



**UNIVERSIDAD NACIONAL AUTÓNOMA DE
MÉXICO**

FACULTAD DE QUÍMICA

**ESTUDIO DEL PROCESAMIENTO POST-TRANSCRIPCIONAL DE
MIR-21 EN LÍNEAS CELULARES DE CÁNCER DE MAMA.**

TESIS

**PARA OBTENER EL TÍTULO DE
QUÍMICA FARMACÉUTICO BIÓLOGA**

PRESENTA

XIADANI GUAJARDO BARRETO



CDMX

AÑO 2021



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

JURADO ASIGNADO:

PRESIDENTE: Profesor: Plasencia de la Parra Francisco Javier

VOCAL: Profesor: Carmona Salazar Laura

SECRETARIO: Profesor: Arriaga Canon Cristian Gabriel Oliverio (ASESOR)

1er. SUPLENTE: Profesor: García Velázquez Lizbeth Esmeralda

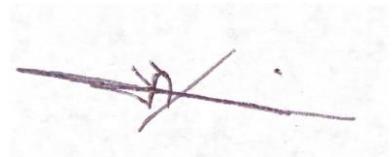
2° SUPLENTE: Profesor: Plata Ramos Maria del Consuelo

**SITIO DONDE SE DESARROLLÓ EL TEMA:
INSTITUTO NACIONAL DE CANCEROLOGÍA (INCAN)**

**ASESOR DEL TEMA:
DR. CRISTIAN GABRIEL OLIVERIO ARRIAGA CANON**



**SUSTENTANTE (S):
XIADANI GUAJARDO BARRETO**



Agradecimientos

Agradezco al Dr. Cristian Canon quién confió en mi para poder desarrollar este proyecto, y me motivó a siempre dar un extra, lo cual dio como fruto esta tesis de licenciatura que dará pauta a un gran proyecto de maestría. Gracias por ser un gran tutor y nunca dejarme sola en el proceso.

Asimismo, agradezco a todo el equipo de ncRNAs que tuvieron toda la paciencia en enseñarme y llevarme de la mano, siendo mis guías tanto en el laboratorio como en el aspecto personal, Pau, Lau, Rogelio, Sara, les agradezco toda la paciencia, apoyo, y sobre todo por su disposición a siempre ayudarme a ser mejor, sin ustedes no se habría podido lograr este gran resultado. También me gustaría dar un especial agradecimiento a Diana Bautista, gracias por empezar con una gran idea para proyecto, el cual pude desarrollar bajo tus primeras ideas.

A mis padres, Norma y Hugo, quienes me han apoyado en todas las decisiones que he tomado a lo largo de la carrera, y que siempre me motivaron a seguir con la investigación, sabiendo que era una de mis grandes pasiones, les debo la vida, el quien soy, y el quien llegaré a ser.

A mis amigos que, sin ellos, la facultad hubiera sido algo interminable, y la investigación algo inalcanzable, gracias por acompañarme en este camino, los quiero con mi alma: Berenice, Adriana, Octavio, Emely y Axel.

Gracias a todos por acompañarme en este camino.

Abreviaturas y sus siglas

°C: Grado Celsius.

CaMa: Cáncer de Mama.

cDNA: DNA complementario.

DMEM/F-12: *Modified Eagle Nutrient Mixture F-12*

DMEM/F-12: Modified Eagle Nutrient
Mixture F-12

DNA: Ácido desoxirribonucleico.

dsRNA: RNA de doble cadena

EDTA: Ácido Etilaminotetraacético.

ENCODE: Enciclopedia de elementos del DNA (del inglés *The Encyclopedia of DNA Elements*).

INCan: Instituto Nacional de Cancerología.

Kb: Kilobase.

lncRNAs: RNA largo no codificante.

MEBM: *Mammary Epithelial Cell Basal Medium*.

miRNA: MiRNA.

mRNA: RNA
mensajero.

nt: Nucleótidos.

