



CONAHCYT

CONSEJO NACIONAL DE HUMANIDADES
CIENCIAS Y TECNOLOGÍAS



Centro de Investigación Científica de Yucatán, A.C.

Posgrado en Ciencias Biológicas

Análisis de una biblioteca (*Stenocereus queretaroensis*)
de RNAs no codificantes con interacción a fosfolípidos
(PA y PIP2)

Tesis que presenta

Susana Guerra Gómez

En opción al título de

MAESTRO EN CIENCIAS

(Ciencias Biológicas: Opción en Bioquímica y Biología Molecular)

Mérida, Yucatán, México

2023

CENTRO DE INVESTIGACIÓN CIENTÍFICA DE YUCATÁN, A. C.
POSGRADO EN CIENCIAS BIOLÓGICAS



RECONOCIMIENTO

Por medio de la presente, hago constar que el trabajo de tesis de Susana Guerra Gómez titulado **Análisis de una biblioteca (*Stenocereus queretaroensis*) de RNAs no codificantes con interacción a fosfolípidos (PA y PIP2)**, fue realizado en la Unidad de Bioquímica, en la línea Morfogénesis y Regulación Génica, Laboratorio 23 del Centro de Investigación Científica de Yucatán, A.C. bajo la dirección del Dr. Enrique Castaño de la Serna, dentro de la opción de Opción Bioquímica y Biología Molecular, perteneciente al Programa de Posgrado en Ciencias Biológicas de este Centro.

Atentamente

Dra. Cecilia Hernández Zepeda
Directora de docencia

Mérida, Yucatán, México, a 22 de junio del 2023

DECLARACIÓN DE PROPIEDAD

Declaro que la información contenida en la sección de Materiales y Métodos, los Resultados y Discusión de este documento proviene de las actividades de investigación realizadas durante el período que se me asignó para desarrollar mi trabajo de tesis, en las Unidades y Laboratorios del Centro de Investigación Científica de Yucatán, A.C., y que a razón de lo anterior y en contraprestación de los servicios educativos o de apoyo que me fueron brindados, dicha información, en términos de la Ley Federal del Derecho de Autor y la Ley de la Propiedad Industrial, le pertenece patrimonialmente a dicho Centro de Investigación. Por otra parte, en virtud de lo ya manifestado, reconozco que de igual manera los productos intelectuales o desarrollos tecnológicos que deriven o pudieran derivar de lo correspondiente a dicha información, le pertenecen patrimonialmente al Centro de Investigación Científica de Yucatán, A.C., y en el mismo tenor, reconozco que si derivaren de este trabajo productos intelectuales o desarrollos tecnológicos, en lo especial, estos se regirán en todo caso por lo dispuesto por la Ley Federal del Derecho de Autor y la Ley de la Propiedad Industrial, en el tenor de lo expuesto en la presente Declaración.

Firma _____



Nombre: Susana Guerra Gómez

Este trabajo se llevó a cabo en la Unidad de Bioquímica del Centro de Investigación Científica de Yucatán, A.C., y forma parte del proyecto titulado: Análisis de una biblioteca (*Stenocereus queretaroensis*) de RNAs no codificante con interacción a fosfolípidos (PA y PIP2), bajo la dirección del Dr. Enrique Castaño de la Serna.

AGRADECIMIENTOS

Al Consejo Nacional de Humanidades, Ciencia y Tecnología (CONAHCYT) y a la dirección adjunta de Posgrados y becas, por haberme otorgado la beca nacional con el número de becario gracias a la cual me fue posible completar mis estudios de posgrado.

Al Centro de Investigación Científica de Yucatán (CiCY) por las instalaciones y recursos prestados para realizar mi trabajo de tesis de maestría.

A mi asesor de tesis, el Dr. Enrique Castaño de la Serna por haberme brindado la oportunidad de formar parte de su equipo de trabajo, por su apoyo y asesoramiento a lo largo de este proyecto y en general en mi formación profesional. Por la paciencia y la ayuda a crecer y madurar en mi carrera científica.

A la Dra. Luisa Alhucema López Ochoa, al Dr. Luis Carlos Rodríguez Zapata, al Dr. Alejandro Pereira Santana y Dr. Luis Figueroa Yáñez por ser parte de mi comité durante estos dos años siempre tan propositivo y paciente, así como su disposición a revisar y atender mis dudas para darle forma a este proyecto y por hacerme disfrutar de cada tutorial como una oportunidad para mejorar mi trabajo y contribuir de forma tan grande a mi formación.

A la Ing. Wilma Aracely González Kantun por su apoyo técnico y paciencia durante el tiempo que trabajé en el laboratorio.

A los compañeros del laboratorio 23 de la unidad de Bioquímica y Biología molecular, Laura, Andrea, Angélica, Stefano, Jaime y Jany por su guía, consejos y ayuda durante mi trabajo.

LISTA DE LOS PRODUCTOS GENERADOS

1. Artículo titulado "RNA no codificante en la oncogénesis" sometido en la Revista Ciencia.
2. Artículo titulado "LIRNAseq: A method to discover lipid interacting RNAs by sequencing" publicado en la revista Molecular Biology Reports.

DEDICATORIAS

Les dedico esta tesis con todo el trabajo que ha llevado tanto dentro como sobre todo fuera de ella:

A mi madre Mayeli Gómez y mi padre Oreste Guerra, por su infinita paciencia y comprensión, por su acompañamiento y apoyo incondicional sobre todo en los momentos donde la distancia era irremediable. Ante todo, agradecerles por enseñarme que en la vida hay que aspirar a ser útil y buena persona, por creer en mí y apoyarme incondicionalmente en mis decisiones. A ustedes que son mi ejemplo diario de superación personal y mi motivo para seguir.

A mi familia, principalmente a mis abuelos, Sonia, Mirtha, Onelio y a mis tías Maylen y Mirian, por estar pendientes de mí.

A Yeli y Nayla, por su amistad que superó las fronteras, su apoyo tanto personal como por haber estado siempre para mi familia, por hacerme sentir que la distancia son solo kilómetros cuando de amor se trata.

A Gema y Wendy, por su amistad, ayuda y apoyo personal y profesional durante estos dos años, por las tardes y mañanas de café.

A Verónica, por permanecer junto a mí en todo momento. Por su motivación buena voluntad, perseverancia y positividad ante la vida, que inspira y contagia. Gracias por creer en mí y alegrarte de mis éxitos y disfrutarlos como tuyos. Gracias por ser mi familia y abrirme las puertas de la tuya.

ÍNDICE

INTRODUCCIÓN	2
ANTECEDENTES.....	4
1. El RNA no codificantes (ncRNA).....	4
1.1. Clasificación de los RNA no codificantes (ncRNA)	4
1.2. Características de los RNA largos no codificantes (lncRNAs)	5
1.2.1 Localización de los RNA largos no codificantes (lncRNAs)	6
1.2.2 Biogénesis de RNA largos no codificantes (lncRNAs).....	6
1.2.3 Función y Mecanismo de acción de los RNA largos no codificantes (lncRNAs) en células humanas	8
1.2.5 Interacciones de los RNA largos no codificantes (lncRNAs).....	13
1.2.6 Plantas y RNA largos no codificantes (lncRNAs)	15
1.2.7 Análisis <i>in silico</i> de RNA largos no codificantes (lncRNAs).....	17
1.3 Fosfolípidos.....	20
1.3.1. Fosfoinosítidos	20
1.3.2 Fosforilación de los fosfoinosítidos.	21
1.3.3 Localización y función de los fosfoinosítidos	23
1.3.4 Fosfatidilinositol 4,5-bisfosfato (PtdIns (4,5) P2).....	25
1.3.5 Ácido Fosfatídico	27
1.4 Modelo de planta utilizado <i>Stenocereus queretaroensis</i>	29

JUSTIFICACIÓN	31
HIPÓTESIS.....	31
OBJETIVO GENERAL	31
OBJETIVOS ESPECÍFICOS.....	31
ESTRATEGIA EXPERIMENTAL	32
CAPÍTULO II	33
MATERIALES Y MÉTODOS	33
2.1 Material Biológico	33
2.2 Extracción de RNA total de <i>Stenocereus queretaroensis</i>	33
2.2.1 Ensayo de interacción RNA-fosfolípidos (PA/PIP2) mediante Pull-down	34
2.3. Obtención y caracterización de la librería de RNA-seq de <i>S. queretaroensis</i>	34
2.4. Análisis <i>in silico</i> de las secuencias	35
2.5. Obtención de las secuencias seleccionadas por RT-PCR	36
2.5.1. Extracción de RNA total de <i>Stenocereus queretaroensis</i>	36
2.5.2. Diseño de oligonucleótidos.....	37
2.5.3 Generar cDNA específico.	37
2.5.4 Amplificación de secuencias por RT-PCR	38
2.6 Protocolo de unión de RNA-lípido.	39
2.7. Predicción de estructura secundaria y posibles sitios de unión a PA y PIP2	40
2.7. Análisis filogenético	40
2.8 Motivos y estructura secundaria conservados en los lncRNAs.....	41

Capítulo III	42
Resultados	42
3.1. Construcción de una librería de RNA-seq de <i>Stenocereus queretaroensis</i>.	42
3.1.1 Preparación de muestras (<i>Stenocereus queretaroensis</i>) para la secuenciación.	42
3.2. Obtención de datos de la librería de RNA-seq de genes de <i>S. queretaroensis</i>..	44
3.3. Anotación funcional de genes	45
3.4. Análisis de genes expresados diferencialmente	46
3.4.1 Análisis de diagramas de Venn de DBG.....	47
3.4.2. Análisis de heat maps de los DBG.....	48
3.4.3. Análisis de gráficos de Volcano de DBG.....	49
3.4.4 Análisis de la ontología de genes (GO) y vías KEGG de los DBG	50
3.5 Análisis y selección de lncRNA	53
3.6. Obtención de los lncRNA seleccionadas por RT-PCR.	59
3.7. Análisis del protocolo de unión de lncRNA a PA y PIP2.....	60
3.8. Estructura secundaria y posibles sitios de unión	61
3.9. Análisis filogenético de los lncRNA	67
3.10 Conservación de elementos estructurales en los lncRNAs Sec_1_PA y Sec_2_PIP2	68
CAPÍTULO IV	71
DISCUSIÓN, CONCLUSIONES GENERALES Y PERSPECTIVAS	71
4.1. DISCUSIÓN.....	71

4.2. CONCLUSIONES GENERALES	74
4.3. PERSPECTIVAS	74
Bibliografía	75
ANEXO(S)	87

ÍNDICE DE FIGURAS

Figura 1.1. Clasificación de los ncRNA	5
Figura 1.2. Biogénesis de los lncRNA	7
Figura 1.3. Mecanismos de acción de los lncRNAs.....	9
Figura 1.4. Representación de la estructura secundaria del lncRNA Bvht	11
Figura 1.5. Interacción de los lncRNA.....	13
Figura 1.6. Funciones biológicas de los lncRNAs vegetales	16
Figura 1.7. Estructura de los fosfoinosítidos	21
Figura 1.8. Ruta de fosforilaciones y desfosforilaciones del fosfatidilinositol	22
Figura 1.9. Distribución celular de los fosfoinosítidos presentes en varios organelos animales	23
Figura 1.10. Estructura química del fosfatidilinositol 4,5-bifosfato (PtdIns-4,5-P ₂)	25
Figura 1.11. Islotes de lípidos nucleares (NLI)	26
Figura 1.12. Representación esquemática del ácido fosfático (PA).....	25
Figura 1.13. Representación esquemática de las vías bioquímicas para la síntesis y el metabolismo del PA	28
Figura 1.14. Morfología vegetativa de <i>Stenocereus Queretaroensis</i>	30
Figura 3.1. Cultivo de células rojas de <i>Stenocereus Queretaroensis</i>	42
Figura 3.2. Extracción de RNA de <i>Stenocereus Queretaroensis</i>	43
Figura 3.3. Representación esquemática del flujo de trabajo desde la preparación de las muestras hasta la obtención de datos	43
Figura 3.4. Resultados de la evaluación de Busco de las transcripciones empalmadas	45
Figura 3.5. Diagramas de Venn de los DBG	48
Figura 3.6. Heat maps de los DBG	49
Figura 3.7. Diagramas de volcán de los DBG	50
Figura 3.8. Diagramas de GO muestran los genes enriquecidos diferencialmente para las perlas de sefarosa de control y las recubiertas de PIP2	51

Figura 3.9. Diagramas de GO muestran los genes enriquecidos diferencialmente para las perlas de sefarosa de control y las recubiertas de PA	52
Figura 3.10. Diagramas de vías KEGC muestran los genes enriquecidos diferencialmente	53
Figura 3.11. Representación del flujo de trabajo para la selección de los lncRNA	54
Figura 3.12. Diagramas de Venn de los 44 lncRNA	55
Figura 3.13. Heat maps generado con los datos de los 44 lncRNA	56
Figura 3.14. Diagramas de volcán de los 44 lncRNAs	57
Figura 3.15. Resultados obtenidos del cálculo del potencial de codificación	58
Figura 3.16. Amplificación de las secuencias por RT-PCR	60
Figura 3.17. Interacción entre los lncRNAs Sec_1_PA y Sec_3_PIP2 con los lípidos PA y PIP2 mediante ensayos de pull-down <i>in vitro</i>	61
Figura 3.18. Representación de la estructura secundaria y la MFE de la Sec_1_PA	62
Figura 3.19. Representación de la estructura secundaria y la MFE de la Sec_3_PIP2	62
Figura 3.20. Estructuras secundarias de los sitios de unión de lncRNA a lípidos reportados en investigaciones anteriores	63
Figura 3.21. Alineamiento de Sec_1_PA con las secuencias de los sitios de unión de lncRNA a lípidos reportados en investigaciones anteriores, y la estructura secundaria de los posibles sitios de unión de Sec_1_PA a lípidos	64
Figura 3.22. Alineamiento de Sec_3_PIP2 con las secuencias de los sitios de unión de lncRNA a lípidos reportados en investigaciones anteriores, y la estructura secundaria de los posibles sitios de unión de Sec_3_PIP2 a lípidos	64
Figura 3.23. Representación de los posibles sitios de unión a PA y PIP2 en las secuencias Sec_1_PA y Sec_3_PIP2	66
Figura 3.24. Localización según un predictor de los lncRNA: Sec_1_PA y Sec_3_PIP2.....	66
Figura 3.25. Árbol filogenético de la secuencia Sec_1_PA.	67
Figura 3.26. Árbol filogenético de la secuencia Sec_3_PIP2.....	68
Figura 3.27. Resultados del alineamiento entre el lncRNA Sec_1_PA y los IEs de algunos ncRNA como BCAR4, Bvht y Xist	69
Figura 3.28. Resultados del alineamiento entre el lncRNA Sec_1_PA y los IEs de algunos ncRNA como BCAR4, Bvht y Xist	70

Índice de Cuadros

Cuadro 2.1. Características de los primers específicos de las secuencias seleccionadas	37
Cuadro 3.1. Datos crudos y limpios de la secuenciación.....	44
Cuadro 3.2. Resumen del número de transcritos y unigenes en diferentes intervalos de longitud	45
Cuadro 3.3. Representación funcional de los genes anotados con éxito por cada base de datos	46
Cuadro 3.4. Visualización de los resultados estadísticos de los genes diferenciales.	47
Cuadro 3.5. Resume el ID, largo del gen y ruta biológica GO en que pueden estar involucradas las secuencia Sec_1_PA (Cluster_18316.12632), Sec_2_PA(Cluster_27261.1), Sec_3_PIP2 (Cluster_2025.8) y Sec_4_PIP2(Cluster_30630.0).	58

ABREVIATURAS

ceRNA (RNA endógeno competidor)

cisRNA (RNA regulador de Cis)

DNA (Ácido desoxirribonucleico)

IEs (elementos interactores)

lincRNAs (RNA no codificante intergénico largo)

lncRNA (Ácido ribonucleico no codificante largo)

miRNA (micro-RNA)

mRNA (Ácido ribonucleico mensajero)

ncRNAs (Ácidos ribonucleicos no codificante)

PIP2 (Fosfatidilinositol 4,5-bisfosfato)

PIP3 (fosfatidilinositol (3,4,5)-trisfosfato)

PA (Ácido fosfatídico)

piRNA (RNAs asociados a Piwi)

rRNA (RNA ribosomal)

tRNA (RNA de transferencia)

SEs (elementos estructurales)

siRNA (pequeños RNA de interferencia)

snRNA (RNA nuclear pequeño)

sRNA (RNA pequeño bacteriano)

tiRNA (RNA pequeño derivado de tRNA)

RESUMEN

Los ncRNA son moléculas de RNA que no codifican para proteínas, presentan una alta diversidad en cuanto a su longitud, estructura y función celular, lo que ha llevado que sean clasificados en dos grandes categorías: estructurales y reguladores según sus funciones, a su vez estos últimos se dividen en sRNA (menos de 200 nt) y lncRNA (más de 200 nt) en cuanto a su longitud. Los lncRNAs se pueden localizar tanto en el núcleo como en el citoplasma, donde desempeñan distintas funciones y/o actúan como mediadores en algunos procesos celulares como la transcripción, organización espacial del compartimento nuclear mediante el andamiaje, transducción de señales, control postranscripcional. Su participación en estas funciones puede ser de manera directa o indirecta al interactuar con diferentes socios moleculares como DNA, RNA, péptidos, microRNA, proteínas y componentes lipídicos como PIP2 y PA moléculas que participan en la señalización celular.

Los RNA no codificantes largos (lncRNA) son un campo de investigación en rápida expansión, cada año se identifican nuevos transcritos, pero el estudio de los lncRNA vegetales es incipiente incluso menos el de los que interactúan con fosfolípidos. En este trabajo nos proponemos el estudio de nuevos lncRNAs vegetales a partir de cultivos celulares de células rojas de la cactácea, *Stenocereus queretaroensis*, y su posible interacción con los componentes lipídicos PA y PIP2.

ABSTRACT

The ncRNAs are RNA molecules that do not code for proteins, they present a huge diversity in terms of their length, structure and cellular function, which has led them to be classified into two main categories: structural and regulatory according to their functions, the latter are divided into sRNA (less than 200 nt) and lncRNA (more than 200 nt) in terms of their length. The lncRNAs can be located both in the nucleus and in the cytoplasm, where they perform different functions and/or act as mediators in some cellular processes such as transcription, spatial organization of the nuclear compartment by scaffolding, signal transduction and post-transcriptional control. Their participation in these functions can be directly or indirectly by interacting with different molecular partners such as DNA, RNA, peptides, microRNA, proteins and lipid components such as PIP2 and PA molecules involved in cell signaling.

Long non-coding RNAs (lncRNAs) are an expanding field of research, new transcripts are identified every year, but the study of plant lncRNAs is still emerging, even less their interaction with phospholipids. In this work we propose the study of new plant lncRNAs from cell cultures of red cells of the cactus, *Stenocereus queretaroensis*, and their possible interaction with the lipid components PA and PIP2.

INTRODUCCIÓN

Los ncRNA son moléculas de RNA que no codifican para proteínas, esto no significa que no contengan información o alguna función. Los datos obtenidos en el proyecto ENCODE, estimaron que entre el 70%-80% del genoma humano se transcribe, pero sólo el 2% de este codifica proteínas (ENCODE Project Consortium *et al.*, 2007). Los ncRNA son heterogéneos en cuanto a su longitud, estructura y función celular, siendo clasificados en dos grandes categorías: estructurales y reguladores según sus funciones, a su vez estos últimos se dividen en sRNA (menos de 200 nt) y lncRNA (más de 200 nt) en cuanto a su longitud.

Los lncRNAs se pueden localizar tanto en el núcleo (Bridges *et al.*, 2021; Fazal *et al.*, 2019; Kaewsapsak *et al.*, 2017) como en el citoplasma (Bridges *et al.*, 2021; Benoit Bouvrette *et al.*, 2018), donde desempeñan distintas funciones y/o actúan como mediadores en algunos procesos celulares como la transcripción, organización espacial del compartimento nuclear mediante el andamiaje, transducción de señales, control postranscripcional (Bridges *et al.*, 2021). Su participación en estas funciones puede ser de manera directa o indirecta al interactuar con diferentes socios moleculares como DNA, RNA, péptidos, microRNA, proteínas y componentes lipídicos como PIP2 y PA moléculas que participan en la señalización celular (Yamazaki *et al.*, 2018).

Las investigaciones sobre la interacción entre los lncRNA y los fosfolípidos han sido creciendo en los últimos años, arrojando importantes resultados como fue el lncRNA LINK-A descubierto por Lin *et al.*, (2017) que interactuaba con PIP3, posteriormente en el 2021 realizan un estudio donde un otro lncRNA llamado SNHGA interactuaba con PA. Los fosfolípidos PIP2 y PA son, ambas investigaciones utilizando modelos de referencia.

El estudio de los lncRNAs vegetales es incipiente, y se han ido identificando gradualmente en estudios de especies como *Arabidopsis*, *Oryza sativa*, *Medicago sativa*, *Zea mays*, *Solanum lycopersicum*, *Gossypium barbadense*, *Gossypium hirsutum* y otras plantas herbáceas, revelado que los lncRNAs están ampliamente implicados en el crecimiento y desarrollo y en procesos de respuesta al estrés mediando la transmisión y expresión de la información genética (Rai *et al.*, 2019)

Los lncRNA son un campo de investigación en rápida expansión, cada año se identifican nuevos transcritos. Sin embargo, hasta ahora sólo se ha caracterizado funcionalmente un pequeño subconjunto de estos. Para ayudar a investigar los mecanismos de acción por los que actúan los

nuevos lncRNAs, así como su localización subcelular, conformación estructural y socios de interacción; el análisis *in silico* resulta útil para estudiar un lncRNA recién descubierto, ya que los resultados obtenidos *in silico* son la guía para orientar la investigación experimental (Pinkney *et al.*, 2020)

En este trabajo nos proponemos el estudio de nuevos lncRNAs vegetales y su posible interacción con los componentes lipídicos PA y PIP2. El modelo que utilizaremos es la cactácea, *Stenocereus queretaroensis*, comúnmente llamada pitaya salvaje, es una de las ocho cactáceas columnares congéneres nativas de México. Se conoce que sus frutos son una fuente de betalaínas, pigmentos enormes potencial debido a su alta hidrosolubilidad, su estabilidad en un amplio rango de condiciones físico-químicas y sus propiedades bioactivas como antioxidantes.

CAPÍTULO I

ANTECEDENTES

1. El RNA no codificantes (ncRNA)

El RNA se consideraba el intermediario imprescindible para la obtención de proteínas; sin embargo, el empleo de la secuenciación masiva cambió drásticamente este panorama, en el 2007 los datos obtenidos durante el proyecto ENCODE estimaron que el 70-80% del genoma humano se transcribe a RNA, pero que sólo el 2% de estos codifican para proteínas (Djebali *et al.*, 2012; Maher., 2012), surgiendo de esta manera un término poco conocido para ese entonces los ncRNA.

Los ncRNA son aquellas moléculas de RNA que no codifican para proteínas, pero esto no significa que no contengan información o realicen una función. Debido a esto se les llama la "materia oscura" del genoma, (Bridges *et al.*, 2021), aunque en los últimos años el número de investigaciones y publicaciones que demuestran el importante papel que tienen en diferentes procesos celulares incluso llegando a jugar un rol vital en el desarrollo y progresión de enfermedades humanas como en plantas.

1.1. Clasificación de los RNA no codificantes (ncRNA)

Los ncRNA son muy heterogéneos en cuanto a su longitud, conformación y función celular, la diversidad que presentan ha hecho que sean clasificados en dos grandes categorías, no codificantes estructurales y reguladores. Los ncRNA estructurales comprenden los RNA ribosómico (rRNA) y los RNA de transferencia (tRNA). Los ncRNA reguladores se dividen a su vez en tres clases en función de su tamaño: ncRNA pequeños (sRNA) (20-50 nt) dentro de estos se encuentran los microRNA (miRNA), los pequeños RNA de interferencia (siRNA), los RNA que interactúa con Piwi (PiRNA), los RNA pequeño específico de telómeros (TelsRNA), los RNA pequeño de interacción asociado a repeticiones de centrómero (CrsiRNA); la segunda clase son los ncRNA medianos (50-200 nt) que incluyen a los pequeños RNA nucleolares (SnoRNA), los transcritos ascendentes del promotor (PROMPTs), los RNA de iniciación de la transcripción (tiRNA), los RNA nucleares pequeños (SnRNA), el RNA citoplasmático pequeño (ScrRNA); por último los lncRNA (<200 nt). Estos a su vez se agrupan en clases y subclases según: su mecanismo de acción donde se encuentran los RNA reguladores en cis y en trans (CisRNA y TransRNA), los RNA endógenos competidores (CeRNA); biogénesis y estructura que a su vez encontramos los lncRNA lineales y circulares (Figura1.1.) (Dahariya *et al.*, 2019).

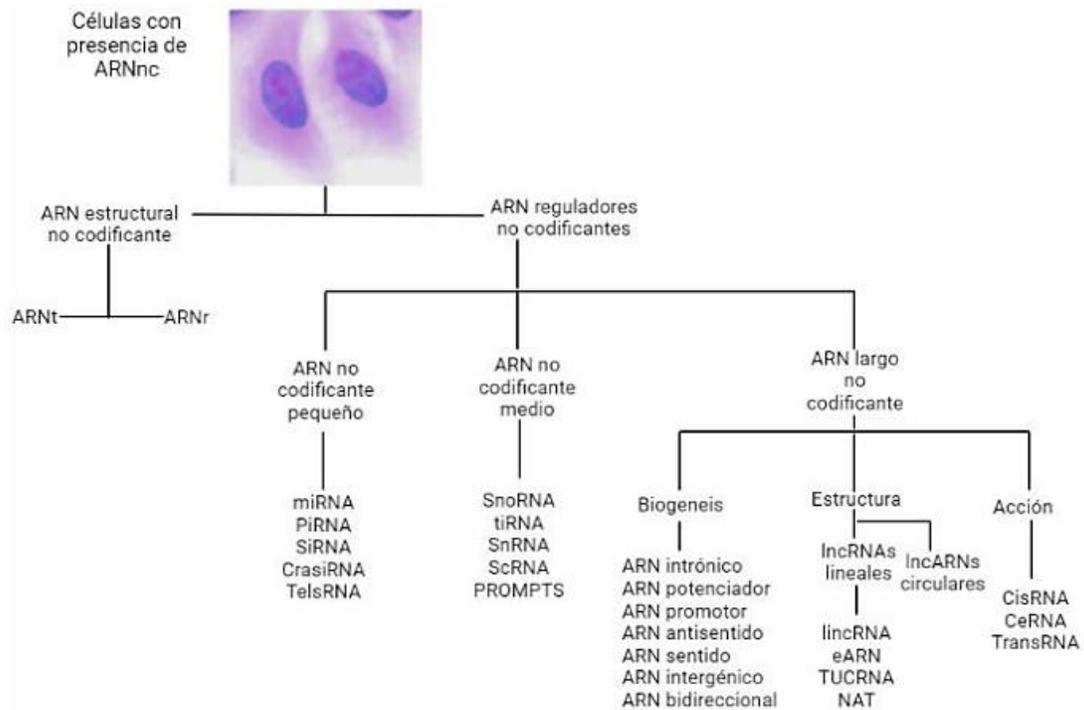


Figura 1.1. Clasificación de los ncRNA (Modificado sobre la base de Dahariya *et al.*, 2019)

1.2. Características de los RNA largos no codificantes (lncRNAs)

Los lncRNA constituyen una parte importante del genoma no codificante, su longitud es variable y se encuentran clasificados en cuanto a su biogénesis, estructura y mecanismo de acción en clases y subclases. Son sintetizados generalmente por la RNA polimerasa II (Pol II) bajo el control de los activadores transcripcionales del complejo SWItch/Sacarosa no fermentable (SWI/SNF) y otros son transcritos por la RNA polimerasa III (Pol III). Se pueden localizar en varios compartimentos y/o subcompartimentos celulares como el núcleo ejemplo de ello son los ncRNA Xist (ncRNA en el cromosoma X) y lncRNA Uchl1 (lncRNA enriquecido nuclearmente que es antisentido de la ubiquitina carboxi-terminal hidrolasa L1 (Uchl1) de ratón); y/o citoplasma. Comparten similitudes con el mRNA como la presencia de una caperuza 5' situada en el extremo 5', pueden estar empalmados y poliadenilados, aunque en comparación con el mRNA muestran una modesta conservación de su secuencia, la razón de esto puede ser que están libres de restricciones de codones (Dahariya *et al.*, 2019). Los lncRNAs se pueden originar a partir de sus promotores propios o compartidos con genes codificantes o no codificantes de transcripción divergente o de secuencias potenciadoras. Estas moléculas no codificantes constituyen en sí mismos la unidad funcional a diferencia de los mRNA que necesitan traducirse en proteínas que desempeñen funciones celulares específicas (Wang *et al.*, 2018).

1.2.1 Localización de los RNA largos no codificantes (lncRNAs)

Los primeros estudios sobre la localización de los lncRNAs demostraron su mayor abundancia en el núcleo (Bridges *et al.*, 2021; Fazal *et al.*, 2019; Kaewsapsak *et al.*, 2017) aunque también se encuentran presentes en el citoplasma (Bridges *et al.*, 2021; Benoit Bouvrette *et al.*, 2018) incluso con mayor estabilidad que sus homólogos nucleares, dicha inestabilidad nuclear refleja su papel en la regulación de la expresión génica, facilitando el ajuste dinámico de sus niveles en respuesta a los estímulos (Bridges *et al.*, 2021).

Las funciones de los lncRNAs en el núcleo son diferentes a las que desempeñan en el citoplasma, en el núcleo funcionan para modular programas transcripcionales a través de las interacciones de la cromatina y remodelación de la misma (Bridges *et al.*, 2021) estableciendo la organización espacial del compartimento nuclear mediante el andamiaje. En el citoplasma median las vías de transducción de señales y el control postranscripcional de la expresión génica. El microambiente subcelular permite que un mismo lncRNA tenga funciones distintas en dependencia de que compartimento celular se encuentre, debido a las interacciones que se establecen con los diferentes socios moleculares funcionales. Un ejemplo de esto es el lncRNA PYCARD-AS1 (lncRNA antisentido del gen proapoptótico PYCARD2) descrito por Miao *et al.*, (2019) presente tanto en el núcleo como en el citoplasma, pero con funciones diferentes en cada compartimento. El lncRNA antisentido del gen proapoptótico PYCARD, denominado PYCARD-AS1, presenta una distribución dual, nuclear y citoplasmática, y es necesario para el silenciamiento de PYCARD en células de cáncer de mama. En el núcleo recluta a la enzima DNA metiltransferasa 1 (DNMT1) y la histona metiltransferasa G9a al promotor de PYCARD para facilitar la metilación del DNA y la modificación de la dimetilación de la lisina 9 de la histona H3 (H3K9me), mientras que en el citoplasma interactúa con el mRNA para inhibir el ensamblaje del ribosoma y la traducción de la proteína PYCARD (también conocida como ASC) (Miao *et al.*, 2019)

1.2.2 Biogénesis de RNA largos no codificantes (lncRNAs)

La biogénesis de los lncRNAs es controlada por estímulos específicos y depende del tipo de célula y su estadio (Akerman *et al.*, 2017). Algunos tipos de lncRNAs en genomas eucarióticos se transcriben a partir de varios elementos de DNA como potenciadores, promotores y regiones intergénicas (Wu *et al.*, 2020). Los lncRNAs se encuentran intercalados en el genoma con varias ubicaciones posibles en función de su origen transcripcional (Figura 1.2) como los lncRNAs

asociados al potenciador o promotor, lncRNAs intrónicos/exónicos/sentidos aquellos que se transcriben en la misma dirección que los genes codificantes, lncRNAs intergénico largo, lncRNAs bidireccional/divergente, lncRNAs antisentido (Wu *et al.*, 2020; Dahariya *et al.*, 2019). Son diferentes los mecanismos implicados en la biogénesis de lncRNA como el corte por la ribonucleasa P (RNasa P) para generar extremos maduros, la formación de complejos de pequeños RNA nucleolares (snoRNA) y de pequeñas partículas de ribonucleoproteínas nucleolares (snoRNP) en sus extremos, y de estructuras circulares.

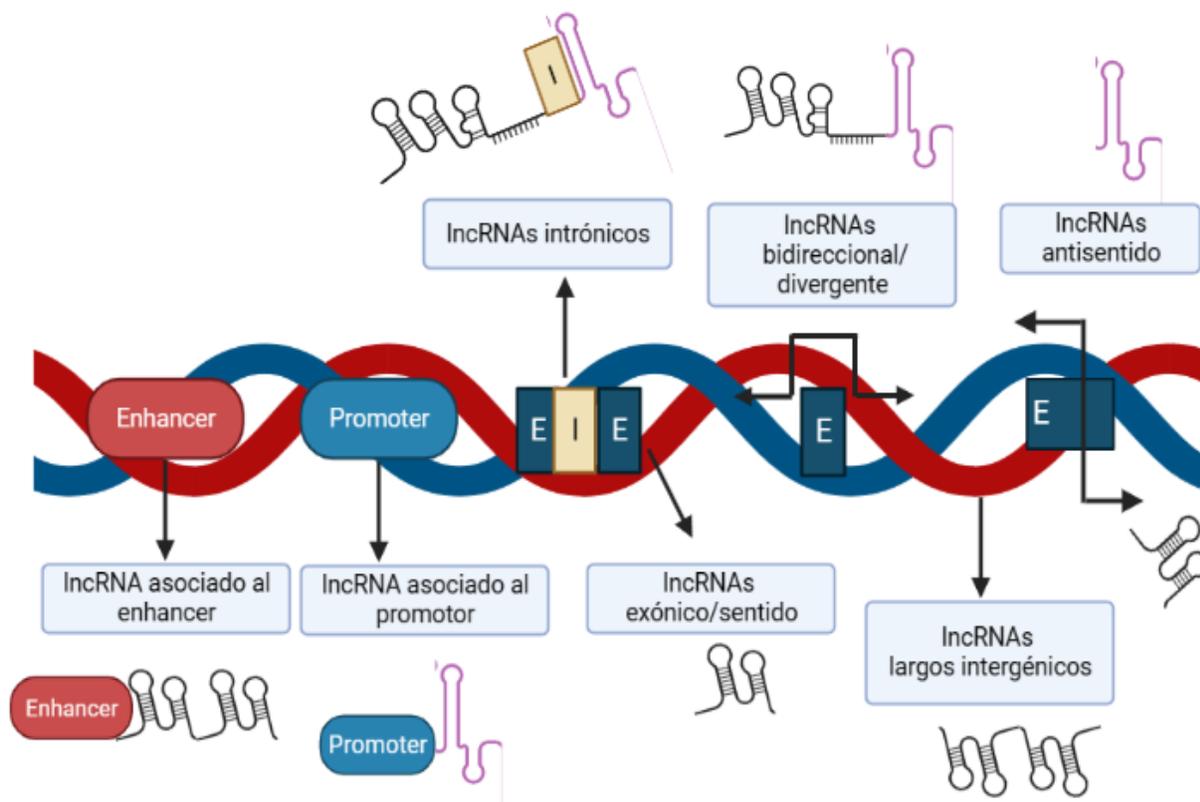


Figura 1.2 Biogénesis de los lncRNAs. La figura resume la diversa gama de lncRNAs en función de su origen transcripcional: transcritos naturales en antisentido (NAT) completos o parciales, genes codificantes, **región intergénica** y dentro de intrones (Modificado sobre la base de Dahariya *et al.*, 2019).

El empleo de técnicas de RNAi (interferencia de RNA) ha permitido identificar la presencia de estructuras subnucleares conocidas como "paraspeckles" que se encuentran alrededor de los lncRNAs y son importantes para su formación durante su biogénesis (Dahariya *et al.*, 2019). En general, los mecanismos de síntesis y regulación de los diferentes lncRNAs no están completamente descubiertos, sin embargo, en los próximos años con la ayuda de diversas técnicas, como ChIRP-Seq (aislamiento de cromatina por purificación de RNA),

inmunoprecipitación por entrecruzamiento (CLIP), la técnica de edición que permite introducir modificaciones genéticas “repeticiones palindrómicas cortas, agrupadas y regularmente interespaciadas” (CRISPR) (Dahariya *et al.*, 2019; Yoon y Gorospe, 2016; Li *et al.*, 2014) entre otras, ayudarán a profundizar nuestra comprensión de la biogénesis y funciones de los lncRNAs.

1.2.3 Función y Mecanismo de acción de los RNA largos no codificantes (lncRNAs) en células humanas y animales

La implicación funcional de los lncRNAs en diferentes tipos de células y enfermedades (Li *et al.*, 2014a; Neguembor *et al.*, 2014) ha sido demostrada en varias investigaciones. Un gran número de estas moléculas están presentes en numerosas enfermedades, como la diabetes, la artritis reumatoide, los trastornos neurodegenerativos, la enfermedad aterosclerótica de las arterias coronarias, y varios tipos de cáncer (Wu *et al.*, 2014; Li *et al.*, 2013); quedando evidenciado su rol tanto en el estado normal como en el agresivo de la enfermedad. También participan en la regulación de procesos biológicos como la diferenciación celular, la determinación del linaje celular, la organogénesis y la homeostasis tisular junto con la inducción de la pluripotencia celular, la inactivación del cromosoma X y la impronta genética (Kazemzadeh *et al.*, 2015).

Los lncRNAs participan tanto en la activación como en la inhibición de la expresión génica a través de cuatro mecanismos moleculares: señal, señuelo, guía y andamio (Quinn y Chang, 2016) (Figura.1.3). Pueden actuar como señuelo al unirse a factores de transcripción y/o proteínas, de este modo, apartan a los factores proteicos de su acción sobre el DNA diana. También funcionan como esponjas moleculares para varias proteínas fijadoras del retinol (RBP) (He *et al.*, 2016) o microRNAs (miRNAs) (Liu *et al.*, 2017^a; Wang *et al.*, 2016; Xiao *et al.*, 2015), ejemplo de esto son los ncRNA: pequeños RNA nucleolares (snoRNAs, ncRNA asociado a metástasis del adenocarcinoma de pulmón (MALAT) y el lncRNA de la transcripción 1 del ensamblaje nuclear de paraspeckles (NEAT1), importante regulador de la respuesta en células humanas al estrés. Algunos lncRNAs median cambios en el transcriptoma de las células animales actuando negativamente sobre moléculas efectoras como factores de transcripción, receptores celulares, factores de crecimiento y reguladores de empalme (Fu, 2014; Cooper *et al.*, 2014) un ejemplo de ello es el lncRNA Gas5 (lncRNA específico de detención del crecimiento 5) que funciona como supresor tumoral en células humanas dirigiéndose a miRNAs promotores del cáncer como miR21 (microRNA de mamíferos), en la leucemia de células B y células T.

