

Tabla S2: defensinas identificadas en el transcriptoma de *A. sellwoiana*.

Nombre	Secuencia*
AsDf1	MKRISICSVLLLVLLLIAGSGSMMAQV SGLCSKKLQLPAHRCIKDECISYCKTTYGDGATGVCE SDTQPEDICLCQYPC
AsDf2a	MDFSKRLIPAAFVVM LLLVATEMEPMAVEARTCESQSHRFKGACVRDSNCAAVCQTEGFQG GHCRCGFRRRRCFCTKHC
AsDf2b	MDFSKRLIPAAFVVM LLLVATGMEPMAVEARTCESQSHRFKGACVRDSNCAAVCQTEGFQG GHCRCGFRRRRCFCTKHC
AsDf3	MDFSKRLIPAAFVVM LLLVATEMEPMAVEARTCESQSHRFKGACVRDSNCASVCQTESFQG GHCRCGFRRRRCFCTKHC
AsDf4	MKTHSTSLILLFVMLISFGNELGMKQGAE AQCQETLYGGGCQDKPCNDACVQKFGPKAYGF CFYKPSDTCVCRHPC
AsDf5	MESLAFPRILVLVLLS GTCMVMVPQANAQKRCVEVLDPHSCNLYNCQKACYEKHTAAQQP RGECQHDA SFNYSCACFFNC
AsDf6	MEKSKMRCCMGLFLM LLLVLAEEAEGRMCSQSHGFHGSCVRSHNCALVCRNEGFSGGR CRGFRRRRCFCTKLC
AsDf7	MEKSKMRCCMGLFLM LLLVLAQAEGRMCSQSHGFHGSCVRSHNCALVCRNEGFSGG RCRGFRRRRCFCTKLC
AsDf8	MIEIPKVEAYCSEGIGLCGKGNCEQRCYAIRGLGSQGSQDYTIKPLPCTCYDCEPELLPEGYC GSAGSCPEPRLKECNGGGGLCTIACTDQCCSDRCA
AsDf9a	MEPMAVEARTCESQSHRFKGACVRDSNCASVCQTESFQGGHCRGFRRRRCFCTKHC
AsDf9b	MEPMAVEARTCESQSHRFKGACVRDSNCASVCQTESFQGGHCRGFRRRRCFCTKHC
AsDf10	MLLLFVATAEMEPMVAVEGGRMGKLSHEYRGICIDRQCAAVCVTEGFWSGYCGGFRRHCI CSIPCHPPRPEFHLLHYILSCLISIMK
AsDf11	MASGRVCSFIGLLCLISL CFAEKSVAREINEETVRYVDIGPCSKFPDCNKACKEMLGLNVGRF CKDNGAGSERCYCVT
AsDf12	LSISDSFVAAWMLPKASAPRACSVELDTSGCSPQDCVARCASQYSGIGECAHQEYVSNFHC VCLYACL
AsDf13	MATIKLPLSSLSLLLLL VLLLSAAWMLPKASAPRACSVELDTSGCSPQDCVARCASQYSGI GECALHQEYVSNFHCVCLYACL
AsDf14	MSKTYQICVLLVVL CITMVSVRTQGTGCLQLSQDFCDVITCDQNCYEFICIAKGCAPQRCRSEG KCSCR
AsDf15	DDCEQRCSAKHGPSQGTCDAAVSPPLCSCLFSCAPSA
AsDf16	MAEVFRKCCLLIVLL VFAGWMIEIPKVEAFCSAGLGLCGKGNDCQQRCSASHGPSQGSQD YTIN
AsDf17	MAKYFTLCFLLLLLL LVFPWDEKRVPLAEAKDCYKVVWTKGGDRCWEDCRNRFNGRG
AsDf18	MELMVAVAGGRKHESLSRKFQIGICFSKRNCTAICKPEGFESGRRCRGFRRRRFCTRPCPPSP
AsDf19a	MVAVAGGRKHESLSRKFQIGICFSKRNCTAICKPEGFESGRRCRGFRRRRFCTRPCPPSP
AsDf19b	MVAVAGGRKHESLSRKFQIGICFSKRNCTAICKPEGFESGRRCRGFRRRRFCTRPCPPSP
AsDf20	MAMVDGKLCQRRSKTWSGFCGNSGNCGRQCRNWEGARNGACHVSSQDS

Tabla S3: esnaquinas identificadas en el transcriptoma de *A. sellwoiana*

Nombre	Secuencia*
AsSn1	MMKKFALATFLLVALLSSFFDAATADSSLCDSKCAVRCSKAGRMDRCLQYCGICCEKCQC VPSGTYGHKDECPCYRDMKNSKGPKCP
AsSn2	MRFVSVFVCFEGLCDSKCAVRCSKAGRMDRCLQYCGICCEKCQC VPSGTYGHKDECPCYRDMKNSKGPKCP
AsSn3	MKMVQATVFLLAACLIVSSSLDSTMAAGSSFCGSKCSARCSKAGLKDRCIKYCRICCDKCKCV PSGTYGNKHECPCYRDLKNSKGPKCP
AsSn4	MKFSRGFCGSKCSARCSKAGLKDRCIKYCRICCDKCKCVPSGTYGNKHECPCYRDLKNSKGK PKCP
AsSn5	TAFCESKRARCAKAKVKARCAKYCNICCAECKCVPSGTYGNKHQCPCYRDKKNSKGPKCP
AsSn6	MKPLFALLLCLLLLTSIFMEPVLA KPAFCESKRARCAKAKVKARCAKYCNICCAECKCVPSG TYGNKHQCPCYRDKKNSKGPKCP
AsSn7	MKLAFATLLLVSLLVTPSLHAMAGGLSPQPAPSSCDAKCGGRCQNAGVKDRCLKYCGICCCQ KCQCVPSTYGNKSECPCYRDMKNSKGTAACP
AsSn8	MRLLFRPAGSCDAKCGGRCQNAGVKDRCLKYCGICCCQKCQCVPSTYGNKSECPCYRDMK NSKGTAACP
AsSn9	MKLAFATLLLASLVLASSLHTMAGGSPQPAPSWCDTKCNVRCQRSGWKDRCLKYCGICCCQ KCHCVPSGPYVNKSECPCYRDMKPNKGGKDKCP
AsSn10	MHIHLSTLLCTRQSGFQIHFFCFPTGWCDTKCNVRCQRSGWKDRCLKYCGICCCQKCHCV SGPYVNKSECPCYRDMKPNKGGKDKCP
AsSn11	AGVPSIFSRVSRPRVLAALIVVIRRPCVVEYCGNKCEGRCAKAGVMDRCIKYCNICCEECR CVPSGTYGNKHECPCYRDKRSSKGPKCP
AsSn12	MKLLLAPALLCSLLLSFALLGPSSTMSQYCGNKCEGRCAKAGVMDRCIKYCNICCEECRCV PSGTYGNKHECPCYRDKRSSKGPKCP
AsSn13	MSLSKVLACLILLVNLVQLVEAAEATELSSVESVSAKKIDCGAACAAARCLASRQKICKRA CGTCCARCNCVPPGTAGNRDVCPCYASMTTHGGRLKCP
AsSn14a	SSADCESRCKYRCSKSSRHKMCIRACNTCCQRCNCVPPGTSGNEDVCPCYANMTTHGGRHK CP
AsSn14b	SSADCESRCKYRCSKSSRHKMCIRACNTCCQRCNCVPPGTSGNEDVCPCYANMTTHGGRHK CP
AsSn15	MASAKTSILFAILCLVLVHELLVFGGEQFQAEAKKIDCESRCKYRCSKSSRHKMCIRACNTCC QRCNCVPPGTSGNEDVCPCYANMTTHGGRHKCP
AsSn16	GLHLRCVDCGGACAARCQLSSRPRLCQRACGTCCARCNCVPPGTSGNLDVCPCYANMTTRG NKRKCP
AsSn17	MAVHKLLIASILVSVLLLQLAEADQMGGVGNAGDSSKLPKMDCGGACAARCQLSSRPRLCQR ACGTCCARCNCVPPGTSGNLDVCPCYANMTTRGNKRKCP
AsSn18	MDCGGACAARCQLSSRPRLCQRACGTCCARCNCVPPGTSGNLDVCPCYANMTTRGNKRK CP
AsSn19	ITNSCPSYEPFSDCGAACAAARCQLASRQLCHRACGTCCNRCNCVPPGTSGNRDVCACYAT MTTHGGRLKCP
AsSn20a	MRGFLTDKLPPLQCADCGLCKERCSSLHSRPNVCARACGTCCFRCKCVPPGTSGNRELCGK CYTDMTTHGNRTKCP
AsSn20b	MRGFLTDKLPPLQCADCGLCKERCSSLHSRPNVCARACGTCCFRCKCVPPGTSGNRELCGK CYTDMTTHGNRTKCP
AsSn21	QLHFSLTKNLQVSSDIDRENSHTQIVKGANRRLQLTDCGGGLCKERCSSLHSRPNVCARACGTC