NTC: (poner en inglés: No template control), Control Negativo.

pb: Pares de bases.

piRNAs: RNA asociados a PIWI.

PCR: Reacción en Cadena de la Polimerasa.

Pol II: RNA Polimerasa II.

RIN: Número de Integridad del RNA (del inglés *RNA Integrity number*). **RNA:** Ácido ribonucleico.

rRNA: RNA ribosomal.

RNA-Seq: Secuenciación masiva en paralelo
de RNA.

RPM: Revoluciones por
minuto.

RT: Enzima reverso-transcriptasa.

s: Segundo.

SFB: Suero Fetal Bovino.

siRNAs: RNA pequeño de interferencia.

snoRNAs: RNA pequeño nucleolar.

snRNAs: RNA pequeño nuclear.

ssRNA: RNA de cadena sencilla.

TPM: Transcritos por Millón.

tRNAs: RNA de transferencia.

Resumen

Los miRNAs son pequeños RNAs no codificantes (19-25 nucleótidos), que regulan la expresión de un gen mediante la degradación de su RNAmensajero, así como la prevención de su traducción. Se ha establecido que cada transcrito precursor de miRNA (pre-miRNA) se compone de dos hebras maduras (5p y 3p), y, aunque se ha demostrado que los miRNAs se encuentran involucrados en vías oncogénicas (oncomiRs y miRNAs supresores de tumor), muchas de sus funciones a nivel de regulación postranscripcional se desconocen aún. Las evidencias experimentales han establecido un modelo de selección de hebras canónico, en el cual únicamente una sola hebra de cada pre-miRNA es funcional, sugiriendo que la segunda hebra es generalmente degradada. Por lo anterior, a la fecha no se ha estudiado a profundidad el modelo no canónico de selección de hebras, en el cual ambas hebras tienen un papel funcional en la regulación postranscripcional de sus genes blanco. En particular, miR-21 fue uno de los primeros oncomiRs identificados; además, ha sido relacionado con varios procesos celulares oncogénicos, específicamente a la hebra miR-21-5p, considerada como la hebra guía. miR-21-5p se encuentra sobreexpresado en varios tipos de cáncer y ha sido asociada a procesos oncogénicos como proliferación, apoptosis y metástasis. Sin embargo, la caracterización funcional de este oncomiR, se ha enfocado únicamente en el estudio de la hebra 5p; descartando la posibilidad de que la hebra descrita originalmente como pasajera, miR-21-3p, sea considerada como una hebra funcional, como se sugiere en el modelo de selección de hebras no canónico. Es por ello que en este estudio el objetivo principal fue buscar alteraciones en los niveles de expresión de miR-21-3p en líneas celulares de cáncer de mama. A partir de este análisis, se encontró una sobreexpresión de la hebra 3p nunca antes reportada en este modelo, sugiriendo que, en cáncer de mama, durante la última etapa de su biogénesis, miR-21 sigue un modelo de selección de hebras no canónico, en el cual, la hebra antes considerada como hebra pasajera, miR-21-3p, también es una hebra funcional, la cual lleva a cabo un mecanismo de regulación post-transcripcional que conduce a una regulación alterada de sus blancos génicos, y, por lo tanto, una alteración de las vías de señalización oncogénicas. Por lo tanto, mediante una búsqueda utilizando servidores web de predicción de blancos génicos, se establecieron 2 genes blanco candidatos para miR-21-3p: *MAT2A* y *MAT2B*. A partir de un análisis *in silico*, se determinó que *MAT2A* y *MAT2B* contienen sitios de unión en sus extremos 3'UTR que son identificados por miR-21-3p. Adicionalmente,

MAT2A y *MAT2B* se encuentran dentro de los blancos génicos (*interactors*) con mayores puntajes de predicción dentro de las redes de interacción estudiadas, validando a *MAT2A* y *MAT2B* como genes blanco potenciales para miR-21-3p en cáncer de mama. Por lo tanto, los resultados encontrados en este estudio, son la primera evidencia en la cual se reporta a miR-21-3p como una posible hebra funcional en cáncer de mama, en donde se sugiere que miR-21-3p regula de manera postranscripcional a *MAT2A* y *MAT2B*, modificando así a vías de señalización asociadas a procesos oncogénicos.