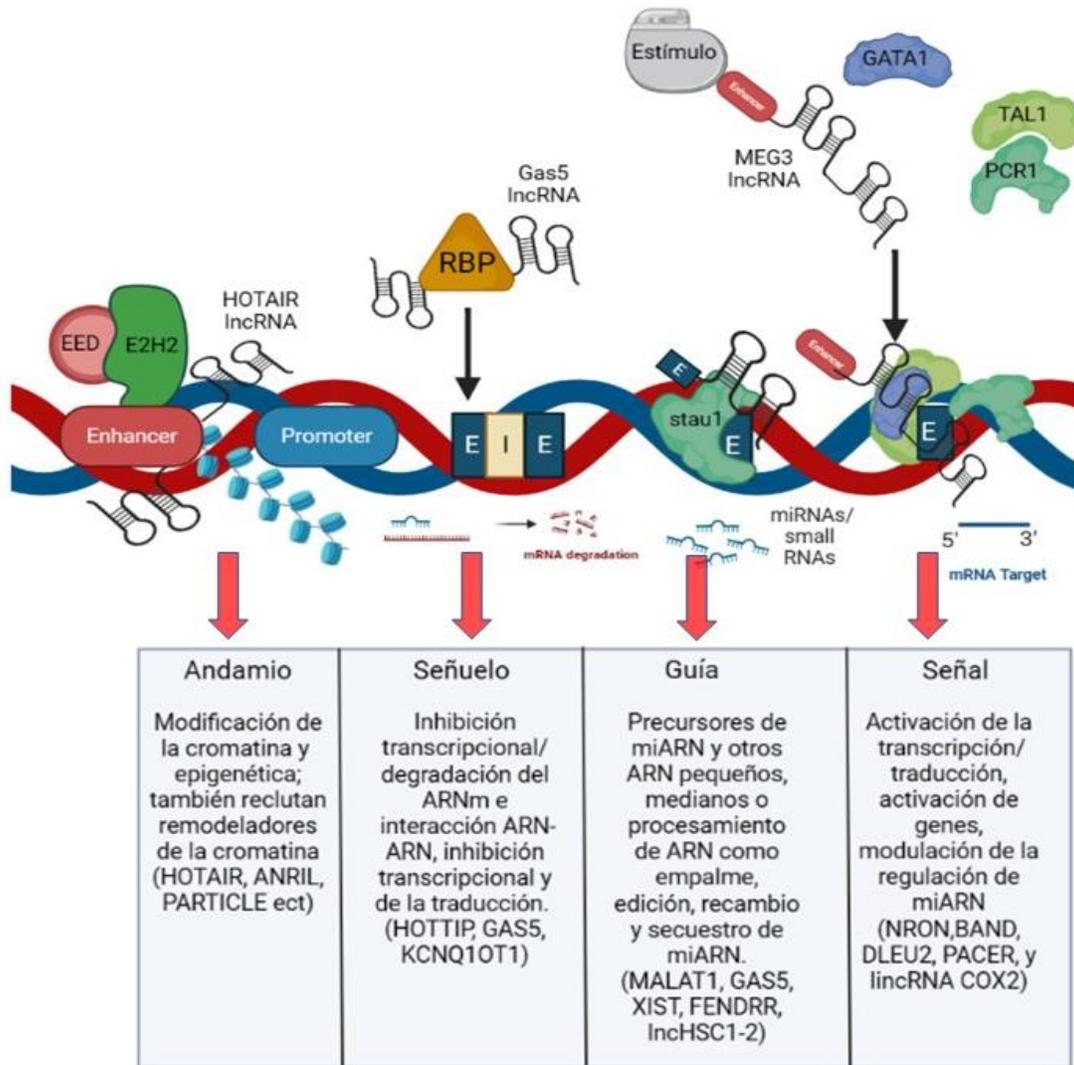


Figura 1.3. Mecanismos de acción de los lncRNAs (Modificado sobre la base de Dahariya *et al.*, 2019).

Otros lncRNAs actúan como reguladores cruciales del epigenoma como es el caso de lncRNA HOTAIR (lncRNA intergénico antisentido de la transcripción HOX) y linc-p21 (transcrito de lncRNA descendente de p53). Los lncRNAs, lncHSC1 y lncHSC2, están implicados en la autorrenovación y en las decisiones tempranas de la línea específicamente para la diferenciación mieloide en células hematopoyéticas humanas. Los lncRNAs HOTAIR y lncRNAs ANRIL (ncRNA antisentido en el locus INK4) han sido estudiados en tumores de vejiga no musculares invasivos y en la patogénesis de la esclerosis múltiple, donde sirven como andamios al modular directamente la actividad de factores de transcripción y múltiples moléculas efectoras para su ensamblaje (Quinn y Chang, 2016).

Algunos estudios demuestran que muchos lncRNAs están implicados en la remodelación de la cromatina en células humanas, ejemplos de ello es el caso del lncRNA HOTAIR que se asocia con complejos de remodelación de la cromatina al unirse a PRC2 (complejo represor de poli-comb clase 2.) y al complejo lisina desmetilasa específica 1 lisina que regula la expresión génica como parte del complejo CoREST junto con el correpresor de REST (LSD1/CoREST/REST) tanto en el dominio 5' y 3', respectivamente (Bhan y Mandal, 2015) y los lncRNAs como ANRIL o CDKNB2-AS1 (ncRNA antisentido1 inhibidor de la cinasa dependiente de ciclina 2B) que interactúan estrechamente con múltiples componentes del complejo represivo poli-comb tanto de clase 1 (PCR1) como 2 (PCR2) formando la heterocromatina alrededor del locus de la familia de inhibidores de la cinasa dependiente de ciclina (CKI) (INK4) resultando en el silenciamiento del gen INK4 a través de la metilación de la cromatina (Aguilo *et al.*, 2011). También otros ncRNAs como XIST (Corbel *et al.*, 2013), DACOR1 (lncRNA asociado al cáncer de colon), HOTAIR, Fendrr (ncRNA regulador del desarrollo fetal letal), LUNAR1 (ncRNA asociado a la leucemia 1) y PARTICLE (lncRNA que sirve de puente entre la histona y la metilación del ADN) a través de la metilación de histonas median la expresión génica de remodelación de la cromatina (Schmitz *et al.*, 2016).

Muchos son los lncRNAs que interactúan con factores epigenéticos para regular la expresión génica y modular la arquitectura nuclear (Hanly *et al.*, 2018; Heo *et al.*, 2013; Mercer y Mattick, 2013). Los lncRNAs también ayudan en los procesos de desarrollo mediante la correulación de su propia expresión génica y la de los genes codificantes de proteínas vecinos (Quinn y Chang, 2016; Luo *et al.*, 2015). Algunos lncRNAs participan en la transducción de señales como regulador para iniciar, alargar o terminar acciones combinatorias de factores de transcripción (Long *et al.*, 2017). Como ha quedado evidenciado son muchos los ejemplos sobre las funciones en las que se encuentran implicados diferentes tipos de lncRNAs actuando a través de distintos mecanismos, aunque aún queda mucho por investigar sobre las funciones y modos de acción de estas moléculas no codificantes.

1.2.4 Estructura de los RNA largos no codificantes (lncRNAs)

Los lncRNAs al igual que muchas proteínas de gran tamaño se organizan de manera espacial y funcional a través de una arquitectura multidominio, por lo que puede asegurarse que su estructura está estrechamente vinculada con la función que realizan. En su estructura se pueden distinguir dos clases de elementos: "elementos interactores" (IEs) necesarios para la interacción física y directa con diferentes socios celulares y en segundo lugar los "elementos estructurales"

(SEs), que dirigen sus interacciones funcionales con otras moléculas mediante la aparición de estructuras secundarias y/o terciarias; dichos elementos son los responsables del modo en que actúan los lncRNAs: *cis* (dentro de la misma molécula de RNA) o *trans* (con otras moléculas: RNA, DNA, proteínas, lípidos). Un transcrito no codificante puede tener uno o múltiples SEs o IEs, los cuales se pueden localizar por separado o encontrarse superpuestos dentro de la secuencia de RNA (Fabbri *et al.*, 2019; Helder *et al.*, 2016).

Los SEs presentes en los lncRNAs son más diversos y menos conocidos en comparación con los de los miRNA, principalmente porque las estructuras de los lncRNAs son más grandes y versátiles debido a que el número de emparejamiento entre los nucleótidos que lo conforman tanto locales como de largo alcance llevan a plegar la cadena en formas dinámicas y flexibles; dando lugar a diferentes motivos estructurales como: hélices, bucles terminales o internos, uniones, pseudonudo, terciaria o G-Cuádruplex. Un ejemplo de IEs que se observa en la **Figura 1.4** es el motivo denominado AGIL rico en G presente en el lncRNA Bvht (lncARN también conocido como Braveheart) que interviene en la interacción específica de este con el factor de transcripción de la proteína de unión a ácidos nucleicos con dedos de zinc de tipo CCHC (CNBP/ZNF9); la supresión de dicho motivo se comprobó que afecta drásticamente a la diferenciación de los cardiomiocitos (Fabbri *et al.*, 2019; Helder *et al.*, 2016).

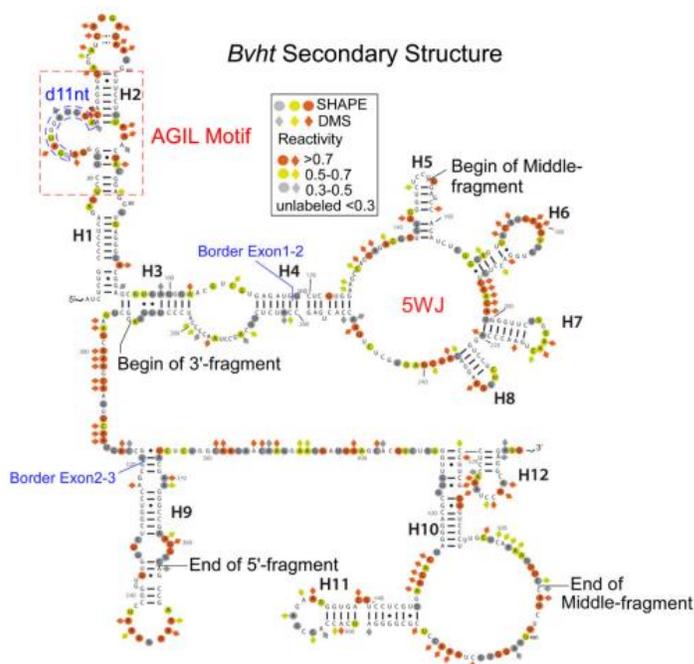


Figura 1.4. Representación de la estructura secundaria del lncRNA Bvht. Se puede observar el motivo denominado AGIL rico en G presente que interviene en la interacción específica de este con el factor de

transcripción de la proteína de unión a ácidos nucleicos con dedos de zinc de tipo CCHC (CNBP/ZNF9) (Tomado sobre la base de Xue *et al.*, 2016).

Los IEs dentro de los lncRNAs están determinados por la evolución, algunos se conservan de una especie a otra como es el caso del IE en el ultraconservado ncRNA uc.63 (ncRNA muy conservado entre los genomas del ratón, la rata y el ser humano, lo que podría indicar un papel definitivo de estos elementos en la salud y la enfermedad) en el oncogén miR-155 que promueve la supervivencia de las células cancerosas y la angiogénesis tumoral al regular directamente a la baja la expresión de genes diana (Calin *et al.*, 2007) o el sitio de unión al precursor de microRNAs miR-7 dentro del lncRNA OIP5-AS1 (lncRNA asociado a tumores que contribuye a intrincados mecanismos celulares durante la evolución de tumores malignos) (Ulitsky *et al.*, 2011). Mientras que otros son específicos para secuencia de algunas especies como es el caso de la secuencia IE situada en el motivo pyknon del lncRNA N-BLR (lncRNA específico de los primates que modula la transición epitelio-mesénquima, facilita la migración celular y aumenta la invasión del cáncer colorrectal) que interactúa con miembros de la familia miR-200, molécula corta de ARN que regula los niveles de expresión de otros genes uniéndose y escindiendo ARNm o inhibiendo la traducción (Rigoutsos *et al.*, 2017).

Los IEs presentes en los lncRNA podrían ser los responsables de la interacción de estos con diferentes socios moleculares como proteínas y lípidos, confiriendo especificidad. Se sabe que muchos lncRNAs interactúan con proteínas localizadas en el mismo compartimento subcelular (Chen *et al.*, 2022), donde un mismo lncRNA puede interactuar con varias proteínas en diferentes tipos de células ejemplo de esto es la interacción resultante entre el complejo PRC2 específicamente la enzima EZH2 (potenciador de zeste homólogo 2) una histona metiltransferasa que puede cambiar la expresión de genes diana aguas abajo, y varios lncRNAs (Cifuentes-Rojas *et al.*, 2014) y otras más específicas como la interacción del lncRNA CCAT2 (lncRNA específico del adenocarcinoma de pulmón) solo con el complejo de factores de clivaje (CFIm) (Lin *et al.*, 2017).

También los lncRNA pueden interactuar de manera específica con lípidos, un ejemplo de esto es el estudio llevado a cabo por Lin *et al.*, (2017), donde demuestran la interacción directa de un lncRNA denominado lincRNA (LINK-A) con el fosfatidilinositol (3,4,5)-trisfosfato (PIP3) a través del dominio de homología de pleckstrin que participa en la vía de señalización de la enzima serina/treonina quinasa (AKT). Las deleciones genómicas de este motivo sensibilizan a las células de cáncer de mama a los inhibidores de la enzima AKT, mientras que su hiperactivación dependiente de LINK-A favoreció la resistencia a los inhibidores de dicha enzima y un aumento

en la tumorigénesis (Lin *et al.*, 2017). Es posible que esta IE sea de hecho una IE-SE combinada, ya que este elemento podría tener también una estructura definida que facilita la interacción.

1.2.5 Interacciones de los RNA largos no codificantes (lncRNAs)

Los lncRNAs pueden modular la transcripción, las modificaciones epigenéticas, la estabilidad de las interacciones proteínas con RNAs, la traducción y las modificaciones postraduccionales al interactuar con el DNA, RNAs, mRNA, miRNA, péptidos y/o proteínas (**Figura 1.5.**) (Kazimierczyk *et al.*, 2020). También se ha demostrado recientemente que interactúan directamente con receptores de señalización como receptores acoplados a proteína G, receptores ligados a enzimas e intracelulares que se encuentran en la superficie celular o en el interior de la célula y que se unen a moléculas de señalización para iniciar una respuesta celular específica (Schmidt *et al.*, 2020).

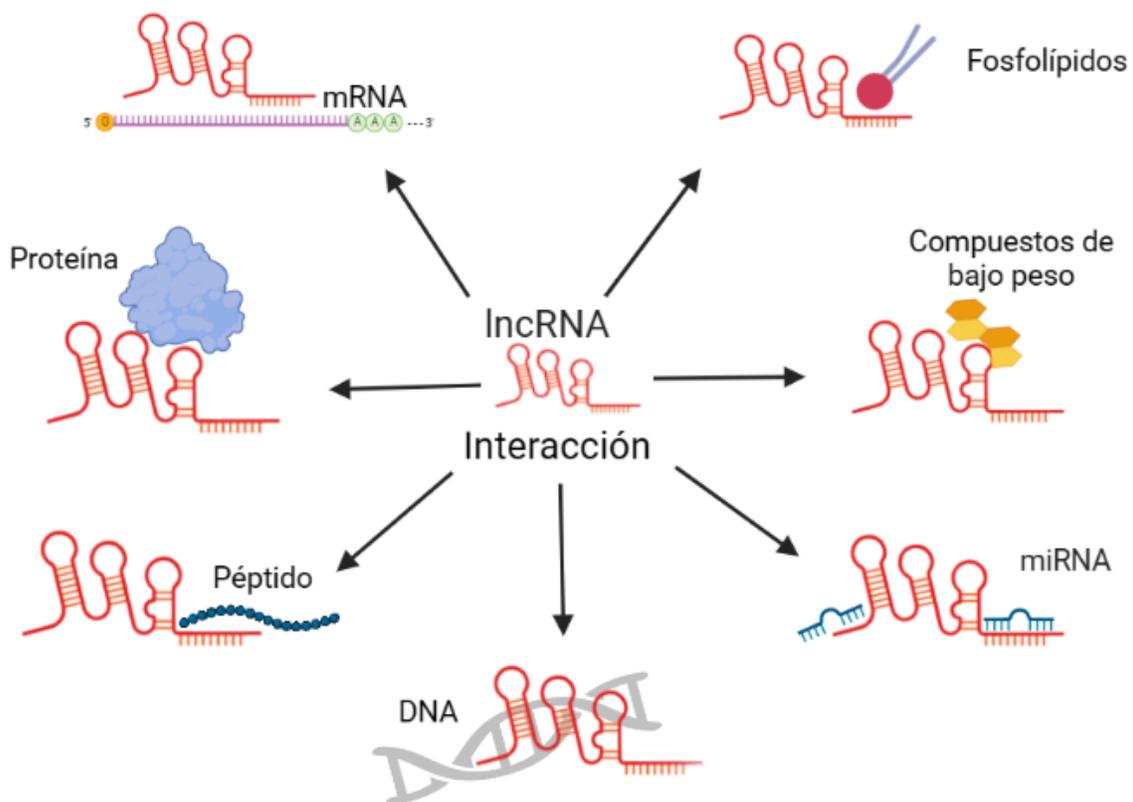


Figura 1.5. Interacciones de los lncRNAs (Modificado sobre la base de Schmidt *et al.*, 2020).

Las interacciones lncRNA-DNA pueden ocurrir de manera directa como el lncRNA FENDRR (lncRNA regulador del desarrollo letal para el feto) que forma una estructura tríplex con los

promotores de genes FOXF (genes reguladores transcripcionales que contienen un dominio de unión al DNA denominado dominio forkhead) y PITX2 (familia de genes homeobox que actúan durante el desarrollo embrionario temprano) creando sitios de unión con PRC2 regulando así la expresión de sus genes diana.

Las interacciones lncRNA-pre-mRNA pueden desempeñar un papel importante en el empalme alternativo ya que casi el 90% de los genes humanos se empalman de esta forma, el control de este proceso es importante para el desarrollo de los organismos (Romero *et al.*, 2018). Los análisis informáticos indican que de unos 24.500 genes 21.000 pueden verse afectados por la formación de dúplex lncRNA-mRNA donde el lncRNA puede afectar al empalme del pre-mRNA de dos maneras: bloqueando el ensamblaje del espliceosoma que implica la unión intrón-exón o convirtiéndose en el objetivo de los factores de empalme, un ejemplo de ello es el lncRNA MALAT que regula el empalme alternativo del pre-mRNA mediante el control de los niveles funcionales de los factores de transcripción. También el dúplex lncRNA-mRNA es un buen sustrato para la adenosina desaminasa del RNA, una enzima específica que convierte la adenina en inosina, modificación en el RNA que influye en el emparejamiento de bases. Los lncRNAs influyen en la estabilidad del mRNA, como es el caso del lncRNA LAST (lncARN que estabiliza los niveles de mRNA del oncogén ciclina D1 (CCND1) que se une a la región 5' no traducida (5' UTR) del CCND1mRNA (CCND1 es un gen que codifica para la proteína ciclina D1) y lo protege contra un posible ataque de nucleasas en células cultivadas (Kazimierczyk *et al.*, 2020; Romero *et al.*, 2018).

Las moléculas de RNA son los andamios perfectos para la unión de proteínas (Ribeiro *et al.*, 2018) existen lncRNAs que se unen específicamente a una o varias proteínas; como es el caso del lncRNA HOTAIR; mientras el lncRNA lnc-DC (lncRNA que se encuentra exclusivamente en células dendríticas convencionales humanas (DCs)) sólo lo hace con la proteína STAT 3 (transductor de señales y activador de la transcripción 3) (Kazimierczyk *et al.*, 2020; Romero *et al.*, 2018). Los complejos lncRNA-proteína regulan procesos celulares como la transcripción y el empalme, así como la expresión génica, es el caso de los lncRNAs HOTAIR y KCNQ1 (lncRNA antisentido en el locus KCNQ1 que regula genes crucialmente involucrados en el desarrollo y en el comportamiento postnatal) al unirse a los complejos de proteínas represivas de PRC1 y PRC2 (Romero *et al.*, 2018); que a su vez interactúan con la cromatina y están involucrados en la remodelación de la misma, así como en la regulación epigenética de la expresión génica a través de la modificación del DNA e histonas. Las interacciones entre lncRNAs y los complejos de ribonucleoproteínas nucleares heterogéneas (hnRNP) juegan un papel importante en muchas

funciones celulares como: la expresión genética, el procesamiento del pre-mRNA, la estabilidad del mRNA y la traducción (Li *et al.*, 2019); un ejemplo concreto es la interacción entre los lncRNAs LincRNA-p21 (lncRNA intergénico largo que es una diana transcripcional de p53 y HIF-1 α .) y PNKY y los complejos hnRNP que regula la transcripción y el empalme alternativo, respectivamente (Li *et al.*, 2019).

Los lncRNA también pueden interactuar con componentes lípidos como fosfolípidos, aunque menos se sabe de esto. Lin *et al.*, (2017) encontraron un lncRNA denominado LINK-A con la capacidad de unirse a fosfoinosítidos como PIP3, PIP2 y PA. Además, Donia *et al.*, (2019) realizaron una investigación donde un aptámero de lncRNA se une al fosfoinosítido PIP3. También Li *et al.*, (2021) describen que el lncRNA SNHG9 interactúa con el ácido fosfatídico (PA). La mayoría de las investigaciones acerca de la interacción lncRNA con lípidos han sido estudiadas en humanos o levaduras, pero poco en plantas.

1.2.6 Plantas y RNA largos no codificantes (lncRNAs)

La investigación sobre los lncRNAs vegetales es también un área emergente, y se han ido identificando gradualmente en estudios de especies como: *Arabidopsis*, *Oryza sativa*, *Medicago sativa*, *Zea mays*, *Solanum lycopersicum*, *Gossypium ssp.* (*Gossypium barbadense* y *Gossypium hirsutum*), y otras plantas herbáceas. Estas investigaciones han revelado que los lncRNAs también desempeñan un importante papel funcional en el crecimiento y desarrollo de las plantas, entre los que se incluyen Enod40 (lncARN que es reconocido directamente por la proteína nuclear RBP1 y participa en el tráfico nucleocitoplasmático de esta, induciendo su acumulación en gránulos citoplasmáticos, modulando el empalme dependiente de RBP1), COOLAIR (conjunto de lncRNAs antisentido empalmados alternativos y poliadenilados), COLDAIR (lncRNA que se transcribe desde la región promotora del gen de floración del locus C (FLC)), LDMAR (lncRNA que regula la fertilidad masculina sensible al fotoperiodo en el arroz), IPS1 (lncRNA que se expresa en caso de inanición de fosfato), HID1 (HIDDEN TREASURE I es un lncRNA), etc.; implicados en varios procesos como la formación de nódulos, la vernalización, absorción de fosfato, fotomorfogénesis y en los procesos de respuesta al estrés mediando la transmisión y expresión de la información genética (**Figura 1.6.**) (Sanchita *et al.*, 2020)

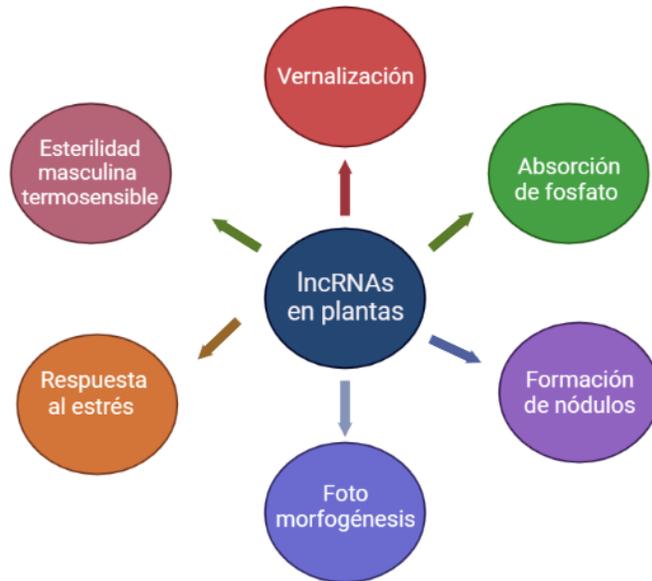


Figura 1.6. Funciones biológicas de los lncRNAs vegetales (Modificado sobre la base de Sanchita *et al.*, 2020).

Los lncRNA COLDAIR y COOLAIR encontrados en *A. thaliana* están involucrados en el proceso de la floración de esta planta, reclutando el complejo proteico PRC2 para activar H3K27me3 (modificación epigenética de la proteína de empaquetamiento del DNA histona H3) y mantener la expresión de bajo nivel del gen FLC (Heo *et al.*, 2013). Otros lncRNAs descritos en esta especie son: ELENA que promueve la liberación de la fibrilarina 2 (FIB2) de la proteína PR1 relacionada con la patogénesis que se induce en respuesta a diversos patógenos (Seo *et al.*, 2019; 2017), lncRNA APOLO (lncRNA regulador de auxina a través de un bucle promotor) que modula el desarrollo posterior de la raíz a través de la formación de un bucle de cromatina (R-loop) y suprime la transcripción de PID (Ariel *et al.*, 2014).

En cuanto a los lncRNAs en plantas superiores, la literatura más antigua informa de que el lncRNA Enod40 (lncRNA que se aisló como marcador temprano de la organogénesis de nódulos en plantas) en *M. truncatula* participa en funciones biológicas mediante la producción de péptidos cortos durante el proceso de formación de nódulos radiculares y que el lncRNA ASCO (ncRNA competidor de empalme alternativo) de *A. thaliana* puede competir con el mRNA para unirse a la proteína RBP, afectando así los niveles de expresión de los genes de respuesta a la auxina y regulando aún más el desarrollo radicular en *A. thaliana* (Bardou *et al.*, 2014).

Los resultados muestran que el lncRNA HID1 modulado por luz roja continua reprime transcripcionalmente la expresión del gen que codifica a la proteína del factor 3 de interacción

con fitocromos (PIF3) y promueve la fotomorfogénesis de las plántulas en *Arabidopsis* (Wang *et al.*, 2014). Además, los lncRNA DRIR (lncRNA que se activa tras los tratamientos contra la sequía, el estrés salino y el ABA, mejorando la tolerancia de las plantas a la sequía y al estrés salino regulando los genes relacionados con la resistencia al estrés) y SVALKKA (lncARN transcrito en sentido descendente) participan en los procesos de respuesta al estrés en *A. thaliana* (Kindgren *et al.*, 2018; Qin *et al.*, 2017) y lncRNAs como GhlncNAT-ANX2, GhlncNAT-RLP7, lncRNA16397 y lncRNA33732 desempeñan papeles importantes en la resistencia del algodón y el tomate a la infección por patógenos (Qin *et al.*, 2017). IPS1 en *A. thaliana* y PILNCR1 en el maíz, inducidos por estrés de bajo contenido en fósforo, contienen elementos de respuesta a miRNA y pueden unirse al microRNA miR399, sin ser escindidos por este, protegiendo así los genes diana de miR399 y participando en la regulación de la homeostasis del fosfato. El lncRNA de manzana, MSTRG.85814.11, actúa como potenciador transcripcional de SAUR32 un miembro de la familia de RNA auxiliar pequeño (SAUR) en plantas y contribuye a la respuesta a la deficiencia de hierro (Sun *et al.*, 2020).

Los estudios centrados en plantas modelo herbáceas han revelado las funciones y mecanismos biológicos de al menos 3000 lncRNAs. Se ha reconocido gradualmente que los lncRNAs median ampliamente la transmisión y expresión de la información genética, regulando así el crecimiento y desarrollo de las plantas, el metabolismo secundario y la adaptación al estrés. Los mecanismos de crecimiento y respuesta al estrés de los árboles son más complejos que los de las plantas herbáceas. En los últimos años, también se han identificado lncRNAs en árboles gracias a la secuenciación de alto rendimiento y la predicción de información biológica (Lin *et al.*, 2019; Wang *et al.*, 2017) en varias etapas del crecimiento y en procesos de respuesta a estrés (Liu *et al.*, 2019; 2018; Wang *et al.*, 2017). La autenticidad, identificación y anotación de estos lncRNAs han supuesto un enorme reto, y aún queda un largo camino por recorrer para su verificación y validación experimental. Hasta la fecha, no se ha publicado sobre las funciones biológicas y los mecanismos de los lncRNAs en árboles.

1.2.7 Análisis *in silico* de RNA largos no codificantes (lncRNAs)

Los ARN no codificantes largos (lncRNA) son un campo de investigación en rápida expansión, cada año se identifican nuevos transcritos. Sin embargo, hasta ahora sólo se ha caracterizado funcionalmente un pequeño subconjunto de estos. Para investigar los mecanismos de acción por los que actúan, así como su localización subcelular, conformación estructural y socios de interacción las herramientas bioinformáticas y las bases de datos tienen un valor incalculable.

El análisis *in silico* puede ser útil para estudiar un lncRNA recién descubierto, ya que los resultados obtenidos *in silico* pueden ser apropiados para orientar la metodología experimental. Por ejemplo, si un predictor de localización indicara que un lncRNA se encuentra en el núcleo, y un predictor de interacción devolviera motivos de unión al DNA, investigar el potencial de regulación de la expresión génica podría ser un punto de partida útil para el lncRNA de interés.

En los últimos años, se ha producido un aumento exponencial de la investigación sobre los lncRNA y del número de nuevos transcritos no codificantes identificados (Kung *et al.*, 2013). Se necesitan bases de datos precisas y de fácil acceso de genes putativos de lncRNA, sobre todo si se tiene en cuenta que determinar nuevos genes de lncRNA no es una tarea sencilla, la identificación de estos suele basarse en RNA-seq (secuenciación del RNA) y cada transcrito anotado como lncRNA puede que no sea un lncRNA. Existen muchas bases de datos de lncRNAs, sin embargo, la mayoría de ellas son subconjuntos de otras bases como es el caso de NONCODE (base de datos de conocimientos integrada dedicada a los RNA no codificantes excluidos los tRNA y los rRNA), FANTOM cap (método de identificación de sitios de inicio de la transcripción (TSS) en genomas eucarióticos). Muchas se construyeron con un propósito específico, como por ejemplo para una enfermedad concreta. Lnc2Cancer es una base de datos curada manualmente que proporciona asociaciones exhaustivas apoyadas experimentalmente entre lncRNA y cáncer humano (Gao *et al.*, 2019). Otras bases de datos son de especies particulares, como GreenC (base de datos verde de no codificación) (Paytuy-Gallart *et al.*, 2019). Existen dos "meta-bases de datos" actuales de lncRNAs, que integran datos de diversas fuentes y además están bien curadas y los integran en un formato de fácil acceso como: LNCipedia y LNCBook ambas enfocadas a genes lncRNA humanos (Pinkney *et al.*, 2020).

Una de las herramientas *in silico* más recomendada es identificar el potencial de codificación de un nuevo transcrito, es el primer paso vital para una anotación precisa y un análisis posterior de un lncRNA (Kong *et al.*, 2007). Determinar el potencial de codificación mediante metodologías *in vitro* puede requerir mucho tiempo y recursos, especialmente cuando se ensayan múltiples transcritos simultáneamente. En los últimos años, se han lanzado una serie de herramientas predictivas que pueden determinar la probabilidad de que un transcrito produzca una proteína, basándose únicamente en la secuencia. Una de las primeras herramientas establecidas es CPC (Coding Potential Calculator) (Kong *et al.*, 2007), que se publicó por primera vez en 2007. CPC utiliza seis características en su modelo predictivo, incluyendo la cobertura del marco de lectura abierto (ORF) predicho y la similitud de la secuencia con proteínas (Kong *et al.*, 2007). Estas características se incorporaron a un clasificador de aprendizaje automático SVM (Support Vector

Machine) (Kong *et al.*, 2007). Aunque CPC fue muy bien recibido en su lanzamiento, desde entonces se han publicado herramientas más precisas y mayor facilidad de uso como: CPPred, CNIT y CPC2 (Pinkney *et al.*, 2020).

La localización subcelular de los lncRNAs es un factor importante para comprender su posible función. Los que se localizan en el núcleo están a menudo implicados en la regulación de la expresión y/o empalme de genes (Engreitz *et al.*, 2016) o en la inactivación del cromosoma X (Wutz *et al.*, 2011) y/o el empalme (Engreitz *et al.*, 2016), la impronta de genes o la inactivación del cromosoma X (Wutz *et al.*, 2011). Por otro lado, los lncRNA exportados al citoplasma pueden modular la estabilidad del mRNA (Kim *et al.*, 2007) y la traducción (Huarte *et al.*, 2010), regular la modificación de proteínas (Liu *et al.*, 2015) o competir con miRNA (Cesana *et al.*, 2011). Algunas de las bases de datos de lncRNA y herramientas para predecir la localización de nuevos transcritos no codificantes largos son: LncSLdb, LncATLAS y LncLocator (Pinkney *et al.*, 2020).

La estructura secundaria y terciaria de un lncRNA puede revelar información sobre posibles socios de interacción y la función del transcrito, como en el caso del lncRNA MEG3 (gen expresado maternalmente que está implicado en la regulación de la proliferación celular y funciona como supresor tumoral en numerosos tejidos), que comprende tres motivos estructurales predichos que se conservan en sus múltiples isoformas (Zhang *et al.*, 2018). Dos de estos tres motivos son necesarios para la activación del supresor tumoral p53 mientras que el otro está implicado en la supresión de la síntesis de DNA (Zhang *et al.*, 2010). Sin embargo, la investigación experimental de las estructuras de los lncRNA es difícil por varias razones. En primer lugar, los RNA pueden cambiar de conformación como en el caso del lncRNA Braveheart, que cambia su estructura 3D al unirse a la proteína de unión al ácido nucleico celular (CNBP) (Kim *et al.*, 2020). En segundo lugar, las estructuras de los lncRNAs tienden a contener regiones dinámicas, como las encontradas en el lncRNA HOTAIR, cuya estructura se determinó recientemente mediante microscopía de fuerza atómica (Spokoini-Stern *et al.*, 2020). Es difícil determinar las regiones dinámicas de los RNA utilizando métodos tradicionales como la cristalografía de rayos X (Obayashi *et al.*, 2007). Por último, un gen lncRNA puede dar lugar a múltiples isoformas de empalme, dando lugar a varias estructuras de RNA diferentes, como el lncRNA SRA (lncRNA activador del RNA del receptor de esteroides 1) (Novikova *et al.*, 2012). Existen dos herramientas de predicción: RNAfold y DMfold (Wang *et al.*, 2018) para la investigación de estructuras 2D de lncRNA.

La identificación de las interacciones moleculares de los lncRNA es fundamental para comprender su función y, por tanto, su posible impacto en los procesos celulares. Por ejemplo, el lncRNA XIST, que inactiva uno de los dos cromosomas X en las mujeres, requiere una interacción física con el PRC2 para lograr su función (Sun *et al.*, 2015). Además, se sabe que el lncRNA TINCR controla la diferenciación de tejidos somáticos interactuando con el mRNA (Kretz *et al.*, 2013), mientras que el lncRNA FENDRR se une directamente al DNA para regular la expresión génica (Grote *et al.*, 2013). Se han descrito bases de datos y herramientas para identificar la posible interacción de los lncRNAs con socios moleculares; algunas de ellas son: RNAInter, RISE, IntaRNA, LncRRISearch, LnChrom, Triplexator y SFPEL-LPI (Pinkney *et al.*, 2020).

Predecir la función de un lncRNA a partir de la secuencia ha sido una tarea difícil en el campo de la biología de los lncRNAs (Kirk *et al.*, 2018). Una herramienta predictiva para determinar la función de los lncRNAs es SEEKR, que utiliza la clasificación basada en k-mer para comparar lncRNAs e inferir su función. La predicción *in silico* de los lncRNAs es un campo emergente con pocas herramientas aún. A medida que se disponga de más datos experimentales los algoritmos de predicción mejorarán su precisión y también cuando se publiquen más herramientas y bases de datos cada año, será importante hacer un seguimiento de las más precisas.

1.3 Fosfolípidos

1.3.1. Fosfoinosítidos

Los fosfoinosítidos se derivan del fosfatidilinositol (PI) un fosfolípido con carga negativa que comprende una fracción de diacilglicerol unida a un anillo de D-mioinositol mediante un enlace fosfodiéster (**Figura 1.7**). Presentan una cola de acil hidrofóbica y una cabeza de inositol hidrofílica, su composición hidrofóbica depende del organismo, así como del tejido, en los animales el rasgo característico es un alto contenido de ácidos esteárico unido a la posición sn-1 y araquidónico a la posición sn-2; mientras que en plantas la composición principal de la cola es de ácido palmítico, principal ácido graso saturado en la posición sn-1, mientras que los ácidos linoleico y linolénico son los componentes insaturados en la posición sn-2. En levadura, el ácido palmítico se encuentra en la posición sn-1, y los ácidos oleico y palmitoleico en posición sn-2 (Castano *et al.*, 2019; Phan *et al.*, 2019).

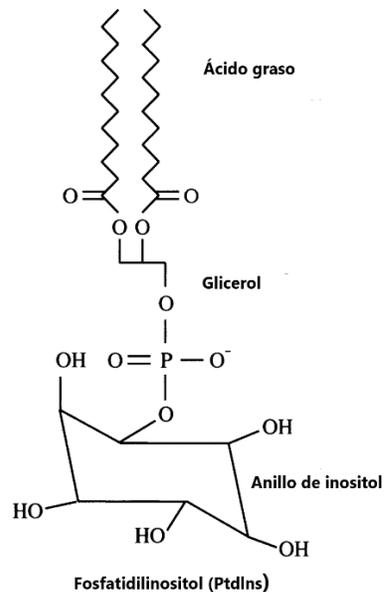


Figura 1.7. Estructura de los fosfoinosítidos. Estructura química del fosfatidilinositol, que comprende moléculas de diacilglicerol y D-mio-inositol enlazadas por fosfodiéster. R1 y R2 representan cualquier cadena de acilo (ácido graso). Los hidroxilos en las posiciones 3, 4 y 5 del D-mio-inositol se fosforilan fácilmente, dando lugar a siete fosfoinosítidos. (Castano *et al.*, 2019)

1.3.2 Fosforilación de los fosfoinosítidos.

El anillo de inositol puede ser fosforilado de forma reversible en las posiciones 3, 4 o 5 por enzimas quinasas y fosfatasas dando lugar a siete fosfoinosítidos fosforilados diferentes (**Figura 1.8**): fosfatidilinositol 3-fosfato (PtdIns(3)), fosfatidilinositol 4-fosfato (PtdIns(4)), fosfatidilinositol 5-fosfato (PtdIns(5)) monosfosforilados, fosfatidilinositol 3,4-bisfosfato (PtdIns(3,4)P₂), fosfatidilinositol 3,5 bisfosfato (PtdIns(3,5)P₂), fosfatidilinositol 4,5-bisfosfato (PtdIns(4,5)P₂) bifosforilados, y un fosfoinosítidos trisfosfato el fosfatidilinositol 1,3,4-trisfosfato (PtdIns(1,3,4)P₃) (Castano *et al.*, 2019; Phan *et al.*, 2019).