	CFRCKCVPPGTSGNRELCGKCYTDMTTHGNRTKCP
AsSn22	LRLSVTSVLLYFSQIVKGANRRLLQTLDCGGLCKERCSLHSRPNVCARACGTCCFRCKCVPPGTSGNRELCGKCYTDMTTHGNRTKCP
AsSn23a	MRGFLTDKLPPLQCADCGLCKERCSLHSRPNVCARACGTCCFRCKCVPPGTSGNRELCGKCYTDMTTHGNRTKCP
AsSn23b	MRGFLTDKLPPLQCADCGLCKERCSLHSRPNVCARACGTCCFRCKCVPPGTSGNRELCGKCYTDMTTHGNRTKCP
AsSn23c	MRGFLTDKLPPLQCADCGLCKERCSLHSRPNVCARACGTCCFRCKCVPPGTSGNRELCGKCYTDMTTHGNRTKCP
AsSn24	MALRLLLLVAFLLVCTIQVSSDIDRENSHTQIVKGANRRLLQTLDCGGLCKERCSLHSRPNVCARACGTCCFRCKCVPPGTSGNRELCGKCYTDMTTHGNRTKCP
AsSn25	MAFKAMFLVATFLLINTRVSSSDEELKMKAPSPPYAQVPVAAPAPAKLPKPPTSPPSKATPPAKPPTPTLPPVKSADCIPLCDQRCKLHSRKRVCMRACMTCCDRCKCVPPGTGYNREMGKCYTDMTTHGNRVKCP
AsSn26a	ESDYTELGEPLYVAADCIPLCDQRCKLHSRKRVCMRACMTCCDRCKCVPPGTGYNRERCGKCYTNMTTHGNRIKCP
AsSn26b	ESDYTELGEPLYVAADCIPLCDQRCKLHSRKRVCMRACMTCCDRCKCVPPGTGYNRERCGKCYTNMTTHGNRIKCP
AsSn27	MASKTVLLLLGLLLLVTTQVSSDDNEEQILIKAPYAKVPTSPVKAPAPAKVPVKALPTLAPVTKTPPPLPPVTKLPPPLPPVAKPPPTLPPFTKPLPPPLPPVAKPPLPPVAKPPPTLPPFTKPLPPPLPPVTKPPLPPVANPPPTLPPFTKPLPPPLPPVAKPPLPPVAKPPPTLPPFTKPLPPPLPPVTKPPPTLPPYTKLPPPLPPVMKPPPTLPPVMPPAAPLPPVRSKADCIPLCDQRCKLHSRKRVCMRACMTCCDRCKCVPPGTGYNRERCGKCYTNMTTHGNRIKCP
AsSn28	MKPPPTLPPVMPPAAPLPPVRSKADCIPLCDQRCKLHSRKRVCMRACMTCCDRCKCVPPGTGYNRERCGKCYTNMTTHGNRIKCP
AsSn29	MASKTVLLLLGLLLLVTTQVSSDDNEEQILIKAPYAKVPTSPVKAPAPAKVPVKALPTLAPVTKTPPPLPPVTKLPPPLPPVAKPPPTLPPFTKPLPPPLPPVTKPPLPPVAKPPPTLPPFTKPLPPPLPPVTKPPPTLPPYTKLPPPLPPVMKPPPTLPPVMPPAAPLPPVRSKADCIPLCDQRCKLHSRKRVCMRACMTCCDRCKCVPPGTGYNRERCGKCYTNMTTHGNRIKCP
AsSn30a	FTRYLECGPRCKERCSKTAYKKPCMFFCQKCCAKCLCVPPGTGYNKQLCPCYNNWTKRGGPKCP
AsSn30b	FTRYLECGPRCKERCSKTAYKKPCMFFCQKCCAKCLCVPPGTGYNKQLCPCYNNWTKRGGPKCP
AsSn30c	FTRYLECGPRCKERCSKTAYKKPCMFFCQKCCAKCLCVPPGTGYNKQLCPCYNNWTKRGGPKCP
AsSn31	MFPGDDHRGPITAPGNSRQSARGNVFLECQLSDLYIFASVDFFFLFFFSQYGAREGSLKPQECGPRCKERCSKTAYKKPCMFFCQKCCAKCLCVPPGTGYNKQLCPCYNNWTKRGGPKCP
AsSn32	MARKVSFVILSLFTLLLLIPENQATITEAPSPQPQATPGNLPAYGAREGSLKPQECGPRCKERCSKTAYKKPCMFFCQKCCAKCLCVPPGTGYNKQLCPCYNNWTKRGGPKCP
AsSn33	LVRKLLCFQATITEAPSPQPQATPGNLPAYGAREGSLKPQECGPRCKERCSKTAYKKPCMFFCQKCCAKCLCVPPGTGYNKQLCPCYNNWTKRGGPKCP
AsSn34	TSTGLVGICLQHGTQGSQSLKQDCGPRCTSRCSQTSYKKPCMFFCQKCCAKCLCVAGTYGNKQTCPCYNNWTKRGGPKCP
AsSn35	MAKAQRILVLCIALLLVLAQENEVFGSVELGFDVTVVEYDVRASQGSQMHGTTQGSQSLKQDCGPRCTSRCSQTSYKKPCMFFCQKCCAKCLCVAGTYGNKQTCPCYNNWTKRGGPKCP
AsSn36	GLVFLQLFFADCGPRCTSRCSQTSYKKPCMFFCQKCCAKCLCVAGTYGNKQTCPCYNNWTKRGGPKCP
AsSn37	MAMTKLACVLLLALLGISMVATQVMAKEAQYHLDAYSQGSLSYQCSGQCSRRCKQTQYKPCCLFFCNKCCAKCLCVPPGFYGNKQVCPCYNNWTKRGGPKCP

AsSn38	GNSNQKAFDENLTECSGQCSRRCKQTQYKKPCLFFCNKCCAACLCVPPGFYGNKQVCPCYN NWKTRGGPKCP
AsSn39	LVEQQVHGKLRPSDCPKKCKFRCSATSHKKPCMFFCQKCCAACLCVPPGTGYNKQVCPCYN SWKTKEGGPKCP
AsSn40	MARRPYTSFLVLSLLLLIALSNVAEVHGKLRPSDCPKKCKFRCSATSHKKPCMFFCQKCCAAC LCVPPGTGYNKQVCPCYNSWKTKEGGPKCP
AsSn41	AFNAKTLTYDVVDCPKKCKFRCSATSHKKPCMFFCQKCCAACLCVPPGTGYNKQVCPCYNS WKTKEGGPKCP
AsSn42	MASFLLRSSLFFLVASMCFVEVSLAGGEGSLRPDQCAAACDYRCSETSHRKPCLFFCNK CAKCLCVPSGTGKHKEEPCYNNWKTQEGKPKCP

Tabla S4: tioninas identificadas en el transcriptoma de *A. sellwoiana*

Nombre	Secuencia*
AsTn1	MDKVKS FVVCLVLGLFVGQPSAG GIKPCYKRCYFHCVIDPRRSLPYCHDKCMKHCLPPSSD NLKDRYFCKLGCAYSSTTFSSKDDPSVRGMP
AsTn2	MDKVKS FVVCLVLGLFVGQSRAD FEDCYQICFILCAITPGNSLFSCSVKCLKDCIVPPSDSLGD TQYFCKLGCAASSCTNLSSKDNPGNKSLSLSLSLSLYVCVCVCGSSYFCSCLCRSGEKRVA GCVDSCESETCANHV
AsTn3	MDKVKS FVVCLVLGLFVGQSRAD FEDCYQICFILCAITPGNSLFSCSVKCLKDCIVPPSDSLGD TQYFCKLGCAASSCTNLSSKDNPGNKSLSLSLSLSLYVCVCVCGSSYFCSCLCRSGEKRVA GCVDSCESETCANHV
AsTn4	MDKVKS FVVCLVLGLFVGQSRAD FEDCYQICFILCAITPGNSLFSCSVKCLKDCIVPPSDSLGD TQYFCKLGCAASSCTNLSSKDNPGNKGKRVAGCVDSCESETCANH
AsTn5	MEIKPFDKIMTLFLMMLLGVSS QESQSTDCISHCVLSCANGHVALPLSCFYDCLKECQDNSP NPAPINQCDLDCVTETCLDDTDARKARVCADECSDICKWPSKPH
AsTn6	MDVKSCAVLVLMLVATIALQASA TEVKSFKSDLCKVQCLVDCAAVPNKRACYLTCYRNCVGA SEVETFESQCNLDCIVSKCFDLPVKKVEGCVSSCSKNCKRS
AsTn7	MEVEASKKMVMMLMVGLIISRSFVEITA TDCDKLCRDFCAGSFTPSFCYLICMRGCVKRDALVP QELGHFQLGCVASTCIDHRSDVKKVEADKEAEACVDSCSENCQKSYGPH