Abstract

MiRNAs are small non-coding RNAs (19-25 nucleotides), capable of regulate gene expression through degradation of mRNAs, and/or inhibiting its translation. It has been established that each miRNA precursor (pre-miRNA) is composed of two mature strands (5p and 3p), and, although miRNAs have been shown to be involved in oncogenic pathways (oncomiRS and tumor suppressor miRNAs), many of its functions at a post-transcriptional level are still unknown. Experimental evidence has established a canonical strand selection model, in which only a single strand of each pre-miRNA is functional (guide strand), suggesting that the second strand is generally degraded (passenger strand). Therefore, to date, the non-canonical strand selection model has not been studied in depth, in which both strands have a functional role in the post-transcriptional regulation of their target genes. In particular, miR-21 was one of the first oncomiRS identified; Furthermore, it has been associated to several oncogenic cellular processes, specifically the miR-21-5p strand, considered as the guide strand. miR-21-5p is overexpressed in various types of cancer and has been associated with oncogenic processes such as proliferation, apoptosis, and metastasis. However, the functional characterization of this oncomiR has focused solely on the study of the 5p strand; ruling out the possibility that the strand originally described as passenger, miR-21-3p, is considered a functional strand, as suggested in the non-canonical strand selection model. Therefore, in this study the main objective was to search a possible alteration in the expression levels of miR-21-3p in breast cancer cell lines. From this analysis, an overexpression of the 3p strand never before reported in this model was found, suggesting that, in breast cancer, during the last stage of its biogenesis, miR-21 follows a non-canonical strand selection model, in which, the strand previously considered as a passenger strand, miR-21-3p, is also a functional strand, which carries out a post-

transcriptional regulation mechanism that leads to an altered regulation of its gene targets, and, therefore hence, an alteration of oncogenic signaling pathways. Moreover, a search using gene target prediction web servers, 2 candidate target genes for miR-21-3p were established: *MAT2A* and *MAT2B*. From an *in silico* analysis, it was determined that *MAT2A* and *MAT2B* contain binding sites at their 3'UTR ends that are identified by miR-21-3p. Additionally, *MAT2A* and *MAT2B* are among the gene targets (interactors) with the highest prediction scores within the interaction networks studied, validating *MAT2A* and *MAT2B* as potential target genes for miR-21-3p in breast cancer. Therefore, the results found in this study are the first evidence in which miR-21-3p is reported as a possible functional strand in breast cancer, where it is suggested that miR-21-3p regulates post-transcriptionally to *MAT2A* and *MAT2B*, thus modifying signaling pathways associated with oncogenic processes.