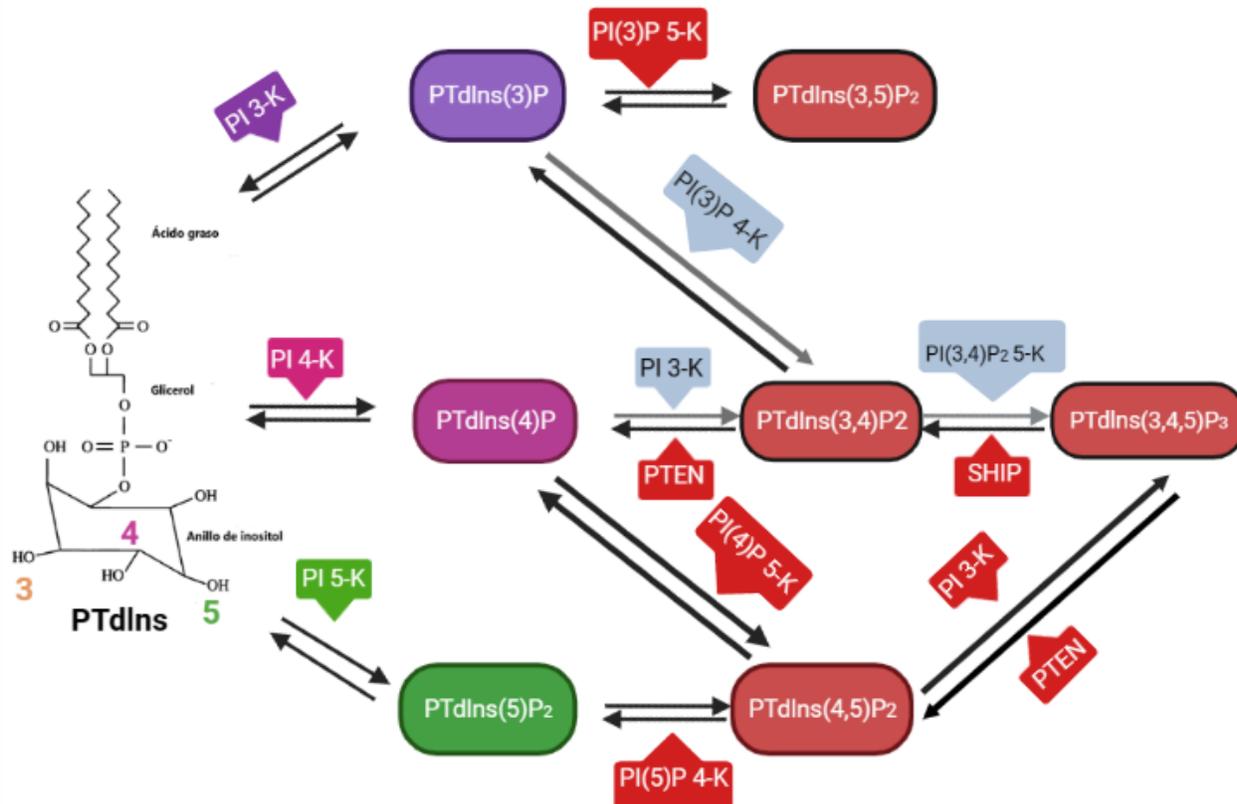


Figura 1.8. Ruta de fosforilaciones y desfosforilaciones del fosfatidilinositol. Las flechas en color negras indican las fosforilaciones y las flechas grises indican las fosforilaciones desconocidas. (Modificado sobre la base de Castano *et al.*, 2019)

Los fosfoinosítidos vegetales se forman de igual manera que en animales por la fosforilación en las mismas posiciones del anillo de inositol del fosfatidilinositol (PtdIns); aunque a diferencia de otras especies en las plantas sólo se ha demostrado la presencia de cinco (PtdIns3P, PtdIns4P, PtdIns5P, PtdIns(3,5)P₂ y PtdIns(4,5)P₂) de las siete especies conocidas. La presencia de PtdIns(3,4)P₂ o PtdIns(3,4,5)P₃ en las plantas no se ha confirmado de forma fiable (Heilmann *et al.*, 2016).

En general, las plantas contienen una compleja red de enzimas que intervienen en la formación y degradación dinámica de los PI. Al igual que en animales, los PI vegetales parecen formarse y degradarse con una alta tasa de recambio, lo que da lugar a un equilibrio dinámico que se altera fácilmente cuando las plantas están expuestas a condiciones de estrés. Este equilibrio dinámico permite que los eventos de señalización transitorios o localizados en respuesta a señales o estímulos del desarrollo o estimulación.

1.3.3 Localización y función de los fosfoinosítidos

Los PIP pueden servir como segundos mensajeros o modular el reclutamiento y/o la actividad de proteínas de membrana, desempeñan importantes funciones en la arquitectura de la misma, la dinámica del citoesqueleto, el tráfico vesicular, la transducción de señales, modulación de canales iónicos, regulación de microtúbulos y microfilamentos, la neurotransmisión; también están implicados en varios procesos como endocitosis, fagocitosis, macropinocitosis y exocitosis. Sus funciones exactas dependen de su localización y estado de fosforilación (Phan *et al.*, 2019; Dickson y Hille 2019; Ratti *et al.*, 2018; Hamann y Blind, 2018).

Los fosfoinosítidos fosforilados están presentes en varias partes de la membrana plasmática y de orgánulos diferentes, donde cada una de ellos participa en varios procesos celulares (Phan *et al.*, 2019) (**Figura 1.9**).

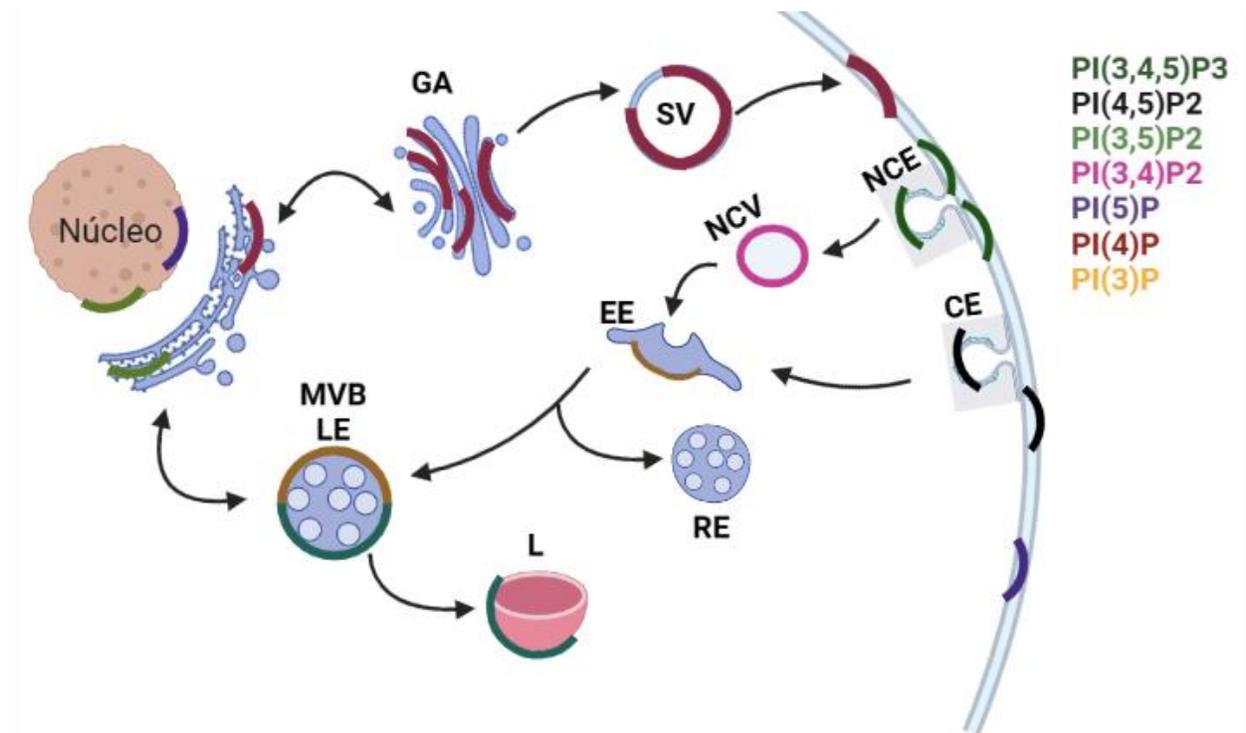


Figura 1.9. Distribución celular de los fosfoinosítidos presentes en varios organelos animales. N: nucleolo, ER: Retículo endoplasmático, CE: Endocitosis dependiente de clatrina, GA: Aparato de Golgi, MVB: Cuerpos multivesiculares, LE: Endosomas tardíos, L: Lisosomas, EE: Endosomas tempranos, SV: Vesículas secretoras NCV: vesículas endocíticas no clatrina NCE: endocitosis independiente de la clatrina (Modificado sobre la base de Phan *et al.*, 2019)

Los fosfoinosítidos generalmente se encuentran en el citoplasma, aunque estudios han revelado la presencia de dos isoformas de proteínas que median el transporte citoplasmático de PI al

núcleo, por tanto, se pudiera especular que PI es importado al núcleo donde posteriormente es fosforilado; incluso muchas enzimas implicadas en el metabolismo de los fosfoinosítidos como cinasas, fosfatasas y fosfolipasas están presentes en el núcleo. Otra hipótesis es que se trasladan al núcleo justo antes de que la envoltura nuclear se reensamble durante la telofase, y luego quedan encerrados en la nueva envoltura nuclear; un ejemplo es PIP2 nuclear que permanece unido a las regiones organizadoras nucleolares (NOR) durante la mitosis. Además, se sabe que la enzima que participa en la regulación de PIP2 de las motas nucleares se transloca al citoplasma y se acumula en los gránulos de intercromatina mitótica (MIG), y permanecen aquí incluso después de la formación de la envoltura nuclear en las células hijas (Castano *et al.*, 2019).

La visualización *in vivo* de la localización de los PI supone un reto aún, el uso de herramientas como microscopía óptica y electrónica, así como de anticuerpos y dominios de unión específicos han permitido su identificación en varias regiones nucleares (Tsuji *et al.*, 2019; Kalasova *et al.*, 2016) utilizando el dominio proteico FYVE del receptor tirosina quinasa, un marcador específico para PIP3, detectaron la presencia de este fosfoinosítido en el componente fibrilar denso (DFC) del nucléolo y a su vez la localización de PIP4 en las motas nucleares con el anticuerpo específico anti-PI(4)P y el dominio OSH1-PH de unión. Otros estudios relevaron la presencia de PIP2 en la membrana nuclear, motas nucleares, nucleolos y en la interfaz entre la heterocromatina y el espacio intercromático en el núcleo; utilizando los anticuerpos anti-PIP2 y el dominio PH específico de PIP2 (Castano *et al.*, 2019; Ulicna *et al.*, 2018; Sobol *et al.*, 2018; Kalasova *et al.*, 2016).

Los PI vegetales se localizan predominantemente en la cara citosólica de las membranas donde no se distribuyen de manera uniforme (Tejos *et al.*, 2014; Ischebeck *et al.*, 2014). Dentro de la membrana de las plantas, los PI actúan como reguladores de las proteínas integrales de membrana o marcan zonas de la membrana para el reclutamiento de proteínas de membrana periféricas (Takenawa, 2010). Los PI vegetales ejercen diferentes funciones interactuando con una variedad de socios de unión, se enriquecen en determinados dominios de membrana, formando microdominios que desempeñan un papel en la mediación de los efectos específicos de los PI (Ischebeck *et al.*, 2014). En los últimos años, los PI vegetales se han descrito como reguladores centrales de la función y el desarrollo de las plantas, jugando un papel importante en la función y desarrollo de las mismas, (Rodríguez-Villalón *et al.*, 2015; Golani *et al.*, 2013). La desregularización en el metabolismo de los PI en plantas puede inducir una variedad de cambios fenotípicos que afectan a múltiples tipos de tejidos y aspectos del desarrollo de la planta, por

ejemplo, la sobreexpresión de PI-sintasa en maíz resulta en una tolerancia alterada a la sequía (Liu *et al.*, 2013).

1.3.4 Fosfatidilinositol 4,5-bisfosfato (PtdIns (4,5) P2)

El fosfatidilinositol 4,5-bisfosfato (PtdIns (4,5) P2) actúa como segundo mensajero lipídico en la traducción de señales que regula diversos procesos celulares, sintetizado por la familia de enzimas cinasas PI4K/ PI5K que fosforila la posiciones 4 y 5 respectivamente del anillo de inositol donde PI (4)P y PI (5)P (**Figura 1.10**) sirven como precursores para la síntesis de PI (4,5) P2.

Este fosfoinosítido se localiza tanto en la membrana plasmática, donde constituye el 1% de los fosfolípidos totales, desempeñando un papel importante en el transporte de iones, el tráfico de membrana y en la regulación del citoesqueleto de actina. También se ha descrito la presencia de PI (4,5) P2 en las motas nucleares, en los nucléolos y en el nucleoplasma siendo uno de los fosfoinosítidos más abundantes descritos en el núcleo, donde juegan un papel importante en la regulación de procesos como la remodelación de la cromatina, empalme, transcripción, reparación del DNA y modificaciones epigenéticas (Sobol *et al.*, 2018).

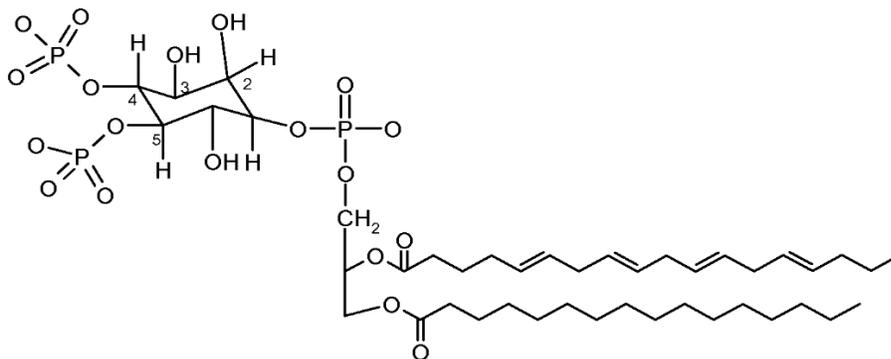


Figura 1.10. Estructura química del fosfatidilinositol 4,5-bisfosfato (PtdIns-4,5- P2).

Estudios recientes han descrito el papel de PIP2 como un lípido determinante en la diversificación espacial y funcional de la arquitectura nuclear; específicamente Sobol *et al.*, (2018) identificaron la presencia de PIP2 mediante anticuerpos anti-PI (4,5) P2 dentro de compartimientos nucleares específicos llamados islotes lipídicos nucleares (NLI) (**Figura 1.11 (A) (B)**). Los NLIs son estructuras globulares lipídicas de aproximadamente 40-100nm de diámetro, en su superficie se encuentran principalmente moléculas de PIP2 ubicadas con cabezas hidrofílicas dirigidas hacia el exterior, mientras que las colas hidrofóbicas se dirigen hacia el interior (Sztacho *et al.*, 2019).

El núcleo de los NLIs se encuentra compuesto por compuestos ricos en carbono, probablemente lípidos, mientras que su periferia está asociada con ácidos nucleicos y proteínas; incluyendo la

subunidad grande de la RNA Pol II, factores de transcripción y la miosina nuclear 1 (NM1) que contiene un dominio de homología de pleckstrin (PH) específico de PIP2 en el extremo C terminal. Ensayos por inmunoprecipitación con anticuerpos anti-PI(4,5)P2 y experimentos de pull-down con perlas acopladas a PI(4,5)P2 (Toska *et al.*, 2012), sugieren que PIP2 puede unirse a la maquinaria de transcripción de la RNA Pol II a través de una interacción directa con NM1, formando un complejo con la forma activa de esta enzima; esta interacción es vital para mantener el proceso de transcripción (Sobol *et al.*, 2018); demostrando que PIP2 se asocia con genes activamente transcritos, la Pol II, los factores de transcripción de la Pol II, la NM1 y transcritos nacientes en la superficie de los NLIs (Sobol *et al.*, 2018) (**Figura 1.11 (C)**). Entonces PIP2 puede servir como plataforma de andamiaje que facilita la formación de fábricas de transcripción, participando así en la conformación de la arquitectura nuclear competente para este proceso (Sobol *et al.*, 2018).

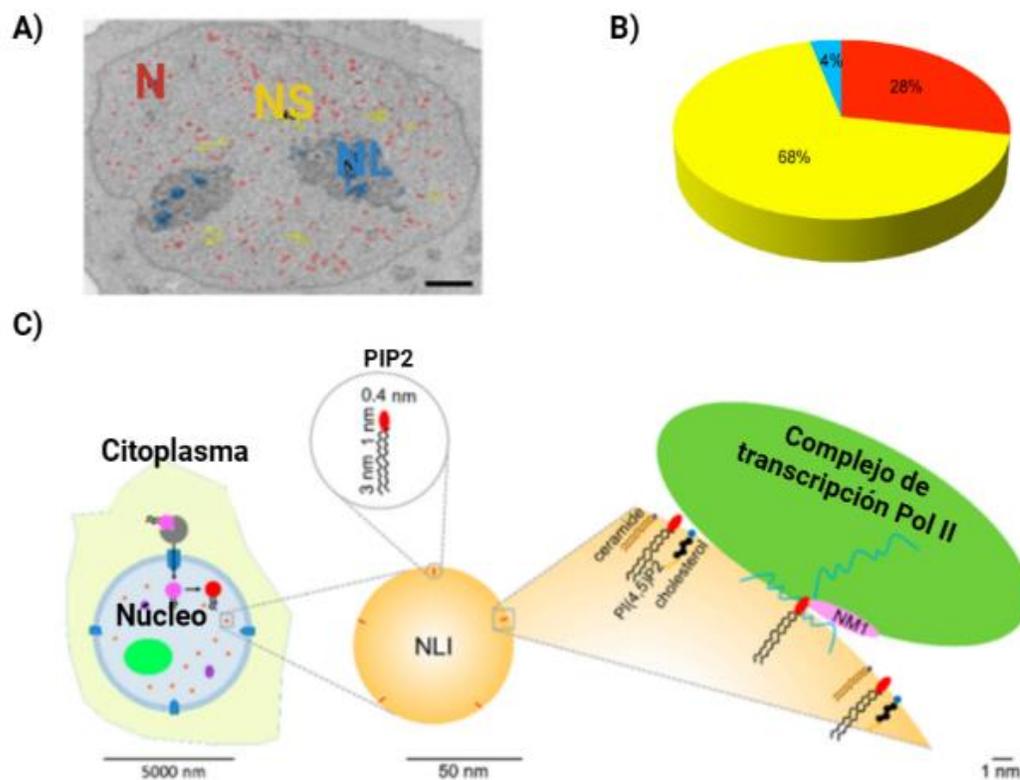


Figura 1.11. Islotes de lípidos nucleares (NLI). (A) La microscopía electrónica de transmisión (TEM) demuestra la localización del marcaje nuclear de PI (4,5) P2. (B) Se determinó la distribución relativa de PIP2 en motas nucleares, nucléolos y NLIs. La imagen codificada por colores y el diagrama muestran la PIP2 en las motas nucleares en amarillo, en los NLIs en rojo, y PIP2 nuclear en azul. (C) Modelo detallado

de la superficie del NLI. El modelo se representa aproximadamente en proporción al tamaño real de las moléculas, basándose en los datos experimentales de Sobol *et al.*, (2018), se puede observar una parte de la superficie de los NLI, compuesta de PIP2, ceramida, colesterol y RNA, con el complejo de transcripción Pol II, la NM1 y transcritos nacientes (Modificado sobre la base de Castano *et al.*, 2019)

El PIP2 en las plantas actúa igual que en animales como una importante molécula de señalización. Interviene en diversos procesos como el crecimiento del tubo polínico, el estrés salino y osmótico, el tráfico de vesículas, la organización de la actina, la actividad de los canales iónicos, los movimientos de las células de guarda y en la apertura estomática (Lee *et al.*, 2007).

1.3.5 Ácido Fosfatídico

El ácido fosfatídico (PA), o 1,2-diacil-sn-glicero-3-fosfato, es uno de los glicerofosfolípidos y, en general, uno de los fosfolípidos de membrana más simple. La molécula de PA está formada por un esqueleto de glicerol de tres carbonos, a la que se le unen dos ácidos grasos en las posiciones sn-1 y sn-2 los cuales pueden variar en cuanto a su longitud (14 a 22 carbonos) y grado de saturación y en la posición sn-3 un grupo fosfato (**Figura 1.12**) (Zhukovsky *et al.*, 2019).

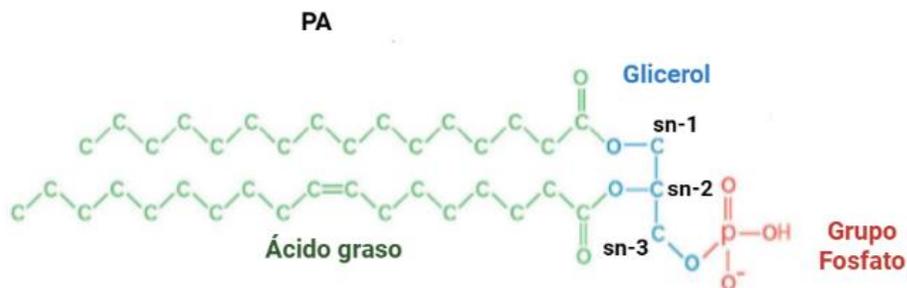


Figura 1.12. Representación esquemática del ácido fosfatídico (PA)

El PA posee propiedades únicas (grupo de cabeza pequeño, cargado negativamente y fosfomonoéster) que le permiten establecer interacciones funcionales con proteínas, lípidos y también con sí mismo. En condiciones fisiológicas este lípido adopta forma de cono y por tanto posee una curvatura espontánea negativa; en respuesta a cambios relativamente leves de pH en el rango fisiológico, la molécula de PA cambia notablemente su estado de protonación, propiedad que le permite actuar como biosensor de pH. El papel de PA como biosensor de pH permite a este lípido regular varios procesos biológicos, como la detección de nutrientes y la señalización del crecimiento celular en la levadura. (Zhukovsky *et al.*, 2019).

La tasa de renovación de PA en la célula es elevada (Zhukovsky *et al.*, 2019; Potocký *et al.*, 2014), está controlada mediante el equilibrio entre su síntesis y degradación (Zhukovsky *et al.*,

2019; Thakur *et al.*, 2019). Esta regulación es importante para la función celular y se consigue a través de la unión mediada por PA y la de una amplia gama de proteínas efectoras (Zhukovsky *et al.*, 2019; Barneda *et al.*, 2015; Foster *et al.*, 2014) en un determinado momento y en una región de la membrana subcelular concreta (Zhukovsky *et al.*, 2019; Kassas *et al.*, 2017). La producción de PA en la célula se obtiene por tres vías principales (Zhukovsky *et al.*, 2019; Kassas *et al.*, 2017; Vodicka *et al.*, 2015; Foster *et al.*, 2014) (**Figura 1.13**), la síntesis de novo a través de la acilación del ácido lisofosfatídico (LPA) (Körbes *et al.*, 2016; Vodicka *et al.*, 2015), la fosforilación del diacilglicerol (DAG) (Ma *et al.*, 2019; Li *et al.*, 2013) y la hidrólisis de fosfolípidos (Panda *et al.*, 2018).

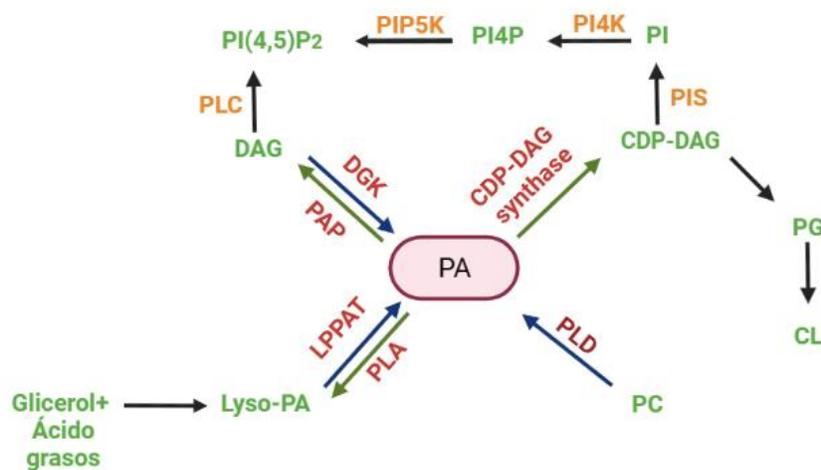


Figura 1.13. Representación esquemática de las vías bioquímicas para la síntesis y el metabolismo del PA. Las flechas azules indican la síntesis de PA mientras que las flechas verdes indican el recambio. Las enzimas implicadas están marcadas en naranja y rojo; las que directamente influyen en los niveles de PA son las rojas. Las especies lipídicas están marcadas en verde: Diacilglicerol (DAG), Diacilglicerol Difosfato de Citidina (CDP-DAG), Fosfatidilinositol (PI), Fosfatidilinositol-4-fosfato (PI4P), fosfatidilinositol-4,5-bis-fosfato (PI (4,5)P₂), ácido liso-fosfatídico (liso-PA) fosfatidilcolina (PC), fosfatidilglicerol (PG), cardiolipina (CL), diacilglicerol cinasa (DGK), PA fosfatasa (PAP), liso-PA acil transferasa (LPAAT), fosfolipasa A (PLA), fosfolipasa D (PLD), fosfatidilinositol-4-quinasa (PI4K), fosfatidilinositol-4-fosfato-5-quinasa (PIP5K), fosfolipasa C (PLC) (Modificado sobre la base de Zhukovsky *et al.*, 2019)

La señalización de PA en la célula es generada principalmente por dos familias de enzimas, la fosfolipasa D (PLD) y la diacilglicerol quinasa (DGK), cada una de las cuales utiliza diferentes sustratos de fosfolípidos. Los dos miembros PLD1 y PLD2 de la familia PLD, hidrolizan el fosfolípido de membrana más abundante, la fosfatidilcolina (PC), para generar PA mientras que los diez miembros de la familia DGK generan PA a través de la fosforilación de diacilglicerol (Ma *et al.*, 2019; Raben y Barber, 2017; Li *et al.*, 2013). En las células vegetales y de levadura, el PA puede ser fosforilado por la PA quinasa (PAK) en diacilglicerol pirofosfato (DAGPP), y el PA puede ser producido a partir del DAGPP por la DPP (DAGPP fosfatasa) (Zegarłinska *et al.*, 2018;

Ufer *et al.*, 2017). Por lo tanto, directa o indirectamente, el PA está implicado en la biosíntesis y la homeostasis de la mayoría de los fosfolípidos (Zhukovsky *et al.*, 2019; Kassas *et al.*, 2017).

También regula directamente y promueve la actividad de una variedad de proteínas de señalización como las PIPKs que generan PIP2. Por el contrario, la actividad de PLD1 y PLD2 también depende en gran medida de la interacción PIP2. Se ha propuesto que la estimulación recíproca de PLD (PA) y PIPKI [PI (4,5) P2] permite un bucle de regulación positiva que es crítica para la rápida generación de lípidos de señalización en localizaciones subcelulares específicas; tras las estimulaciones PA activa una enzima, la PtdIns4P5K cuyo producto PtdIns (4,5) P2 está implicado en la fisión y fusión de membranas (Zhukovsky *et al.*, 2019).

El PA es una molécula de señalización que regula varios procesos celulares como la proliferación celular, la supervivencia, la organización del citoesqueleto, el tráfico vesicular, la secreción, la reproducción, respuesta a las hormonas y el estrés abiótico (Zhukovsky *et al.*, 2019; Thakur *et al.*, 2019; Tanguy *et al.*, 2019; Zegarlinka *et al.*, 2018; Raben y Barber, 2017)

1.4 Modelo de planta utilizado *Stenocereus queretaroensis*

Las cactáceas pertenecen a un grupo de plantas con características morfológicas variadas: arborescentes, columnares, globulares articuladas y barreloides. Sus características las catalogan como angiospermas suculentas y su peculiar metabolismo fotosintético CAM les confiere una gran adaptación a ambientes extremos, particularmente aquellos donde hay escasez de agua estacionalmente (Ortega-Baés *et al.*, 2010).

En México existen cactus con alto potencial agrobiotecnológico; entre ellos destaca el aprovechamiento de sus metabolitos secundarios para la obtención de pigmentos naturales, uno de ellos es *Stenocereus queretaroensis* (F.A.C. Weber ex Mathes.) Buxb, comúnmente llamada pitaya salvaje, es una de las ocho cactáceas columnares congéneres nativas de México, distribuida en el centro del país. *Stenocereus queretaroensis* es una cactácea columnar arborescente con tronco bien definido que puede llegar a medir hasta ocho metros de altura. Sus ramas son cilíndricas y poseen ocho costillas prominentes. Sus flores se diferencian en las areolas de la mitad superior de las ramas. La forma de sus frutos varía entre ovoide o globosa y están cubiertos por un exocarpio suave con areolas espinosas las cuales se desprenden en su madurez fenológica durante la primavera; estos llegan a pesar entre 100 y 200 gramos y poseen semillas piriformes muy pequeñas, de color negro, su mesocarpio es jugoso y blando y van desde

una coloración roja hasta formas amarillas y rara vez verdes, dependiendo de la variedad (Figura 1.14).



Figura 1.14. Morfología vegetativa de *Stenocereus queretaroensis*. A. Planta arborescente de *S. queretaroensis*. B. Flores de *S. queretaroensis* atrayente de polinizadores. C. Frutos ovoides de *S. queretaroensis*, con mesocarpio blando y dulce, cubiertos por areolas espinosas. (Tomado de Cejas *et al.*, 2021 Fotografías: Carlos Valenzuela).

Esta cactácea se ha utilizado durante mucho tiempo como fuente de alimento en el centro y norte de México. Sus frutos, comúnmente llamados pitayos, biosintetizan grandes cantidades de betalainas. Estas moléculas son compuestos nitrogenados solubles en agua; que en comparación con a otros pigmentos, como antocianinas o carotenoides, destacan por su estabilidad fisicoquímica en los procesos industriales (Cejas *et al.*, 2022).

1.5. JUSTIFICACIÓN

La interacción de los lncRNAs con componentes lipídicos como PA Y PIP2 importantes moléculas de señalización celular ha sido estudiada y demostrada utilizando células humanas vinculadas a enfermedades como el cáncer. Sin embargo, en plantas estas siguen sin identificarse. Determinar la unión entre estos dos tipos de biomoléculas supone un hallazgo en el mundo vegetal abre una nueva puerta de preguntas biológicas. Los resultados de este trabajo contribuirán a reflejar procesos moleculares que se ven favorecidos por esta interacción, a su vez permitirán conocer si los lncRNAs son moléculas que presentan estructuras conservadas en plantas.

1.6 HIPÓTESIS

Existen lncRNAs de *Stenocereus queretaroensis*. que interactúan específicamente con PIP2 y PA.

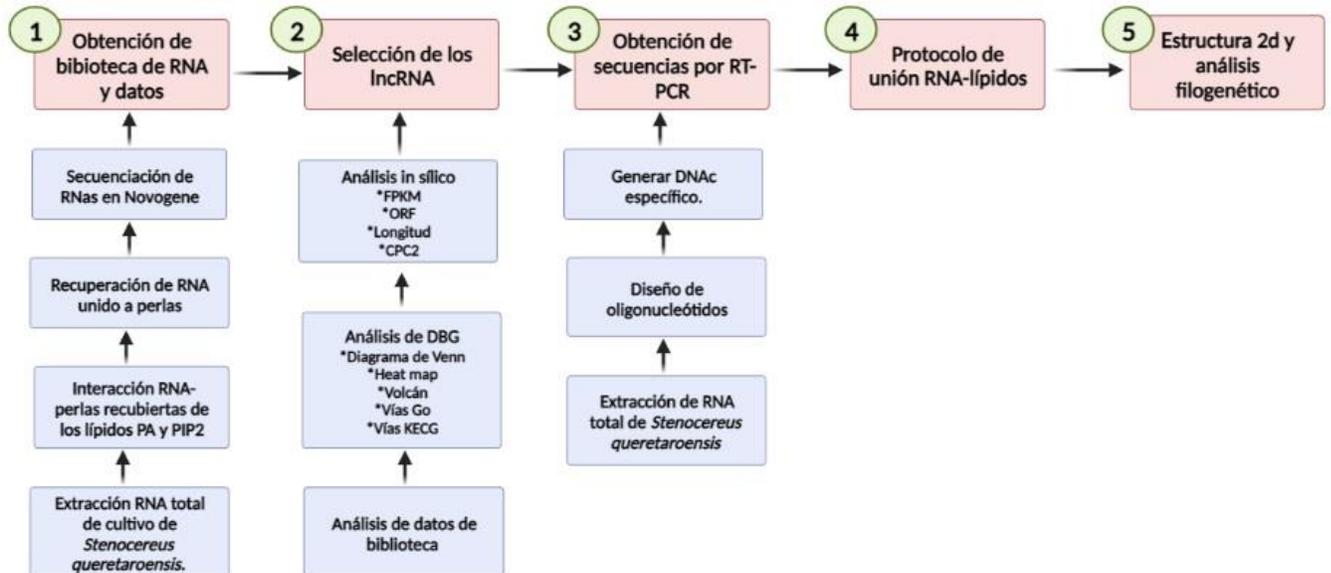
1.7 OBJETIVO GENERAL

Determinar lncRNAs con interacción a PIP2 y PA a partir del análisis de secuenciación de RNA (RNA-seq) de cultivos celulares de *Stenocereus queretaroensis*.

1.8 OBJETIVOS ESPECÍFICOS

- Generar una biblioteca de RNA-seq y manejar los datos de la misma.
- Seleccionar lncRNA de la biblioteca de RNA-seq de *Stenocereus queretaroensis*.
- Comprobar la interacción de los lncRNA con PIP2 y PA.

1.9 ESTRATEGIA EXPERIMENTAL



CAPÍTULO II

MATERIALES Y MÉTODOS

2.1 Material Biológico

Se utilizó la línea celular sqR1 de *Stenocereus queretaroensis* (Weber) Buxbaum previamente establecidas a partir del fruto inmaduro de una pitaya traído de Querétaro [Kim *et al.*, 2001]. La línea celular se mantuvo en el medio Murashige y Skoog (MS) adaptado para cactus, a base de sales inorgánicas de MS (Sigma M5524), con una concentración de calcio total de 4,4 μM y sin ácido nicotínico ni glicina. Además, contenía sacarosa al 8%, ácido 1-naftalenacético (Sigma-Aldrich) (5,3 μM), ácido 2,4-diclorofenoxiacético (Sigma-Aldrich) (4,5 μM) y kinetina (Sigma-Aldrich) (4,6 μM). El pH del medio se ajustó a $5,8 \pm 0,02$ y se distribuyó a razón de 40 mL en matraces Erlenmeyer de 250 mL. Posteriormente, éstos se esterilizaron en autoclave a una presión de 1 kg/cm² y una temperatura de 121°C durante 20 minutos.

Las suspensiones celulares se propagaron mediante subcultivo cada 15 días y se mantuvieron en las siguientes condiciones: temperatura $25 \pm 2^\circ\text{C}$, luz (4700 lúmenes) y continua agitación (120 rpm). Cada subcultivo se realizó transfiriendo 15-20 mL de las células en suspensión a Erlenmeyer con 40 mL de medio de cultivo fresco. Para determinar el ciclo de crecimiento de las líneas celulares, se preparó un caldo inicial vertiendo 15 frascos de suspensiones en un matraz de Erlenmeyer de un litro (en condiciones estériles). Una vez homogeneizado el caldo celular, se vertieron 10 mL de la suspensión, equivalentes a $1,8 \pm 0,2$ g de peso fresco, se transfirieron a cada uno de los Erlenmeyer de 250 mL con 40 mL de medio de cultivo MS modificado para cactus. Tras la inoculación los matraces se mantuvieron en la sala de cultivo bajo luz continua a 4700 lúmenes, una temperatura de $25 \pm 2^\circ\text{C}$ y 120 rpm, durante todo el ciclo de cultivo.

2.2 Extracción de RNA total de *Stenocereus queretaroensis*

La extracción de RNA total se realizó utilizando como modelo investigativo células en suspensión de *Stenocereus queretaroensis* usando el paquete de purificación PureLink® RNA Mini Kit (n° 12183020) Invitrogen (Anexo. 1). Se filtraron 100 ml de cultivo celular de la línea celular sqR1 de *Stenocereus queretaroensis* (Weber) Buxbaum y luego se pulverizaron utilizando nitrógeno líquido y un mortero. Se evaluó la calidad y cantidad del RNA extraído utilizando un Nanodrop y se trató con DNasa I libre de RNasa; (New England No. M0303S).

2.2.1 Ensayo de interacción RNA-fosfolípidos (PA/PIP2) mediante Pull-down

Para el ensayo de interacción RNA-PIP (PA/PIP2) mediante pull-down, se añadieron 200 µg de RNA a 100 µl de perlas de agarosa (también conocidas como perlas de control) e inhibidor de RNAsas RNase Block. Tras 5 min de incubación, se limpió el RNA, a continuación, se prelavaron 100 µl de perlas lipídicas (PA/PIP2) con PBS RNase Free y se cambiaron a tubos limpios. Las perlas utilizadas para los pull-down descritos en este estudio son las perlas de control #P-B000, las perlas recubiertas de PI (4,5) P2 #P-B045a (también descritas como PIP2), y las perlas recubiertas de ácido fosfatídico (PA) #P-B0PA de Echelon Biosciences Inc... El RNA prelavado se añadió a las perlas limpias y se incubó durante 30 min en rotación en cámara fría. Los tubos se centrifugaron a baja velocidad para eliminar el RNA no unido, se lavaron tres veces con 100 µl de PBS RNase Free y se cambiaron de tubo durante cada lavado. Para la extracción del RNA unido, se añadieron 200 µl de Ribozol y 200 µl, se mezclaron y se centrifugaron durante 10 min a 14000 x g. Finalmente, se tomó la parte superior y se precipitó con 700 µl de isopropanol, se lavó la pastilla con etanol al 70% y se almacenó a -80 °C, quedando así preparada la muestra para su posterior envío y secuenciación por Novogene Company.