Tabla S5: *hevein-like* identificadas en el transcriptoma de *A. sellwoiana*

Nombre	Secuencia*
AsHv1	MRISTLAFLSLVSGALAGNC CPDNGNKKCAATECCSQYGWCGTTKDHCDLPTCIKAFSGASSP CEPSGSTPTTLKTSVVKQATSPATTFPSTVPNIDVCGSAQGGVTCPGVGADGYFYRCCSSAG HCGPKNDIQDQALYCGDGCQAGFGKCDSMAGPPVPTQTPGTAGDGETCGPIVNRKCGSLCC SGSNFCGSGVDFCGAANWCQSKWGRCD
AsHv2	ISKSYITSVGRRRSSSTMRTTILGVSALIGLATANIS FTWVQPTCDITSDFDPCLTGQHCTEENTC AADIHSRDKV/SARSRSSIRFVPEKRQAQYSLDGKCGPANGNLLCDPNSVNYQGTCCSEYGW CGNTPTHCGTGCQSGCNNGAAPTQAQ
AsHv3	MHFSNALAVLVASPLVA AHSGIPGAPKIFGLPRDSKAGAPVVPVIRHPTEPKGGLETRQGGQ NGRCGPNFGGASCAAGYCCSGAGYCGTTKDHCAAPDCLLDYGPDCDANTTPSGASTRNDPR PQKGNVAYGGEGVYTCQTPGTVAITYDDGPYSFTNGVLDQFASYGFKATFFVTGINIGKAIDTT ASWSNVIKRMVAEGHQVASHWHDLSAITDAQRYDQMVKNEMAIRNIIGKYPTYMRPPYSS CNDACQNMKNLGYVVSYFDLTDYDYNQLSNIQVAQNNFKTILDQTEGGSASGDRLAIADHIE VSANQLTAYMLSULKERYGVTMGECMGEPESNWYRTSGEGGGSTTPPTTGGSVSTDGT CGVQGGKTCTNSSFGNCCSAFGWCGSSTDHCGTGCNSSFGTCSNGGGSTTPPAPSKPVST DGTCTGQGGATCAGSTFGDCCSQYGWCGSSTGHC GSGCDSSSGTCS
AsHv4	MHIPLIHHYKYPSSLSSSKLQSNIPPFKMMTTLKLLKLLTAAALTAIVVTALPHTASAQNCGCSAD QCCSQWGYCGTGDDYCGTGCQAGPCNAPPTNGVSVADVVEAFFDGIIEGAAESCEGKGFY TREKFLEAVGYPQFGKVGSVDDSKREIAAFFAHVTHEIHFYIEEIDGSPSKDYCDEKNTQYPC NPNKGYGRGPIQLSWNFNYGPAGESIGFDGLNSPETVATDPLIAFKTALWYWMNFVRPVINQ

	GFGATIRVINGALECDGGNPATVQARVQYYTEYCNQFGVAPGDNLTC
AsHv5	MAKSVKIARALLNVLVVSGLIAGFLPAHVFGQNCGCAADLCCSQYGYCGTGDQYCGTGCCQSG PCSSSSTPTTPSSGTSVADVVTDSFFNSIIGQAAASCAGKSFYTRQAFLDAAGAYPEFGTVGSA DDSKREIAAFFAHVTHETGRKNLNMFKVL
AsHv6	MAKSVKIARALLNVLVVSGLIAGFLPAHVFGQNCGCAADLCCSQYGYCGTGDQYCGTGCCQSG PCSSSSTPTTPSSGTSVADVVTDSFFNSIIGQAAASCAGKSFYTRQAFLDAAGAYPEFGTVGSA DDSKREIAAFFAHVTHETGHFCYIEEIDGASKDYCDETNTQYPCPNKGYGRGPIQISWNYNY GPAGQSIGFDGLGSPETVANDAIVSFKTAFWFWMNNVHSIITSGQGFGATIRAINGGECGGGNS GAVQARVQYYTDYCNQLGVSPGDNLTC
AsHv7	MDRNALTDILVACLAVLALVPHVTAQNCGCAAGLCCSRYGCGTSHDYCGPGCQAGPCDPA AAPPATNAVVGIVTDAFFNGIINQAAASCAGKRFYSRKAFLDAIGSFPRFGRVGSVDNSKREI AAFFAHVTHETGHFCYIEEIDGRDTPKTYCDPNVPQYPCPKPKKYFGRGPLQISWNYNYGPA GESLGFDDLAPETVANNPVAFKALWFWRNDLQSKISSQGFATIRAINGMECNGGNPGA VQARVTYNNYCRQFGVAPGGNLSC
AsHv8	HKEGALEKRQGDAQCGAGIGSCAAGQCCSQAGYCGTDPDYCYSPGCNYQYGPGPCDPAVPS GTNTSTIARTQLGSVAVGGEGIYNCQQVGTIALTYDDGPAKTLTAHILDMKSYNAKATFFITGNN INKGQIDITQEYIDVIKRMDESEGHQIASHTWTHLDLSAISQLDRKNQMWKNEMALRNIVGKIPTYM RPPYSSCTGQCETDMAALGYHVIYFDVTDQDDPTLIQHSKDWFKGNITAGGATAAKNTW LDIQHDIHVQTANLLEFMLQTITSLG
AsHv9	TTDHCCTGCGTGFQTCGSAVSSAAASPSAVKTSAAAASASAANKVSTDGSCAGTSGFTCS GSSFGNCCSQYGWCGTTTGHCGTGCQTSFGTCT
AsHv10	QISTSGKCGVANGGLTCTGSSFGECSSQFQFCGRSSAHCNTGCQPAFGKCKNPVGSIGGLQI SPDGSCGGGTGYTCQGSQFQGNCCS
AsHv11	PECGEGIGSPAGKCCSRAGYCGTDDAFCYSPGCKYQYGPGPCENPPSGTNTSSVTRDKLG SIAYGGDGIFRCVTPGTVALTFDDGPQTNFTDRILDIFKSYDAKGTFFVTGNINIKGQIDLKHAAT LKRIDAEGHQVASHWTHLDLSQLSLDRKNQMMNEMAIRNVLGKIPTYMRPPYSSCTKDSG CQQDMADLGYHITNFVNDTDDYQITAENIQKSKDWFKGNITKDGATSQKGDKWLISISHDIEQT AN
AsHv12	VKSSAGVSATSSANKVSTDGTCAGTGKQTCQGSTFGNCCSQYNYCGSTGDHCGTGCQSGFG TCGSSSASKAAAVTTTKATPTNATTS

Tabla S6: LTPs identificadas en el transcriptoma de *A. sellwoiana*

Nombre	Secuencia*
AsLTP1	MACSKSTASLALFFTLNLLFFALVSSCGTSPRPRPRPRPRPRPTPTPTPTVPTPTPTVPT TPSTGACPRDALKLGICADVLGSLNVTIGTPPVTPCCSLLAGLVDEAAVCLCTAIRANILGINLNI PLSLTLLLNVCSTVPPGFQCA
AsLTP2	MASRTLATTALLSLNLLFFTMVSSTYCPPPPKHHPKHPKPTPTPTPAKATCPKDTLKLGVCA DVLKSLHLVVGTPPKTPCCSLIGGLADLEAAVCLCTAIKANVLGIHLNVPVLSLLLNVCYCKKVP GFQCA
AsLTP3	MASKTYAALALVLSLNVFASLVAAGPPKLISNSGASCPRDALKLGICANLLGSLNITIGTPPKPEPC CSLIAGLADLEAAVCLCTAIKANVLGINLVPLSLSLLLNVCTKKVPPGFQCA
AsLTP4	MASKTVATGAILLVNLLFFTMVSSTHVPCLPPPKTLKHPSPSSPAAAACPKDTLKLGVCANLLND LLHLVVGTPPKTPCCSLIEGLADLEAAVCLCTAIKANILGINLVPISSLNLCYCKGKVPKGFQCA
AsLTP5	MARSTPILISLLLLAAVSNACNTCMLPSPPPPTSCPTPATACPRDTLKLGVCDLLGGVGVVV GTPASSPCCALLKGLAGLDVALCLCTAIKANVLGINLVPIALSTIVSACGESIPPGFKCE
AsLTP6	TPRRIPAAQASSGVVPCPPPPAPAAPQTCPIDTLKLGACVDVLGGLVHVIGISSAKDTCCPVL QGLLDLDAVCLCTTIRAKLLNISIVIPALQVLVDCGKTTTPGFQCPA
AsLTP7	MFPKCPRTTVTLVLLNIFFTCVSSHNLPCPPKVPSPSPPPPAPEGLGKCPKDKLFGVCGSWL GLVTENIGAKPSKECCTLVAGLADLEAALCLCTAIKANVLGLVKIKVPVALSLLVNGCGKVPQGF VCA
AsLTP8	MGSETMATFFISMLLLLLSLPIIYACGYCAPRPPYHGPTIPRRPGPRPHGPRHHGGKSGGSG