TABLA DE CONTENIDO

1	Introducción	8
1.1	Generalidades de los miRNA.....	8
1.2	Biogénesis: la transcripción y la maduración del miRNA.....	9
1.2.1	Biogénesis: Selección de hebras.....	16
1.3	miRNAs y su asociación con cáncer.....	21
1.4	miR-21.....	23
1.4.1	Vías metabólicas intracelulares reguladas por miR-21-5p en cáncer.	26
2	Justificación	30
3	Planteamiento del problema	31
4	Pregunta de investigación	32
5	Hipótesis	33
6	Objetivo general	34
6.1	Objetivos particulares.....	34
7	Metodología	35
7.1	Etapa A. Análisis bioinformático.....	35
7.1.1	Análisis de expresión diferencial de pri-miRNA-21 mediante RNA-Seq.....	35
7.1.2	Análisis de expresión diferencial de miR-21-5p y miR-21-3p por RNA-Seq..	36
7.1.3	Búsqueda de genes blanco de miR-21-3p.....	37
7.1.4	Expresión diferencial de mRNAs que son blancos génicos de miR-21-5p y miR-21-3p.....	38
7.1.5	Elaboración de gráficos de caja de expresión <i>in silico</i> e <i>in vitro</i>	39
7.1.6	Determinación de vías metabólicas intracelulares asociadas a cáncer de los genes <i>PTEN</i> , <i>PDCD4</i> , <i>MAT2A</i> y <i>MAT2B</i>	39
7.2	Etapa B. Validación experimental.....	40
7.2.1	Purificación de RNA total de líneas celulares de cáncer de mama.	40
7.2.1.1	Cultivo de las líneas celulares de cáncer de mama.....	40
7.2.1.2	Purificación de RNA total.....	41
7.2.1.3	Cuantificación de RNA.....	42
7.2.1.4	Determinación de la calidad e integridad del RNA.....	42
7.2.1.4.1	Gel Agarosa	42
7.2.1.4.2	Bioanalizador	43
7.2.2	Expresión de los transcritos pri-miRNA-21, miR-21-5p, miR-21-3p, y sus genes blanco mediante PCR en tiempo real en líneas celulares de cáncer de mama.	43
7.2.2.1.1	Síntesis de cDNA para ensayo de expresión de pri-miRNA-21	43
7.2.2.1.2	Determinación de la calidad de cDNA por PCR punto final	44
7.2.2.1.3	Cuantificación de expresión diferencial de pri-miRNA-21.	46

7.2.2.2	Determinación de la expresión relativa de miR-21-5p y miR-21-3p mediante PCR tiempo real.....	46
7.2.2.2.1	Síntesis de cDNA para ensayo de expresión de miRNAs	46
7.2.2.2.2	Cuantificación de expresión diferencial de miR-21-5p y miR-21-3p por RT-qPCR.	49
7.2.2.3	Determinación de la expresión relativa de los transcritos PTEN, PDCD4, MAT2A y MAT2B, mediante PCR en tiempo real	49
7.2.2.3.1	Diseño de oligonucleótidos	49
7.2.2.3.2	Determinación de la temperatura de desnaturalización (TM) de cada par de oligonucleótidos	50
7.2.2.3.3	Determinación de la eficiencia de amplificación de los oligonucleótidos mediante PCR tiempo real	51
7.2.2.3.4	Cuantificación de expresión diferencial de miR-21-5p y miR-21-3p por RT-qPCR.	52
7.2.2.4	Análisis de resultados de expresión obtenidos mediante qPCR, por medio de $\Delta\Delta CT$	52
8	Resultados	54
8.1	Análisis bioinformáticos.....	54
8.1.1	Determinación de la expresión diferencial de pri-miRNA-21, miR-21-5p y miR-21-3p a partir de datos de RNA-Seq de bases públicas de secuenciación.	54
8.1.2	Determinación de los mRNAs blancos para miR-21-5p y miR-21-3p.....	57
8.1.3	Determinación de la asociación de <i>PTEN</i>, <i>PDCD4</i>, <i>MAT2A</i> y <i>MAT2B</i> en vías oncogénicas.	65
8.2	Validación experimental	69
8.2.1	Calidad e integridad de RNA total de líneas celulares de cáncer de mama..	69
8.2.2	Determinación de la expresión de los RNA maduros de miR-21-5p, miR-21-3p, mediante PCR en tiempo real con sondas de hibridación en líneas celulares de CaMa.	71
8.2.3	Determinación de la expresión diferencial de los mRNA de <i>PTEN</i>, <i>PDCD4</i>, <i>MAT2A</i> y <i>MAT2B</i> mediante RT-PCR en tiempo real en líneas celulares de Cáncer de Mama.	74
9	Discusión	78
10	Conclusiones	85
11	Perspectivas	86
12	Bibliografía	88
13	Anexo	100
	Apéndice A: Habilidades aprendidas durante la realización de este proyecto de tesis	103
	Apéndice B: Artículo en el que se participó como co-autora: Saliva is a reliable and accessible source for the detection of SARS-CoV-2 publicado en 2021.	103