2.3. Obtención y caracterización de la librería de RNA-seq de *S. queretaroensis*

Para la preparación y construcción de bibliotecas, se hizo uso de los servicios de secuenciación de RNA-seq de Novogene Company. La primera etapa consistió en evaluar la calidad de las muestras de RNA total que enviamos, una vez comprobado que este cumplía con los requisitos de calidad, cantidad e integridad se pasó a la preparación de la librería. En esta segunda etapa se eliminó el rRNA utilizando el kit Ribo-Zero™ de NEB imprescindible cuando se trata de lncRNAs, se fragmentó el RNA y se sometió a transcripción inversa para obtener cDNA, y se preparó una biblioteca específica de cadena la cual permite realizar una cuantificación más precisa de lncRNA. Este tipo de librería se diferencia de una librería de cDNA normal que durante el paso de síntesis de la segunda cadena de cDNA se reemplazó el dNTPs por dUTPs, de esta forma la segunda cadena de cDNA estará formada por uracilo en lugar de timina, posteriormente se degrada selectivamente la segunda hebra utilizando la uracil DNA glicosilasa entonces de esta manera la cadena restante es la que se amplifica para generar una librería de cDNA listo para secuenciar; de esta forma sabremos tras la secuenciación si la lectura procede de la primera cadena de cDNA o de la segunda, los beneficios de hacer una librería específica de hebra incluyen el de identificar con precisión las transcripciones antisentido, determinar la cadena transcrita de los ncRNAs como los lncRNA y distinguir los límites de los genes superpuestos. Luego que quede la librería se pasó a la etapa de secuenciación para lo que se utilizó la

plataforma Illumina Novaseq 6000 con una longitud de lectura paired-end de 150 pb (dos lecturas de 150 pb de bases cada una).

Obtenidos los datos crudos se les realizó un control de calidad para garantizar la calidad del análisis bioinformático, Novogene establece varios criterios para obtener datos limpios: quitar las lecturas con adaptador; eliminar las lecturas que contienen más del 10 % de N (N representa la base incierta mientras se llama a la base) y eliminación de lecturas de baja calidad (una lectura que contiene más del 50 % de las bases con un puntaje de calidad bajo: puntaje de calidad ≤ 20). Seguidamente como no disponemos de genoma de referencia, se generó un transcriptoma *de novo* para generar un transcriptoma de referencia, posteriormente todas las lecturas cortas se mapearon con los unigenes utilizando el software HISAT2. Se utilizó el software RSEM para realizar la cuantificación de genes de cada muestra. Para el análisis bioinformático se utilizó el método de normalización de recuentos de lecturas TMM (Trimmed Mean of M-values). El análisis de enriquecimiento diferencial se realizó utilizando el paquete EdgeR R. Se utilizó $\log_2(\text{FoldChange}) > 1$ y $p_{\text{adj}} < 0,001$ como estándar de cribado de genes diferencialmente enriquecidos.

2.4. Análisis de los DBG y selección de lncRNA

El análisis de los genes expresados diferencialmente se realizó tomando como punto de partida el total de genes secuenciados y como criterios de selección los valores de un $q_{\text{val}} < 0.005$ donde “ q_{val} ” es el valor de p ajustado que se utiliza después de aplicar una corrección de Bonferroni para múltiples comparaciones entre diferentes tratamientos y “ < 0.005 ” significa que el valor de p ajustado es menor que 0.005, lo que indica una alta confianza en la significación estadística de nuestros resultados. Por otra parte, el segundo criterio que utilizamos fue un $|\log_2(\text{FoldChange})| > 1$, que permite determinar los cambios en la expresión génica entre los grupos en estudio, donde “ > 1 ” significa que el cambio en la expresión génica es mayor que 2 veces (o menor que 0.5 veces) entre los tratamientos.

Obtenidos los DBG el siguientes criterios tomados en cuenta para seleccionar las secuencias dentro de este grupo que coincidieran con las características de lncRNAs fueron: longitud de las secuencias ≥ 200 debido a que los lncRNAs tienen una longitud superior a 200 nucleótidos y el marco abierto de lectura (ORF) < 100 , generalmente estos ncRNA no tienen marcos de lectura abiertos (ORF) que codifican proteínas, sin embargo, se ha encontrado que algunos lncRNA tienen ORF pequeños y desconocidos que pueden codificar péptidos funcionales.

Además, se utilizó la herramienta de CPC2 (Coding Potential Calculator 2) para conocer el potencial de codificación de estas secuencias y determinar si codificaban para alguna proteína funcional. Este programa utiliza seis características en su modelo predictivo: longitud del péptido putativo (secuencia de aminoácidos que se traduce a partir del ORF), puntuación Fickett (probabilidad de que una secuencia de DNA sea un exón), TESTCODE, punto isoelectrónico putativo, la integridad y cobertura del marco de lectura abierto (ORF) predicho y la similitud de la secuencia con genes de proteínas. Estas características se incorporan a un clasificador de aprendizaje automático SVM (Support Vector Machine) para arrojar como resultado si una secuencia se clasifica como codificadora o no. Simultáneamente se realizó una búsqueda en bases de datos de secuencias de proteínas y nucleótidos no redundantes del NCBI (Nr/Nt) y en las bases de datos de proteínas Swiss-Prot y PFAM y Swiss-Prot para de esta manera descartar aquellas secuencias que fueran coincidentes con algunas reportadas como codificadoras para proteínas funcionales. Además, se realizó un análisis de los diagramas de Venn, heat-maps y gráficos de volcán de los genes enriquecidos diferencialmente ante la presencia de PA y PIP2 respecto al control. También se realizó un análisis general de las vías de señalización KEGG y ontología génica (GO), con el objetivo de identificar las vías biológicas y los procesos moleculares que se encontraban favorecidos.

De esta manera quedarán seleccionadas las secuencias que sean estrictamente lncRNAs, posteriormente se escogerán dentro de este conjunto de lncRNAs aquellos genes que se encuentren más enriquecidos ante las condiciones de interés, presencia de los lípidos PA/PIP2. Esta selección se basó en un análisis de los gráficos de Venn, heat-maps y gráficos de volcán para estos lncRNAs, los cuales fueron generados utilizando los programas Prism-GraphPad y el paquete Rstudio. Los criterios que se tuvieron en cuenta para esta selección fueron un $\text{padj} < 0,005$ donde el término “padj” se refiere al valor de ajuste de p y un valor de p ajustado (padj) $< 0,05$ indica que la probabilidad de que el resultado observado sea un hallazgo falso es inferior al 5% y un $\log_2(\text{Fold Change}) > 1$ que hace referencia al cambio de expresión génica mayor a 2 veces.

2.5. Obtención de los lncRNA seleccionados por RT-PCR.

2.5.1. Extracción de RNA total de *Stenocereus queretaroensis*

La extracción de RNA total se realizó usando el protocolo de Chomczynski y Sacchi (1987). Previamente a la extracción se filtraron las células rojas del cultivo de *Stenocereus queretaroensis* y se pesaron 0.1 g de estas, las cuales fueron pulverizadas con nitrógeno líquido.

La muestra se colocó en tubos de microcentrífuga de 1.5 mL y se le añadieron 750 µL del reactivo Trizol (Life Technologies™). Se homogeneizó manualmente con 300 µL de cloroformo y se incubó en hielo durante 3 min. A continuación, las fases se separaron centrifugando a 12 000 rpm durante 15 min a 4°C, la fase acuosa se recuperó en un tubo nuevo al que se le añadieron 500 µL de isopropanol, se agitó e incubó a temperatura ambiente durante 10 min. La mezcla se centrifugó a 12 000 rpm durante 15 min a 4°C, desechándose cuidadosamente el sobrenadante. Se lavó la pastilla de RNA en 1000 µL de etanol al 70%, se agitó en vortex y se centrifugó a 7500 rpm durante 5 min. Se realizaron 3 lavados consecutivos de la pastilla, añadiendo 100 µL de etanol al 70% y centrifugando a 8000 rpm durante 5 min a 4°C cada vez. Finalmente, la pastilla se resuspendió en 300 µL de agua con dietilpirocarbonato (DEPC) y se almacenó hasta su uso.

2.5.2. Diseño de oligonucleótidos

Se diseñaron oligonucleótidos específicos para las secuencias seleccionadas (Sec_1_PA, Sec_2_PA, Sec_3_PIP2 y Sec_4_PIP2) (Cuadro 2.1) utilizando la página web Primer designing tool (nih.gov) y la secuencia completa del transcrito. Las secuencias completas fueron tomadas de la base de datos de la librería RNA-seq de *Stenocereus queretaroensis*. (Anexo.2).

Cuadro 2.1. Características de los cebadores específicos de las secuencias seleccionadas.

Secuencias Primer	Longitud	T°C	[GC %	Amplifican
Fw_Sec_1_PA 5'- AGC TCA TAC ATG GAA GGC GG -3'	20	57.2 °C	55%	230
RV_Sec_1_PA 3'- GCT CAT GAT CCC ACG GAT GT -5'	20	57.2 °C	55%	230
FwSec_2_PA 5'- CGG TAG GAT TTG GGG TTG GT -3'	20	57.2 °C	55%	220
RV_Sec_2_PA 3'- ACG AGC AAG AGC TTC GGA TT -5'	20	56.7 °C	50%	220
FwSec_2_PIP2 5'- GCA TTG TTC GGT <u>GGT</u> GCT TT -3'	20	56.8 °C	50%	231
RvSec_2_PIP2 3'- GGC ACA CTG TCA CCG AAG TA -5'	20	57.1 °C	55%	231
FwSec_3_PIP2 5'- TTG CAG GTG CAT GTA GGG AG -3'	20	57.4 °C	55%	426
RVSec_3_PIP2 3'- CTG GCG AAC AGA CTG GGA TT -5'	20	57.4 °C	55%	426

2.5.3 Generar cDNA específico.

La obtención de cDNA se realizó utilizando el kit comercial transcriptasa reversa (RT) SuperScript™ II (Invitrogen Cat. No. 18064-022), para ello primero añadimos a un tubo de microcentrífuga libre de nucleasas: 2 pmole del oligonucleótido reverso específico de cada secuencia (GSP) (1 µL), 5 µg de RNA total (2 µL), 1 µL de dNTPs Mix (10 mM cada uno) y 8 µL

de agua destilada estéril, alcanzando un total de 12 μ L. Seguidamente la mezcla se incubó a 65°C durante 5 minutos, posteriormente se puso en hielo para enfriar rápidamente. Luego se recoge el contenido del tubo por centrifugación leve y se le añade: 2 μ L de amortiguador RT 10X, 2 μ L 0.1M DTT y 1 μ L de RNAasa OUT™ y se mezcló suavemente el contenido del tubo. Posteriormente se incubó a 42°C durante 2 min, se añadió 1 μ L (200 unidades) de SuperScript™ II RT y se mezcló pipeteando suavemente. Finalmente se incubó a 42°C durante 50 min y se calentó a 70°C durante 15 min para inactivar la reacción

2.5.4 Amplificación de secuencias por RT-PCR

La amplificación de las secuencias Sec_1_PA, Sec_2_PA, Sec_3_PIP2 y Sec_4_PIP2, se realizó a partir de cDNA previamente sintetizado; para ello se utilizó el paquete DreamTaq PCR Máster Mix (Pub. No MAN0012702) de Thermo Scientific el cual contiene una mezcla de amortiguador, enzima DreamTaq DNA polimerasa, MgCl₂ y dNTPs. El RT-PCR se realizó con un volumen de reacción de 25 μ L: 12.5 μ L de DreamTaq Máster Mix (2x), 2.5 μ L de cada uno de los cebadores forward y reverse, 2 μ L de cDNA y 5.5 μ L de agua. Las condiciones experimentales utilizadas fueron 24, 26, 28 y 30 ciclos de reacción todos a una temperatura de alineamiento de 57°C. Una vez terminado el proceso se corrieron las muestras en un gel de agarosa 1X a 95V por 30 minutos.

2.6 Clonación de los lncRNAs amplificados

Las secuencias previamente amplificadas fueron ligadas al vector de clonación pJET1.2 un vector de clonación linealizado, que acepta insertos de 6 pb a 10 kb. El kit de clonación utilizado fue el PCR CloneJET un sistema avanzado de selección positiva con una alta eficiencia de clonación de productos de PCR generados por Thermo Scientific™ DreamTaq™ DNA polimerasa pueden ligarse directamente con el vector en sólo 5 minutos. Todas las cepas comunes de *E. coli* de laboratorio pueden transformarse directamente con el producto de ligación. Lo primero que se realizó fue la reacción de ligación, para ello se añadió en un microtubo Eppendorf: 2 μ L de buffer 2X, 1 μ L de vector pJET1.2 (50ng/ μ L), 4 μ L de producto de PCR, 1 μ L T4 DNA ligasa y 2 μ L de agua libre de nucleasas, para un volumen total de 10 μ L; posteriormente se puso en el vortex brevemente y se centrifugó de 3 a 5 segundos, se incubó por 5 min a temperatura ambiente quedando lista la reacción de ligación.

Luego se pasó a la transformación de las células de *E. Coli* DH5 α con la reacción de ligación, para ello se añadieron 10 μ L de la reacción de ligación en 50 μ L de células competentes *E. Coli* DH5 α (en hielo), las cuales se sometieron a choque térmico a 42°C por 2 min y luego se mantuvo

en hielo por 5 min, después se le añadió 1 mL de medio LB sin ampicilina y se incubó a 37° C por una hora con agitación constante de 200 rpm. Transcurrido el tiempo de inoculación se plaquearon 100 µL de bacterias transformadas en cajas Petri con medio LB con ampicilina sólido y se incubó durante toda la noche a 37° C

El siguiente paso fue la extracción del DNA plasmídico (Miniprep), para ello se picaron las colonias y se inocularon en 3 mL de medio LB líquido con ampicilina 100 µg mL⁻¹ y se dejaron incubando toda la noche a 37° C con agitación a 200rpm. Al día siguiente se centrifugó a 12 000 rpm por un min a temperatura ambiente. Luego la pastilla bacteriana fue resuspendida con 200 µL de la suspensión de lisis alcalina I (glucosa 50nM, Tris Base pH:8 25 mM, EDTA pH:8 10mM) y se mantuvo en hielo. Posteriormente se agregaron 200 µL de la solución de lisis alcalina II (NaOH 0,2 N, 1% w/v, recién preparada) a cada suspensión bacteriana y se homogenizó por suspensión, luego se mantuvo a temperatura ambiente por 5 min y seguidamente en hielo por 2 min. Después se le añadieron 300 µL de la solución fría de lisis alcalina III (acetato de potasio 3 M, ácido acético glacial 2. 01 M), se homogenizó por inversión y mantuvo en hielo por 15 min.

Por último, el lisado bacteriano obtenido fue centrifugado a 12, 000 rpm por 10min, el sobrenadante fue transferido a un tubo estéril, al cual se le añadieron 700 µl de isopropanol, se mezcló por inversión y se mantuvo a temperatura ambiente por 2 min. Luego se centrifugó a 12 000 rpm por 10 min hasta que se obtuvo una pastilla, se descartó el sobrenadante, la pastilla fue lavada dos veces con etanol al 70 % y recuperada a 12 000 rpm por 2 min. Se descartó el sobrenadante y la pastilla se secó y se disolvió en 50µL de agua libre de nucleasas.

2.7 Protocolo de unión de RNA-lípido.

Para el análisis de la unión las secuencias de interés: Sec_1_PA, Sec_2_PA, Sec_3_PIP2 y Sec_4_PIP2, se realizaron ensayos de pull-down utilizando perlas de agarosa recubiertas con los lípidos PA y PIP2 y perlas de control sin ellos. Las perlas utilizadas para los pull-down descritos en este estudio son las perlas de control #P-B000, las perlas recubiertas de PI (4,5) P2 #P-B045a (también descritas como PIP2) y las perlas recubiertas de ácido fosfatídico (PA) #P-B0PA de Echelon Biosciences Inc. Las microesferas de control son microesferas de agarosa bloqueadas utilizadas como control negativo, las microesferas recubiertas de lípidos están compuestas de agarosa con 10 nanomol de lípido unido por 1 mL de microesferas y el tamaño oscila entre 45 y 165 micrómetros de diámetro. Para el ensayo realizamos un paso previo de limpieza, se añadieron 200 µg de ARN a 100 µl de perlas de control con 1 ml de RNase Block

(Agilent Technologies) durante 5 min de incubación. A continuación, 100 µl de perlas recubiertas de lípidos se lavaron previamente con PBS RNase Free y se cambiaron a tubos limpios, y el ARN previamente aclarado se añadió a las perlas limpias y se incubó durante 30 min en rotación en cámara fría; este procedimiento se realizó para cada tipo de perlas recubiertas de lípidos y perlas de control sin lípidos. Los tubos de microcentrifuga se centrifugaron a 1000 g durante 1 minuto para eliminar el ARN no unido, se lavaron tres veces con 500 µl de PBS RNase Free y se repitió el proceso de centrifugación para evitar coger las perlas y cambiar de tubo durante cada lavado. Para la extracción de ARN unido al fosfoinosítido (PI), se añadieron 100 µl de PBS RNase Free, después se añadieron 200 µl de Ribozol y 200 µl de cloroformo, se mezclaron y se centrifugaron durante 10 min a 14000 x g. Por último, se tomó la parte superior y se precipitó con 700 µl de isopropanol por 1 ml del volumen resultante, y el pellet se lavó con etanol y se almacenó a -80 °C. Todas las soluciones declaradas como libres de ARNasa se añadieron con RNase Block (Agilent Technologies) 1µl por ml. El reactivo Ribozol™ es una solución monofásica de fenol utilizada en el protocolo de extracción de ARN.

2.7. Predicción de estructura secundaria y posibles sitios de unión a PA y PIP2

La estructura secundaria de los lncRNAs seleccionados se predijo utilizando los software RNAfold (RNAfold web server (univie.ac.at)) y RNAstructure (Welcome to the Mathews Lab RNAstructure Web Servers (rochester.edu)), los cuales predicen la estructura de mínima energía libre (MFE) utilizando un modelo de energía basado en bucles y el algoritmo de programación dinámica introducido por (Gruber *et al.*, 2008) además proporciona la posibilidad de calcular las probabilidades de emparejamiento de bases para cada posible par de bases; conocer la estructura secundaria de los lncRNAs es importante para el estudio de su interacción con socios moleculares.

Para conocer los posibles sitios de unión a lípidos dentro de las estructuras predichas se realizó un alineamiento utilizando el programa Jalview, donde se compararon las secuencias de lncRNA seleccionadas con los sitios de unión reportados por Lin *et al.*, (2017) y Donia *et al.*, (2019).

2.7. Análisis filogenético

El análisis filogenético se realizó primeramente haciendo una búsqueda en BLAST, luego se utilizó los programas de alineamientos múltiples Jalview y Molecular Evolutionary Genetic Analysis (MEGA) para la construcción del árbol filogenético, y conocer de esta manera cuan conservadas están los lncRNAs.

2.8 Motivos y estructura secundaria conservados en los lncRNAs

Para determinar si en los lncRNAs hallados existían IEs e IIs conservados, se hizo un alineamiento múltiple de las secuencias de interés utilizando el programa Jalview con los IEs investigados experimentalmente reportados en lncRNAs como: BCAR4, Bvht y Xist.

Capítulo III

Resultados

3.1. Construcción de una librería de RNA-seq de *Stenocereus queretaroensis*.

3.1.1 Preparación de muestras (*Stenocereus queretaroensis*) para la secuenciación.

La preparación, construcción y secuenciación de la librería de lncRNA-seq (mRNA/lncRNA) de la interacción RNA-lípido se realizó a partir de la extracción del RNA total de células en suspensión de Pitaya (*Stenocereus queretaroensis*) como se explicó en la sección 2.1 y 2.2 de Materiales y Métodos. En la **Figuras 3.1 y 3.2** se muestran los resultados de los cultivos en suspensión de Pitaya, y la evaluación de la calidad y cantidad del RNA total extraído utilizando para esto el equipo Nanodrop.



Figura 3.1. Cultivo de células rojas de *Stenocereus queretaroensis*.

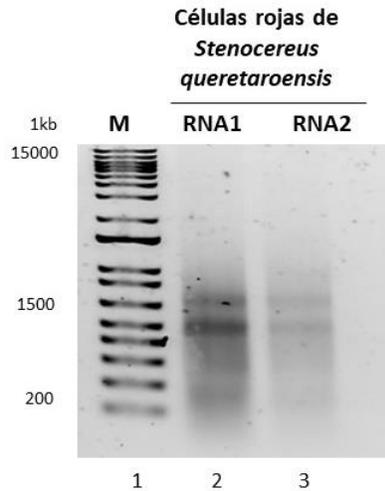


Figura 3.2. Extracción de RNA total de *Stenocereus queretaroensis*

En la **Figura 3.3.** se muestra el flujo de trabajo desde la preparación de las muestras hasta la obtención de los datos de RNA-seq.



Figura 3.3. Representación esquemática del flujo de trabajo desde la preparación de las muestras hasta la obtención de los datos. Comenzando con la extracción de RNA de cultivos celulares de *Stenocereus queretaroensis* (1), seguido del prelavado del RNA para deshacerse del RNA de unión inespecífica (2), el ensayo de pull-down con perlas de sefarosa como control y perlas recubiertas de los lípidos: ácido fosfatídico y fosfatidilinositol 4, 5 bisfosfato (3), el lavado de cualquier RNA no ligado y la repetición de este paso dos veces (4), y finalmente la recuperación del RNA ligado a lípidos mediante extracción con TRizol (5). Las muestras se enviaron a Novogene Co para el procesamiento de secuenciación de lncRNA (6) y para la obtención de la biblioteca de RNA (7) y la obtención de datos utilizados para análisis posteriores (8).

3.2. Obtención de datos de la librería de RNA-seq de genes de *S. queretaroensis*

Se obtuvo una librería con el número de ID H202SC21112222 (**Anexo. 3**) de genes pertenecientes a la planta *Stenocereus queretaroensis* expresados bajo las condiciones experimentales propuestas presencia/ausencia de PIP2/PA. Esta biblioteca fue construida siguiendo la metodología expuesta en la sección 2.3 de Materiales y Métodos haciendo usos de los servicios de secuenciación por Novogene Company.

Luego de realizada la secuenciación utilizando la plataforma Illumina Novaseq 6000 con una longitud de lectura paired-end de 150 pb (dos lecturas de 150 pb de bases cada una) se obtuvieron los datos crudos y se realizó un control de la calidad para obtener las lecturas limpias para cada una de las condiciones bajo los criterios de calidad establecidos por Novogene expuestos en la sección 2.3 de Materiales y Métodos, garantizando de esta manera la calidad y fiabilidad del análisis bioinformático. En el **Cuadro 3.1** se muestran el recuento de lecturas para cada secuenciación de los datos en crudo y limpios, datos filtrados, que son procesados por el control de calidad de datos. También se muestran los porcentajes de la tasa de error de bases de toda la secuenciación (% error), el porcentaje de GC (porcentaje de números de bases G&C del total de bases (Número de bases G&C) / (Número total de bases) *100.) y la puntuación de calidad de secuenciación de una base determinada como $Q = -10\log_{10}(e)$, donde e es la probabilidad estimada de que la llamada de la base sea errónea. Una puntuación de calidad de 20 (Q20) representa una tasa de error de 1 en 100, con una precisión de llamada correspondiente del 99%, del mismo modo, Q30 significa una precisión de la llamada del 99,9%; donde Q20 y Q30 se calcula como: $(\text{Número de bases con valor } Q \text{ Phred} > 20/30) / (\text{Número de bases totales}) * 100$ siendo el porcentaje de bases cuyo valor de Q Phred es superior a 20/30.

Cuadro 3.1. Datos crudos y limpios de la secuenciación.

Muestra	ID_librería	Lecturas_crudas	Bases_crudas	Lecturas_limpias	Bases_limpias	% error	Q20	Q30	% GC
SQ_total	BRRL210021516-1A	45014736	13.5	43504328	13.1	0.03	96.81	91.2	45.24
SQ_PA	FRRL210235793-1a	12383600	3.7	7371455	2.2	0.03	96.83	91.49	51.62
SQ_PIP2	FRRL210235794-1a	18459588	5.5	18065516	5.4	0.03	97.27	92.69	52.45
SQ_control	FRRL210235792-1a	14278063	4.3	13962678	4.2	0.03	97.06	92.36	55.37

Posteriormente se realizó la reconstrucción del transcriptoma necesario para muestras sin genoma de referencia, para esto las lecturas limpias fueron ensambladas *de novo* utilizando el software Trinity.fasta, unigene.fasta y cluster.fasta (Grabherr *et al.*, 2011) para obtener el transcriptoma ensamblado. Luego se realizó una agrupación jerárquica para eliminar redundancias utilizando Corset, donde los transcritos más largos de cada cluster serán

seleccionados como unigenes. En el **Cuadro 3.2** se observan el número de transcritos y unigenes en diferentes intervalos de longitud.

Cuadro 3.2. Resumen del número de transcritos y unigenes en diferentes intervalos de longitud

Transcript length interval	300-500bp	500-1kbp	1k-2kbp	>2kbp	Total
Number of transcripts	53946	46365	35601	27503	163415
Number of Unigenes	34103	25868	13031	9134	82136

Seguidamente se realizó el análisis de los resultados de empalme Trinity.fasta, unigene.fasta y cluster.fasta para ello se utilizó Busco (benchmarking universal de originales de copia única) para evaluar la calidad, precisión e integridad de los resultados de empalme. En la **Figura 3.4** se muestran los resultados de la evaluación de los transcritos empalmados (diferentes colores representan diferentes tipos de transcritos empalmados), donde n representa el total de grupos de BUSCO buscados, S BUSCOs completos de una sola copia, D : BUSCOs duplicados completos, F : BUSCO fragmentados y M : BUSCOs ausentes

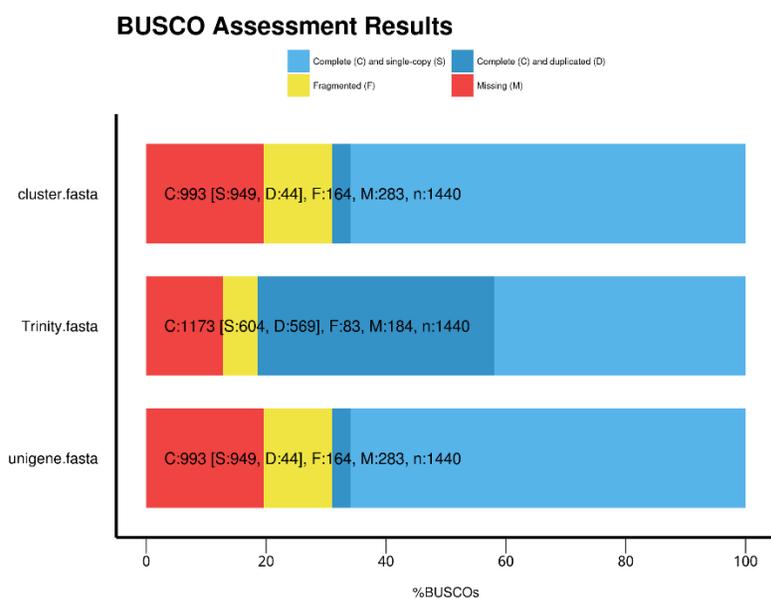


Figura 3.4. Resultados de la evaluación de Busco de las transcripciones empalmadas

3.3. Anotación funcional de genes

Se obtuvieron un total de 82 136 genes con anotación funcional completa, para esto Novogene utilizó siete bases de datos. En el **Cuadro 3.3** se representan los genes anotados con éxito por cada base de datos: se obtuvieron 32 333 genes anotados en la base de datos de secuencias de proteínas no redundantes del NCBI (Nr), 18 432 en la base de secuencias de nucleótidos del NCBI (Nt), 9 809 en la base de datos de genes de ortólogos (KO), 23 6667 y 23 886 en las bases

de datos de proteínas Swiss-Prot y PFAM respectivamente, 23 885 para la base de ontología de genes (GO) y 6 450 en la base de grupos de ontología de eucariotas (KOG). Además, resultaron un total de 3 548 genes anotados en todas las bases de datos y 39 989 genes anotados al menos en una base de datos.

Cuadro 3.3. Representación funcional de los genes anotados con éxito por cada base de datos

Genes anotados en bases de datos	Número de Genes	Porcentaje de genes anotados(%)
Anotados en NR	32333	39.36
Anotados en NT	18432	22.44
Anotados en KO	9809	11.94
Anotados en SwissProt	23667	28.81
Anotados en PFAM	23886	29.08
Anotados en GO	23885	29.07
Anotados en KOG	6450	7.85
Anotado en todas las bases de datos	3548	4.31
Anotado en al menos una base de datos	39989	48.68
Total Unigenes	82136	100

3.4. Análisis de genes expresados diferencialmente

El análisis de genes expresados diferencialmente resultó en un total de 22 348 secuencias del total de genes anotados que fueron 82 136 como se había mencionado en la sección anterior. En el **Cuadro 3.4** se muestran el número de genes expresados diferencialmente (incluidos los genes regulados al alza y baja). Este análisis se realizó utilizando la herramienta de expresión génica edgeR que nos permitió identificar los genes que se mostraban diferencialmente expresados para cada uno de los tratamientos (presencia/ausencia de perlas recubiertas de PA/PIP2), los criterios de selección para este análisis fueron un $qval < 0.005$ y un $|\log_2 \text{FoldChange}| > 1$.

Cuadro 3.4. Visualización de los resultados estadísticos de los genes diferenciales.

Comparación entre tratamientos	Total de genes	up	down	Criterios de selección
SQ_PAvsSQ_control	15397	11634	3763	edgeR qval<0.005 log2FoldChange >1
SQ_PAvsSQ_PIP2	16324	8196	8128	edgeR qval<0.005 log2FoldChange >1
SQ_PIP2vsSQ_control	13644	10760	2884	edgeR qval<0.005 log2FoldChange >1

3.4.1 Análisis de diagramas de Venn de DBG

El análisis de los diagramas de Venn de los DBG nos permitió analizar la especificidad y posible interacción de los RNA en estudio por PA y PIP2 o bien por ninguno de estos lípidos. En la **Figura 3.5.A** se observa que, de un total de 23 915 secuencias, 6598 secuencias fueron específicas para PA (cian) mientras 5601 resultaron para PIP2 (púrpura) y 1341 no resultaron específica para ninguno de los lípidos (amarillo). El solapamiento para control-PA y control-PIP2 fue de 924 y 1748 secuencias completas de un total de 18 314 y 17 317, respectivamente, estas secuencias pueden representar ligantes no específicos o bien RNA en grandes cantidades que no se pueden seleccionar fácilmente. En cuanto al solapamiento PA-PIP2 resultó en 4531 secuencias de un total de 22 574, que mostraban especificidad para ambos lípidos esto resulta importante para ensayos de competencia, que nos permita conocer cuantas de estas secuencias se muestran mayor especificidad por un lípido u otro, o también pueden indicar RNA que no pueden distinguir la unión a estos dos tipos de lípidos.

En las **Figuras 3.5.B** y **3.5.C** se observa de manera más detalla la especificidad de 11129 secuencias para PA (cian), y 10132 secuencias para PIP2 (púrpura) respecto al Control (amarillo), mientras que 4096 y 4920 secuencias, para PA y PIP2 mostraron solapamiento con el control, lo que puede ser debido a poca afinidad hacia el fosfolípido específico.

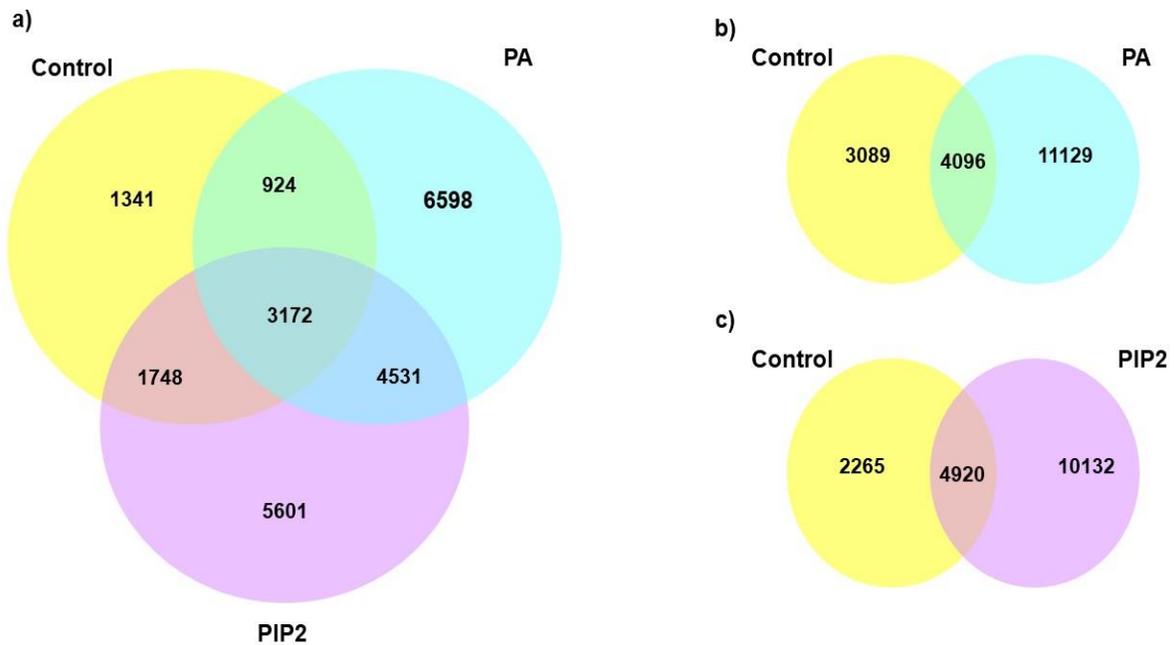


Figura 3.5. Diagramas de Venn de los DBG **A)** Se muestra la distribución de los genes presentes en el estudio con la interacción el control y las perlas recubiertas de lípidos (PA y PIP2). **B)** Se muestra la distribución de la interacción de las secuencias respecto al ácido fosfórico. **C)** Se muestra la distribución de la interacción de las secuencias respecto al fosfatidilinositol 4,5 bisfosfato

3.4.2. Análisis de heat maps de los DBG

El análisis de los mapas de calor (heat maps) de los DBG nos permitió observar que en nuestra librería existen genes con patrones de expresión diferenciales ante las condiciones experimentales propuestas. En la **Figura 3.6** se observa una agrupación jerárquica en los RNA (mRNA y lncRNA) asociados a las perlas recubiertas de lípidos. Las áreas de diferentes colores denotan distintos grupos (agrupaciones) de genes, y los genes dentro de cada agrupación pueden tener funciones similares o participar en el mismo proceso biológico. Las zonas de color rojo más alta indican una población enriquecida de los RNA para una condición particular entre las distintas condiciones, lo que a su vez refleja una unión preferencial por un lípido específico (PA/PIP2), demostrando que la unión no se debía a una precipitación no específica con las microesferas, sino al tipo particular de microesferas utilizadas.

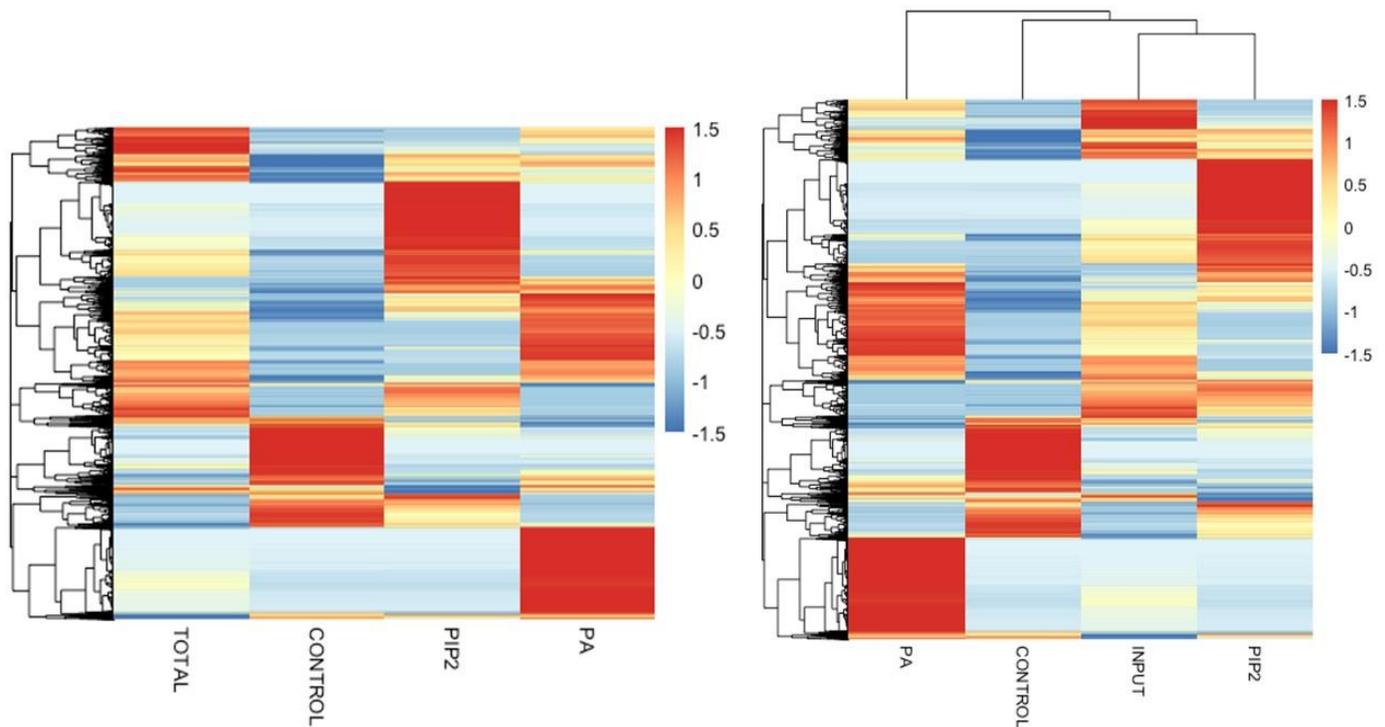


Figura 3.6. Heat maps de los DBG generado con los datos de RNA-seq a partir de las células de Pitaya muestra el patrón de enriquecimiento distintivo de los lncRNAs que interactúan en el ensayo pulldown con control (perlas de sefarosa sin lípidos) y perlas recubiertas de lípidos (PA y PIP2). Los resultados globales del análisis de clusters FPKM, utilizando el valor $\log_{10}(\text{FPKM}+1)$. El rojo denota genes muy enriquecidos y el azul genes con niveles de enriquecimiento bajos. La gama de colores del rojo al azul representa el valor $\log_{10}(\text{FPKM}+1)$ de mayor a menor.

3.4.3. Análisis de gráficos de Volcano de DBG

El análisis de volcán de los DBG nos permitió cuantificar los genes significativamente expresados de forma diferencial ante la presencia de cada lípido respecto al control. En la **Figura 3.7** se observa en el gráfico de la izquierda que 11 634 secuencias con unión más abundante y preferente por PA mientras 3763 muestran menor afinidad por este lípido. En el gráfico de la derecha correspondiente a la condición presencia de perlas recubiertas con PIP2 se observa una abundancia similar donde 10 760 muestran especificidad por este lípido respecto a las 2884 que lo hacen menos. Los criterios tomados en cuenta para definir genes significativamente expresados de forma diferencial (DBG) fueron un $\text{padj} < 0,005$ y $\log_2(\text{Fold Change}) > 1$. El valor de p ajustado (padj) es una medida de la significación estadística de este resultado en el análisis de expresión génica. Un valor de $\text{padj} < 0,005$ indica que la probabilidad de que el resultado sea un falso positivo es menor al 0,5%. Mientras el \log_2 del Fold Change se utiliza para expresar el cambio en la expresión génica entre dos condiciones, un $\log_2(\text{Fold Change}) > 1$ indica que la

expresión génica se ha incrementado en más de dos veces en la condición experimental en comparación con la condición de control.