	GSRGGGGGRGGGRGGGYTPRVPTPHPPVLPPIVNPPPVTNPPGIVPPITNPPGGIVPPIINPPGI LPPITNPPPTPPSSPCPPYGGSPGGGSKGGGVNPPPTTPTTCPINALKGLCLDVLGGLVHIGIGN PIENVCCPVLQGLLELEAAICLCTAIRLKLNLNIFIPLALQVLATCGITPPPGFVCPPL
AsLTP9	MGSETMATFFISMLLLLLSLPPIYACGYCAPRPPYHGPTIPRRPGPRPHGPRHHGKSGGSG GSRGGGGGRGGGRGGGYTPRVPTPHPPVLPPIVNPPPVTNPPGIVPPITNPPGGIVPPIINPPGI LPPITNPPPTPPSSPCPPYGGSPGGGSKGGGVNPPPTTPTTCPINALKGLCLDVLGGLVHIGIGN PIENVCCPVLQGLLELEAAICLCTAIRLKLNLNIFIPLALQVLATCGITPPPGFVCPPL
AsLTP10	MGSETMATFFISMLLLLLSLPPIYACGYCAPRPPYHGPTIPRRPGPRPHGPRHHGKSGGSG GSRGGGGGRGGGRGGGYTPRVPTPHPPVLPPIVNPPPVTNPPGIVPPITNPPGGIVPPIINPPGI LPPITNPPGGIVPPIINPPGILPPITNPPPTPPSSPCPPYGGSPGGGSKGGGVNPPPTTPTCPINA LKLGLCLDVLGGLVHIGIGNPIENVCCPVLQGLLELEAAICLCTAIRLKLNLNIFIPLALQVLATCGIT PPPGFVCPPL
AsLTP11	MGSETMATFFISMLLLLLSLPPIYACGYCAPRPPYHGPTIPRRPGPRPHGPRHHGKSGGSG GSRGGGGGRGGGRGGGYTPRVPTPHPPVLPPIVNPPPVTNPPGIVPPITNPPGGIVPPIINPPGI LPPITNPPPTPPSSPCPPYGGSPGGGSKGGGVNPPPTTPTTCPINALKGLCLDVLGGLVHIGIGN PIENVCCPVLQGLLELEAAICLCTAIRLKLNLNIFIPLALQVLATCGITPPPGFVCPPL
AsLTP12	MDSSKIHLSLILMLCLSSAAPILGCGSCGKAAPKHKKPGHKHHPKGPVVVPPITLPPIVKPPVTL PPIVKPPITLPPIVKPPITLPPVTVPPIVPPVINPPVAKPCPPPPATCPIGTLKLGACVDLLGGLVHI GLGDPVVNECCPVLQGLVELEAAVCLCTTLKLLNLSIYVPLALQLLVTCGKTPPPGYTCSL
AsLTP13	MDSSKIHLSLILMLCLSSAAPILGCGSCGKAAPKHKKPGHKHHPKGPVVVPPITLPPIVKPPVTL PPIVKPPITLPPIVKPPITLPPVTVPPIVPPVINPPVAKPCPPPPATCPIGTLKLGACVDLLGGLVHI GLGDPVVNECCPVLQGLVELEAAVCLCTTLKLLNLSIYVPLALQLLVTCGKTPPPGYTCSL
AsLTP14	MDSSKIHLSLILMLCLSSAAPILGCGSCGKAAPKHKKPGHKHHPKGPVVVPPITLPPIVKPPVTL PPIVKPPITLPPIVKPPITLPPVTVPPIVPPVINPPVAKPCPPPPATCPIGTLKLGACVDLLGGLVHI GLGDPVVNECCPVLQGLVELEAAVCLCTTLKLLNLSIYVPLALQLLVTCGKTPPPGYTCSL
AsLTP15	MESNHLILVFLIWASIACLSSTSNLPSEYSIVGENPEKSLLLPEGRVLELFRRWQEKHHKVKYKHA EAEKRELENFRKNLRHVIERSSRQKSPAASRVGLNKFADLSNEEFREVYLSKVKKPASKKWNHKK MSLKGKMKTCDA PSSLDWRNYGIVTAVKDQGECSWAFSSTGAMEGINALVNGDLVSLSEQE LVSCDTTNYGCEGGYMDYAFEWVINNGIDSEADYPYTSSDGETGSCVTSKEENKVV SIDGYED VAESDSALLCAVVQQPISVGMGDSALDFQLYTGGIYNGSCSDDPDDIDHAVLIVGYGSESGEDY WIVKNSWGTDWGIDGYFYLRNTDIEYGVCAINAMASYPTKESSAPSPTSPTPPSPATPPPPP PPVTTPPPPPPPSPSPSECDFSYCPTDETCCCIFEFFDYCLVYGCCEYENAVCCTGTEYCCP SDYPICDVDEGLCLKVWKS P S S S F
AsLTP16	MESNHLILVFLIWASIACLSSTSNLPSEYSIVGENPEKSLLLPEGRVLELFRRWQEKHHKVKYKHA EAEKRELENFRKNLRHVIERSSRQKSPAASRVGLNKFADLSNEEFREVYLSKVKKPASKKWNHKK MSLKGKMKTCDA PSSLDWRNYGIVTAVKDQGECSWAFSSTGAMEGINALVNGDLVSLSEQE LVSCDTTNYGCEGGYMDYAFEWVINNGIDSEADYPYTSSDGETGSCVTSKEENKVV SIDGYED VAESDSALLCAVVQQPISVGMGDSALDFQLYTGGIYNGSCSDDPDDIDHAVLIVGYGSESGEDY WIVKNSWGTDWGIDGYFYLRNTDIEYGVCAINAMASYPTKESSAPSPTSPTPPSPATPPPPP PPVTTPPPPPPPSPSPSECDFSYCPTDETCCCIFEFFDYCLVYGCCEYENAVCCTGTEYCCP SDYPICDVDEGLCLKVWKS P S S S F
AsLTP17	MESNHLILVFLIWASIACLSSTSNLPSEYSIVGENPEKSLLLPEGRVLELFRRWQEKHHKVKYKHA EAEKRELENFRKNLRHVIERSSRQKSPAASRVGLNKFADLSNEEFREVYLSKVKKPASKKWNHKK MSLKGKMKTCDA PSSLDWRNYGIVTAVKDQGECSWAFSSTGAMEGINALVNGDLVSLSEQE LVSCDTTNYGCEGGYMDYAFEWVINNGIDSEADYPYTSSDGETGSCVTSKEENKVV SIDGYED VAESDSALLCAVVQQPISVGMGDSALDFQLYTGGIYNGSCSDDPDDIDHAVLIVGYGSESGEDY WIVKNSWGTDWGIDGYFYLRNTDIEYGVCAINAMASYPTKESSAPSPTSPTPPSPATPPPPP PPVTTPPPPPPPSPSPSECDFSYCPTDETCCCIFEFFDYCLVYGCCEYENAVCCTGTEYCCP SDYPICDVDEGLCLKNSGDYLGVAARRKMAKHKFPWTKIEETQKSYHPLQWKRNRFAAAMR
AsLTP18	MESNHLILVFLIWASIACLSSTSNLPSEYSIVGENPEKSLLLPEGRVLELFRRWQEKHHKVKYKHA EAEKRELENFRKNLRHVIERSSRQKSPAASRVGLNKFADLSNEEFREVYLSKVKKPASKKWNHKK MSLKGKMKTCDA PSSLDWRNYGIVTAVKDQGECSWAFSSTGAMEGINALVNGDLVSLSEQE LVSCDTTNYGCEGGYMDYAFEWVINNGIDSEADYPYTSSDGETGSCVTSKEENKVV SIDGYED VAESDSALLCAVVQQPISVGMGDSALDFQLYTGGIYNGSCSDDPDDIDHAVLIVGYGSESGEDY

	WIVKNSWGTDWGIDGYFYILRNTDIEYGVCAINAMASYPTKESSAPSPTSPTPPSPATPPPPP PPVTPPPPPPPPPSPSPSECDFSICYPTDETCCCIFEFFDYCLVYGCCEYENAVCCTGTEYCCP SDYPICDVDEGLCLKNSGDYLGVAARRRMAKHKFPWTKIEETQKSYHPLQWKRNRFAAAMR
AsLTP19	MASYPTKESSAPSPTSPTPPSPATPPPPPPVTPPPPPPPSPSPSECDFSICYPTDETCC CIFEFFDYCLVYGCCEYENAVCCTGTEYCCPSDYICDVDEGLCLKNSGDYLGVAARRRMAKH KFPWTKIEETQKSYHPLQWKRNRFAAAMR
AsLTP20	MASYPTKESSAPSPTSPTPPSPATPPPPPPVTPPPPPPPSPSPSECDFSICYPTDETCC CIFEFFDYCLVYGCCEYENAVCCTGTEYCCPSDYICDVDEGLCLKVWKSPPSSSF
AsLTP21	MLDLRKVSPLLLFLTLIFGFGTPTTSAATAAADKSGASTTPALAAVSRDQVCTMCAACENP CQVPVSPPPPPPEVECPSPQPPPSALPPSPPPPEAECPPPPMPACGACESPGTGSYIT SPPPPATSGGGGGGVYPPPYGGYYPMQPPNPVYFYFYNNPPQASVASSSLHLETTAAGF ATSLVLSGLLYVLLRF
AsLTP22	MAPSALLKLACVLLACMVAAPAAAVTCGQVTGALAPCIPYARSGSGVPASCCNGIRSLNSA AQTPDRRATCRCLKSASGSISGLNYGVVSGLPKCGVSIPYKISPSTNCDSVK
AsLTP23	MARYGLSRLVCVTLCLVVASSVAEAAVTCGQVASALAPCISYARSGRGPVAGCCNGIRSLNN AARTTPDRQATCRCLKAAAGGISGINYGVAAPGKCGVSIPYKISPSTNCNSVK
AsLTP24	MAPSALLKLACVLLACMVAAPAAAVTCGQVTGALAPCIPYARSGSGVPASCCNGIRSLNSA AQTPDRRATCRCLKSASGSISGLNYGVVSGLPKCGVSIPYKISPSTNCDRYVLF
AsLTP25	MAPSALLKLACVLLACMVAAPAAAVTCGQVTGALAPCIPYARSGSGVPASCCNGIRSLNSA AQTPDRRATCRCLKSASGSISGLNYGVVSGLPKCGVSIPYKISPSTNCDRYVILVSRSEGVTK PYFSL
AsLTP26	MADLSKLSCFLLLAQLVAPLQPIGSAAVTCSQVSTSMAPCIGYLRGGGPLPDACCSGVQSLNGAA RTTADRQAACRCLQSAAGSIPGINLGLAAGLPKCGVNVPIYKISPSTNCNSVK
AsLTP27	MARYGLSRLVCVTLCLVVASSVAEAAVTCGQVASALAPCISYARSGRGPVAGCCNGIRSLNN AARTTPDRQATCRCLKAAAGGISGINYGVAAPGKCGVSIPYKISPSTNCNRSRYSTKALTYKRH
AsLTP28	MADLSKLSCFLLLAQLVAPLQPIGSAAVTCSQVSTSMAPCIGYLRGGGPLPDACCSGVQSLNGAA RTTADRQAACRCLQSAAGSIPGINLGLAAGLPKCGVNVPIYKISPSTNCNS
AsLTP29	MASLKLCLALVVGLAAVAPLAASGVNCGQVASYVAPCLGYLRVGGGLIPQACCNGIRSLNTAAKD TPNRQATCRCLQAAAGNIRGLKGLVLSALPGKCGINVPYKISPSTDCSRVR
AsLTP30	MANPGLLKTAALAVAVVCMLVANAPAAAQMSSENQVNTLMPCASYVLNNGGAVPAACCNGIR SLYSAAKTTADRQGVCKCLESAINGIPYNAYNAGLAEGLPGKCGVNIPYKISPSTDCSKSVQ
AsLTP31	MANPGLLKTAALAVAVVCMLVANAPAAAQMSSENQVNTLMPCASYVLNNGGAVPAACCNGIR SLYSAAKTTADRQGVCKCLESAINGIPYNAYNAGLAEGLPGKCGVNIPYKISPSTDCRQKHRHFI ERH
AsLTP32	MANPGLLKTAALAVAVVCMLVANAPAAAQMSSENQVNTLMPCASYVLNNGGAVPAACCNGIR SLYSAAKTTADRQGVCKCLESAINGIPYNAYNAGLAEGLPGKCGVNIPYKISPSTDCRQVLLHHL
AsLTP33	MKGAALAVALAAVVAVVALMASQVEGITCGQVNSLGPAPYLTKGGDPSACCNGVRELKNS TPYPADRRACVCIKELASRVHNIKADAAAALPGKCGVSIGVPIRTDVCNSIS
AsLTP34	MRPSSFAVLALLVALFLLARFPVSES AVSCSDVLKDLKPCVSYLKSGSGMPPAACAGVSALAN AATSSADKKAACVCIKNAQAQMNPNLAQALPGNCKITLPVAVSANVDCSKVG
AsLTP35	MRPSSFAVLALLVALFLLARFPVSES AVSCSDVLKDLKPCVSYLKSGSGMPPAACAGVSALAN AATSSADKKAACVCIKNAQAQMNPNLAQALPGNCKITLPVAVSANVDCSKYVFNLALKFFEKE QIITHREKNLRELPLVLSYASELFLSF
AsLTP36	MPDGQAACKCLKSASESISGLNYGVVAGLPKCGVSIPYKISPSTNCDR
AsLTP37	MAARRMEAALVLAAVAMLWVRVAAQSSCTNVLISLSPCLNYITGNSSNPSPNCCAQLASVVRSQ PQCLCQVLSGGGSSLGINVNQTQALALPGACNVQTPPISRCNGKQIIQMMMVFSEFKIQKG
AsLTP38	MAARRMEAALVLAAVAMLWVRVAAQSSCTNVLISLSPCLNYITGNSSNPSPNCCAQLASVVRSQ PQCLCQVLSGGGSSLGINVNQTQALALPGACNVQTPPISRCNGKQIIQMMMVFSEFKIQKG
AsLTP39	MAARRMEAALVLAAVAMLWVRVAAQSSCTNVLISLSPCLNYITGNSSNPSPNCCAQLASVVRSQ