ÍNDICE DE FIGURAS Y TABLAS

Figura 1. Esquema general de la biogénesis de miRNAs en organismos eucariontes.....	10
Figura 2. Estructura tallo-asa del transcrito primario de miRNA y su reconocimiento por el complejo microprocesador.....	12
Figura 3. Figura representativa del efecto biológico, antagónico y sinérgico de ambas hebras de un mismo miRNA que se presentan en condiciones neoplásicas.	20
Figura 4. Localización y caracterización genómica de miR-21.....	25
Figura 5. Blancos génicos que se encuentran asociados a procesos oncogénicos.	26
Figura 6: Importancia del proyecto de investigación	29
Tabla 1. Claves de acceso en GEO Datasets para los archivos de RNA-Seq de líneas celulares de cáncer de mama.....	36
Figura 7. Expresión diferencial de pri-miRNA-21 en líneas celulares de cáncer de mama en distintos subtipos moleculares.	55
Figura 8. Expresión diferencial de las hebras maduras de miR-21-5p y miR-21-3p en distintos tipos de cáncer en la base de acceso público TACCO.	56
Figura 10. Sitios de unión predichos por análisis in silico entre miRNA y sus mRNAs blanco ..	61
Figura 11. Análisis in silico de expresión miRNA-mRNA blanco en una cohorte de tejidos de cáncer de mama.....	63
Figura 12. Análisis in silico de expresión de MAT2A, MAT2B, PTEN y PDCD4 en líneas celulares neoplásicas de diferentes subtipos moleculares.....	64
Figura 13. Redes moleculares de asociación e interacción de miR-21-5p y miR-21-3p.....	66
Tabla 2. Vías metabólicas oncogénicas asociadas a PTEN, PDCD4, MAT2A y MATB.	68
Figura 16. Expresión diferencial de la hebra guía y pasajera de miR-21 en líneas celulares de Cáncer de mama.....	73
Figura 17. Expresión diferencial de los mRNAs de PTEN, PDCD4, MAT2A y MAT2B en líneas celulares de cáncer de mama.....	76
Tabla A1. Información de los oligonucleótidos diseñados para los experimentos de PCR en tiempo real.	100
Figura A1. Integridad del RNA total de líneas celulares de mama evaluada mediante un gel agarosa.	101
Figura A2. Determinación de la calidad del cDNA sintetizado a partir del RNA extraído de líneas celulares de mama.....	101
Figura A3. Eficiencia de amplificación de los oligonucleótidos de los genes PTEN, MAT2A, MAT2B y PDCD4 blanco.	102