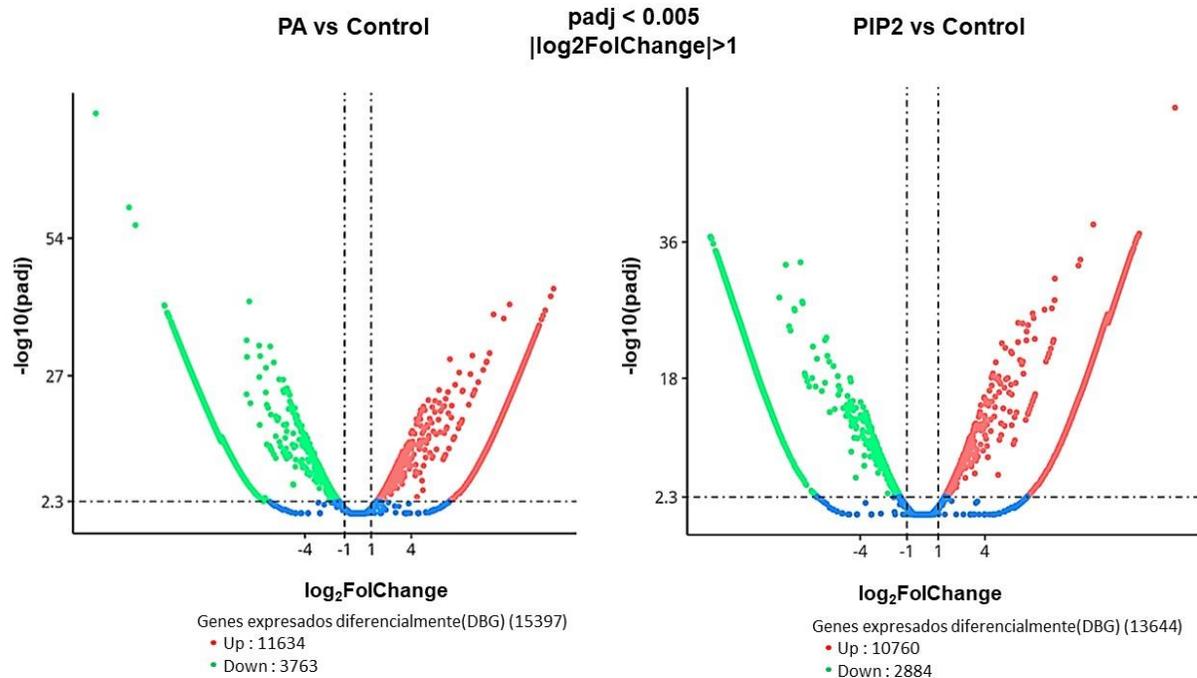


Figura 3.7. Diagramas de volcán de los DBG muestran los genes enriquecidos diferencialmente entre las perlas de sefarosa de control y las recubiertas de PA y PIP2. Los puntos rojos representan los genes más enriquecidos diferencialmente y los verdes los menos. El eje horizontal indica el cambio de pliegue de los genes y la coordenada vertical indica el grado estadísticamente significativo de los cambios en los niveles de enriquecimiento de los genes. Los criterios significativos de los genes diferencialmente enriquecidos son $\log_2(\text{FoldChange}) > 1$ y $\text{padj} < 0,005$.

3.4.4 Análisis de la ontología de genes (GO) y vías KEGG de los DBG

El análisis de las vías de enriquecimiento GO nos permitió determinar la sobreexpresión y abundancia de genes en una categoría biológica específica: proceso biológico (BP), componente celular (CC) y función molecular (MF) ante las condiciones experimentales del estudio. Además, a través de estos análisis se podría deducir la funcionalidad de los RNA de nuestra librería

En la **Figura 3.8** se muestran los principales BP, CC y FM en que se observan más abundantes y enriquecidos los genes ante la presencia de perlas recubiertas con PIP2 en reacciones y vías químicas como aquellas en las que intervienen diversos compuestos nitrogenados; las que dan lugar a la formación de sustancias; normalmente, la parte del metabolismo que requiere energía y en la que las sustancias más simples se transforman en otras más complejas, la formación de

moléculas pequeñas y transporte transmembrana. Otros resultados visibles en términos GO son la presencia de genes específicos en componentes celulares como organelos intracelulares, ribosomas, citoplasma y complejos macromoleculares. Si hablamos en términos de funciones moleculares se ven favorecidas las reacciones de oxidación-reducción (redox), la integridad estructural de un complejo y ribosoma; en la catálisis de la reacción: $GTP + H_2O$ y unión a proteínas desplegadas.

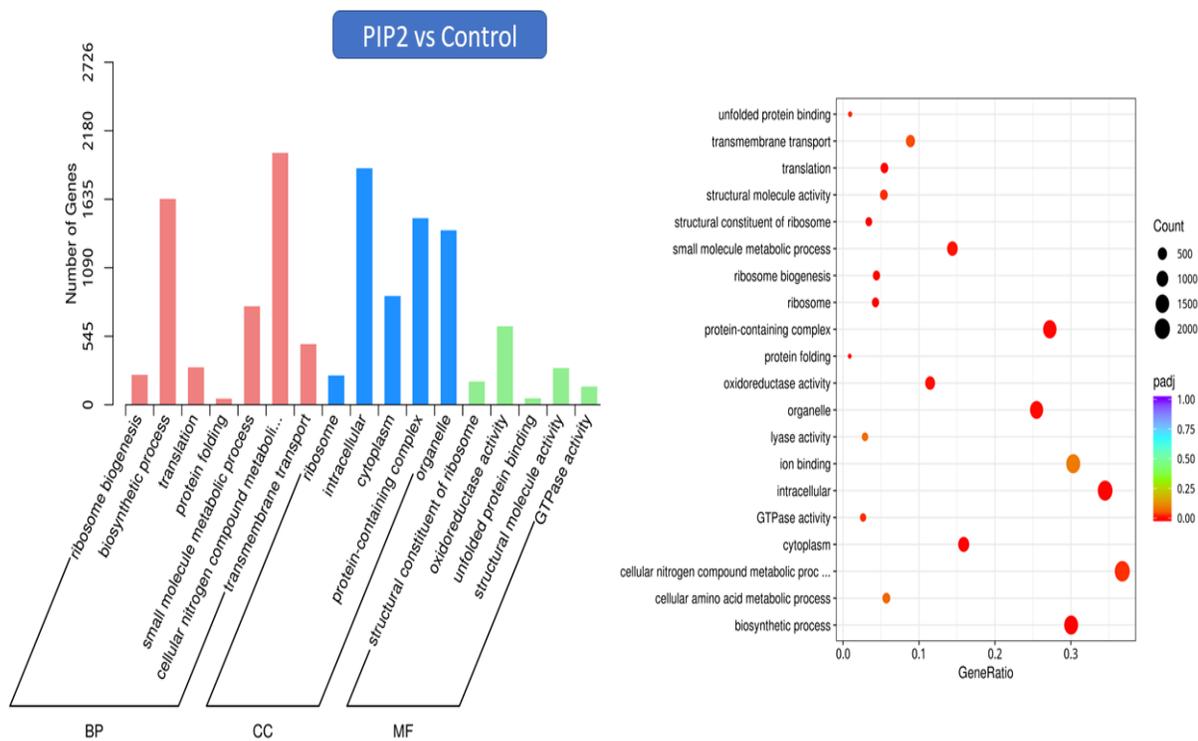


Figura 3.8. Diagramas de GO muestran los genes enriquecidos diferencialmente para las perlas de sefariosa de control y las recubiertas de PIP2, se observa los cinco principales procesos biológicos (BP), componentes celulares (CC) y funciones moleculares (MF), que presentan mayor número de genes. El rojo denota genes muy enriquecidos y el azul genes con niveles de enriquecimiento más bajos. La gama de colores del rojo al azul representa el valor padj de mayor a menor.

A continuación, en la **Figura 3.9** se muestran de igual manera la cantidad de genes y su expresión ante la segunda condición experimental la presencia de perlas recubiertas con PA, observándose que el proceso metabólico más favorecido y en el que participan mayor cantidad de genes es el metabolismo de moléculas pequeñas; de igual manera que para PIP2 los

componentes celulares que presentan mayor cantidad de genes son los intracelulares y los complejos de unión a proteínas así como la actividad oxidorreductasa sigue prevaleciendo.

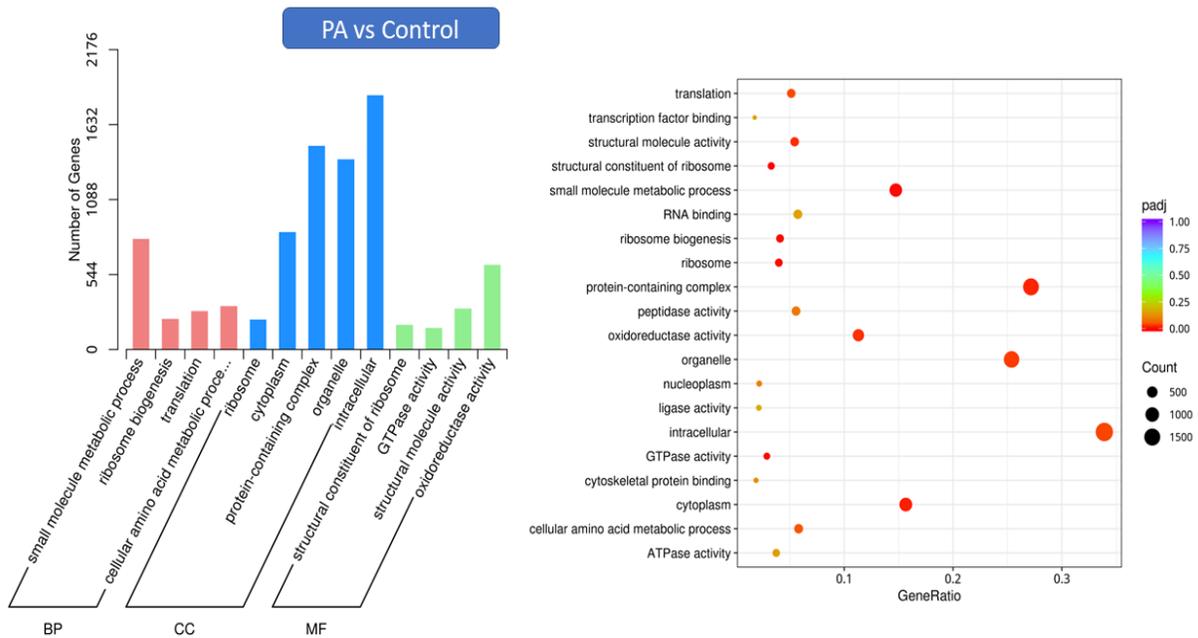


Figura 3.9. Diagramas de GO muestran los genes enriquecidos diferencialmente para las perlas de sefariosa de control y las recubiertas de PA, se observa los cinco principales procesos biológicos (BP), componentes celulares (CC) y funciones moleculares (MF), que presentan mayor número de genes. El rojo denota genes muy enriquecidos y el azul genes con niveles de enriquecimiento más bajos. La gama de colores del rojo al azul representa el valor padj de mayor a menor.

Los resultados correspondientes al análisis KEGG (**Figura 3.10**) revelaron que ante la presencia de los fosfolípidos PIP2 y PA los genes enriquecidos participaban principalmente en la vía de señalización PI3K-AKT, señalización del fosfatidilinositol, en la traducción, la regulación del ciclo celular, la diferenciación celular, las vías de transducción de señales, el procesamiento de antígenos para respuestas inmunitarias adecuadas, la señalización del estrés, las respuestas inflamatorias y la apoptosis. También se ven favorecidos los procesos de metabolismo y energía como la fosforilación oxidativa.

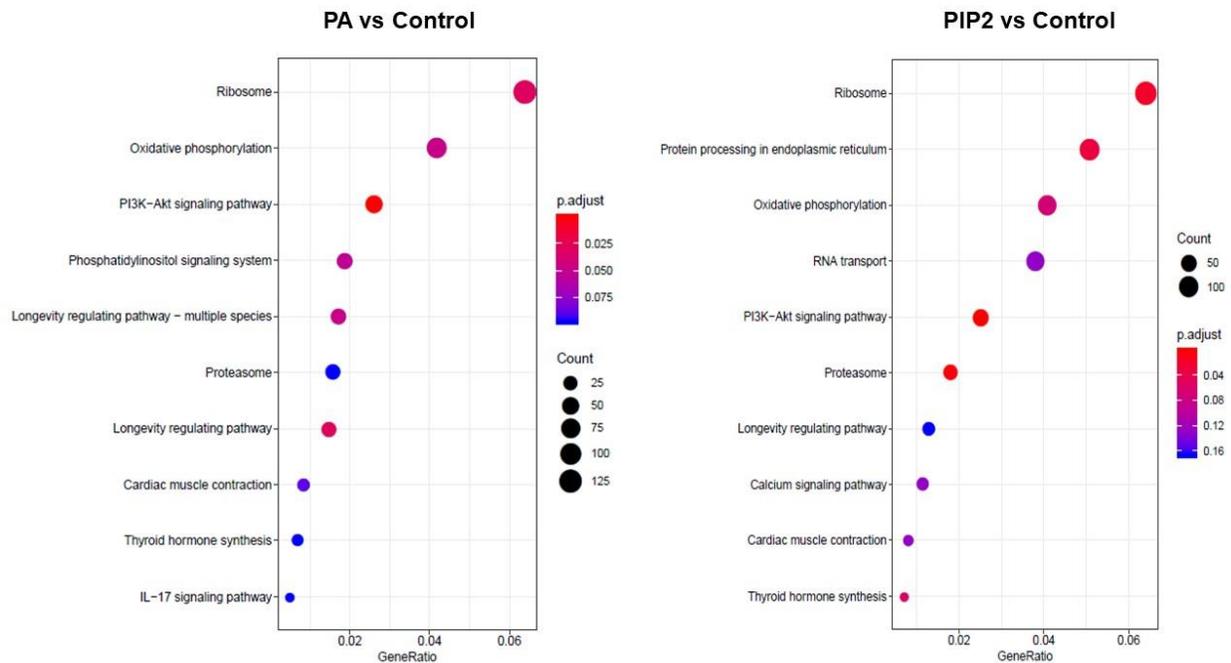


Figura 3.10. Diagramas de vías KEGG muestran los genes enriquecidos diferencialmente. La abscisa son los genes anotados, la ordenada es el término de la vía.

3.5 Análisis y selección de lncRNA

Obtenidos los genes que se encuentran expresados de manera diferencial para cada condición pasamos a la selección de aquellas secuencias que cumplieran con las características de lncRNAs. En la **Figura 3.11** se muestra una representación esquemática de los pasos seguidos para esta selección; los números representan el total de transcripciones identificadas en cada filtro. A partir de los 22 348 se descartaron los transcritos de menos de 200 pb de longitud. Los marcos de lectura abiertos (ORF) de los transcritos filtrados se analizaron mediante el servidor web Orf Predictor (Min *et al.*, 2005) (<http://proteomics.yzu.edu/tools/OrfPredictor.html>) y se descartaron las transcripciones con ORF que codifica ≥ 100 residuos de aminoácidos de cadena polipeptídica larga. Las secuencias filtradas se utilizaron para la búsqueda BLAST frente a bases de datos de proteínas NCBI-nr, NCBI-nt, Swiss-Prot y PFam (Finn *et al.*, 2014) con un valor e de 10^{-3} para eliminar las transcripciones coincidentes con cualquier proteína reportada y dominio de la familia de proteínas. Las transcripciones restantes se sometieron al cálculo del potencial de codificación mediante la Calculadora de Potencial de Codificación 2 (CPC2) (<http://cpc.cbi.pku.edu.cn>) siguiendo el procedimiento estándar (Kong *et al.*, 2007). Los transcritos con una puntuación de potencial de codificación superior a 0.05 se descartaron. Luego de aplicados los diferentes filtros seleccionamos 44 secuencias que cumplieran con características de lncRNAs. Seguidamente utilizando el programa Prism-GraphPad y el paquete Rstudio

generamos un diagrama de Venn, heat maps y gráfico de volcán de estos 44 lncRNAs para seleccionar dentro de este grupo, los lncRNA que mostraban mayor especificidad por PA o PIP2., quedando 4 secuencias de lncRNA elegidas.

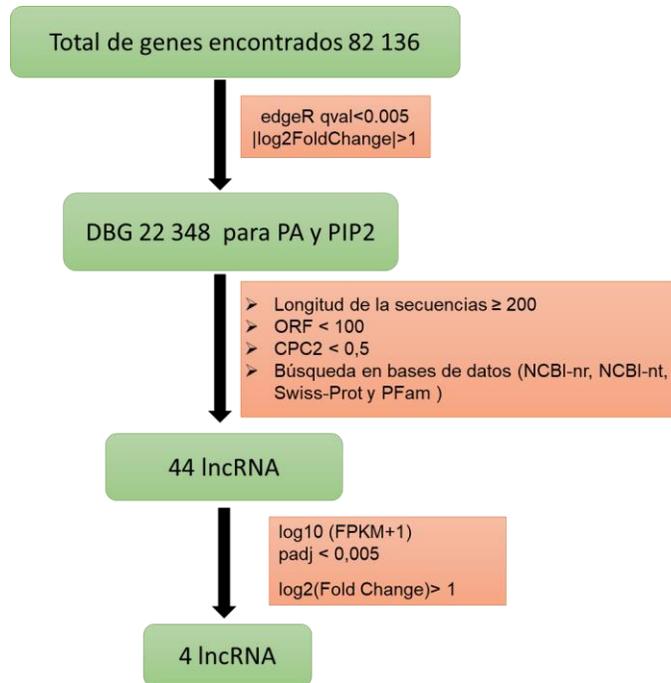


Figura 3.11. Representación del flujo de trabajo para la selección de los lncRNA. A partir del total de genes mapeados se utilizaron diferentes criterios para determinar las secuencias que cumplieran con las características de lncRNA. Los cuadros en color rosa representan los filtros mientras los cuadros en color verde la cantidad de secuencias restantes después del filtro.

La representación de los 44 lncRNAs como diagramas de Venn nos permitió cuantificar la distribución de estas secuencias entre los diferentes tratamientos y a su vez conocer la especificidad por un lípido u otro respecto al control. En la **Figura 3.12** se observa que, de un total de 44 secuencias de lncRNA, 14 fueron específicas para PA (cian) mientras 10 resultaron para PIP2 (púrpura). Además 9 secuencias no resultaron específicas para ninguno de los lípidos, de estas 5 se mostraron solo para el Control (amarillo) y 4 se encontraron tanto en Control como en PA/PIP2, lo que puede ser debido a poca afinidad hacia el fosfolípido específico. En cuanto al solapamiento PA-PIP2 resultó en 11 secuencias que mostraban especificidad para ambos lípidos esto resulta importante para ensayos de competencia, que nos permita conocer cuantas de estas secuencias se muestran mayor especificidad por un lípido u otro, o también pueden indicar RNA que no pueden distinguir la unión a estos dos tipos de lípidos

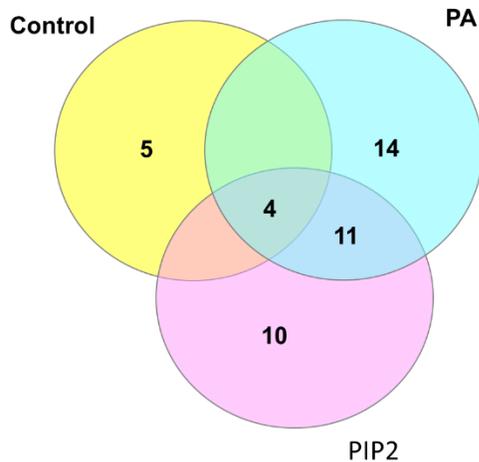


Figura 3.12. Diagramas de Venn de los 44 lncRNA. Se muestra la distribución y especificidad para cada tratamiento: Control y las perlas recubiertas de lípidos (PA y PIP2) de los 44 lncRNA seleccionados.

En el mapa de calor generado, utilizando los valores de $\log_{10}(\text{FPKM}+1)$, para los 44 lncRNA que se muestra en la **Figura 3.13** se puede observar zonas de diferentes colores entre los distintos tratamientos, esto significa que existen lncRNA que se expresan de manera diferencial para una condición respecto a la otra. Es el caso de los lncRNA identificados como Cluster-27261 y Cluster-20205.8 que mostraron especificidad para los tratamientos con perlas recubiertas con los lípidos PA y PIP2 respectivamente, mientras que otros lncRNA como los denominados Cluster-18316.2123, Cluster-18316.8708 y Cluster-18316.11481 no presentaron especificidad. Estos resultados reflejan una unión preferencial de algunos lncRNA por un lípido específico (PA/PIP2), demostrando que la unión no se debe a una precipitación no específica con las microesferas, sino al tipo particular de microesferas utilizadas. Las flechas en color azul representan los 4 lncRNAs elegidos.

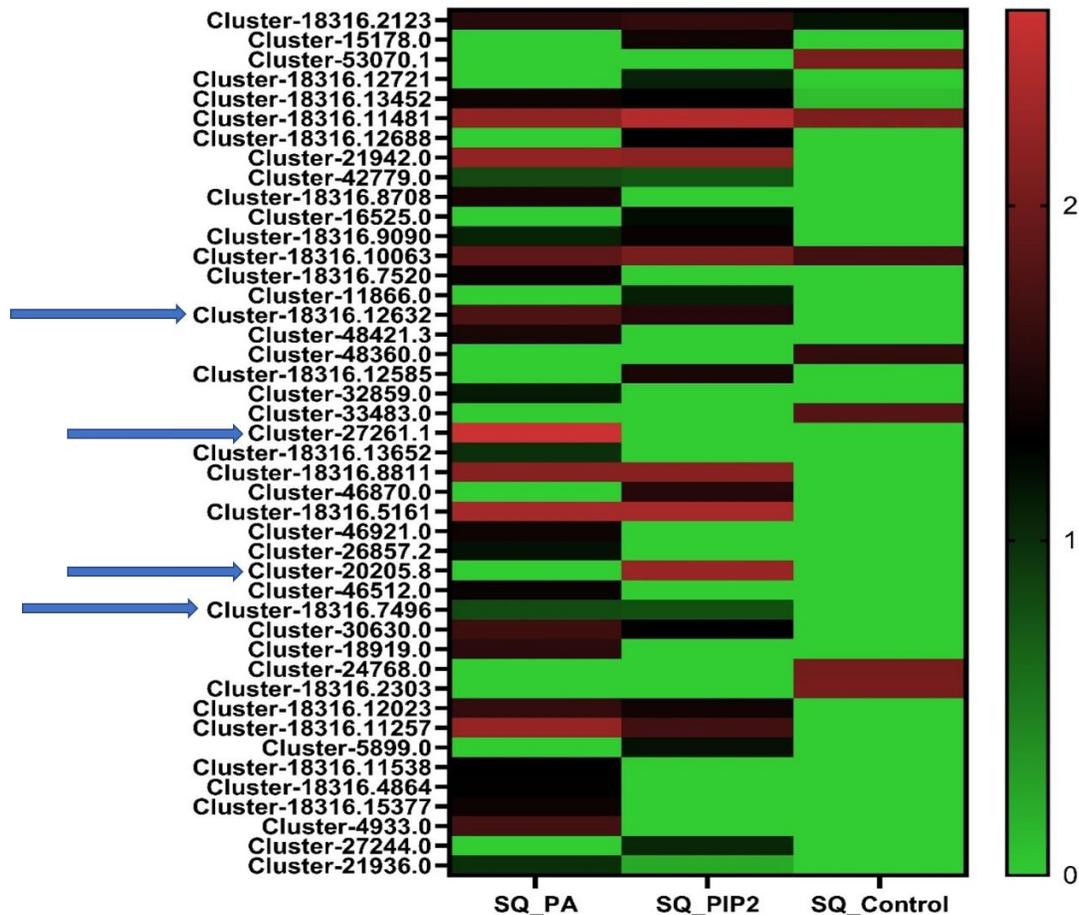


Figura 3.13. Heat maps generados con los datos de los 44 IncRNA muestran el patrón de enriquecimiento distintivo entre estos respecto al control (perlas de sefaroza sin lípidos) y perlas recubiertas de lípidos (PA y PIP2). El rojo denota genes muy enriquecidos y el verde genes con niveles de enriquecimiento bajos. La gama de colores del rojo al verde representa el valor $\log_{10}(\text{FPKM}+1)$ de mayor a menor, las flechas en color azul representan los 4 IncRNA elegidos

En la **Figura 3.14** se muestra el gráfico de volcán de los 44 IncRNA donde se observa que 39 secuencias presentaron una unión más abundante y preferente por los lípidos PA y PIP2 mientras 5 secuencias presentaron menor especificidad, cuantificando de esta manera los IncRNA que se encuentran significativamente expresados de forma diferencial para cada lípido. Los criterios tomados en cuenta fueron un $\text{padj} < 0,005$ y $\log_2(\text{Fold Change}) > 1$ de cada secuencia, en azul se muestran los genes regulados a la baja y en verde los que se encontraban regulados al alza, y dentro de estos los genes señalados con puntos rojos son los 4 IncRNA seleccionados dentro de los 44 IncRNAs encontrados en la librería.

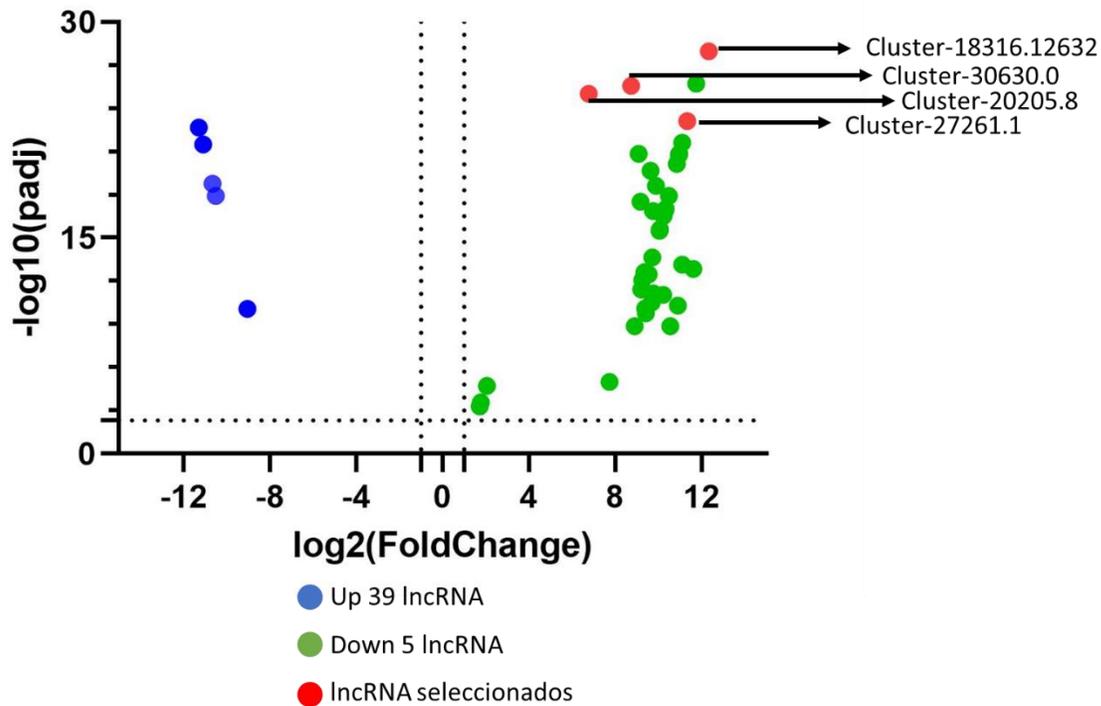


Figura 3.14. Diagramas de volcán de los 44 lncRNAs. Se representaron las 44 secuencias de lncRNA permitiendo observar dentro de este conjunto cuales genes se encontraban más enriquecidos de los genes diferencialmente entre las perlas de sefaroza de control y las recubiertas de PA y PIP2. Los puntos verdes representan los genes más enriquecidos diferencialmente, los azules los menos y los rojos las secuencias seleccionadas. El eje horizontal indica el cambio de pliegue de los genes y la coordenada vertical indica el grado estadísticamente significativo de los cambios en los niveles de enriquecimiento de los genes. Los criterios significativos de los genes diferencialmente enriquecidos son $\log_2(\text{FoldChange}) > 1$ y $\text{padj} < 0,005$.

De esta manera quedaron seleccionadas las cuatro secuencias de lncRNAs a las que llamamos Sec_1_PA (Cluster_18316.12632), Sec_2_PA(Clúster_27261.1), Sec_3_PIP2 (Cluster_20205.8) y Sec_4_PIP2(Cluster_30630.0) después del análisis de los 44 lncRNA previamente seleccionados a partir del total de genes mapeados. En la **Figura 3.12** se muestran los resultados del CPC2 para estas cuatro secuencias, quedando comprobado que se trata de ncRNA. Mientras en el **Cuadro 3.5** se muestran el ID, longitud de la secuencia, así como la posible función molecular (FM), proceso biológico (PB) y componente celular (CC) en el que pudieran estar involucrados y/o encontrarse las secuencias elegidas.

CPC 2.0 [Home](#) [Batch](#) [Download](#) [Help](#) [About](#)

Download the result Sort by ID

ID	Label	Coding probability	Peptide length(aa)	Fickett score	Isoelectric point	ORF integrity	Details
Sec_1_PA	noncoding	0.226341	122	0.3227	9.92388916016	complete	View
Sec_2_PA	noncoding	0.0581687	70	0.36168	9.51617431641	complete	View
Sec_3_PIP2	noncoding	0.057593	86	0.2663	8.44622802734	complete	View
Sec_4_PIP2	noncoding	0.145323	121	0.2816	9.50286865234	complete	View

« 1 » Go

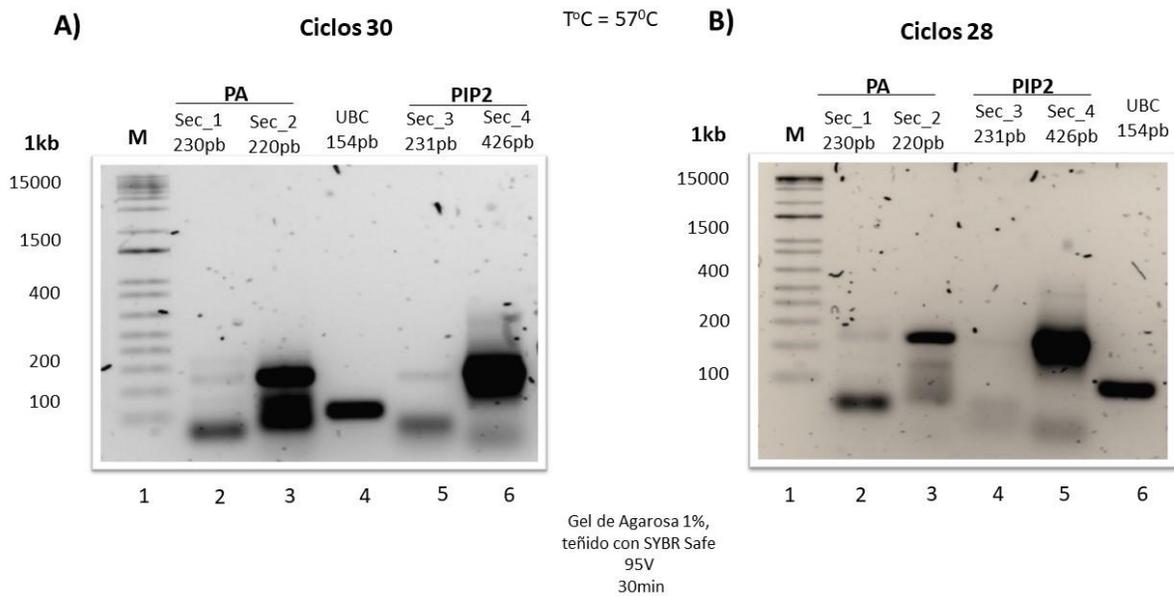
Figura 3.15. Resultados obtenidos del cálculo del potencial de codificación mediante la calculadora de potencial de codificación ([CPC2 @ CBI, PKU \(gao-lab.org\)](#)), donde se muestran características (longitud del péptido putativo, puntuación Fickett TESTCODE y punto isoelectrico putativo) de las secuencias introducidas en formato fasta obteniendo como resultado el poder de codificación.

Cuadro 3.5. Resume el ID, longitud de la secuencia y ruta biológica GO en que pueden estar involucradas las secuencia Sec_1_PA (Cluster_18316.12632), Sec_2_PA(Cluster_27261.1), Sec_3_PIP2 (Cluster_2025.8) y Sec_4_PIP2(Cluster_30630.0).

Secuencias	ID_GEN	Largo_GEN	Ruta biológica GO (PB/FM/CC)
Sec_1_PA	Cluster-18316.12632	3020	PB _Integración del ADN/regulación de la transcripción, ADN-templado FM _Actividad del factor de transcripción de unión al ADN CC _Complejo regulador de la transcripción//componente integral de la membrana
Sec_2_PA	Cluster-27261.1	664	PB _ Transporte de protones acoplado a la síntesis de ATP FM _ Actividad ATP sintasa transportadora de protones, mecanismo de rotación CC _Complejo ATP sintasa mitocondrial transportador de protones, sector catalítico F(1)//complejo ATP sintasa transportador de protones
Sec_3_PIP2	Cluster-20205.8	4100	PB _Protección de la horquilla de replicación//respuesta celular al estímulo de daño en el ADN FM _ Unión de iones de zinc CC _Núcleo
Sec_4_PIP2	Cluster-30630.0	3303	PB _ Vía de señalización de los receptores acoplados a proteínas G//respuesta a la feromona/percepción sensorial del estímulo químico FM _ actividad del receptor de feromonas del factor de apareamiento CC _membrana

3.6. Obtención de los lncRNA seleccionadas por RT-PCR.

Una vez seleccionadas nuestros 4 lncRNA, el siguiente paso fue la amplificación de estas mediante RT-PCR utilizando el cDNA sintetizado para cada una, así como los oligos diseñados, las especificaciones de estos se muestran en el **Cuadro 2.1** de la sección Materiales y Métodos. Se logró obtener las secuencias al tamaño esperado donde Sec_1_PA amplificó a 230pb, Sec_2_PA a 220pb, Sec_3_PIP2 a 231pb y Sec_4_PIP2 a 426 utilizando como control al gen de Ubiquitina (UBC) que amplifica a un tamaño de 154 pb como se observa en la **Figura 3.16 A y B** estos resultados se obtuvieron bajo las condiciones de RT-PCR de una temperatura de alineamiento de 57°C utilizando 30 y 28 ciclos. Mientras en la **Figura 3.16 C** se observa que cuando se disminuyeron la cantidad de ciclos a 26 y 24 solo amplificaron las Sec_2_PA y Sec_4_PIP2. Todas las muestras se corrieron en Gel de Agarosa 1%, teñido con SYBR Safe a 95 V por 30 minutos.



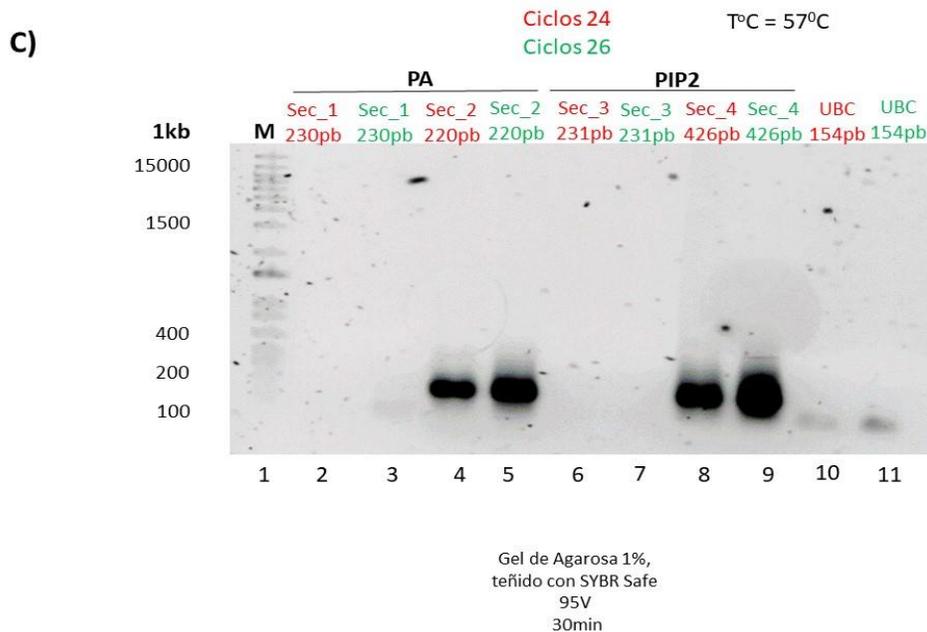


Figura.3.16. Amplificación de las secuencias por RT-PCR. A) Amplificación a 30 ciclos de las secuencias Sec_1_PA (230pb), la Sec_2_PA (220pb), Sec_3_PIP2 (231pb) y Sec_4_PIP2 (431pb). B) Amplificación a 28 ciclos de las secuencias Sec_1_PA (230pb), la Sec_2_PA (220pb), Sec_3_PIP2 (231pb) y Sec_4_PIP2 Amplificación de la Secuencia snR63 a partir de cDNA. C) Amplificación a 26 (color verde) y 24 (color rojo) ciclos de las secuencias Sec_1_PA (230pb), la Sec_2_PA (220pb), Sec_3_PIP2 (231pb) y Sec_4_PIP2. Todas fueron amplificadas a una temperatura de 57oC y corridas a 95 V por 30 min en Gel de Agarosa 1%, teñido con SYBR Safe.

3.7. Análisis del protocolo de unión de lncRNA a PA y PIP2

Obtenidas las secuencias por RT-PCR se clonaron utilizando el vector pJET1.2 y mandaron a secuenciar, comprobando que realmente eran los amplicones esperados para después realizar el ensayo de interacción lncRNA-lípidos (PA/PIP2). El análisis de la posible interacción de los lncRNA Sec_1_PA y Sec_3:PIP2 a los fosfolípidos PA y PIP2 se realizó mediante ensayos de pull-down utilizando perlas de agarosa recubiertas con lípidos (PA y PIP2) y perlas de control sin ellos, seguido de un ensayo de RT-PCR, donde se confirmó la interacción lncRNA-Lípido. Los resultados obtenidos se muestran en la **Figura 3.17** donde se observa como el lncRNA Sec_1_PA que amplifica para 230 pb se une a las perlas recubiertas con PA, de igual manera lo hace el lncRNA Sec_3_PIP2 de un tamaño de 231pb con las perlas recubiertas con PIP2.