	PQCLCQVLSGGGSSSLGINVNQTQALALPGACNVQTPPISR
AsLTP40	MTHKVFLTPPQFTDDNPSPTAAMTAPRSAMLVTVLVAVAVALLWAGAEAQTSNALTSLSPCLSFITGNSSAPSPACCTQLASVVNSEPLCLCAVLNGSASSALGVNINQTQALQLPPACNVQTPPTSSCNAALSPAGSPAGSPTSTPGTRASLSLSLP
AsLTP41	MTHKVFLTPPQFTDDNPSPTAAMTAPRSAMLVTVLVAVAVALLWAGAEAQTSNALTSLSPCLSFITGNSSAPSPACCTQLASVVNSEPLCLCAVLNGSASSALGVNINQTQALQLPPACNVQTPPTSSCNAALSPAGSPAGSPTSTPGSSSDGSSASLRWVSLVFLTFMASHEAVMLMQNNLEPI
AsLTP42	MRYVVRHGKGSRHCLPTCCSGLQQLARNVQSVNDKQDICRCLKAGAKFVGGGLQDKFSPSIPTACHINVGFPVSVANVRCEKIHSVWR
AsLTP43	MGSKTSNGQLAAKVLFFSIIMFMSSNPVNSITCQDALKVVLPCGPFALGAVPPPPSAECCSGAQLASMANNTTEARRTLCCQYKDIPPSVGIKPERVQEIPIQYCKVNVTIPTDPHVDCSSIP
AsLTP44	MAARRMEAALVLAAVAMLWVRVAAQSSCTNVLIISLSPCLNYITGNSSNPSPNCCAQLASVVRSPQPCLCQVLSGGGSSSLGINVNQTQALALPGA
AsLTP45	MAARRMEAALVLAAVAMLWVRVAAQSSCTNVLIISLSPCLNYITGNSSNPSPNCCAQLASVVRSPQPCLCQVLSGGGSSSLGINVNQTQALALPGA
AsLTP46	LGSIVQSQRCLCLLLNGTASSLGYNINQTLALALPGACNVNMPASLCKSPSPSSAPTTSPPADDNTPVAPTASSTPTVPSGSKTPTTSGSKGIGVSLGSILFALILASFGSNPLKF
AsLTP47	MALKNLEIMGLALLTVLAAALLARPGLAQSTSGCTQVLMTLTPCLNYVTGNSSAPSTCCSTLGSIVQSQRCLCLLLNGTASSLGYNINQTLALALPGACNVNMPASLCKSPSPSSAPTTSPPADDNTPVAPTASSTPTVPSGNNLTSLLVQQ
AsLTP48	MALKNLEIMGLALLTVLAAALLARPGLAQSTSGCTQVLMTLTPCLNYVTGNSSAPSTCCSTLGSIVQSQRCLCLLLNGTASSLGYNINQTLALALPGACNVNMPASLCKCNGLTSKNYSSFFVFTSVLG
AsLTP49	MGKICYLIFTSFFLALANCQSPSLVSSSSCSQIIYEMIDCISFLTGAGGSAKPDALCCSGFKSVLKADAHICQAIQREAPQMGIRLNKTRAIELPSDCGVSHFSALKICGCNCVFNFSQ
AsLTP50	MGKICYLIFTSFFLALANCQSPSLVSSSSCSQIIYEMIDCISFLTGAGGSAKPDALCCSGFKSVLKADAHICQAIQREAPQMGIRLNKTRAIELPSDCGVSHFSALKICGFLDSPGTPSAESPKTPVPPPTKTPTPSPLPPTMSPASSPRRTKPAPSPTPPVPSPKSLVSPAPSPTPTTNTVPSPKSPVSSLPKSTSSAPSSLERGISQVPAPAPENLKSGGYSLRACFAVVISLCAIHVS
AsLTP51	MAQTKMRIVALILAMVAAGAAAQSSSCTNALISMSPCLNYITGNSSSPSSSSCCSQLANVVRSEPCCLCQALNNGGSSSLGISINQTRALAVPSACNIQTQPISRCNGECVCTRKTSKLTFSISSKSFDDLTFCTD
AsLTP52	MAQTKMRIVALILAMVAAGAAAQSSSCTNALISMSPCLNYITGNSSSPSSSSCCSQLANVVRSEPCCLCQALNNGGSSSLGISINQTRALAVPSACNIQTQPISRCNGECVCTRKTSKLTFSISSKSFDDLTFCTD
AsLTP53	MAQTKMRIVALILAMVAAGAAAQSSSCTNALISMSPCLNYITGNSSSPSSSSCCSQLANVVRSEPCCLCQALNNGGSSSLGISINQTRALAVPSACNIQTQPISRCNAASPSASPTGAPDTPNAVPSDSGTKDVPTTERDGSSAGSMAELSIPLLFALLFAASYASANT
AsLTP54	MASKGDEMGLGRVMALVVTMAALCAGAVAQSGCTSVMGLAPCLSFVTGSSSTPSSSSCCSQLASVVQSQRCLCMVLNNGGASSSLGVTLNQLALALPGACKVQTPPVSKCNASGGTVAPAGSPESPSDVSOGGTSSPSGS
AsLTP55	MASKGDEMGLGRVMALVVTMAALCAGAVAQSGCTSVMGLAPCLSFVTGSSSTPSSSSCCSQLASVVQSQRCLCMVLNNGGASSSLGVTLNQLALALPGACKVQTPPVSKCNASGPESSPSDVSASKEVSSTAGSSTSHGIAKMPPTLPLAVLSVFMATFGYAPMGMAI
AsLTP56	MASKGDEMGLGRVMALVVTMAALCAGAVAQSGCTSVMGLAPCLSFVTGSSSTPSSSSCCSQLASVVQSQRCLCMVLNNGGASSSLGVTLNQLALALPGACKVQTPPVSKCNAGEFERKLISSFLLSWILMLQLVGSFRRNGCSSRLAREFAERRIRWYF
AsLTP57	MASKGDEMGLGRVMALVVTMAALCAGAVAQSGCTSVMGLAPCLSFVTGSSSTPSSSSCCSQLASVVQSQRCLCMVLNNGGASSSLGVTLNQLALALPGACKVQTPPVSKCNASGGTVAPAGSPESPSDVSOGASKEVSSTAGSSTSHGIAKMPPTLPLAVLSVFMATFGYAPMGMAI