1 Introducción

1.1 Generalidades de los miRNA

Los miRNAs (miRNAs) son pequeños RNAs no codificantes endógenos que regulan post-transcripcionalmente la expresión de mRNAs, mediante vías de degradación y/o inhibición de su traducción. La importancia del estudio de los miRNAs se dio a conocer gracias al descubrimiento inicial de la función de dos genes, *lin-4* y *let-7*, los cuales no presentaban potencial codificante, pero se unían por complementariedad de RNA con las regiones 3'UTRs, y así regulaban los niveles de expresión de sus mRNAs blancos, reportando al miRNA *lin-4* como un regulador post-transcripcional del gen *LIN-14*, y al miRNA *let-7* como un regulador pos-transcripcional de los genes *lin-14*, *lin-28*, *lin-41*, *lin-42* and *daf-12*, controlando el tiempo de desarrollo embrionario en el nemátodo *Caenorhabditis elegans*(Lee et al., 1993; Reinhart et al., 2000). Además, en 2002, Reinhart y colaboradores describieron la existencia de miRNAs in *Arabidopsis thaliana*, dando origen a la teoría de que los miRNAs surgieron en una etapa temprana de la evolución eucariota(Reinhart, 2002). Después de estos descubrimientos, aumentó el número de miRNAs identificados en varias especies(Lai et al., 2003; Lim, 2003a, 2003b). En particular, en mamíferos, se ha estimado que aproximadamente el 30% de los genes que codifican para proteínas son regulados por miRNAs(Bartel, 2004), donde un solo miRNA puede tener como blancos génicos diferentes mRNAs, y un mRNA puede ser regulado de forma coordinada por múltiples miRNAs. De esta manera, el papel de molécula reguladora de los miRNAs se ha identificado en diferentes procesos celulares como lo son el desarrollo, la diferenciación, la apoptosis, y la tumorigénesis, entre otras funciones celulares. Finalmente, los miRNAs participan en la modulación génica de los distintos procesos celulares de homeostasis celular resaltando la importancia del estudio de este tipo de RNAs no codificantes pequeños.

Desde su descubrimiento, los miRNAs han sido cada vez más estudiados. Los miRNAs pueden ser localizados en el genoma tanto en regiones intergénicas como intragénicas(Bartel, 2004). Los miRNAs intergénicos derivan de un transcrito primario generado de regiones de DNA situadas entre dos genes, y los miRNAs intragénicos son codificados a partir de intrones de genes que codifican para proteínas. La biogénesis de los miRNAs comienza a partir de los transcritos primarios (pri-miRNA), estos transcritos

no codificantes pueden compartir el mismo promotor con un gen codificante o pueden ser sintetizados de manera independiente a partir de sus propios promotores (Bartel, 2009; B. Wang, 2013). Una vez sintetizado, el pri-miRNA sufre un procesamiento, donde el producto final, son dos hebras de miRNAs maduros de alrededor de 22 nucleótidos en longitud (dúplex de miRNA). Ambas hebras tienen la capacidad de regular la expresión génica a nivel post-transcripcional de diversos mRNAs (Bagga et al., 2005; Breving & Esquela-Kerscher, 2010; Zeng et al., 2003) mediante el reconocimiento de blancos por medio de emparejamiento complementario entre el miRNA y diversas regiones del mRNA, entre ellas se encuentran el promotor, los exones, las regiones no traducidas que se conocen como 3'UTR y 5'UTR (del inglés, *Untranslated Region*), llevando al mRNA a diferentes vías de regulación mediadas por el miRNA, como son, el arresto de la traducción del mRNA, la degradación del mRNA blanco y/o el almacenamiento de los mRNA en los cuerpos P (del inglés, *p-bodies*) (Chu & Rana, 2006; Zeng et al., 2003). Por lo tanto, en el último paso de la biogénesis se generan miRNAs maduros, que pertenecen al biotipo de RNAs pequeños no codificantes regulatorios, los cuales han demostrado estar asociados en la regulación de varias vías metabólicas intracelulares (Chu & Rana, 2006; Filipowicz et al., 2008).

1.2 Biogénesis: la transcripción y la maduración del miRNA

La biogénesis de los miRNAs inicia cuando la RNA polimerasa II o la RNA polimerasa III (únicamente miRNAs intergénicos) sintetizan al pri-miRNA^{15,16}. Este pri-miRNA presenta tanto la adición del nucleótido modificado 7-metil guanósina (7mGDP) localizado en la región 5' como la modificación bioquímica del extremo 3', la cual consiste en la adición de una larga serie de nucleótidos de adenina a la que se conoce como "cola de poli A", la cual incluye múltiples adenosín monofosfatos (Figura 1).