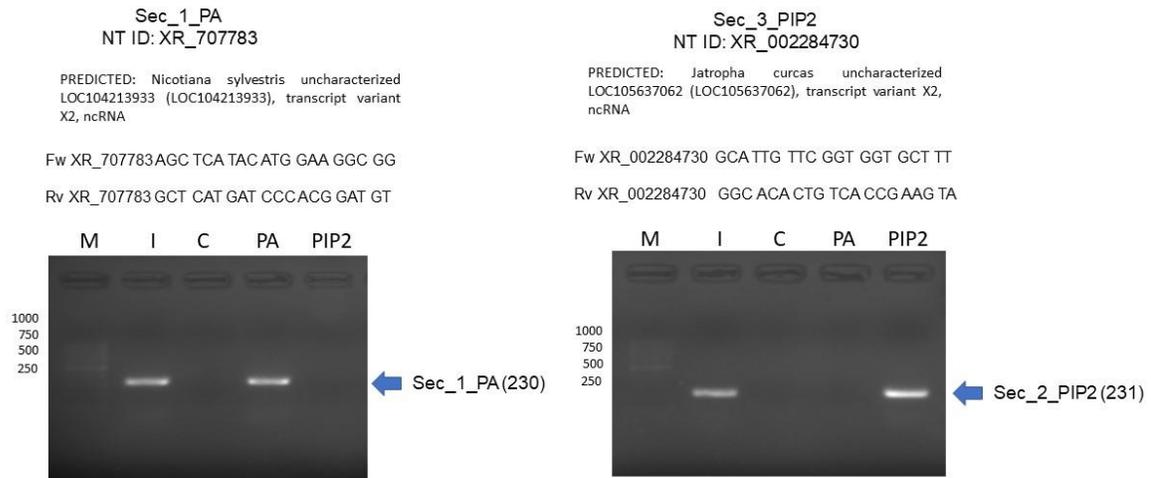


Figura 3.17. Interacción entre los lncRNAs Sec_1_PA y Sec_3_PIP2 con los lípidos PA y PIP2 mediante ensayos de pull-down *in vitro*. A) PCR de Sec_1_PA con perlas de agarosa sin lípidos (control) y perlas recubiertas con PA. B) PCR de Sec:3_PIP2 con perlas control y perlas de agarosa recubiertas de PIP2, respectivamente. M; Marcado de 1Kb; Input, C; Control: perlas de agarosa sin PA y PIP2; PA: perlas de agarosa recubiertas de PA; PIP2: perlas de agarosa recubiertas de PIP2.

3.8. Estructura secundaria y posibles sitios de unión

Una vez realizado el ensayo de interacción y conocidas las secuencias (Sec_1_PA y Sec_3_PIP2) con posible unión con los fosfolípidos PA y PIP2, se determinó las estructuras secundarias de estas: utilizando los software RNAfold (RNAfold web server (univie.ac.at)) y RNAstructure (Welcome to the Mathews Lab RNAstructure Web Servers (rochester.edu)) basados en una técnica de programación dinámica (PD) en el que una estructura secundaria se descompone en varias subestructuras como bucles de vecino más próximo, bucles de horquilla, bucles internos, bucles de abombamiento, apilamientos de pares de bases externos. El mínimo de energía libre de una estructura secundaria de ARN se calcula sumando la energía libre de los bucles descompuestos del vecino más próximo a través del algoritmo Zuker *et al* (1981). Las **Figuras 3.18** y **3.19** muestran las estructuras secundarias de los lncRNA Sec_1_PA y Sec_3_PIP2 basados en dichos algoritmos.

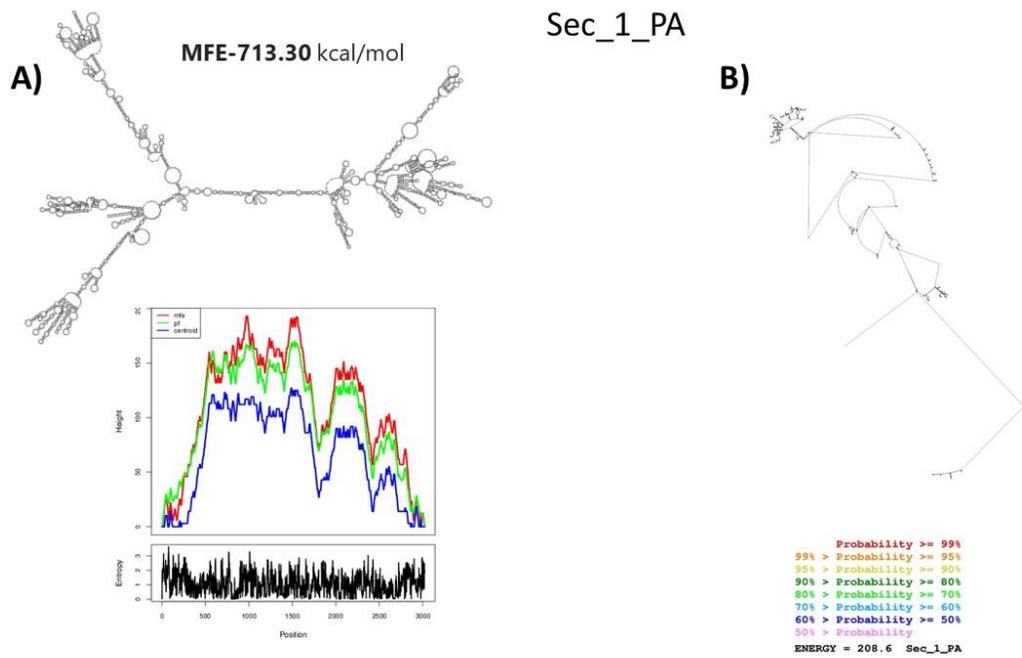


Figura 3.18. Representación de la estructura secundaria y la MFE de la Sec_1_PA. A) Representación de la estructura secundaria utilizando RNAfold y el diagrama de montaña que representa la posición de la secuencia frente al número de pares de bases que encierran esa posición. B) Representación de la estructura secundaria utilizando RNAstructure.

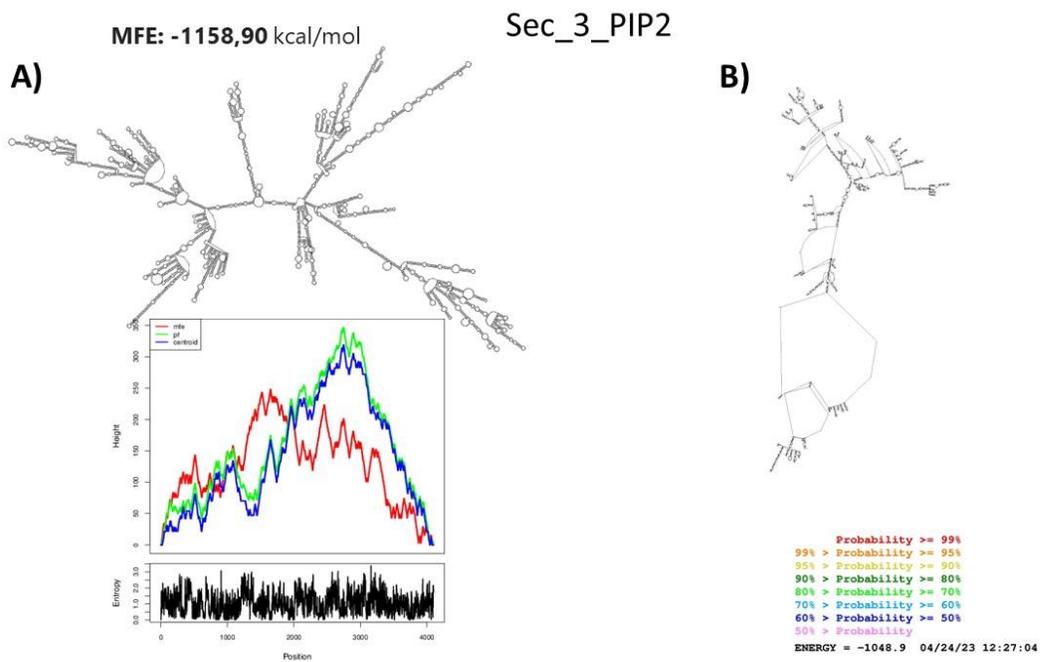


Figura 3.19. Representación de la estructura secundaria y la MFE de la Sec_3_PIP2. A) Representación de la estructura secundaria utilizando RNAfold y el diagrama de montaña que representa la posición de la secuencia frente al número de pares de bases que encierran esa posición. B) Representación de la estructura secundaria utilizando RNAstructure

Para determinar los posibles sitios de unión en los lncRNA Sec_1_PA y Sec_3_PIP2 a los lípidos PA y PIP2 a partir de las estructuras secundarias conocidas de estas secuencias, primeramente, se representaron las estructuras secundarias utilizando el programa MXfold2 Server de los sitios de unión de lncRNAs a lípidos reportados en las investigaciones de Lin *et al.*, (2017) donde un lncRNA denominado LINK-A interactúa a través de una región de 18 nt de su estructura con el fosfoinosítido PIP3 (**Figura 3.20.A**); también Donia *et al.*, (2019) donde un aptámero de ARN de 40 nucleótidos se une a PIP3 (**Figura 3.20.B**) y el motivo IV de lncRNA SNHG9 que se une a PA descrito por Li *et al.*, (2020) (**Figura 3.20.C**).

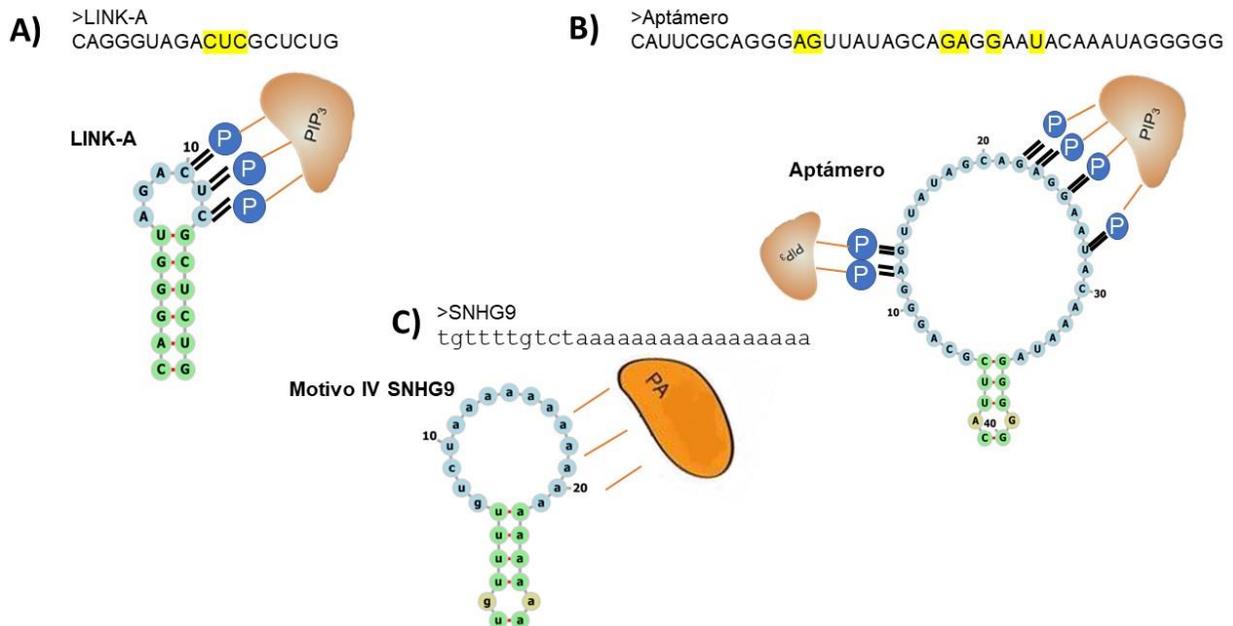


Figura 3.20. Estructuras secundarias de los sitios de unión de lncRNA a lípidos reportados en investigaciones anteriores. A) Lin *et al.*, (2017) donde un lncRNA denominado LINK-A interactúa a través de una región de 18 nt de su estructura con el fosfoinosítido PIP3 B) Aptámero de RNA de 40 nucleótidos se une a PIP3 Donia *et al.*, (2019) C) Domino IV de lncRNA SNHG9 que se une a PA descrito por Li *et al.*, (2020).

Seguidamente se realizó un alineamiento múltiple utilizando las plataformas Jalview y Clustal X Multiple Sequence Alignment (MEGAX), entre los lncRNA Sec_1_PA y Sec_3_PIP2 con los sitios de unión a lípidos reportados, donde para la Sec_1_PA se obtuvo un porcentaje de identidad (%)

ID) = 50.00 con los 18 nt correspondientes al sitio de unión de LINK-A a PIP3 (**Figura 3.21. A**), respecto a los 40 nt correspondiente al aptámero con interacción a PIP3 el %ID par Sec_1_PA = 55.81 (**Figura 3.21. B**), y en el caso del motivo IV del lncRNA SNHG9 con PA el %ID= 67.57 para la misma secuencia (**Figura 3.21. C**). Además, se representaron las estructuras secundarias de estos posibles sitios de unión de la secuencia Sec_1_PA utilizando el programa MXfold2 Server (**Figura 3.21**).

Posibles sitios de unión de Sec_1_PA con lípidos

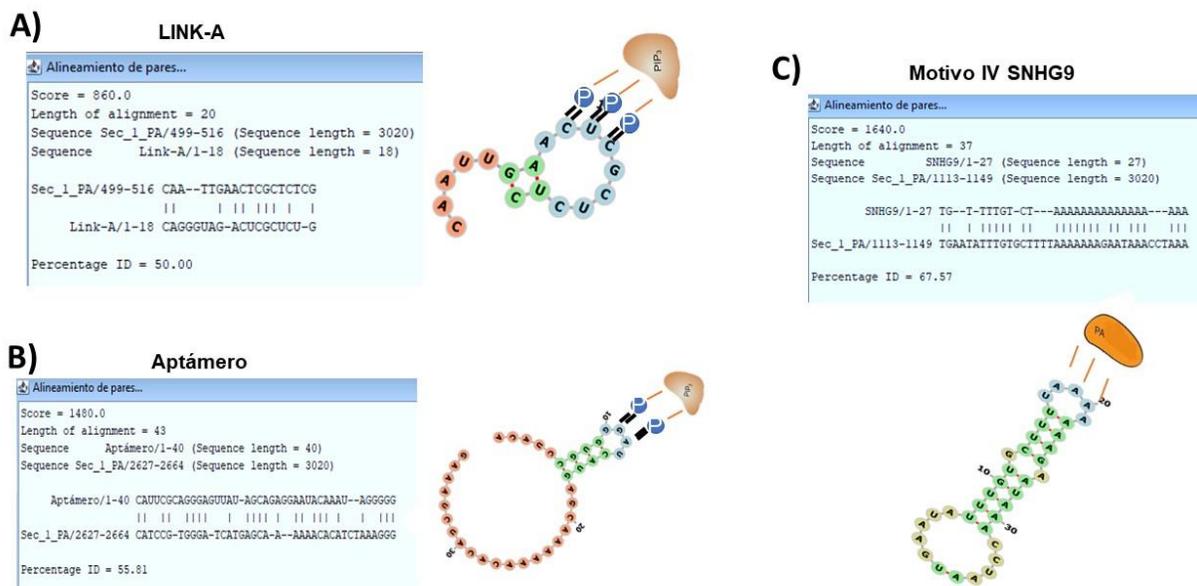


Figura 3.21. Alineamiento de Sec_1_PA con las secuencias de los sitios de unión de lncRNA a lípidos reportados en investigaciones anteriores, y la estructura secundaria de los posibles sitios de unión de Sec_1_PA a lípidos. A) Se muestra el alineamiento de Sec_1_PA con los 18 nt correspondientes al lncRNA LINK-A que se une con PIP3 Lin *et al.*, (2017) B) Se muestra el alineamiento de Sec_1_PA con los 40 nt correspondientes al aptámero de RNA que se une a PIP3 Donia *et al.*, (2019). C) Se muestra el alineamiento de Sec_1_PA con el dominio IV de lncRNA SNHG9 que se une a PA descrito por Li *et al.*, (2020). En los tres casos se muestran las estructuras secundarias correspondientes a los posibles sitios de unión de Sec_1_PA a lípidos.

El mismo análisis se realizó con la Sec_3_PIP2, donde se obtuvo un porcentaje de identidad (% ID) = 44.00 con los 18 nt correspondientes al sitio de unión de LINK-A a PIP3 (**Figura 3.22. A**), respecto a los 40 nt correspondiente al aptámero con interacción a PIP3 el %ID par Sec_3_PIP2 = 44.00 (**Figura 3.22. B**), y en el caso del motivo IV del lncRNA SNHG9 con PA el %ID= 53.66

para la misma secuencia (**Figura 3.22. C**). Además, se representaron las estructuras secundarias de estos posibles sitios de unión de la secuencia Sec_3_PIP2 utilizando el programa MXfold2 Server (**Figura 3.22**).

Posibles sitios de unión de Sec_3_PIP2 con lípidos

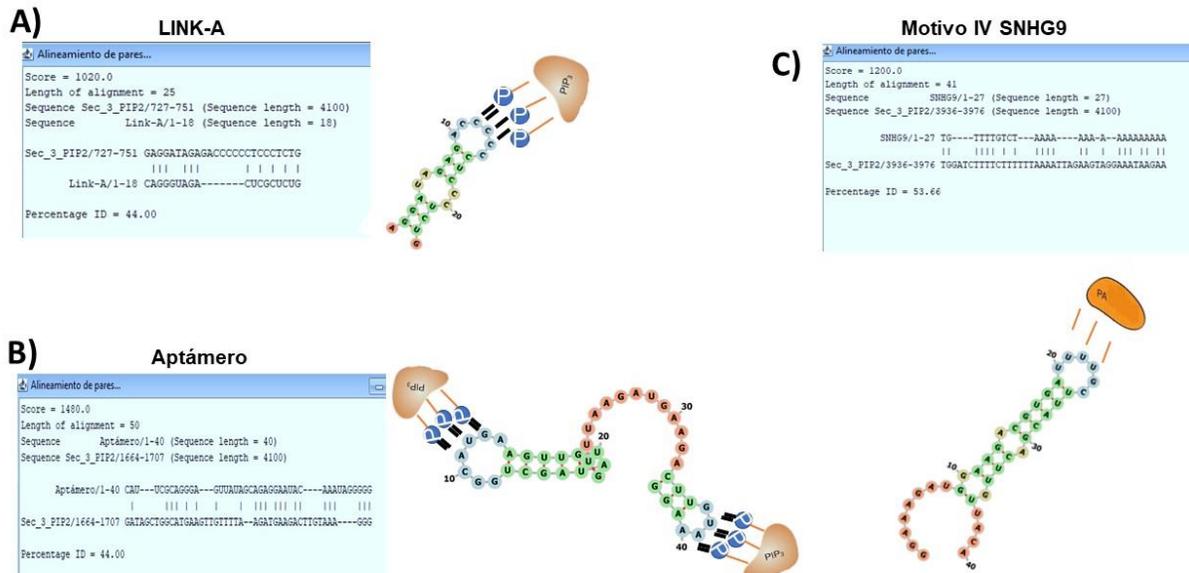


Figura 3.22. Alineamiento de Sec_3_PIP2 con las secuencias de los sitios de unión de lncRNA a lípidos reportados en investigaciones anteriores, y la estructura secundaria de los posibles sitios de unión de Sec_3_PIP2 a lípidos. A) Se muestra el alineamiento de Sec_3_PIP2 con los 18 nt correspondientes al lncRNA LINK-A que se une con PIP3 Lin *et al.*, (2017) B) Se muestra el alineamiento de Sec_3_PIP2 con los 40 nt correspondientes al aptámero de RNA que se une a PIP3 Donia *et al.*, (2019). C) Se muestra el alineamiento de Sec_3_PIP2 con el dominio IV de lncRNA SNHG9 que se une a PA descrito por Li *et al.*, (2020). En los tres casos se muestran las estructuras secundarias correspondientes a los posibles sitios de unión de Sec_3_PIP2 a lípidos.

Luego de realizado los análisis de alineamiento entre los lncRNA Sec_1_PA y lncRNA Sec_3_PIP2 con los sitios de unión descritos en las investigaciones de Lin *et al.*, (2017), Donia *et al.*, (2019) y Li *et al.*, (2020) para los lncRNA LINK-A, el aptámero de RNA de 40 nucleótidos y lncRNA SNHG9 que interactúan con los fosfolípidos PIP3, PIP2 y PA descrito por Li *et al.*, (2020) se puede concluir que los lncRNA Sec_1_PA y lncRNA Sec_3_PIP2 presentan similitudes en su secuencias con estos sitios, lo que entonces corrobora la posible interacción de estos con los lípidos PA y PIP2. En la **Figura 3.23** se observa la estructura secundaria de estos dos lncRNA (Sec_1_PA y Sec_3_PIP2) y la ubicación de los posibles sitios de unión representados anteriormente, además se puede apreciar la presencia de estructuras en forma de bucles lugares

por donde también podría darse la interacción con especies lipídicas, es importante tener en cuenta que la estructura secundarias de estos lncRNA es variable, y su conformación va a depender tanto de la longitud de su secuencia, de las características del medio así como de su localización celular.

Posibles sitios de unión con PIP2 y PA

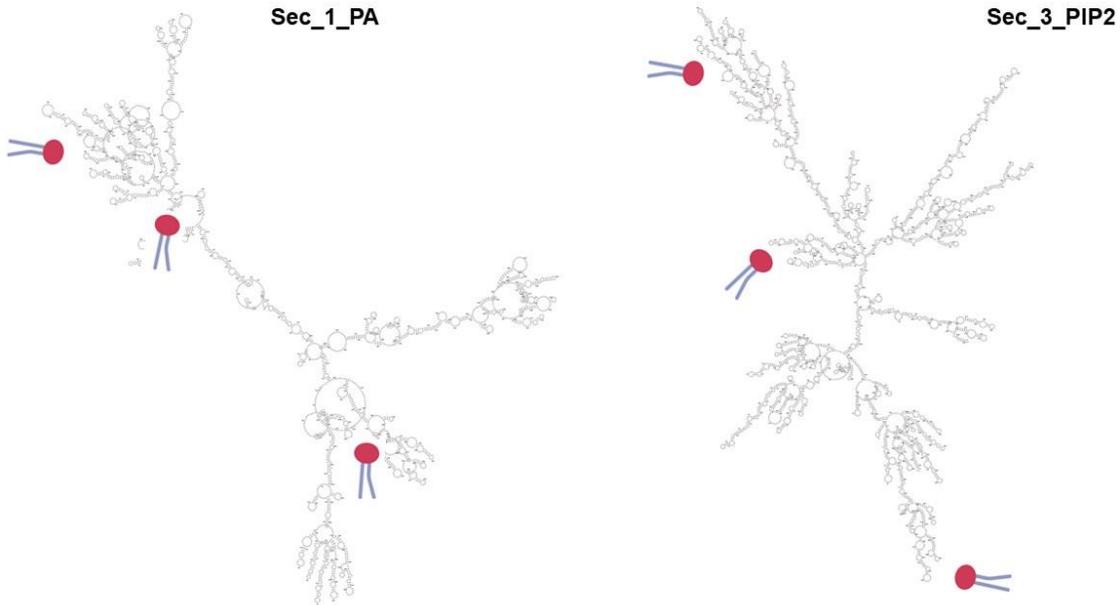


Figura 3.23. Representación de los posibles sitios de unión a PA y PIP2 en las secuencias Sec_1_PA y Sec_3_PIP2.

Como habíamos dicho la localización celular de los lncRNA puede inferir en la conformación estructural de estos, así como en su interacción con los lípidos, en la **Figura 3.24** utilizando el programa IncLocator: long non-coding RNA subcellular localization predictor se predijo la posible localización de los lncRNAs Sec_1_PA y Sec_3_PIP2, resultando para ambas secuencias el núcleo como posible sitio de localización.

Sec_1_PA	
Subcellular locations	score
Cytoplasm	0.24879348331
Nucleus	0.660029681922
Ribosome	0.0162765338353
Cytosol	0.02335870301
Exosome	0.0515415979222

Sec_3_PIP2	
Subcellular locations	score
Cytoplasm	0.288862705117
Nucleus	0.667847060726
Ribosome	0.00444753299727
Cytosol	0.0143909698807
Exosome	0.0244517312791

Predicted location
Nucleus

Predicted location
Nucleus

Figura 3.24. Localización según un predictor de los lncRNA: Sec_1_PA y Sec_3_PIP2.

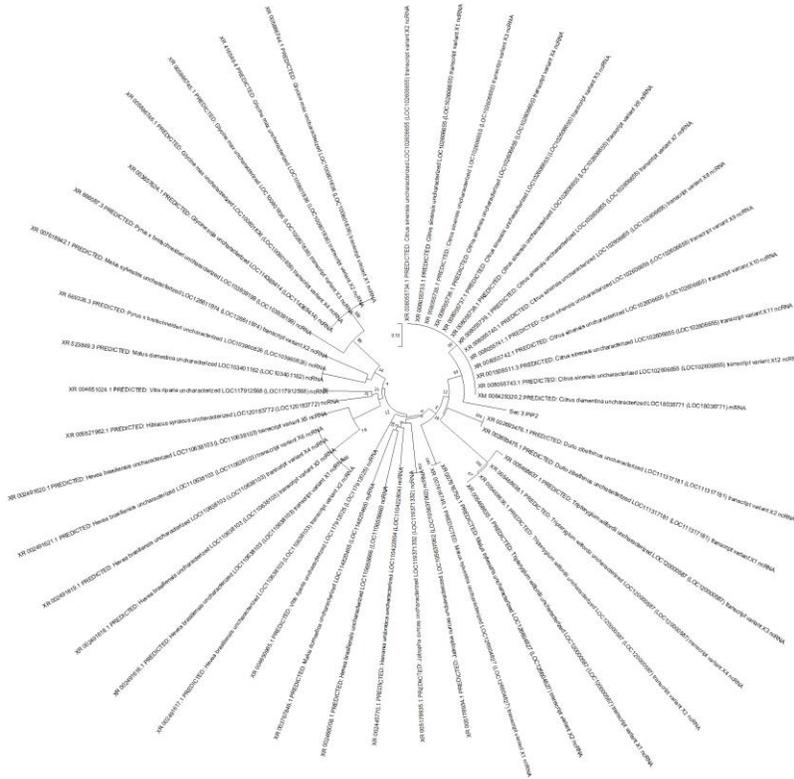


Figura 3.26. Árbol filogenético de la secuencia Sec_3_PIP2.

3.10 Conservación de elementos estructurales en los lncRNAs Sec_1_PA y Sec_2_PIP2

En la **Figura 3.27** se muestran los resultados del alineamiento entre el lncRNA Sec_1_PA y los IEs de algunos ncRNA como BCAR4 involucrado en la metástasis del cáncer de mama (Xing *et al.*, 2014), Bvht que participa en la diferenciación de los cardiomiocitos humanos (Xue *et al.*, 2016) y Xist que actúa en la inactivación del cromosoma X (Pintacuda *et al.*, 2017; Sun *et al.*, 2019), descritos experimentalmente. El lncRNA BCAR4 perteneciente a *Homo sapiens* (human), presenta dos IEs uno en la posición 235-288nt para la interacción con SNIP1 y un segundo en la posición 991-1044nt para la interacción con PPP1R10, la Sec_1_PA presentó un %ID=65.67 (**Figura 3.27. A**) para el primer IEs, mientras que con el segundo fue de %ID=47.92 (**Figura 3.27. B**), el lncRNA BCAR4 actúan en respuesta a la citoquina CCL21 uniéndose a SNIP1 y PPP1R10 a través de estos IES mecanismo que activa un programa transcripcional no canónico Hedgehog/GLI2 que promueve la migración celular (Xing *et al.*, 2014). En la **Figura 3.27.C** se observan los resultados del alineamiento de Sec_1_PA con el motivo AGIL correspondiente al ncRNA Bvht *Homo sapiens* (human), donde el %ID=81.82, este IEs rico en G en el ncRNA braveheart (Bvht) permite la interacción con el factor de transcripción CNBP/ZNF9 y afecta drásticamente a la diferenciación de los cardiomiocitos (Xue *et al.*, 2016). También en la **Figura**

3.27.D y **Figura 3.27.E** se muestran los alineamientos de Sec_1_PA con los IEs del lncRNA XIST perteneciente a *Mus musculus* (ratón doméstico) que actúan como esponjas para varios miRNAs (Pintacuda *et al.*, 2017; Sun *et al.*, 2019), como el sitio GTACTGT para miR-101 que tuvo un %ID=85.71 con Sec_1_PA o un sitio GCACTTT para miR-106-5p que presentó un %ID= 100 con Sec_1_PA.

Conservación de IEs en Sec_1_PA

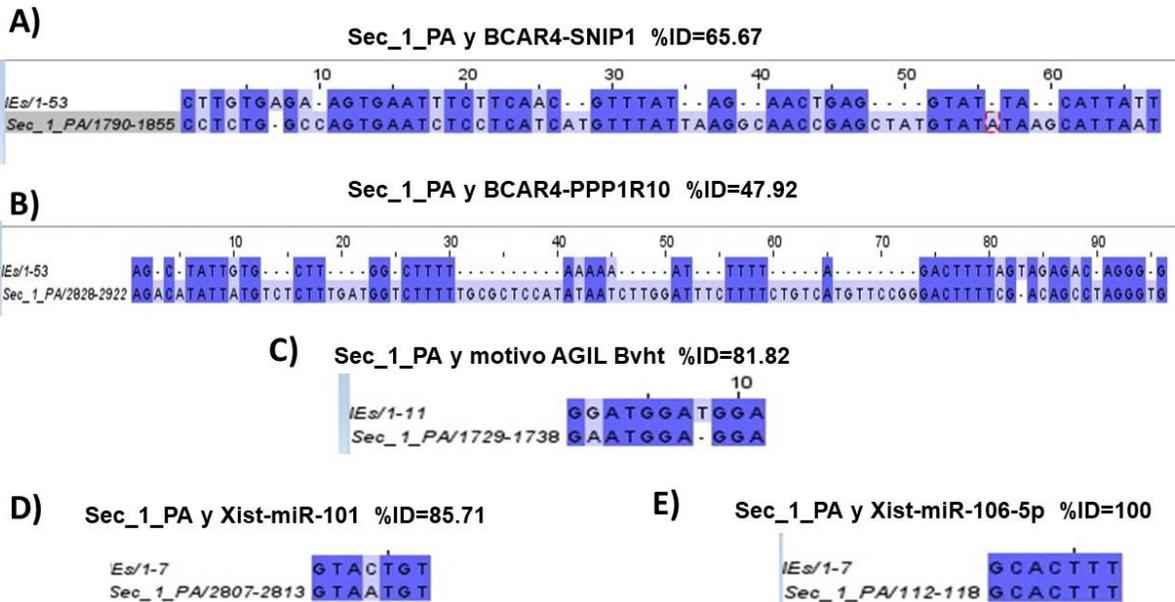


Figura 3.27. Resultados del alineamiento entre el lncRNA Sec_1_PA y los IEs de algunos ncRNA como BCAR4, Bvht y Xist. **A)** Alineamiento entre Sec_1_PA y el IEs del ncRNA BCAR4 en la posición 235-288nt para la interacción con SNIP1, con un %ID=65.67. **B)** Alineamiento entre Sec_1_PA y el IEs del ncRNA BCAR4 en la posición 991-1044nt para la interacción con PPP1R10, con un %ID=47.92. **C)** Alineamiento de Sec_1_PA con el motivo AGIL correspondiente al ncRNA Bvht), donde el %ID=81.82. **D)** Alineamiento de Sec_1_PA con el IEs GTACTGT del lncRNA XIST para miR-101 que tuvo un %ID=85.71. **E)** Alineamiento de Sec_1_PA con el IEs GCACTTT del lncRNA XIST para miR-106-5p que tuvo un %ID=100.

En la **Figura 3.28** se muestran de igual modo los resultados del alineamiento entre el lncRNA Sec_3_PIP2 y los IEs de los ncRNA BCAR4, Bvht y Xist. El elemento IEs presente en el ncRNA BCAR4 en posición 235-288nt presentó un %ID= 62.69 con Sec_3_PIP2 (**Figura 3.28. A**) y el de la posición 991-1044nt para la interacción con PPP1R10, presentó un %ID=54.22 con la Sec_3_PIP2 (**Figura 3.28. B**). En la **Figura 3.28.C** se observa los resultados del alineamiento de Sec_3_PIP2 con el motivo AGIL correspondiente al ncRNA Bvht, donde el %ID=78.57. También en la **Figura 3.28.D** y **Figura 3.28.E** se muestran los alineamientos de Sec_3_PIP2 con los IEs del lncRNA XIST con el sitio GTACTGT para miR-101 tuvo un %ID=87.50 con Sec_3_PIP2 y con el sitio GCACTTT para miR-106-5p presentó al igual que Sec_1_PA %ID= 100.

Conservación de IEs en Sec_3_PIP2

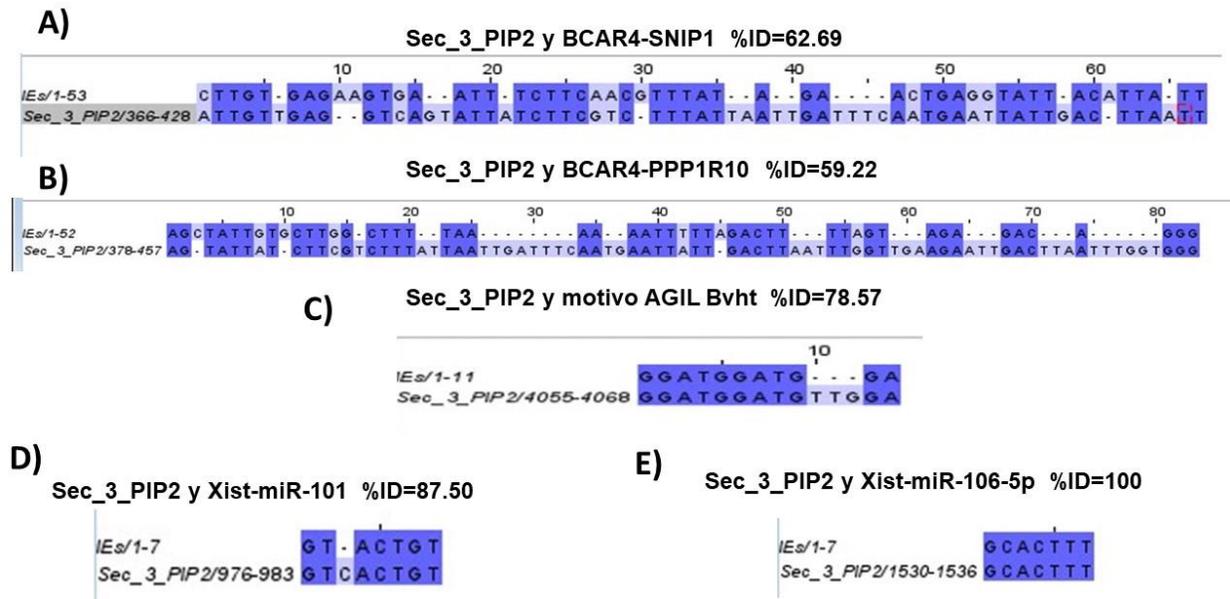


Figura 3.28. Resultados del alineamiento entre el lncRNA Sec_3_PIP2 y los IEs de algunos ncRNA como BCAR4, Bvht y Xist. **A)** Alineamiento entre Sec_3_PIP2 y el IEs del ncRNA BCAR4 en la posición 235-288nt para la interacción con SNIP1, con un %ID=62.69. **B)** Alineamiento entre Sec_3_PIP2 y el IEs del ncRNA BCAR4 en la posición 991-1044nt para la interacción con PPP1R10, con un %ID=54.22. **C)** Alineamiento de Sec_3_PIP2 con el motivo AGIL correspondiente al ncRNA Bvht), donde el %ID=78.57. **D)** Alineamiento de Sec_3_PIP2 con el IEs GTACTGT del lncRNA XIST para miR-101 que tuvo un %ID=87.50. **E)** Alineamiento de Sec_3_PIP2 con el IEs GCACTTT del lncRNA XIST para miR-106-5p que tuvo un %ID=100

Las IEs de los lncRNA son más diversas que las de los miARN y se conocen menos, en parte porque están integradas en estructuras de RNA más grandes. Los resultados obtenidos del alineamiento entre los lncRNA Sec_1_PA y Sec_3_PIP2 con los IEs conocidos experimentalmente de ncRNA como BCAR4, Bvht y XIST, demuestran una clara conservación y con ello evolución entre especies de estos elementos, lo que podría marcar un camino a suponer alguna de las funciones de estos dos lncRNA debido a que por ejemplo el motivo AGIL del lncRNA Bvht está involucrado en el proceso de transcripción y a su vez este IEs comparte con Sec_1_Pa un %ID= 81.82 y con Sec_3_PIP2 %ID=78.57. También podríamos suponer que posiblemente ambos lncRNA hallados en nuestra investigación actúan como esponjas para miRNAs esto al observar que tras el alineamiento con el motivo GCACTTT perteneciente al lncRNA XIST, ambas secuencias tuvieron un %ID =100.

CAPÍTULO IV DISCUSIÓN, CONCLUSIONES GENERALES Y PERSPECTIVAS

4.1. DISCUSIÓN

Los lncRNAs suponen un campo en investigación dinámica y creciente, se sabe que, dependiendo de su localización, mecanismo de acción e interacción con diferentes moléculas pueden estar implicados en la regulación de diferentes procesos celulares. La interacción entre lncRNAs y componentes celulares (proteínas, microRNAs, DNA, péptidos) ha sido demostrada, sin embargo, la posible unión de estos con componentes lípidos como los fosfoinosítidos aún es incipiente. Investigaciones como las de Lin *et al.*, (2017) donde demuestran que el lncRNA LINK A regula la proteína AKT cuando se une a PIP3 y PIP2 utilizando líneas celulares de cáncer de mama humano; de igual manera Donia *et al.*, (2019) encuentran un aptámero de RNA que interactuaba con PIP3 en una línea celular de fibrosarcoma y Li *et al.*, (2020) reportaron que el lncRNA SNHG9 interactúa con el ácido fosfatídico (PA) mejorando la interacción de la proteína LAST1 en líneas celulares de cáncer de mama y de riñón embrionario humano. Estos hallazgos reflejan un nuevo tipo de interacciones moleculares en las células humanas, pero aún se desconoce sobre lncRNAs en plantas que interactúan con fosfolípidos como PA y PIP2.