AsLTP58	MIDMALKRMQIVLLAAVLAASLARPCSAQSTSGCTAALMTLTPCLSYVSGNPSTPSSTCCSTLSS VVQSQRCLCLLLNGGVPRLRYNINQTLALALPGACSVKTPSVSLCNALSPSSPPTTSPPADSSD NTPVTPTASSTPTIPSGSKATPAATSTSRSEASFGSLAFALLASF
AsLTP59	MSDVRQGGSVSDVRHGKGSRHCLPTCCSGLQQLVRNVRSVNDNQDICRCLKAGAKFVGGGLQD KFPSQIPTACHINVGFLVSANVSCEKFH
AsLTP60	MSDVRQGGSVSDVRHGKGSRHCLPTCCSGLQQLVRNVRSVNDNQDICRCLKAGAKFVGGGLQD KFPSQIPTACHINVGFLVSANVSCEKFH
AsLTP61	MCSFTEKMERRNIRAALETAALVLAALLVVDAQADTSCLSRLAPCLNYLNGTRDPPSTCCDPLK SVISSDPGCLCGLISNQGADQAQQAGINVTEAQQLPGRCGQHVNPIACLSGSPGSRNSDENSAG LFGSSLRTPLAIVALSVAAQILWD
AsLTP62	FSNEIHCPVTKCLFFSALFLISLVSAEATITCSVVNAKVAPCAMYTMGKAPDPASPLAAVACNSS PGTFDQSTTSKTFAVA
AsLTP63	MASRGIMMGLVLVLAATLWTRSAAQSGCTAALTSLVPCLNFINGNLSAPSSSCCSQLASVVKASP RCLCSVMNGPPMFGITINQTLALPNACNVQTPPVQCNANAPTASVGAPEVPPSPSDSP AGPSEDEPTAAAPAATTPPSGTTGTGSKAVPSTNGNTSDGSIVKAQLHLAALVLAASIVKRF
AsLTP64	MPPMSHVVRTLLGLLLVSAAAAASGHPTCKYVDEYFPDCLDFLVGTYYIPPQRCCRSVGALNYM ANHLVGPRMICQCIEAMVMRTNPLMASRIESLPKFCATHLSFPISTAMDCSRVT
AsLTP65	MPPMSHVVRTLLGLLLVSAAAAASGHPTCKYVDEYFPDCLDFLVGTYYIPPQRCCRSVGALNYM ANHLVGPRMICQCIEAMVMRTNPLMASRIESLPKFCATHLSFPISTAMDCSR
AsLTP66	MPPMSHVVRTLLGLLLVSAAAAASGHPTCKYVDEYFPDCLDFLVGTYYIPPQRCCRSVGALNYM ANHLVGPRMICQCIEAMVMRTNPLMASRIESLPKFCATHLSFPISTAMDCSR
AsLTP67	MAAFFNCWTCYEVVFLGLVPHFPPRVGRGHHHVQHGQCQGGSVSDVRHGKGSRHCLPTCC SGLQQLVRNVRSVNDNQDICRCLKAGAKFVGGGLQDKFPSQIPTACHINVG
AsLTP68	MVHISIIHLSLFLLLCVSSRGDEFTFCINVYQKFADCMAFLRGLAPVPTGDCCSNLLALNDLAKRK QSSPELLCQCNEGIGYVLDTPYLPRIEDLRSKCQVHLSFPISNAMNCSMGMRQAKLELTNE
AsLTP69	MVHISIIHLSLFLLLCVSSRGDEFTFCINVYQKFADCMAFLRGLAPVPTGDCCSNLLALNDLAKRK QSSPELLCQCNEGIGYVLDTPYLPRIEDLRSKCQVHLSFPISNAMNC
AsLTP70	MESPMGKFICILALLMASMASRGAHAAGECGKSTTPDAEAMKLAPCAVAQAQAKAPVSDSCCA QVKSIGKSPSCLCAVMSLNTAKASGINPEIAITIPKRCNIPNRPVGYKCGAYTLP
AsLTP71	MEAPMKQISVLVFFAILAISGFQTVYGAGECGKSSTPDREAFKLAPCMSAAQDETAPVSSSCCAR VKRIGQNPSCCAAMLSNTAKIAGIKPEVAVTIPKRCNLSDRPAGYKCGGYTVP
AsLTP72	CARVKRIGQNPSCCAAMLSNTAKIAGIKPEVAVTIPKRCNLSDRPAGYKCGGQISLLSLTAEELS FGNLYLQMHPALDVTML
AsLTP73	MASKASGCSCVLTVLFLVSSMASREAIAAGECGKTPISTAASVSLPCLSAAGNAKAVPPACCT KVGALIATAPKCLCAVLLSPLAKQAGINLGIAITIPKRCNIIKRPVGGKCGKYTVP
AsLTP74	MSHNSHVAMFVPLSILMVLAISNGFIWVASAAGECGRTPIRSAAAASLSPCLGAARDARAKVPPAC CAKTNALIRAAPRCLCAVLLSPLAKQAGITPGTAISIPKRCNIRNRPAGKKCGRYTVP
AsLTP75	MARNSLVPLSILMVLAISNGFIWVASAAGECGRTPVRSAAAASLSPCLGAARNARAKVPPACCAKV NALIRTAPRCLCAVLLSPVAKQAGIMPGTAISIPKRCNIRNRPAGKKCGRYTVP
AsLTP76	MSHNSHVAMFVPLSILMVLAISNGFIWVASAAGECGRTPIRSAAAASLSPCLGAARDARAKVPPAC CAKTNALIRAAPRCLCAVLLSPLAKQAGITPGTAISIPKRCNIRNRPAGKKCGSEYKHIPINSPLK
AsLTP77	MGEHQTGMKTHSRSYHALDLLKPLLQPNSTPHKEFRQTQMATLHYISFLPLLLVGIPAALSQE PASSSPAVASCQPRLLSLAPCAPFVQGTASQSPSEPCCDNLQIYTQQPHCLCLVLDNTSLVPLPI NGTLVLELPQLCNLQADASSCSGNNETFKGTRIKIGFNPSEPNIFTKGSGFG
AsLTP78	MGEHQTGMKTHSRSYHALDLLKPLLQPNSTPHKEFRQTQMATLHYISFLPLLLVGIPAALSQE PASSSPAVASCQPRLLSLAPCAPFVQGTASQSPSEPCCDNLQIYTQQPHCLCLVLDNTSLVPLPI NGTLVLELPQLCNLQADASSCSGLGAPASPSAPGPQVSVGARRTNNSTVAASPVVEVTPRPSIIR SGPGLSSGTMLDREGRPTFLVAAAVALLLAQVYF
AsLTP79	MASRGIMMGLVLVLAATLWTRSAAQSGCTAALTSLVPCLNFINGNLSAPSSSCCSQLASVVKASP RCLCSVMNGPPMFGITINQTLALPNACNVQTPPVQCNANAPTASVGAPEVPPSPSDSP

	ANAPTPASVGAPEVPPSPSDSPAGPSDETPTAAAPA
AsLTP80	MRAPRVALCALLAVLLLAEAGVSVEAAVTCKPIELSSCVSAITSASPPTTLCCSKLKEQRPCLCQY LRNPNLKKFVNPNARKVASTCGSPFPKC
AsLTP81	MTKPSNRSPSPRLLAAALVAALVLLSLDRAPVAEAVTCSATELSSCVAAITSSAPPSALCCSKLRE QRPCLCGYIRNPNLRQYVTSNARKVARTCGVPYPTC
AsLTP82	MKAPSIVVCSLLLLLAEAQITMAVTCKPTELSSCANSIVSSTPPSSLCCTKIKEQKPCCLCQYLKPN SLKKFVNPNARKVASTCGTFFPKC
AsLTP83	MTKPSNRSPSPRLLAAALVAALVLLSLDRAPVAEAVTCSATELSSCVAAITSSAPPSALCCSKLRE QRPCLCGYIRDPNLRPYATSPNVKRVASTCGVAYPTC
AsLTP84	MSKYVLCIVLIFLVKLEPSTAVTCDPMELSSCASAISSSPPSAMCCAALKEQKPCCLCQYMKNPN LQKLINSPNSRKVAIQCGSPFPSC
AsLTP85	MAKPSNRSPSPRLLVAALMAALVLLSPDRAPVAEAVTCTINELSSCLPAVTSSAPPSALCCSKLRE QRPCLCGYIRDPNLRPYATSPNVKRVASTCGVAYPTC
AsLTP86	MTKPSNRSPSPRLLVAALMAALVLLSPDQAPVAEAVTCSATQLSSCVSAMTSSAPPSALCCSKM REQRPCLCGYIRNPNLRKYLTSADGKRVMHVCGVPYPTC
AsLTP87	MTRKGPSPFVASFLVTALVSSQLAPGTEAECDATQLTPCLPSLMPPFSPPTDTCCGKLKEQQS CFCDYMKDPRYGKYVKSPRAKKILELCNIPYPQC

*En negrita se marca la secuencia correspondiente al péptido señal

Anexo 3: *primers* y su ubicación en las secuencias del transcriptoma.

Tabla 3: *Primers* utilizados para validación de las secuencias seleccionadas.