Posteriormente, el pri-miRNA es procesado (corte en el compartimento nuclear) por un complejo multipéptidico que se conoce como microprocesador. Este microprocesador se encuentra conformado por una molécula de la proteína Drosha, la cual contiene dos dominios de RNasa III encargados de realizar un corte catalítico en cada extremo del pri-miRNA, se genera un corte en el extremo 5' y otro corte en el extremo 3'. Además, el dímero de la proteína de unión a RNA de cadena doble, DGCR8 (*DiGeorge syndrome critical region 8*), que se incluyen en la subunidad del complejo

microprocesador, funcionan como cofactores de Drosha(Denli et al., 2004; Meijer et al., 2014).

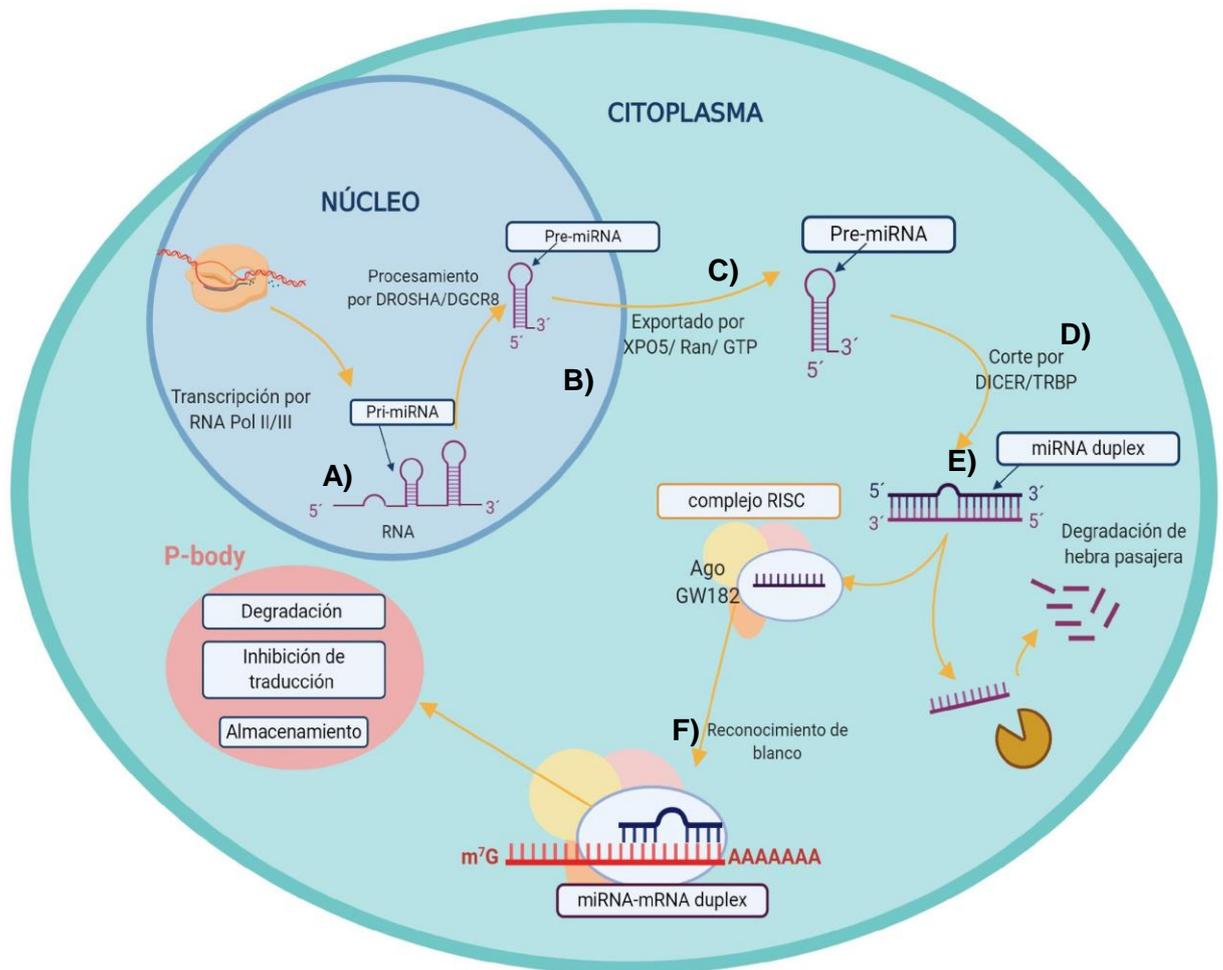


Figura 1. Esquema general de la biogénesis de miRNAs en organismos eucariontes. A) En el núcleo, es sintetizado mediante la RNA-Pol-II o RNA-Pol-III, el pri-miRNA, generando una estructura de tallo-asa. B) El pri-miRNA es reconocido y procesado (corte nuclear) por el complejo microprocesador compuesto por Drosha-DGCR8. Como resultado se obtienen estructuras tallo-asa de menor longitud conocidas como pre-miRNAs (alrededor de 60-80 nucleótidos en longitud). C) El pre-miRNA es transportado fuera del núcleo hacia el citoplasma mediante la exportina 5 (XPO5). D) En el citoplasma, el pre-miRNA se somete a un segundo corte, realizado por la RNasa III, DICER auxiliado por la proteína TRBP, generando un dúplex de miRNAs maduros de ~ 22 nt en longitud (miRNA / miRNA *). E) Mediante el proceso de selección de hebras, el miRNA maduro (hebra guía) se asocia con el complejo silenciador inducido por RNA formando el complejo miRISC, y la hebra pasajera es degradada. F) El complejo miRISC se encarga de presentar a la hebra guía con sus genes blanco, con el fin de regular su expresión génica mediante complementariedad de bases miRNA-mRNA, resultando en un mecanismo de degradación y/o represión de la traducción dentro de los cuerpos de procesamiento P.