En nuestra investigación identificamos dos lncRNAs a los que llamamos Sec_1_PA y Sec_3_PIP2 con posible interacción a los fosfolípidos PA y PIP2. Estos dos lncRNAs fueron seleccionados dentro del total de genes mapeados resultado de la librería de RNA-seq para las condiciones experimentales de perlas recubiertas de los lípidos PA/PIP2 y un control el cual consistió en perlas libres de lípidos; la biblioteca se construyó utilizando el RNA total partiendo del cultivo de células en suspensión de la planta *Stenocereus queretaroensis*. El flujo de trabajo para la selección de los lncRNAs, consistió primeramente en determinar dentro del total de genes cuales se encontraban expresados de forma diferencial ante la presencia de los lípidos PA y PIP2. Obtenidos los DBG, se aplicaron filtros para reconocer cuantas de estas secuencias cumplían con características propiamente de lncRNAs, encontrando 44 lncRNA y dentro de estas se eligieron 4 secuencias que mostraron mayor afinidad por un lípido u otro.

Estas cuatro secuencias de lncRNA llamadas como: Sec_1_PA (Cluster_18316.12632), Sec_2_PA (Cluster_27261.1), Sec_3_PIP2 (Cluster_20205.8) y Sec_4_PIP2 (Cluster_30630.0) fueron amplificadas por RT-PCR a una temperatura de alineamiento de 57°C y utilizando diferentes ciclos (30,28,26 y 24), para ello se diseñaron los oligos y se sintetizó el cDNA específicos para cada una. Los resultados de RT-PCR mostraron que cada secuencia para los ciclos 30 y 28

amplificaron al tamaño esperado: Sec_1_PA amplificó a 230pb, Sec_2_PA a 220pb, Sec_3_PIP2 a 231pb y Sec_4_PIP2 a 426 mientras que para los ciclos 26 y 24 amplificaron las secuencias Sec_2_PA a 220pb y Sec_4_PIP2 a 426. Seguidamente para realizar los ensayos de interacción con los lípidos PA/PIP2 se clonaron las secuencias Sec_1_PA y Sec_3_PIP2 a pb utilizando el vector de clonación pJET1.2, por su alta eficiencia, utilizando para la transformación las células de E. Coli DH5 α , comprobando así que se trataba de las secuencias de interés.

Los resultados del pull-down entre los lncRNAs Sec_1_PA y Sec_3_PIP2 con perlas recubiertas de los lípidos PA y PIP2 seguido de RT-PCR mostraron la unión de las secuencias con los lípidos, en el caso del lncRNA Sec_1_PA se unió a las perlas que presentaban PA, mientras Sec_3_PIP2 se unió a perlas con PIP2, esto corrobora los análisis de expresión diferencial realizados donde precisamente estas secuencias se mostraron enriquecidas en los tratamientos con PA y PIP2 respectivamente. Además, estos resultados obtenidos son similares a los reportados en el pull-down realizado en la investigación de Lin *et al.*, (2017) para el lncRNA LINK-A con el fosfoinositidos PIP3.

Seguidamente se realizó un alineamiento entre los lncRNAs Sec_1_PA y Sec_3_PIP2 con los sitios de unión a lípidos reportados en otros lncRNAs experimental. Los resultados obtenidos demuestran que en el caso del lncRNA Sec_1_PA que interactúa con PA demostrado en los ensayos de pull-down comparte un porcentaje de identidad (% ID) = 67.57 con el motivo IV del lncRNA SNHG9 reportado por Li *et al.*, (2020) que interactúa a través de este motivo con PA, mejorando la interacción de la proteína LAST1, al asociarse al dominio C-terminal de esta proteína, promoviendo la separación de fases de la misma e inhibiendo la fosforilación de YAP mediada por LATS1, es decir la pérdida de SNHG9 suprime el crecimiento del tumor de mama Mientras que el lncRNA Sec_3_PIP2 que demostramos que interactúa con PIP2, se obtuvo un porcentaje de identidad (% ID) =44.00 tanto con el dominio de homología de pleckstrin correspondientes al sitio de unión de LINK-A a PIP3 y PIP2, que participa en la regulación de la proteína AKT y con los 40 nt correspondiente al sitio de interacción del aptámero con PIP3. Estos resultados demuestran que los lncRNAs Sec_1_PA y Sec_3_PIP2 presentan en su estructura secuencias similares a las descritas en otros lncRNAs que propician la interacción de estas moléculas con fosfolípidos.

Los resultados obtenidos del alineamiento de nuestros dos lncRNAs Sec_1_PA y Sec_3_PIP2 con elementos reguladores conservados y que están involucrados en la función de ncRNA reportados en *Homo sapiens* (human) como los lncRNA BCAR4 y Bvht, así como el lncRNA XIST

perteneciente a *Mus musculus* (ratón doméstico), resultaron novedosos. El lncRNA Sec_1_PA presentó un %ID=65.67 y %ID=47.92 con dos IEs ubicados en las posiciones 235-288 nt y 991-1044 nt del lncRNA BCAR4, los cuales permiten la unión de este lncRNA a SNIP1 y PPP1R10 en respuesta a la citoquina CCL21 activando un programa transcripcional no canónico Hedgehog/GLI2 que promueve la migración celular (Xing *et al.*, 2014). Además, Sec_1_PA tuvo %ID=81.82 con el motivo AGIL correspondiente al ncRNA Bvht que permite la interacción de este con el factor de transcripción CNBP/ZNF9 y afecta drásticamente a la diferenciación de los cardiomiocitos (Xue *et al.*, 2016). Los resultados de identidad fueron similares para Sec_3_PIP2, con los IEs de BCAR4 presentó un %ID= 62.69 y %ID=54.22; el motivo AGIL de Bvht, fue de %ID=78.57. Ambos lncRNA Sec_1_PA y Sec_3_PIP2 presentaron % de ID mayor a 80% con el IEs GTACTGT del lncRNA XIST y 100 %ID con otro IEs GCACTTT de este mismo lncRNA, ambos IEs participan en el reclutamiento de miRNAs. Estos resultados demuestran una clara conservación y evolución entre especies, lo que podría marcar un camino a suponer alguna de las funciones de estos dos lncRNA debido a que por ejemplo el motivo AGIL del lncRNA Bvht está involucrado en el proceso de transcripción y a su vez este IEs comparte similitud con Sec_1_Pa y Sec_3_PIP2. También podríamos suponer que posiblemente ambos lncRNA hallados en nuestra investigación actúan como esponjas para miRNAs esto al observar que tras el alineamiento con el motivo GCACTTT perteneciente al lncRNA XIST, ambas secuencias tuvieron un %ID =100.

Los lncRNAs Sec_1_Pa y Sec_3_PIP2 podrían estar involucrados en procesos biológicos como la regulación de la transcripción, protección de la horquilla de replicación y en la respuesta celular al estímulo de daño en el DNA según los análisis de enriquecimientos GO y su localización podría ser nuclear, incluso esto coincide con los resultados obtenidos al utilizar un predictor de localización que arrojó mayor porcentaje de posibilidad de que estos lncRNAs se encuentren en el núcleo

El descubrimiento de lncRNAs que potencialmente se unen a lípidos puede reflejar un proceso molecular particular en la célula. Además, estas interacciones pueden formar parte de un mecanismo dinámico de modificaciones lipídicas y separación de fases líquido-líquido las cuales pueden aumentar la concentración local de determinadas moléculas y dotar a la estructura en la que se forman de papeles funcionales.

4.2. CONCLUSIONES GENERALES

- 1-) Se generó una biblioteca de RNA-seq
- 2-) Se encontraron 22 348 genes expresados diferencialmente para PA y PIP2.
- 3-) Se encontraron 44 secuencias de lncRNAs y se seleccionaron 4 lncRNAs.
- 4-) Se amplificaron por RT-PCR los lncRNA Sec_1_PA, Sec_2_PA, Sec_3_PIP2 y Sec_4_PIP2.
- 5-) Se demostró que los lncRNAs Sec_1_PA y Sec_3_PIP2 interactúan con los fosfolípidos PA y PIP2.
- 6-) Se comprobó que los lncRNAs Sec_1_PA y Sec_3_PIP2 presentan sitios en su secuencia de unión a PA y PIP2.
- 7-) Se comprobó que los lncRNAs Sec_1_PA y Sec_3_PIP2 presentan elementos reguladores conservados entre especies.

4.3. PERSPECTIVAS

- Determinar experimentalmente los sitios de unión de Sec_1_PA y Sec_3_PIP2 a los lípidos PA y PIP2
- Determinar *in vivo* la localización de estos dos lncRNAs

Bibliografía

Aguilo, F., Zhou, M.-M., & Walsh, M. J. (2011). Long Noncoding RNA, Polycomb, and the Ghosts Haunting INK4b-ARF-INK4a Expression. *Cancer Research*, 71(16), 5365-5369. <https://doi.org/10.1158/0008-5472.CAN-10-4379>

Akerman, I., Tu, Z., Beucher, A., Rolando, D. M. Y., Sauty-Colace, C., Benazra, M., Nakic, N., Yang, J., Wang, H., Pasquali, L., Moran, I., Garcia-Hurtado, J., Castro, N., Gonzalez-Franco, R., Stewart, A. F., Bonner, C., Piemonti, L., Berney, T., Groop, L., ... Ferrer, J. (2017). Human Pancreatic β Cell lncRNAs Control Cell-Specific Regulatory Networks. *Cell Metabolism*, 25(2), 400-411. <https://doi.org/10.1016/j.cmet.2016.11.016>

Alvarez-Dominguez, J. R., & Lodish, H. F. (2017). Emerging mechanisms of long noncoding RNA function during normal and malignant hematopoiesis. *Blood*, 130(18), 1965-1975. <https://doi.org/10.1182/blood-2017-06-788695>

Ariel, F., Jegu, T., Latrasse, D., Romero-Barrios, N., Christ, A., Benhamed, M., & Crespi, M. (2014). Noncoding Transcription by Alternative RNA Polymerases Dynamically Regulates an Auxin-Driven Chromatin Loop. *Molecular Cell*, 55(3), 383-396. <https://doi.org/10.1016/j.molcel.2014.06.011>

Bardou, F., Ariel, F., Simpson, C. G., Romero-Barrios, N., Laporte, P., Balzergue, S., Brown, J. W. S., & Crespi, M. (2014). Long Noncoding RNA Modulates Alternative Splicing Regulators in Arabidopsis. *Developmental Cell*, 30(2), 166-176. <https://doi.org/10.1016/j.devcel.2014.06.017>

Barneda, D., Planas-Iglesias, J., Gaspar, M. L., Mohammadyani, D., Prasannan, S., Dormann, D., Han, G.-S., Jesch, S. A., Carman, G. M., Kagan, V., Parker, M. G., Ktistakis, N. T., Klein-Seetharaman, J., Dixon, A. M., Henry, S. A., & Christian, M. (2015). The brown adipocyte protein CIDEA promotes lipid droplet fusion via a phosphatidic acid-binding amphipathic helix. *ELife*, 4, e07485. <https://doi.org/10.7554/eLife.07485>

Benoit Bouvrette, L. P., Cody, N. A. L., Bergalet, J., Lefebvre, F. A., Diot, C., Wang, X., Blanchette, M., & Lécuyer, E. (2018). CeFra-seq reveals broad asymmetric mRNA and noncoding RNA distribution profiles in *Drosophila* and human cells. *RNA*, 24(1), 98-113. <https://doi.org/10.1261/rna.063172.117>

Bhan, A., & Mandal, S. S. (2015). lncRNA HOTAIR: A master regulator of chromatin dynamics and cancer. *Biochemical et Biophysical Acta (BBA) - Reviews on Cancer*, 1856(1), 151-164. <https://doi.org/10.1016/j.bbcan.2015.07.001>

Bridges, M. C., Daulagala, A. C., & Kourtidis, A. (2021). LNCcation: lncRNA localization and function. *Journal of Cell Biology*, 220(2), e202009045. <https://doi.org/10.1083/jcb.202009045>

Calin, G. A., Liu, C., Ferracin, M., Hyslop, T., Spizzo, R., Sevignani, C., Fabbri, M., Cimmino, A., Lee, E. J., Wojcik, S. E., Shimizu, M., Tili, E., Rossi, S., Taccioli, C., Pichiorri, F., Liu, X., Zupo, S., Herlea, V., Gramantieri, L., ... Croce, C. M. (2007). Ultra conserved Regions Encoding ncRNAs Are Altered in Human Leukemias and Carcinomas. *Cancer Cell*, 12(3), 215-229. <https://doi.org/10.1016/j.ccr.2007.07.027>

- Castano, E., Yildirim, S., Fáberová, V., Krausová, A., Uličná, L., Paprčková, D., Sztacho, M., & Hozák, P. (2019). Nuclear Phosphoinositide—Versatile Regulators of Genome Functions. *Cells*, 8(7), 649. <https://doi.org/10.3390/cells8070649>
- Cesana, M., Cacchiarelli, D., Legnini, I., Santini, T., Sthandier, O., Chinappi, M., Tramontano, A., & Bozzoni, I. (2011). A Long Noncoding RNA Controls Muscle Differentiation by Functioning as a Competing Endogenous RNA. *Cell*, 147(2), 358-369. <https://doi.org/10.1016/j.cell.2011.09.028>
- Chen, B., Dragomir, M. P., Yang, C., Li, Q., Horst, D., & Calin, G. A. (2022). Targeting non-coding RNAs to overcome cancer therapy resistance. *Signal Transduction and Targeted Therapy*, 7(1), 121. <https://doi.org/10.1038/s41392-022-00975-3>
- Cifuentes-Rojas, C., Hernandez, A. J., Sarma, K., & Lee, J. T. (2014). Regulatory Interactions between RNA and Polycomb Repressive Complex 2. *Molecular Cell*, 55(2), 171-185. <https://doi.org/10.1016/j.molcel.2014.05.009>
- Cooper, D., Carter, G., Li, P., Patel, R., Watson, J., & Patel, N. (2014). Long Non-Coding RNA NEAT1 Associates with SRp40 to Temporally Regulate PPAR γ 2 Splicing during Adipogenesis in 3T3-L1 Cells. *Genes*, 5(4), 1050-1063. <https://doi.org/10.3390/genes5041050>
- Corbel, C., Diabanguouaya, P., Gendrel, A.-V., Chow, J. C., & Heard, E. (2013). Unusual chromatin status and organization of the inactive X chromosome in murine trophoblast giant cells. *Development*, 140(4), 861-872. <https://doi.org/10.1242/dev.087429>
- Dahariya, S., Paddibhatla, I., Kumar, S., Raghuwanshi, S., Palapati, A., & Gutti, R. K. (2019). Long noncoding RNA: Classification, biogenesis, and functions in blood cells. *Molecular Immunology*, 112, 82-92. <https://doi.org/10.1016/j.molimm.2019.04.011>
- Dickson, E. J., & Hille, B. (2019). Understanding phosphoinositide's: Rare, dynamic, and essential membrane phospholipids. *Biochemical Journal*, 476(1), 1-23. <https://doi.org/10.1042/BCJ20180022>
- Djebali, S., Davis, C. A., Merkel, A., Dobin, A., Lassmann, T., Mortazavi, A., Tanzer, A., Lagarde, J., Lin, W., Schlesinger, F., Xue, C., Marinov, G. K., Khatun, J., Williams, B. A., Zaleski, C., Rozowsky, J., Röder, M., Kokocinski, F., Abdelhamid, R. F., ... Gingeras, T. R. (2012). Landscape of transcription in human cells. *Nature*, 489(7414), 101-108. <https://doi.org/10.1038/nature11233>
- Donia, T., Jyoti, B., Suizu, F., Hirata, N., Tanaka, T., Ishigaki, S., F, P. T. J., Nio-Kobayashi, J., Iwanaga, T., Chiorini, J. A., & Noguchi, M. (2019). Identification of RNA aptamer which specifically interacts with PtdIns (3) P. *Biochemical and Biophysical Research Communications*, 517(1), 146-154. <https://doi.org/10.1016/j.bbrc.2019.07.034>
- Engreitz, J. M., Haines, J. E., Perez, E. M., Munson, G., Chen, J., Kane, M., McDonel, P. E., Guttman, M., & Lander, E. S. (2016). Local regulation of gene expression by lncRNA promoters, transcription and splicing. *Nature*, 539(7629), 452-455. <https://doi.org/10.1038/nature20149>
- Fabbri, M., Girnita, L., Varani, G., & Calin, G. A. (2019). Decrypting noncoding RNA interactions, structures, and functional networks. *Genome Research*, 29(9), 1377-1388. <https://doi.org/10.1101/gr.247239.118>

- Fazal, F. M., Han, S., Parker, K. R., Kaewsapsak, P., Xu, J., Boettiger, A. N., Chang, H. Y., & Ting, A. Y. (2019). Atlas of Subcellular RNA Localization Revealed by APEX-Seq. *Cell*, 178(2), 473-490.e26. <https://doi.org/10.1016/j.cell.2019.05.027>
- Foster, D. A., Salloum, D., Menon, D., & Frias, M. A. (2014a). Phospholipase D and the Maintenance of Phosphatidic Acid Levels for Regulation of Mammalian Target of Rapamycin (mTOR). *Journal of Biological Chemistry*, 289(33), 22583-22588. <https://doi.org/10.1074/jbc.R114.566091>
- Gao, Y., Wang, P., Wang, Y., Ma, X., Zhi, H., Zhou, D., Li, X., Fang, Y., Shen, W., Xu, Y., Shang, S., Wang, L., Wang, L., Ning, S., & Li, X. (2019). Lnc2Cancer v2.0: Updated database of experimentally supported long non-coding RNAs in human cancers. *Nucleic Acids Research*, 47(D1), D1028-D1033. <https://doi.org/10.1093/nar/gky1096>
- Golani, Y., Kaye, Y., Gilhar, O., Ercetin, M., Gillaspay, G., & Levine, A. (2013). Inositol Polyphosphate Phosphatidylinositol 5-Phosphatase9 (At5PTase9) Controls Plant Salt Tolerance by Regulating Endocytosis. *Molecular Plant*, 6(6), 1781-1794. <https://doi.org/10.1093/mp/sst072>
- Grote, P., & Herrmann, B. G. (2013). The long non-coding RNA Fendrr links epigenetic control mechanisms to gene regulatory networks in mammalian embryogenesis. *RNA Biology*, 10(10), 1579-1585. <https://doi.org/10.4161/rna.26165>
- Gruber, A. R., Lorenz, R., Bernhart, S. H., Neubock, R., & Hofacker, I. L. (2008a). The Vienna RNA Website. *Nucleic Acids Research*, 36(Web Server), W70-W74. <https://doi.org/10.1093/nar/gkn188>
- Hamann, B. L., & Blind, R. D. (2018). Nuclear phosphoinositide regulation of chromatin. *Journal of Cellular Physiology*, 233(1), 107-123. <https://doi.org/10.1002/jcp.25886>
- Hanly, D. J., Esteller, M., & Berdasco, M. (2018). Interplay between long non-coding RNAs and epigenetic machinery: Emerging targets in cancer? *Philosophical Transactions of the Royal Society B: Biological Sciences*, 373(1748), 20170074. <https://doi.org/10.1098/rstb.2017.0074>
- He, C., Sidoli, S., Warneford-Thomson, R., Tatomer, D. C., Wilusz, J. E., Garcia, B. A., & Bonasio, R. (2016). High-Resolution Mapping of RNA-Binding Regions in the Nuclear Proteome of Embryonic Stem Cells. *Molecular Cell*, 64(2), 416-430. <https://doi.org/10.1016/j.molcel.2016.09.034>
- Heilmann, I. (2016). Phosphoinositide signaling in plant development. *Development*, 143(12), 2044-2055. <https://doi.org/10.1242/dev.136432>
- Helder, S., Blythe, A. J., Bond, C. S., & Mackay, J. P. (2016). Determinants of affinity and specificity in RNA-binding proteins. *Current Opinion in Structural Biology*, 38, 83-91. <https://doi.org/10.1016/j.sbi.2016.05.005>
- Heo, J. B., Lee, Y.-S., & Sung, S. (2013). Epigenetic regulation by long noncoding RNAs in plants. *Chromosome Research*, 21(6-7), 685-693. <https://doi.org/10.1007/s10577-013-9392-6>
- Huarte, M., Guttman, M., Feldser, D., Garber, M., Koziol, M. J., Kenzelmann-Broz, D., Khalil, A. M., Zuk, O., Amit, I., Rabani, M., Attardi, L. D., Regev, A., Lander, E. S., Jacks, T., & Rinn, J. L.

- (2010). A Large Intergenic Noncoding RNA Induced by p53 Mediates Global Gene Repression in the p53 Response. *Cell*, 142(3), 409-419. <https://doi.org/10.1016/j.cell.2010.06.040>
- Ischebeck, T., Werner, S., Krishnamoorthy, P., Lerche, J., Meijón, M., Stenzel, I., Löffke, C., Wiessner, T., Im, Y. J., Perera, I. Y., Iven, T., Feussner, I., Busch, W., Boss, W. F., Teichmann, T., Hause, B., Persson, S., & Heilmann, I. (2014). Phosphatidylinositol 4,5-Bisphosphate Influences PIN Polarization by Controlling Clathrin-Mediated Membrane Trafficking in Arabidopsis. *The Plant Cell*, 25(12), 4894-4911. <https://doi.org/10.1105/tpc.113.116582>
- Kaewsapsak, P., Shechner, D. M., Mallard, W., Rinn, J. L., & Ting, A. Y. (2017). Live-cell mapping of organelle-associated RNAs via proximity biotinylation combined with protein-RNA crosslinking. *ELife*, 6, e29224. <https://doi.org/10.7554/eLife.29224>
- Kalasova, I., Fáberová, V., Kalendová, A., Yildirim, S., Uličná, L., Venit, T., & Hozák, P. (2016a). Tools for visualization of phosphoinositides in the cell nucleus. *Histochemistry and Cell Biology*, 145(4), 485-496. <https://doi.org/10.1007/s00418-016-1409-8>
- Kalasova, I., Fáberová, V., Kalendová, A., Yildirim, S., Uličná, L., Venit, T., & Hozák, P. (2016b). Tools for visualization of phosphoinositides in the cell nucleus. *Histochemistry and Cell Biology*, 145(4), 485-496. <https://doi.org/10.1007/s00418-016-1409-8>
- Kassas, N., Tanguy, E., Thahouly, T., Fouillen, L., Heintz, D., Chasserot-Golaz, S., Bader, M.-F., Grant, N. J., & Vitale, N. (2017a). Comparative Characterization of Phosphatidic Acid Sensors and Their Localization during Frustrated Phagocytosis. *Journal of Biological Chemistry*, 292(10), 4266-4279. <https://doi.org/10.1074/jbc.M116.742346>
- Kassas, N., Tanguy, E., Thahouly, T., Fouillen, L., Heintz, D., Chasserot-Golaz, S., Bader, M.-F., Grant, N. J., & Vitale, N. (2017b). Comparative Characterization of Phosphatidic Acid Sensors and Their Localization during Frustrated Phagocytosis. *Journal of Biological Chemistry*, 292(10), 4266-4279. <https://doi.org/10.1074/jbc.M116.742346>
- Kazemzadeh, M., Safaralizadeh, R., & Orang, A. V. (2015). LncRNAs: Emerging players in gene regulation and disease pathogenesis. *Journal of Genetics*, 94(4), 771-784. <https://doi.org/10.1007/s12041-015-0561-6>
- Kazimierczyk, Kasprowicz, Kasprzyk, & Wrzesinski. (2020). Human Long Noncoding RNA Interactome: Detection, Characterization and Function. *International Journal of Molecular Sciences*, 21(3), 1027. <https://doi.org/10.3390/ijms21031027>
- Kim, D. N., Thiel, B. C., Mrozowich, T., Hennelly, S. P., Hofacker, I. L., Patel, T. R., & Sanbonmatsu, K. Y. (2020). Zinc-finger protein CNBP alters the 3-D structure of lncRNA Braveheart in solution. *Nature Communications*, 11(1), 148. <https://doi.org/10.1038/s41467-019-13942-4>
- Kim, Y. K., Furic, L., Parisien, M., Major, F., DesGroseillers, L., & Maquat, L. E. (2007). Stauf1 regulates diverse classes of mammalian transcripts. *The EMBO Journal*, 26(11), 2670-2681. <https://doi.org/10.1038/sj.emboj.7601712>
- Kindgren, P., Ard, R., Ivanov, M., & Marquardt, S. (2018). Transcriptional read-through of the long noncoding RNA SVALKKA governs plant cold acclimation. *Nature Communications*, 9(1), 4561. <https://doi.org/10.1038/s41467-018-07010-6>

- Kirk, J. M., Kim, S. O., Inoue, K., Smola, M. J., Lee, D. M., Schertzer, M. D., Wooten, J. S., Baker, A. R., Sprague, D., Collins, D. W., Horning, C. R., Wang, S., Chen, Q., Weeks, K. M., Mucha, P. J., & Calabrese, J. M. (2018). Functional classification of long non-coding RNAs by k-mer content. *Nature Genetics*, 50(10), 1474-1482. <https://doi.org/10.1038/s41588-018-0207-8>
- Kong, L., Zhang, Y., Ye, Z.-Q., Liu, X.-Q., Zhao, S.-Q., Wei, L., & Gao, G. (2007a). CPC: Assess the protein-coding potential of transcripts using sequence features and support vector machine. *Nucleic Acids Research*, 35(suppl_2), W345-W349. <https://doi.org/10.1093/nar/gkm391>
- Kong, L., Zhang, Y., Ye, Z.-Q., Liu, X.-Q., Zhao, S.-Q., Wei, L., & Gao, G. (2007b). CPC: Assess the protein-coding potential of transcripts using sequence features and support vector machine. *Nucleic Acids Research*, 35(suppl_2), W345-W349. <https://doi.org/10.1093/nar/gkm391>
- Körbes, A. P., Kulcheski, F. R., Margis, R., Margis-Pinheiro, M., & Turchetto-Zolet, A. C. (2016). Molecular evolution of the lysophosphatidic acid acyltransferase (LPAAT) gene family. *Molecular Phylogenetics and Evolution*, 96, 55-69. <https://doi.org/10.1016/j.ympev.2015.12.001>
- Kretz, M., Siprashvili, Z., Chu, C., Webster, D. E., Zehnder, A., Qu, K., Lee, C. S., Flockhart, R. J., Groff, A. F., Chow, J., Johnston, D., Kim, G. E., Spitale, R. C., Flynn, R. A., Zheng, G. X. Y., Aiyer, S., Raj, A., Rinn, J. L., Chang, H. Y., & Khavari, P. A. (2013). Control of somatic tissue differentiation by the long non-coding RNA TINCR. *Nature*, 493(7431), 231-235. <https://doi.org/10.1038/nature11661>
- Kung, J. T. Y., Colognori, D., & Lee, J. T. (2013). Long Noncoding RNAs: Past, Present, and Future. *Genetics*, 193(3), 651-669. <https://doi.org/10.1534/genetics.112.146704>
- Lee, Y., Kim, Y.-W., Jeon, B. W., Park, K.-Y., Suh, S. J., Seo, J., Kwak, J. M., Martinoia, E., Hwang, I., & Lee, Y. (2007). Phosphatidylinositol 4,5-bisphosphate is important for stomatal opening. *The Plant Journal*, 52(5), 803-816. <https://doi.org/10.1111/j.1365-313X.2007.03277.x>
- Li, D., Lyons, J. A., Pye, V. E., Vogeley, L., Aragão, D., Kenyon, C. P., Shah, S. T. A., Doherty, C., Aherne, M., & Caffrey, M. (2013a). Crystal structure of the integral membrane diacylglycerol kinase. *Nature*, 497(7450), 521-524. <https://doi.org/10.1038/nature12179>
- Li, D., Lyons, J. A., Pye, V. E., Vogeley, L., Aragão, D., Kenyon, C. P., Shah, S. T. A., Doherty, C., Aherne, M., & Caffrey, M. (2013b). Crystal structure of the integral membrane diacylglycerol kinase. *Nature*, 497(7450), 521-524. <https://doi.org/10.1038/nature12179>
- Li, J., & Liu, C. (2019). Coding or Noncoding, the Converging Concepts of RNAs. *Frontiers in Genetics*, 10, 496. <https://doi.org/10.3389/fgene.2019.00496>
- Li, J., Xuan, Z., & Liu, C. (2013). Long Non-Coding RNAs and Complex Human Diseases. *International Journal of Molecular Sciences*, 14(9), 18790-18808. <https://doi.org/10.3390/ijms140918790>
- Li, J., Xue, Y., Amin, M. T., Yang, Y., Yang, J., Zhang, W., Yang, W., Niu, X., Zhang, H.-Y., & Gong, J. (2020). ncRNA-eQTL: A database to systematically evaluate the effects of SNPs on non-coding RNA expression across cancer types. *Nucleic Acids Research*, 48(D1), D956-D963. <https://doi.org/10.1093/nar/gkz711>

- Li, J.-H., Liu, S., Zhou, H., Qu, L.-H., & Yang, J.-H. (2014). star Base v2.0: Decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data. *Nucleic Acids Research*, 42(D1), D92-D97. <https://doi.org/10.1093/nar/gkt1248>
- Lin, A., Hu, Q., Li, C., Xing, Z., Ma, G., Wang, C., Li, J., Ye, Y., Yao, J., Liang, K., Wang, S., Park, P. K., Marks, J. R., Zhou, Y., Zhou, J., Hung, M.-C., Liang, H., Hu, Z., Shen, H., ... Yang, L. (2017). The LINK-A lncRNA interacts with PtdIns (3,4,5) P3 to hyperactivate AKT and confer resistance to AKT inhibitors. *Nature Cell Biology*, 19(3), 238-251. <https://doi.org/10.1038/ncb3473>
- Liu, B., Sun, L., Liu, Q., Gong, C., Yao, Y., Lv, X., Lin, L., Yao, H., Su, F., Li, D., Zeng, M., & Song, E. (2015). A Cytoplasmic NF-κB Interacting Long Noncoding RNA Blocks IκB Phosphorylation and Suppresses Breast Cancer Metastasis. *Cancer Cell*, 27(3), 370-381. <https://doi.org/10.1016/j.ccell.2015.02.004>
- Liu, P., Yang, H., Zhang, J., Peng, X., Lu, Z., Tong, W., & Chen, J. (2017). The lncRNA MALAT1 acts as a competing endogenous RNA to regulate KRAS expression by sponging miR-217 in pancreatic ductal adenocarcinoma. *Scientific Reports*, 7(1), 5186. <https://doi.org/10.1038/s41598-017-05274-4>
- Liu, X., Li, D., Zhang, D., Yin, D., Zhao, Y., Ji, C., Zhao, X., Li, X., He, Q., Chen, R., Hu, S., & Zhu, L. (2018). A novel antisense long noncoding RNA, TWISTED LEAF, maintains leaf blade flattening by regulating its associated sense R2R3-MYB gene in rice. *New Phytologist*, 218(2), 774-788. <https://doi.org/10.1111/nph.15023>
- Liu, X., Zhai, S., Zhao, Y., Sun, B., Liu, C., Yang, A., & Zhang, J. (2013). Overexpression of the phosphatidylinositol synthase gene (ZmPIS) conferring drought stress tolerance by altering membrane lipid composition and increasing ABA synthesis in maize: ZmPIS and maize drought stress tolerance. *Plant, Cell & Environment*, 36(5), 1037-1055. <https://doi.org/10.1111/pce.12040>
- Liu, Y., Liu, B., Jin, G., Zhang, J., Wang, X., Feng, Y., Bian, Z., Fei, B., Yin, Y., & Huang, Z. (2019). An Integrated Three-Long Non-coding RNA Signature Predicts Prognosis in Colorectal Cancer Patients. *Frontiers in Oncology*, 9, 1269. <https://doi.org/10.3389/fonc.2019.01269>
- Liu, Y., Su, Y., & Wang, X. (2013). Phosphatidic Acid-Mediated Signaling. En D. G. S. Capelluto (Ed.), *Lipid-mediated Protein Signaling* (Vol. 991, pp. 159-176). Springer Netherlands. https://doi.org/10.1007/978-94-007-6331-9_9
- Long, Y., Wang, X., Youmans, D. T., & Cech, T. R. (2017). How do lncRNAs regulate transcription? *Science Advances*, 3(9), eaao2110. <https://doi.org/10.1126/sciadv.aao2110>
- Luo, M., Jeong, M., Sun, D., Park, H. J., Rodriguez, B. A. T., Xia, Z., Yang, L., Zhang, X., Sheng, K., Darlington, G. J., Li, W., & Goodell, M. A. (2015). Long Non-Coding RNAs Control Hematopoietic Stem Cell Function. *Cell Stem Cell*, 16(4), 426-438. <https://doi.org/10.1016/j.stem.2015.02.002>
- Ma, Q., Gabelli, S. B., & Raben, D. M. (2019a). Diacylglycerol kinases: Relationship to other lipid kinases. *Advances in Biological Regulation*, 71, 104-110. <https://doi.org/10.1016/j.jbior.2018.09.014>

- Ma, Q., Gabelli, S. B., & Raben, D. M. (2019b). Diacylglycerol kinases: Relationship to other lipid kinases. *Advances in Biological Regulation*, 71, 104-110. <https://doi.org/10.1016/j.jbior.2018.09.014>
- Maher, B. (2012). ENCODE: The human encyclopedia. *Nature*, 489(7414), 46-48. <https://doi.org/10.1038/489046a>
- Mercer, T. R., & Mattick, J. S. (2013). Structure and function of long noncoding RNAs in epigenetic regulation. *Nature Structural & Molecular Biology*, 20(3), 300-307. <https://doi.org/10.1038/nsmb.2480>
- Miao, H., Wang, L., Zhan, H., Dai, J., Chang, Y., Wu, F., Liu, T., Liu, Z., Gao, C., Li, L., & Song, X. (2019). A long noncoding RNA distributed in both nucleus and cytoplasm operates in the PYCARD-regulated apoptosis by coordinating the epigenetic and translational regulation. *PLOS Genetics*, 15(5), e1008144. <https://doi.org/10.1371/journal.pgen.1008144>
- Min, X. J., Butler, G., Storms, R., & Tsang, A. (2005). OrfPredictor: Predicting protein-coding regions in EST-derived sequences. *Nucleic Acids Research*, 33(Web Server), W677-W680. <https://doi.org/10.1093/nar/gki394>
- Neguembor, M. V., Jothi, M., & Gabellini, D. (2014). Long noncoding RNAs, emerging players in muscle differentiation and disease. *Skeletal Muscle*, 4(1), 8. <https://doi.org/10.1186/2044-5040-4-8>
- Novikova, I. V., Hennelly, S. P., & Sanbonmatsu, K. Y. (2012). Structural architecture of the human long noncoding RNA, steroid receptor RNA activator. *Nucleic Acids Research*, 40(11), 5034-5051. <https://doi.org/10.1093/nar/gks071>
- Obayashi, E., Oubridge, C., Krummel, D. P., & Nagai, K. (2007). Crystallization of RNA-Protein Complexes. En J. M. Walker & S. Doublé (Eds.), *Macromolecular Crystallography Protocols* (Vol. 363, pp. 259-276). Humana Press. https://doi.org/10.1007/978-1-59745-209-0_13
- Ortega-Baes P, Aparicio-González M, Galíndez G, del Fueyo P, Sühling S, Rojas-Aréchiga M. (2010). Are cactus growth forms related to germination responses to light? A test using *Echinopsis* species. *Acta Oecologica* 36: 339-342. DOI: <https://doi.org/10.1016/j.actao.2010.02.006>
- Panda, A., Thakur, R., Krishnan, H., Naik, A., Shinde, D., & Raghu, P. (2018). Functional analysis of mammalian phospholipase D enzymes. *Bioscience Reports*, 38(6), BSR20181690. <https://doi.org/10.1042/BSR20181690>
- Paytvi-Gallart, A., Sanseverino, W., & Aiese Cigliano, R. (2019). A Walkthrough to the Use of GreenC: The Plant lncRNA Database. En J. A. Chekanova & H.-L. V. Wang (Eds.), *Plant Long Non-Coding RNAs* (Vol. 1933, pp. 397-414). Springer New York. https://doi.org/10.1007/978-1-4939-9045-0_25
- Phan, T. K., Williams, S. A., Bindra, G. K., Lay, F. T., Poon, I. K. H., & Hulett, M. D. (2019). Phosphoinositides: Multipurpose cellular lipids with emerging roles in cell death. *Cell Death & Differentiation*, 26(5), 781-793. <https://doi.org/10.1038/s41418-018-0269-2>

- Pinkney, H. R., Wright, B. M., & Diermeier, S. D. (2020). The lncRNA Toolkit: Databases and In Silico Tools for lncRNA Analysis. *Non-Coding RNA*, 6(4), 49. <https://doi.org/10.3390/ncrna6040049>
- Potocký, M., Pleskot, R., Pejchar, P., Vitale, N., Kost, B., & Žárský, V. (2014). Live-cell imaging of phosphatidic acid dynamics in pollen tubes visualized by S po20p-derived biosensor. *New Phytologist*, 203(2), 483-494. <https://doi.org/10.1111/nph.12814>
- Qin, T., Zhao, H., Cui, P., Albeshar, N., & Xiong, L. (2017). A Nucleus-Localized Long Non-Coding RNA Enhances Drought and Salt Stress Tolerance. *Plant Physiology*, 175(3), 1321-1336. <https://doi.org/10.1104/pp.17.00574>
- Quinn, J. J., & Chang, H. Y. (2016). Unique features of long non-coding RNA biogenesis and function. *Nature Reviews Genetics*, 17(1), 47-62. <https://doi.org/10.1038/nrg.2015.10>
- Raben, D. M., & Barber, C. N. (2017a). Phosphatidic acid and neurotransmission. *Advances in Biological Regulation*, 63, 15-21. <https://doi.org/10.1016/j.jbior.2016.09.004>
- Raben, D. M., & Barber, C. N. (2017b). Phosphatidic acid and neurotransmission. *Advances in Biological Regulation*, 63, 15-21. <https://doi.org/10.1016/j.jbior.2016.09.004>
- Rai, M. I., Alam, M., Lightfoot, D. A., Gurha, P., & Afzal, A. J. (2019). Classification and experimental identification of plant long non-coding RNAs. *Genomics*, 111(5), 997-1005. <https://doi.org/10.1016/j.ygeno.2018.04.014>
- Ratti, S., Ramazzotti, G., Faenza, I., Fiume, R., Mongiorgi, S., Billi, A. M., McCubrey, J. A., Suh, P.-G., Manzoli, L., Cocco, L., & Follo, M. Y. (2018). Nuclear inositide signaling and cell cycle. *Advances in Biological Regulation*, 67, 1-6. <https://doi.org/10.1016/j.jbior.2017.10.008>
- Ribeiro, D. M., Zanzoni, A., Cipriano, A., Delli Ponti, R., Spinelli, L., Ballarino, M., Bozzoni, I., Tartaglia, G. G., & Brun, C. (2018). Protein complex scaffolding predicted as a prevalent function of long noncoding RNAs. *Nucleic Acids Research*, 46(2), 917-928. <https://doi.org/10.1093/nar/gkx1169>
- Rigoutsos, I., Lee, S. K., Nam, S. Y., Anfossi, S., Pasculli, B., Pichler, M., Jing, Y., Rodriguez-Aguayo, C., Telonis, A. G., Rossi, S., Ivan, C., Catela Ivkovic, T., Fabris, L., Clark, P. M., Ling, H., Shimizu, M., Redis, R. S., Shah, M. Y., Zhang, X., ... Calin, G. A. (2017). N-BLR, a primate-specific non-coding transcript leads to colorectal cancer invasion and migration. *Genome Biology*, 18(1), 98. <https://doi.org/10.1186/s13059-017-1224-0>
- Robinson, M. D., McCarthy, D. J., & Smyth, G. K. (2010). edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, 26(1), 139-140. <https://doi.org/10.1093/bioinformatics/btp616>
- Rodriguez-Villalon, A., Gujas, B., van Wijk, R., Munnik, T., & Hardtke, C. S. (2015). Primary root protophloem differentiation requires balanced phosphatidylinositol-4,5-biphosphate levels and systemically affects root branching. *Development*, dev.118364. <https://doi.org/10.1242/dev.118364>