	F	R
AsDf1	TCAACCAAATAGCAGAAAACAAACA	AATCCTTTTTTCAGATTGGTATTTGCCT
AsDf2	TCACCACTTCTTCACTCGCC	TGATCACGCCGATGTATCCG
AsDf4	TCAAGAACAAGAAGCCACCACT	TCTCAAAAACCACAATAAAGGGCA
AsDf5	CGTACGAGTTCTCCTGTTTGC	TTCTCGACACGATTGCAAATCT
AsDf6.		
AsDf7	CCACCACCACCTCCTCAAAG	ACATTACACCCCGGATCAGG
AsSn1	TGACGTGGAACCCATCGTTT	ATCAGCAACAGCTCGGACAT
AsSn13	CACACCCCTTCATTGATCCCA	TGCCATCCAAATGACACGA
AsSn42	ATTAAGATCGTCGTCGGCGA	GCTTACATGAACTTTTTGTGCA
AsHv1	TCGTTTCGCTTCCCTATCGT	CCCCTTCTCGATCTCCCTCA

Fig S1. Localización de secuencias *primers* en los contigs del transcriptoma

En verde se marcó el codón de inicio de la proteína, en rojo el codón de finalización, en celeste la secuencia codificante y en fucsia la ubicación de los *primers*.

Defensinas

AsDf1

CTATCAACCAATAGCAGAAAACAAACA AAACATGAAGAGAATCTCCATCTGTTCTGTGCTTCTGCTTG
TTCTCTTACTAATTGCTGGCAGTGGAAGCATGATGGCACAGGTATCGGGCCTCTGCAGCAAGAAGTTGC
AACTCCCTGCTCACCGATGCATCAAGGACGAGTGCATTTCTACTGCAAGACTACCTACGGCGATGGGG
CTACGGGAGTTTGTGAGTCAGACACACAGCCGGAAGATATCTGCCTCTGTCAATACCTTGTAGTCGT
ACCCGTTCCCGTCGCTGCACAATAAGTGGCATATATATGTCGCATATGATTTATATTAGTTTCGTTTTCGA
ACTTCTCGTATACCTACTGATTACTCTGCTCAGGCCAAATACCAATCTGAAAAAGGATTT

AsDf2a

GTGGCTGTGAATTAAGTGAATCTGGACCGTCGGATCTCTGTCGGCATCCATGCAAGACCCGACGTTTC
GAAATAATTGAAAA TCACCACTTCTTCACTCGCC TATAAATCGCCTTTCCTGCCGTGCCAATCCTCCATC
CATCTCTCTCAAAATGGATTTCTCGAAGCGATTGATTCCAGCTGCTTTCGTTGTGATGCTGCTGCTCGTG
GCAACTGAGATGGAGCCGATGGCGGTGGAGGCGAGAACTTGTGAGTCGCAAAGCCACCGTTCAAAG
GGGCATGCGTGAGGGACAGCAACTGCGCCGCTGTTTGCCAGACGGAGGGCTTCCAGGGCGGGCATTG
CCGGGGCTTTCGTCGCCGTTGCTTCTGCACCAACATTGTAGAGAGTGC GCGCTCTGAACGCGCTTGT
GAATTGCATCTTGTGATCTCTGCCAGTCTTTTGAATAATCCGGTATTGTACCGGAATGTCATTGTGTG
GTCGCGGATGTAATCTTCTTTGGCTGTTTACTCTTTTTGTCATGGTCGCTTCTGTTTGAGAGTGTGAAAT
CTGTTTTAAGAGCCACTATAGGATTGCGGATACATCGGCGTGATCATGACATCAACTCGACACGACAA
CTTTGCCATCATGGCCCTAGCAACTACGTA CTGTTTGTGAGTCTGGACTCTGTATGTTCTTCACTAGGAT
CAACTATTTGTTAAGTACGCGAACCCTTTTATCTCCATGACGCCGAGGATGGGCTGTGCAGTCATTCC
AAGATTGCAGTCCGATCCAGTTTCTCCACATCAGAGTTCTTGTTCCAAAGTCGGGCGATTGATGGC
TCATTAGACTGTAACCGAAGCCGTA CTGTTTCAACACAAGCAATTAGAGATTTCCCTGAAAAAAAAA
AA

AsDf2b

GTGGCTGTGAATTAAGTGAATCTGGACCGTCGGATCTCTGTCGGCATCCATGCAAGACCCGACGTTTC
GAAATAATTGAAAA TCACCACTTCTTCACTCGCC TATAAATCGCCTTTCCTGCCGTGCCAATCCTCCATC
CATCTCTCTCAAAATGGATTTCTCGAAGCGATTGATTCCAGCTGCTTTCGTTGTGATGCTGCTGCTCGTG
GCAACTGGTCCGTCCTCATCTACAACACTTATGCATGCATCAGTTTCATAGTTCAAATCAATAATACGT
TTTCCAATGTTGGCCTTGTGAACTTATAATAGATGATTCTGAGCACGATTAGAATTGCTTATAGATGAAT
TTTCATGAGAGAACTTTGATTAACCTCGACCCTTTTCTTTCACCTCCTTTTTCGGTATGAGCAGAGATG
GAGCCGATGGCGGTGGAGGCGAGA ACTTGTGAGTCGCAAAGCCACCGTTCAAAGGGGCATGCGTGA
GGGACAGCAACTGCGCCGCTGTTTGCCAGACGGAGGGCTTCCAGGGCGGGCATTGCCGGGGCTTTCG
TCGCCGTTGCTTCTGCACCAACATTGTAGAGAGTGC GCGCTCTGAACGCGCTTGTGAATTGCATCTT
GTCGTATCTGCCAGTCTTTTGAATAATCCGGTATTGTACCGGAATGTCATTGTGTGGTCGCGGATGT
AATCTTCTTTGGCTGTTTACTCTTTTTGTCATGGTCGCTTCTGTTTGAGAGTGTGAAATCTGTTTTAAGA
GCCACTATAGGATTGCGGATACATCGGCGTGATCATGACATCAACTCGACACGACAACTTTGCCATCA
TGCCCTAGCAACTACGTA CTGTTTGTGAGTCTGGACTCTGTATGTTCTTCACTAGGATCAACTATTTGTT
TAACTTACGCGAACCCTTTTATCTCCATGACGCCGAGGATGGGCTGTGCAGTCATTCCAAGATTGCAGT

CCGATCCAGTTTCTCCCCACATCAGAGTCTTGTGTTTTCAAAGTCGGGCGATTGATGGCTCATTAGACTG
TAACCGAAGCCGTA CTGTTTCAACACAAGCAATTAGAGATTTCCCTGAAAAAAAAAAAA

AsDf4

TGTTCTACCTCTCTCTCAAGAACAAGAAGCCACCACTTAGCCATGAAAACACTCACTCAACTCCCTCATT
TTCTGTTGTCATGTTGATCTCCTTTGGGAATGAGCTGGGAATGAAGCAAGGTGCGGAGGCTCAGTGT
CAAGAGACATTGTATGGAGGTGGCTGTCAAGACAAACCTTGAATGATGCTTGTGTGCAGAAGTTCGG
GCCGAAAGCGTATGGATTTGCTTCTTCTACAAGAAACCTAGCGACACTTGTGTGTGCCGTCATCCTGT
TGA TGCTTTTTCTAATGCCCTTTATTGTGGTTTTGAGATGAATAAAGAGAAATAGTGAATCATA C

AsDf5

CTTCTCTCTA CGTACGAGTTCTCCTGTTTTGCTCGAGGTAGCCCTTGAAAAGCACATCTTATTCTCAACC
ACATGGAGTCCCTCGTTTTCTCGTATTCTCGTGCTAGTCTCCTTCTCTCAGGGACGTGCATGGTGT
GGTGCCGCAAGCAAATGCCGAGAAGAGATGTGTAGAGGTGCTGGACCCGCACAGCTGCAACCTCTACA
ATTGCCAAAAGGCGTGTACGAGAAGCACACCGCCGCCAGCAGCCCCGCGGAGAGTGCCAGCACGA
TGCCTCTTCACTACTCATGCGCTTGCTTCTTCAACTGCTGATTTGATGAACGATCTGCTACTCTGAGTG
TCTCTTTTTCTGTCTGTCTCTCCCTCTTTAAGTTGTAAGAGATTTGCAATCGTGTGCGAGAAATAAG

AsDf6

CCTTCGTCACATAAATATATCACCA CCACCACCACCTCCTCAAAGCCCTCCTCCATTCTAAGTTTTGCA
GACAAGTTGCAAGAAATGGAGAAGAGCAAGATGAGATGCTGCATGGGGCTTTCTGATGCTCCTCCT
CGTCCTCGCCGCCGTAAGCCCCGTCTGTCAACCCTATTCTTTACTTTACTTCCTAAATGTTACATCTAT
ATGACCCTACGTCTGTGTGTGTACACATGCGAG*AGGAGGCGGAGGGGAGGATGTGCCAGTCCAGAG
CCACGGCTTCCACGGGTCTTGCCTCCGAGCCACAACCTGCGCCCTCGTCTGCCGCAACGAGGGCTTCTC
CGCGCGCCGCTGCCGTGGCTTCCGCGCCGCTGCTTCTGCACCAAGCTCTGCTGA AAAGCCAATGCA
CGCGCATGCAACAGAATTAGAACGTGTTGACTAGCTTAGGTTGGTGTGTGCTGCATTGTAGCGTGTGA
CCGAGTGAGCCAAATAAACTTATTTGAAGATTCTGGGCTCGTTAGTTTCTCAGCTTTCTAGGACTTGATG
TGCTTTGACCATTCACTGTGCTCTCTCTTCACTTTGTGTAATTTCTGATCCGGGGTGTAAATGTGTT
TTCATGCTTCTGTTTTCACTCTCTCTGCACTTCCATTTGACTTATATCTCGTTCAGACTGGACTTTTCT
ACAGCTTCGGAACAGCTACAGATCATGTCAGTTTCTGGGTAACAATCGAAACACGGGTCTATAACCT
CG

*Subrayado se indica una región que parece ser un inserto con respecto a AsDf7.

AsDf7

CCTTCGTCACATAAATATATCACCA CCACCACCACCTCCTCAAAGCCCTCCTCCATTCTAAGTTTTGCA
GACAAGTTGCAAGAAATGGAGAAGAGCAAGATGAGATGCTGCATGGGGCTTTCTGATGCTCCTCCT
CGTCCTCGCCGCCAGGAGGCGGAGGGGAGGATGTGCCAGTCCAGAGCCACGGCTTCCACGGGTCT
TGCGTCCGAGCCACAACCTGCGCCCTCGTCTGCCGCAACGAGGGCTTCTCCGGCGCCGCTGCCGTGG
CTTCCGGCGCCGCTGCTTCTGCACCAAGCTCTGCTGA AAAGCCAATGCACGCGCATGCAACAGAATTA
GAACGTGTTGACTAGCTTAGGTTGGTGTGTGCTGCATTGTAGCGTGTGACCGAGTGAGCCAAATAAAC
TTATTTGAAGATTCTGGGCTCGTTAGTTTCTCAGCTTTCTAGGACTTGATGTGCTTTGACCATTCACTGTG
CTCTCTTCTTCACTTTGTGTAATTTCTGATCCGGGGTGTAAATGTGTTTTATGCTTCTGTTTTCACT
CTCTCTGCACTTCCATTTGACTTATATCTCGTTCAGACTGGACTTTTCTACAGCTTCGGAACAGCTAC
AGATCATGTCAGTTTCTGGGTAACAATCGAAACACGGGTCTATAACCTCG

Esnaguinas

AsSn1

CACATTACCAACTTTTCTAAATCCATGATATAAGGTAAAAAGAAAAGCAACAGATTAGGGTTGCGGTAG
TAGGGTAAAAGCAACTATAAACTAAAGCAAGAAGCTACGTAGACGACGGTCAAGAAGCTTCTTGTAAAC
CTTTGTATCGTTCCCTGCATGATGACGACGACCAACTTGTTAGCTTCTACGGAATGGCTTCGGTAAT
AGACGCTCGGGAGAATCGATGCTTTTTGACAGATGAGATGAGAACTCCCCGAGGTGAACGAATCACCA
GGCTCCCTGAAGAATTCCTTGAGGCGTCAAGGGCGTGCATTAGTAGCAAGTTTGGTCGTTTCGGCTT
TCGTTTTCTCTTTTTACCCCGTGATCAGCAAACGAGGGATTCCCACGCCGAGAGCGCAGGGCGCTTGTG
CTTCTCATCCAAACCACTAGTAAACACCTCTCTCGCGTGTTCATCATGCGGGCTTCGTCTATTCCAAAG
ATCATTAAACAAATCCGGAGCCCCGGCATATTACCCACACCAGGAATAACCTTATCTAGGACCCACCCCT
TGTCTGTTGAATACGTAATGATGATCCGACGTGGAACCCATCGTTTCTTTGCTTCGTAGGCACTATCT
ATTCCCCTACACGCCCTTCTCTCCATGTATAAATACCCTCTCACTGGCCATTCCATCACCACCTCGCAC
ACATCAACCACGAGACTCTCCATCTCTCTTTGCCTTATTCTTGTAGTAACGATACATTGGGATGAT
GAAGAAGTTTGCTTTGGCACTTTCTTCTGGTGGCTCTTCTCCTCAGCTCCTCCTTTCGACGCCGCCA
CGGCCGATTCAAGCCTTTGTGACTCCAAGTGCAGGTTGAGGTGCTCCAAGGCAGGAAGGATGGACCG
GTGCTCCAGTACTGCGGGATTTGCTGCGAGAAATGCCAGTGCCTCCGTCGGGACTTACGGCCACA
AGGACGAGTGCCTTGTTACCGGGACATGAAGAATCCAAGGGAAAGCCCAAGTGCCTTAAACGTCGT
CCTTTTCTCGTTGATTCGTGGTGGCCTATATGGTACTTTGATGCGTCTATAACTAAGCAGTCGTGGCTAC
TTATGTCCGAGCTGTTGCTGATGATTCTGTGTAATACTGTGTCGTGTATGGATGTTTTTCATTTTTTGTTG
ATGTTGTCGCTGTTATTGTGTATATTGGAGAATAATGAAGAATAAATCGCTACCTACCTTCTCAGATT
ATTGCCCTGTAGCCTTAGGGAAGGAAGCATTGGTAATATTGCTGTCGTTTCGTGGGAAAACCTGGG
CAAATGTGTGTAGAGTACATGAAAGCGACGCATGACATGTGATGCAATTCAATGGATAGATCATAGTG
TGATTTGCACTCTAGAGTGATAATCGTCCGTCTCAAAGCCCTCGAAATGAAAAATAAAAGTAACGGT
TTGGTAGTAGTACTAGTAACTGCTTTGAAAGTCATATAACGGCGGTGTAACGGCTTTCGGATTTTGT
AACGGC

AsSn13

CTGCAAGAGACAAGAAAACATCCTTAAACACACCCCTTCATTGATCCCAATCGTTCAAGCAGTCAATGCTC
TGTTCAAGGTTTTGGCTTGTGCTCCTTTCCCTCCTCGTTCTCAATCAGCTTGTGCAAGCTGCTGAAGCT
ACGGAGCTGAGCTCTGAATCCGTGGAAAGTTCTGCCGCCAAAAAATAGATTGTGGAGCGGCTTGC
CGCCAGGTGCCGGTAGCATCGAGGCAGAAGATTTGCAAGAGGGCATGCGGGACCTGCTGCGCCCG
TGCAACTGCGTGCCGCCGGGACC GCCGCAACCGGGACGTGTGCCCTTGTACGCCAGCATGACCAC
TCATGGAGGCAGACTCAAGTGCCTTAAACCTAGTCAAGCACCCGCTGTGGAATTGGAGTGCAGTTTC
GTGTCATTTGGATGGGCAATATGTAGCCGAGAATAAAG

AsSn42

CATACTTCTATTAAGATTAAGATCGTCGTCGGCGAGGGTGGTCGCTATGGCTAGTTTCTCGTTGCTCCG
CTCTTCTCTGCTATTCTTCTCGTCGCTTCGATGTGCTTCGTCGAAGTTTCGCTGGCTGGCGGAGAAGGA
TCGCTTCGTCGGACCAATGCGCCGCCGCGTGC GACTACCGGTGCTCGGAGACGCTCACC GGAAAGCC
GTGCCTGTTCTTCTGCAACAAGTGTGCGCCAAGTGCCTGTGCGTCCCCTCCGGCACGTACGGACACAA
GGAGGAGTGCCCGTGTACAACAATTGGAAGACGCAGGAGGGAAAACCCAAGTGCCTTGAAGAATT
TCTATCCATTTTCGACCAAAAAAAAAAAAAAATAACGTATATATATATATATATGTCTTTTCAATTTCCAGAA
AATGTAATGGATTTTTTTTTTATTCTCGGAATCTCTAGATATGTTCCAGGCAAATAATATGCACAAAAAG
TTCATGTGAAGCTCG

Hevein-like

AsHv1

ATTCCCTTTTGAAGAAATATTGCTTGGGTCTTTGAGATTCTTACTGTTGTCGTTTCGCTTTCCCTATCGT
ATCACAGCTCACATTCCAAGTTCATCTAATATCGTATTTTCATAATGCGTATCTCTACACTCGCTTTCCTC
TCCTTGGTATCAGGCGCTCTCGCCGGAAACTGCGGCCAGACAATGGCAACAAGAAGTGCCTGCAAC
CGAATGCTGCTACAATACGGCTGGTGCAGAACCAAAAAGACCACTGCGACCTCCCCACCTGCATAAA
GGCCTTCTCAGGCGCATCTCACCTGCGAACCAAGTGGTAGCACACCCACCACAACCCTCAAGACCAG
CGTCGTCAAACAAGCCACTTCCCCGCAACCACGTTCCAAGCACCGTCCCCAATATCGACGTCTGCGG
CTCCGCCAAGGCGGCGTAACCTGTCCCGCGTCGGTGCCGATGGCTACTTCTACCGCTGCTGTTGCTC
GGCCGGCCACTGCGGACCGAAAAACGACATTCAAGACCAAGCACTCTACTGCGGTGATGGATGCCAG
GCTGGGTTTCGGGAAATGCGATTCTATGGCTGGACCGCTGTGCCGACGCAGACTCCGGGGACGGCGG
GGGATGGCGAGACGTGTGGGCCGATTGTGAATAGGAAGTGTGGGGGAAGCTTGTGTTGCTCTGGTAG
TAATTTTTGTGGCTCTGGAGTGGATTTTTGTGGCGCTGCGAATTGGTGTGATGAGTAAGTGGGGGAGGT
GTGATTAGGTGTTGGTTGGCTGGTTGTTGGTTGCGGGTTGTGATGATGGGAAAATTGGTGGGGAGA
TGGGGGTGAGAATCGGGTTTTCAAGAAGGATGAGGGAGATCGAGAAGGGGGTTCAAGGTAGCAGAG
GACGG