- Romero-Barrios, N., Legascue, M. F., Benhamed, M., Ariel, F., & Crespi, M. (2018). Splicing regulation by long noncoding RNAs. *Nucleic Acids Research*, 46(5), 2169-2184. <https://doi.org/10.1093/nar/gky095>
- Sanchita, Trivedi, P. K., & Asif, M. H. (2020). Updates on plant long non-coding RNAs (lncRNAs): The regulatory components. *Plant Cell, Tissue, and Organ Culture (PCTOC)*, 140(2), 259-269. <https://doi.org/10.1007/s11240-019-01726-z>
- Schmidt, K., Weidmann, C. A., Hilimire, T. A., Yee, E., Hatfield, B. M., Schneekloth, J. S., Weeks, K. M., & Novina, C. D. (2020). Targeting the Oncogenic Long Non-coding RNA SLNCR1 by Blocking Its Sequence-Specific Binding to the Androgen Receptor. *Cell Reports*, 30(2), 541-554.e5. <https://doi.org/10.1016/j.celrep.2019.12.011>
- Schmitz, S. U., Grote, P., & Herrmann, B. G. (2016). Mechanisms of long noncoding RNA function in development and disease. *Cellular and Molecular Life Sciences*, 73(13), 2491-2509. <https://doi.org/10.1007/s00018-016-2174-5>
- Seo, D., Kim, D., & Kim, W. (2019). Long non-coding RNA linc00152 acting as a promising oncogene in cancer progression. *Genomics & Informatics*, 17(4), e36. <https://doi.org/10.5808/GI.2019.17.4.e36>
- Seo, J. S., Sun, H.-X., Park, B. S., Huang, C.-H., Yeh, S.-D., Jung, C., & Chua, N.-H. (2017). ELF18-INDUCED LONG-NONCODING RNA Associates with Mediator to Enhance Expression of Innate Immune Response Genes in Arabidopsis. *The Plant Cell*, 29(5), 1024-1038. <https://doi.org/10.1105/tpc.16.00886>
- Sobol, M., Krausová, A., Yildirim, S., Kalasová, I., Fáberová, V., Vrkoslav, V., Philimonenko, V., Marášek, P., Pastorek, L., Čapek, M., Lubovská, Z., Uličná, L., Tsuji, T., Lísa, M., Cvačka, J., Fujimoto, T., & Hozak, P. (2018). Nuclear phosphatidylinositol 4,5-bisphosphate islets contribute to efficient RNA polymerase II-dependent transcription. *Journal of Cell Science*, jcs.211094. <https://doi.org/10.1242/jcs.211094>
- Spokoini-Stern, R., Stamov, D., Jessel, H., Aharoni, L., Haschke, H., Giron, J., Unger, R., Segal, E., Abu-Horowitz, A., & Bachelet, I. (2020). Visualizing the structure and motion of the long noncoding RNA HOTAIR. *RNA*, 26(5), 629-636. <https://doi.org/10.1261/rna.074633.120>
- Sun, M., Gadad, S. S., Kim, D.-S., & Kraus, W. L. (2015). Discovery, Annotation, and Functional Analysis of Long Noncoding RNAs Controlling Cell-Cycle Gene Expression and Proliferation in Breast Cancer Cells. *Molecular Cell*, 59(4), 698-711. <https://doi.org/10.1016/j.molcel.2015.06.023>
- Sun, Z., Jing, C., Xiao, C., & Li, T. (2020). An autophagy-related long non-coding RNA prognostic signature accurately predicts survival outcomes in bladder urothelial carcinoma patients. *Aging*, 12(15), 15624-15637. <https://doi.org/10.18632/aging.103718>
- Sztacho, M., Sobol, M., Balaban, C., Escudeiro Lopes, S. E., & Hozák, P. (2019). Nuclear phosphoinositides and phase separation: Important players in nuclear compartmentalization. *Advances in Biological Regulation*, 71, 111-117. <https://doi.org/10.1016/j.jbior.2018.09.009>
- Takenawa, T. (2010). Phosphoinositide-binding interface proteins involved in shaping cell membranes. *Proceedings of the Japan Academy, Series B*, 86(5), 509-523. <https://doi.org/10.2183/pjab.86.509>

- Tanguy, E., Wang, Q., Moine, H., & Vitale, N. (2019). Phosphatidic Acid: From Pleiotropic Functions to Neuronal Pathology. *Frontiers in Cellular Neuroscience*, 13, 2. <https://doi.org/10.3389/fncel.2019.00002>
- Tejos, R., Sauer, M., Vanneste, S., Palacios-Gomez, M., Li, H., Heilmann, M., van Wijk, R., Vermeer, J. E. M., Heilmann, I., Munnik, T., & Friml, J. (2014). Bipolar Plasma Membrane Distribution of Phosphoinositides and Their Requirement for Auxin-Mediated Cell Polarity and Patterning in Arabidopsis. *The Plant Cell*, 26(5), 2114-2128. <https://doi.org/10.1105/tpc.114.126185>
- Thakur, R., Naik, A., Panda, A., & Raghu, P. (2019a). Regulation of Membrane Turnover by Phosphatidic Acid: Cellular Functions and Disease Implications. *Frontiers in Cell and Developmental Biology*, 7, 83. <https://doi.org/10.3389/fcell.2019.00083>
- Thakur, R., Naik, A., Panda, A., & Raghu, P. (2019b). Regulation of Membrane Turnover by Phosphatidic Acid: Cellular Functions and Disease Implications. *Frontiers in Cell and Developmental Biology*, 7, 83. <https://doi.org/10.3389/fcell.2019.00083>
- The ENCODE Project Consortium. (2007). Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature*, 447(7146), 799-816. <https://doi.org/10.1038/nature05874>
- Toska, E., Campbell, H. A., Shandilya, J., Goodfellow, S. J., Shore, P., Medler, K. F., & Roberts, S. G. E. (2012). Repression of Transcription by WT1-BASP1 Requires the Myristylation of BASP1 and the PIP2-Dependent Recruitment of Histone Deacetylase. *Cell Reports*, 2(3), 462-469. <https://doi.org/10.1016/j.celrep.2012.08.005>
- Tsuji, T., Takatori, S., & Fujimoto, T. (2019). Definition of phosphoinositide distribution in the nanoscale. *Current Opinion in Cell Biology*, 57, 33-39. <https://doi.org/10.1016/j.ceb.2018.10.008>
- Ufer, G., Gertzmann, A., Gasulla, F., Röhrig, H., & Bartels, D. (2017). Identification and characterization of the phosphatidic acid-binding A. thaliana phosphoprotein PLDrp1 that is regulated by PLD α 1 in a stress-dependent manner. *The Plant Journal*, 92(2), 276-290. <https://doi.org/10.1111/tbj.13651>
- Ulicna, L., Kalendova, A., Kalasova, I., Vacik, T., & Hozák, P. (2018). PIP2 epigenetically represses rRNA genes transcription interacting with PHF8. *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids*, 1863(3), 266-275. <https://doi.org/10.1016/j.bbalip.2017.12.008>
- Ulitsky, I., Shkumatava, A., Jan, C. H., Sive, H., & Bartel, D. P. (2011). Conserved Function of lincRNAs in Vertebrate Embryonic Development despite Rapid Sequence Evolution. *Cell*, 147(7), 1537-1550. <https://doi.org/10.1016/j.cell.2011.11.055>
- Vodicka, P., Mo, S., Tousley, A., Green, K. M., Sapp, E., Iuliano, M., Sadri-Vakili, G., Shaffer, S. A., Aronin, N., DiFiglia, M., & Kegel-Gleason, K. B. (2015). Mass Spectrometry Analysis of Wild-Type and Knock-in Q140/Q140 Huntington's Disease Mouse Brains Reveals Changes in Glycerophospholipids Including Alterations in Phosphatidic Acid and Lyso-Phosphatidic Acid. *Journal of Huntington's Disease*, 4(2), 187-201. <https://doi.org/10.3233/JHD-150149>

- Wang, J., Choi, J.-M., Holehouse, A. S., Lee, H. O., Zhang, X., Jahnel, M., Maharana, S., Lemaitre, R., Pozniakovsky, A., Drechsel, D., Poser, I., Pappu, R. V., Alberti, S., & Hyman, A. A. (2018). A Molecular Grammar Governing the Driving Forces for Phase Separation of Prion-like RNA Binding Proteins. *Cell*, 174(3), 688-699.e16. <https://doi.org/10.1016/j.cell.2018.06.006>
- Wang, K., Sun, T., Li, N., Wang, Y., Wang, J.-X., Zhou, L.-Y., Long, B., Liu, C.-Y., Liu, F., & Li, P.-F. (2014). MDRL lncRNA Regulates the Processing of miR-484 Primary Transcript by Targeting miR-361. *PLoS Genetics*, 10(7), e1004467. <https://doi.org/10.1371/journal.pgen.1004467>
- Wang, S., Zhang, W., Wu, X., Weng, M., Zhang, M., Cai, Q., Zhou, D., Wang, J., & Quan, Z. (2016). The lnc RNA MALAT 1 functions as a competing endogenous RNA to regulate MCL -1 expression by sponging miR-363-3p in gallbladder cancer. *Journal of Cellular and Molecular Medicine*, 20(12), 2299-2308. <https://doi.org/10.1111/jcmm.12920>
- Wang, X., Lv, G., Li, J., Wang, B., Zhang, Q., & Lu, C. (2017). lncRNA-RP11-296A18.3/miR-138/HIF1A Pathway Regulates the Proliferation ECM Synthesis of Human Nucleus Pulposus Cells (HNPCs). *Journal of Cellular Biochemistry*, 118(12), 4862-4871. <https://doi.org/10.1002/jcb.26166>
- Wu, Y., Zhao, Y., Huan, L., Zhao, J., Zhou, Y., Xu, L., Hu, Z., Liu, Y., Chen, Z., Wang, L., Huang, S., He, X., & Liang, L. (2020). An LTR Retrotransposon-Derived Long Noncoding RNA lncMER52A Promotes Hepatocellular Carcinoma Progression by Binding p120-Catenin. *Cancer Research*, 80(5), 976-987. <https://doi.org/10.1158/0008-5472.CAN-19-2115>
- Wu, Z., Liu, X., Liu, L., Deng, H., Zhang, J., Xu, Q., Cen, B., & Ji, A. (2014). Regulation of lncRNA expression. *Cellular and Molecular Biology Letters*, 19(4). <https://doi.org/10.2478/s11658-014-0212-6>
- Wutz, A. (2011). Gene silencing in X-chromosome inactivation: Advances in understanding facultative heterochromatin formation. *Nature Reviews Genetics*, 12(8), 542-553. <https://doi.org/10.1038/nrg3035>
- Xiao, H., Tang, K., Liu, P., Chen, K., Hu, J., Zeng, J., Xiao, W., Yu, G., Yao, W., Zhou, H., Li, H., Pan, Y., Li, A., Ye, Z., Wang, J., Xu, H., & Huang, Q. (2015). lncRNA MALAT1 functions as a competing endogenous RNA to regulate ZEB2 expression by sponging miR-200s in clear cell kidney carcinoma. *Oncotarget*, 6(35), 38005-38015. <https://doi.org/10.18632/oncotarget.5357>
- Yamazaki, T., Souquere, S., Chujo, T., Kobelke, S., Chong, Y. S., Fox, A. H., Bond, C. S., Nakagawa, S., Pierron, G., & Hirose, T. (2018). Functional Domains of NEAT1 Architectural lncRNA Induce Paraspeckle Assembly through Phase Separation. *Molecular Cell*, 70(6), 1038-1053.e7. <https://doi.org/10.1016/j.molcel.2018.05.019>
- Yoon, J.-H., & Gorospe, M. (2016). Cross-Linking Immunoprecipitation and qPCR (CLIP-qPCR) Analysis to Map Interactions Between Long Noncoding RNAs and RNA-Binding Proteins. En Y. Feng & L. Zhang (Eds.), *Long Non-Coding RNAs* (Vol. 1402, pp. 11-17). Springer New York. https://doi.org/10.1007/978-1-4939-3378-5_2

Zegarłinska, J., Piaścik, M., Sikorski, A. F., & Czogalla, A. (2018a). Phosphatidic acid – a simple phospholipid with multiple faces. *Acta Biochimica Polonica*, 65(2), 163-171. https://doi.org/10.18388/abp.2018_2592

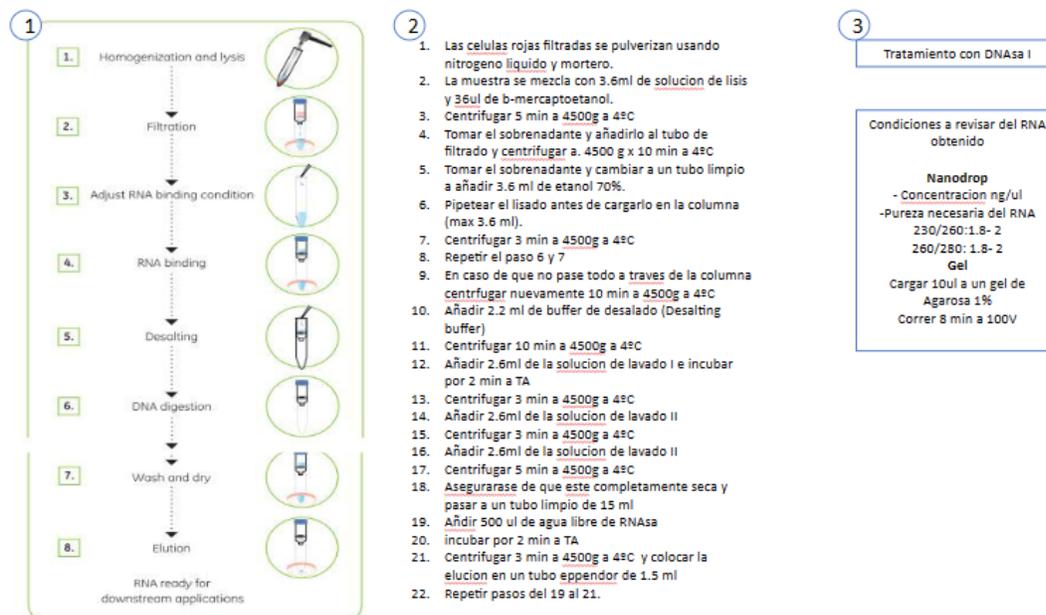
Zegarłinska, J., Piaścik, M., Sikorski, A. F., & Czogalla, A. (2018b). Phosphatidic acid – a simple phospholipid with multiple faces. *Acta Biochimica Polonica*, 65(2), 163-171. https://doi.org/10.18388/abp.2018_2592

Zhang, W., Yue, X., Tang, G., Wu, W., Huang, F., & Zhang, X. (2018). SFPEL-LPI: Sequence-based feature projection ensemble learning for predicting LncRNA-protein interactions. *PLOS Computational Biology*, 14(12), e1006616. <https://doi.org/10.1371/journal.pcbi.1006616>

Zhukovsky, M. A., Filograna, A., Luini, A., Corda, D., & Valente, C. (2019). Phosphatidic acid in membrane rearrangements. *FEBS Letters*, 593(17), 2428-2451. <https://doi.org/10.1002/1873-3468.13563>

Zuker, M., & Stiegler, P. (1981). Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information. *Nucleic Acids Research*, 9(1), 133-148. <https://doi.org/10.1093/nar/9.1.133>

ANEXO(S)

Anexo.1. Protocolo de Extracción de RNA utilizando el PureLink® RNA Mini Kit.**Anexo.2.** Información sobre la biblioteca genómica obtenido (sin genoma de referencia).

Contract Information	Contract Content
Contract ID	H202SC21112222
Batch ID	X202SC21112222-Z01-F004
Species	stenocereus queretaroensis
Report Time	2022-03-04

Anexo.3.**>Sec_1_PA (Clúster-18316.12632_1)**

```

ATGCATTTGGACCATCGTTGATTCTCATGTTTGGCGCACTGAGGCGCAAAGGGTCCCGAG
GCGCGAGGCGAGAGGCGCGAGGCGAAGGCGCACGCATCATGGACGCGAGGCGCACTTT
TACGCTATAATGCTTCCAAATCAATCTTAATGCACAGTTTTGACAAATAAGCTAAATAATGAT
ATCAAAGAAGTGATTAGACAACAGTAAAGTTTAGAAAACATCGCAAATAGTCTTAAGGACA
CGATGATATCGTATCAGAGCACCTGGAATCACACATATAAGTTACCACAGTCTAGAAACATA
CAAGCAAGCACAAACACACAAAGCAAAGAGTCCCTTAAATCTCCCAAAGGTCAGCTAAATG
CCATCCTCATCATCATCCAACACCACATGCTTAGGCTCAACATCATCACCAAACACCTCATC
TTCATTTCCAATGTCTCTTCCTCTCTTCCTCATCCTCTTCATATTCTCTTCCTCATCTTCAT
CAATCAATTGAACTCGCTCTCGGGAAGATTTAGAAGCGGAGGCAGGGGCTTTGCCTTTGTG
CTTTGCCGAAGTCTAGAGGTTGTAGTGCTCTTTGATACCCTACCTCTAGTCCCATAGCGA
GGCTCATTTGCCCCACAAGCTTGGGCAACAGCAGCCCATGTCAAAGTGTATCCTCAAACA
CAAGGTCATCCTCATCCGACTCTCCTTCCATTCTCCCAAGTAACCACTCATTACTTTTCATCA
ATCTCTTTGAGAAGAATGGGATCCCTAGAGGAGCTCTCTTCGCGCCTTCGTTCAAAGCAC

```

GATTCGATTTGACAAAAACCAAGTCATTCAATCGAGATTGTGCTAATCTGTTTCTCTTCTTG
 GTGTGAAGCTGTAAATAAGTACACTTCATTAATCTCTAAAAAGGTAAAAAGTATTTAGCTATC
 AAAGAATAAAATGATTTTGTGGATGTTGCAACTTACATGTTGGAAAACACCCCAATTCCTCT
 CGCATCCAGTTGCACTACAAGTGAGGCTAAGGACCCTTATCGCAAACCTTTTTCAAGTGAGG
 TGCCGAGGATCCATATGAACTCCACCACTCAGCTATTAGAAACCAAGAAAGGAAACAATTT
 AATTGAATATTTGTGCTTTTTAAAAAAGAATAAACCTAAATTGAATATTGTGTTAAATATTCTT
 ACCGGGAGCTAGAGTCTTCCTTTGCCTTACGGCCATGTTGATTCCAAAGAGATCTTGTGCA
 CTCTTGATAAATTCAATTCCTGAAGAATCTTATCTTGGGTCCCAATTCCGGAACCAACCT
 CGAAATGCAATCATACAAGCCCCTCATCACTTCACTACAATCTTCAATAGTTGGGTTTGAAT
 AATAAATATTGGGATTTAGAAAATGTCCGGCCGCATGCAAAGGTTGATGAAGTTGGCATT
 CCATCTTGTATCAATGTATTTAAATGTATCTCATAATGCTCTTCCTTGTGTCCAAAGCTAGT
 AGCAATAGCCTCCTTAGCTCTATCCATGGCCTCATAAATGTAACCCATAGGAGGTTTTTTTT
 CTCCATCTACTATTCGAAGAACCTTAACAATAGGGCCTCTCAATTTAAGAGCATAGAGAATT
 CTCCAATAGCTTTCTTGCAAAGATAAGATGCCATCTTCTTTCCCCCGGAGTCCTTTGACC
 ATTTGCTATCATTCCACTCTTGAGAAGTCACCATTTTTCTCAAATTGTTCTTTTGTGTTGTGAA
 TGGAGGAAAGAGTGATGAAGGATGTAGCAAATCTGGTGACTGCGGGTCTATGCAAGTTCC
 TCTGGCCAGTGAATCTCCTCATCATGTTTATTAAGGCAACCGAGCTATGTATATAAGCATT
 ATGGCCATACTCTTCGGTAAAGCAATCTTGACTTTTTGGAATGTTCTTTCCAATATCTTCTAAA
 ATGAGATCAAGACAATGAGCGGCACATGGAGTCCAATAAAGATGCGGTCTCTTAGCCATTA
 GCAATTTTCTACAATCATTTTTTATAACAATAAAATTAGAATTTAAAAAATAGAAACAAGATA
 CATAATTACATATGTATCATTCAAGAATTACGACAAAGAAAATTAAGGCTCTTAGTTTCTTAC
 CCGCCTTACGTAATTAGATGCGTTGTCGGTCACAACCTTGACCACATTCTCTTCTCCAAC
 TCTTCTACCATATCATCAAGAACTTTGAACAAGAAGTTGGCATCCTTTATAACTTCGGAAAC
 ATCCATAGATCTAATAAAGACCGAACCTTTGGGAGAGTTGACAAGGAAGTTGACAATATCC
 TTTTGCACGGTTGTATCACGCCAACCATCGGATAAGATTGAACATCCTTTATGTGCCCACTC
 TTTCTTGTGCTCTACCATTGCCTTTTCGGTTGCCTCCACTTCTTTCTTGTAGGAAAGGAACTC
 TT**AGCTCATACATGGAAGGCGG**TTTCAATCCGGGACCAAATTGATCAATTGATTCACTCATA
 ATAGCAAACCTTTCTAATGTGGCCGCATGAAAAGTAATTCTGGCATCAGAAAACACCTAG
 CAATGTCTCCACAAGCCTTCTCCCTTAGTTCTTTTCTACACAATTCATTTATTGTTTTTACCT
CCCTTGCTCGTTTTGAGCTTTGAGAACATCCGTGGGATCATGAGC
 AAAAAACACATCTAAA
 GGGCCTTTTTGTCTTGGCTTTTTAGAAGGCATGCTACTACTCCCTTGGAGCCGCCACTTG
 AAATTTTCCACTTCCACCTTTTAAATTCATTTCAACACAATCATCATTATCATCATCATC
 ATCTTGGATTTCTCACCCATGTAATGTTGATGAATGTGTCTAGACATATTATGTCTTTGAT
 GGTCTTTTTGCGCTCCATATAATCTTGGATTTCTTTTCTGTCATGTTCCGGGACTTTTTCGAC
 AGCCTAGGGTGTTCGATGACCACTGGCCATGTGTTGTTAAGTCGAAAAACTCCTCCCGTG
 GCCTCACACGCAAAGTCGCAATACCAAGTGTCTATTGTTTGA

Los oligos **(Morado)** de Sec_1_PA amplifican la secuencia siguiente **(AMARILLO)** de 230 nt.

AGCTCATACATGGAAGGCGGTTTCAATCCGGGACCAAATTGATCAATTGATTCACTCATAAT
 AGCAAACCTTTCTAATGTGGCCGCATGAAAAGTAATTCTGGCATCAGAAAACACCTAGCA
 ATGTCTCCACAAGCCTTCTCCCTTAGTTCTTTTCTACACAATTCATTTATTGTTTTTACCTC
 CCTTGCTCGTTTTGAGCTTTGAGA**ACATCCGTGGGATCATGAGC**

FwSec_1_PA

5'- AGC TCA TAC ATG GAA GGC GG -3'

RvSec_1_PA

3'- GCT CAT GAT CCC ACG GAT GT -5'

>Sec_2_PA (Cluster-27261.1_1)

CGAATTATTATGATAAGCTGTTCTGTTCTATGGCTTTTTGTTTGCTGTGAGATGTTCTTCCTT
 TGTGAATATGCGAGTGCTGTCATCTGACTCATATTTGGGTTTCATGTGATAGCTGGCATGA
 AGTTGTTTTAAGATGAAGACTTGTAAGGGGAAAAATGTAGTATATGCAGTGTTCTAGTGATT
 TCCTGTAGAATCCCCGGTTGAACTGTTGGCAGTTGCTGAGAAAAGTTGAGGTTGTTTATTCT
 GTAAATTGGCCATGTTGATTTTACCACTTGACGACTTGTTGGTGGATGAGATCCCGTTACAT
 ATATGGGCTTCTACATGTATGGGCATATGTATCTTCAGTGCTGAGTAGGTGCTTTCGGCTC
 TTTGTACGGGTCGTTGAGATTAGATGGTGGAACATGGTTTGTAACTATATGGTTTGTAACT
 TCTGTGTCCGTGATGAGAGCTGAGTTGTGGTTGCCGAGTTAGGAGAGTGTCATGCTGCTG
 TGTGTCAACTTTGTGAGATCCTTATGTTTGTATGAGCAAGAATACATTCAGCCGAGAATCC
 TTGTACATGCAATTACCTGCTGCCTGTAAGGCCATTGTGGGTTCAAGTTGATGTCTCCACAT
 GTGATAGGCTTCTGGATGTGTATGAGTTTGAAGTTCTACAGCAATTACCTGCTGCCTGTGA
 GGCCATTGTGAGTTCAAGTTGATGTCTCCACGGGTGACTGGTTTCTGGATGTGTTAGAGTTC
 ACTTCTACAGCACAATGATTGGATTCAAGGAGAATCTCCATGTTTGGTGCCTATGTATGGAA
 CTTGACTGCTCATCCTCATTTTCCATTGCATTGTTCCGGTGGTGCTTTTGTGACCAATGATAG
 AGTGTTCCTATTTTGTCTGCTGTTTGTGGCAGGGTATTAAGAATTGTACCCACATAGCA
 GATGAGACCTTGTAAAGTTGGTGCCAGAATTTGATTTGAACATTCTGGGTGCTTGTGGATCA
 CCATGCTCAGAGGACTGCGTTATTTTCTTCCCTGTGGGTACTGAATTTGGAATACTCCG
 TGACAGTGTGCC TGAATATTTTCATGCTTCTTGGAGTAATCGCGCGTTATTTCCACTGTGTGA
 TGGCATTCTTATGGTATTGTTTGTGTCAGGCTAATGTTGAAACAGTGTGATTATTTGCTTTTGT
 AAGGCACAAGGCAGTCGGGCACATTTGTTGTAAGCTGCGTTTAAATCATCATTGTTACTATT
 ATAAGCCTCTCATTGTATGCTCTTTAGGCTGAGGCACGCTTGTGAGAAGCTACCATTAATC
 ATCGTCGATGCTATTGCATGGAGGTTAATGTGAAATTGAGAATCTATCTGTTCTGGTTCATG
 AATCTTTCTCCCCCATGTAGTGTAACCTTTCGAATTTAATCAAGGAACTAATAAACTATG
 GTGAGATTTTTGTTGGTGTGTAACCTTCGAATTTAATCAAGGTAAGTACTGAGGTGAGATTTTTG
 TGGTGTGTAATGTATTGTGGCGTGAATTTCTGAGGGGGATGTGGTACGTAAAGGTTTTTTCT
 GTTTTATTTTTTTGAACAAATGTTGAGGATGGTAGCCTGGGTTTCTTATTGTGCTGTTTAGAT
 TGGGTCTTATGTAGCGGCTGTATGATGTATTGATCTCTCCGGTGTGCTTTTGGCTGTTAT
 TGACTTTACCGTTTAGATTTGGTTTTAATTTAGTAACTCTGTGATTATTGGTTATTGATTTTT
 TCCAGTCTCTTTGTTTTAGCTATTGTGATGCTCGTTTATTTGTTCCAGTCTCTCATCATTTT
 CGTGGCAGCCGAGGAAGAAAGAAGCTCAGAATAACTGATGAGAGTCTTGATTGGAGCTT
 ACCGCCAGATTGCAATCCTCTGTTTGCAAATGTACTTGTGTTGGCATCGAATTGTGGATT
 AAATATGTCCGTTTAGCTTGTAGATACGGTCTTCGTGCTGCACGCATTGCGGCATCTTCAT
 GGTGGATCAGTTGAAGACATTGAGATATCAGCAATCTTATCTTACCTTGGAGAAGATTCCCT
 TTGTCCGAGGAAGACATTTGTCTATAATGTTGCCATTTATTATGGGAAAGATGGAAGACGT
 GATTTTGTCTTACGACTTGTACAGTTTATTCTATATTGCTATGTGCGACATGAGAGAATTG
 TGGGTATGGATTGATAAGCTCTTTCTGAAGCGGAAAGAGTGCCCGCATTGAGACCAAGTCT
 GGGAGGTTGGTGGTTCGAGTCCGCCTAGATCCTCCTGTTCTTTTCTCTTTTTGTTGAGGGA
 ATTTTGTAGATATTGGGTGGAGTGGAGTGGAGCAAGAGGTGACCAACAATTTGCACTTGTGTTG
 AAGTGTGTTGGAATTTTGTCTTTTTCACACTTTTGGTTTGGATCTTTTCTTTTTTAAATTAGA
 AGTAGGAAATAAGAATTTAAGTTAACCCTCATGAATATAATTGATTAATTGATAATTACATGT
 CCATATTGATAAAGGAAGTAGAATGTTGTTGGATGGATGTTGGATCCACCGCAATGTATGTT
 GTAAAGCGCTGGA

Los oligos (Morado) de Sec_3_PIP2 amplifican la secuencia siguiente (AMARILLO) de 231 nt.

GCATTGTTCCGGTGGTGCTTTTGTGACCAATGATAGAGTGTTCCTATTTTGTCTGCTGTTTTG
 TTGGCAGGGTATTAAGAATTGTACCCACATAGCAGATGAGACCTTGTAAAGTTGGTGCCAG
 AATTTGATTTGAACATTCTGGGTGCTTGTGGATCACCATGCTCAGAGGACTGCGTTATTTTC
 TTCCCTGTGGGTACTGAATTTGGAATACTTCCGGTGACAGTGTGCC

FwSec_3_PIP2

5'- GCA TTG TTC GGT GGT GCT TT -3'

RvSec_3_PIP2

3'- GGC ACA CTG TCA CCG AAG TA -5'

Sec_4_PIP2 (Clúster-30630.0_1)

GTTCAAATAGCCTCGATTTTCCTTGAATTAATTTGGCATCAACCAAAGGAGCATAACACAT
ACATATTTATGGCTCCAAAAGAGAAGTTTCGAGTGGTTTAACACTTGTCCCTTCAATAAGG
TCTCGAGTTTCGATTCCTGTAAATGGAGAAAATTTGTAATGGAAGAGCTCTACTCTGTATTGG
GTCAATAGTAGCTCGACTCCATTAAAACAGAAAATGCTCCTTTTTTTGGCTTTTTGGGAAGGG
AAACAACACAATAACAGTAGAAATAGTCAAAGATACATGCATCCTAAATACATAAGCCTAAA
TTGCCCTTTAGCTAAGCCATGAACAATTACCAAATGGCAACCCTACAGTAACTTCTACTG
ATACTTAGACCCCTACAGCTAGCAAAGAAGCTCTCCAGAAAGCTTTAGTTCTTGTCTTTTT
CAGTTCTGAGTTCTGACCCGACAGAGTTGTCAAGTTCAGTTCCTACTGCATCATCTGATTG
ATTTTGC GGGCTTCTTCATCTTGTAAAGATCCACCTGACCTTGGGGGGAGGCAGGAGCAAG
GGCGCGCAAACGATTTGTTCCATTTGATCCTTGTGCGCGTAGCTAGCATGGTTGCTGCT
AGGTTAACTTGCTTCGAAACACAGTGGATGGTTATTGTCCGCTTCTCTGAAGCTGCTCTC
TGATTCACGCCACCAGATATGCACTAGATCTATCTAGGAAAAGATTGTGGTGGACGGTCGA
AGCTAAATAACATCAAGTGCAAGGAAACCATGTAGCCATGTAGGTATCAGTTCTGGACGGT
AATGCGTAGATCTACAACAGAATATTATCTACGGTCTAAGACCTTCAGAATCTGAGGTAAGC
AGTCAGAGCCACTCCACCCCCAGAAAGGTTGGAGATGCATCTATTTACAATAAAGCTCA
AACAGATGAAAACAAGACAGACACTCAGAACA AAAATGCTAATATTGCATCTCCAATAACCA
GGCAGAGAACCAGTCAAGCTCAAATCAATTAGCCAACTAGAAATGGCTCACACAACCAAA
GTGAAAATAGAGCGGGCCAACTTATCATCTCTGCAGGCAACTACATCCATGAGCCCAGTT
GGGTTGGTGAATTCAGATTCCTATCTCCTATAGGCTGATTCATCCCAATCCATGGAACATG
CACCAATAGTCAGTTCAGCCTGAAGACTGCTTTGAGTGTATCAATCCAATAACTTACTCTGG
CATGCGATGTACCAACCCAATGGCCATCAGTAACCAGTATACAAATTCCCACCTACAAACA
CAGAGAAAGAATTATGACACCCTAAATTACATTCAAGAACTCTTTAAGCTACTTCGAACTTC
AAAGTGATTCAAAGTATGTATTTAAAACCAGAGCCAACAATAATGAAGTGAGCAGGAAAG
CTCAAGGGAAGGATCAGCACAAAGCTTTTTATCAATTGCAGGAAGCA **TTGCAGGTGCATGT**
AGGGAGAAGTCTAGACTGGCATTCAATTCTAGAATGCAGCCACGCCTTTTTATTCAACCC
CATAACCAAGTTGAAGCCTGGCCTTAGTTTCTGTAGAAGAGGTTGCGAATGTCACAGTGCAA
CTCGTGCAAGGACAAGGGCCCTCTTCACGTTGGTTACTACCTGCGACACTATCCTGATCAG
ACCTTTTACTTATCATCATATCCATTCCCCTTTTATGAATCAGTCACAACACTCAATGCA
GCATGGATGATGTA AACACAACGATAAAAAACAAAAGAAAACAGCTACCAAACCCAGTAT
GTGTCCACAGAACAAAGGCCTGTGCCTGTGGATAACCAAGTCTTGGCCCTTGGCAAGTGACA
GAATGCACAACATTTCAACAGATTTAATCCCAGTCTGTTCCGCCAGAATATAGATCCTAGCAA
CCCACAATATTTCAAAAACCGGAAGAATCTCACAACACTACACCATAATGGACATTCAACAA
TCCAACATGGGTATGGCCAGATCCACGTAAGTACGTAACACTAGGCCTGCCATGGATAGC
AGATCAACACTCTACACACAACAACACTGTAAGTGC AAAAGCCTTCTATAAGAAATAGGAA
CATTATGTTACATAGCATATGCATTCAAACCTTACAGAAAAGTGAAAACACAATGGAGTTGA
ACATCTAAGAAGTTAATTGTTTCATCCTCTCCATAGTGATGAAATAATTCAAATTAATCCAAA
GATAAATGACCAACGTAGATAGGGAGGAAAAATCAAAGTAGACACAACCCAGAGATGCAA
CCAAGATTCAACAATGCATAACTCATTAACTTAAACATAAGCCGTACGCCTTAACTATAAG
GAATGAAAACATGTACGATTTTCAGACAAAAGCCAGAGCTCACACTCAATTCTAGAAATGCA
GCCCTTTTTACCTGTCCTGCAATAGAAACCAAAGCACATCAAATGGGGTTAGTTTTATGG
TAGAAGAGTGCGAACGTGCCAATTTAGTGA CTTCACAAAAGACAAAACCTCTTCTTGATTA
AGTTAGCAAAAATAATGCTATTATAAGACATATATAAGGGGGAAAAAACTATGAAAAGTGA
ATAATCAAGATGGTCTCACTGGTTTCACACATTGATTAAGTGTTCAAATAAAAAGGGTAAAAA
TTCCGGAAAGGCATAGTCACCTTCGGATGAATTATGACTTCGAAACTTGAATGATAGGCAT
ACCCACGTTAACTGGAAAAATGCCATAGATTCCAACCCCTCAAATTCACAATTTGCTACCA
ATTAATAAATTACTTCAACCAACAAGACCAATAAAAACAGGGGAAACCATCAACTAACAGAGC

AAGAGCTACACCGACAAGCACACAAAAAGGCTCAATTTTGAAAGCTACATCCAAATCACAG
 ATTGAGAACTCTCAGTTTGTCTTTATCTTGGAGCATTATATCACACCCAATCAAACGAAAT
 TCAAGCAAACAAATCCAGAAGTTCGAGTAGCAAACAGATCACTCACGAATCAAACAGCAAA
 AACCCAGCAACGAAAGCGAAAAATTCATTCTAGCAGATCCACAGCAATCAGATTATTCAAAA
 AAAAGAATAAATACAAACTCAAAAAGCTCGAACAAAGCGAAATCAAGATGAATCCAGGTA
 AGGAAAGACGCAGGAGGAGAGAAAAACGTAGCCTCTTGGTTTCTGGCTGTGTCTTGGCTT
 ATGTAAGGAAAGCTCTAAACCTAGCCAGAAATGGCGGCTTCGCGTTGTGATTGATGACATA
 ACCTTAAAGGTTGGGCCGACATGGGCTGGCTGCTTATGGTCT

Los oligos (Morado) de Sec_4_PIP2 amplifican la secuencia siguiente (AMARILLO) de 426 nt.

TTGCAGGTGCATGTAGGGAGAAGTCTAGACTGGCATTCAATTCTAGAATGCAGCCCACGC
 CTTTTATTCAACCCCATACCAAGTTGAAGCCTGGCCTTAGTTTCTGTAGAAGAGGTTGCG
 AATGTCACAGTGCAACTCGTGCAAGGACAAGGGCCCTCTTACGTTGGTTACTACCTGCGA
 CACTATCCTGATCAGACCTTTTACTTATCATCATATCCATTCCCCTTTTATGAATCAGTC
 ACAACACTCAATGCAGCATGGATGATGTAACACAACGATAAAAAACAAAAAGAAAAACAGC
 TACCAAACCCAGTATGTGTCCACAGAACAAGGCCTGTGCCTGTGGATACCAAGTCTTGCC
 CTTGGCAAGTGACAGAATGCACAACATTTCAACAGATTTAATCCCAGTCTGTTCCGCG

FwSec_4_PIP2

5'- TTG CAG GTG CAT GTA GGG AG -3'

RvSec_4_PIP2

3'- CTG GCG AAC AGA CTG GGA TT -5'