(Modificado de Hajarnis y colaboradores) (Department of Medicine and Division of Nephrology, UT Southwestern Medical Center, Dallas, TX, USA et al., 2015).

Por otro lado, el mecanismo por el cual los sitios de corte catalítico del pri-miRNA son reconocidos por Drosha y DGCR8 ha sido estudiado por diversos grupos de trabajo, en donde se ha propuesto que la eficiencia del corte realizado por el complejo microprocesador, es favorecido por las estructuras tallo-asa con una longitud de $\sim 35 \pm 1$ nt, la cual, a nivel estructural, se divide en tallo superior (~ 22 nt), un sitio de corte de 2 nt y un tallo inferior (~ 11 pb) (Figura 2)(Kwon et al., 2019). Asimismo, la longitud del asa apical demostró ser compatible con el procesamiento cuando la estructura del asa apical se compone de ≥ 10 nt, debido a que se ha demostrado que permite un procesamiento eficiente. Por otro lado, y en el contexto de motivos encontrados dentro de la estructura tallo-asa que regulan y aumentan la velocidad en la tasa de procesamiento, se encuentra el motivo basal UG, localizado en la región de unión basal (posicionado a ~ 11 nt desde la unión basal) en donde, este motivo basal UG es reconocido por Drosha para realizar el corte catalítico nuclear; el motivo CNNC (donde la C se refiere a citosina y la N, significa cualquier nucleótido). En la parte apical del asa, las secuencias UGU/GUG/UGUG (motivo apical UGU), las cuales son reconocidas por DGCR8; y junto con el motivo basal UG potencian la tasa de procesamiento del pri-miRNA mediante Drosha y aseguran la correcta orientación de unión del pri-miRNA en el microprocesador. Finalmente, el motivo mGHG (*mismatched* GHG), localizado dentro de la parte inferior del tallo, el cual, fue demostrado por Kwong y colaboradores en el 2019, corresponde a un motivo de unión crítico y conservado en la determinación de los sitios de corte, el cual es reconocido por el dominio de unión a RNA de doble cadena (dsRBD) de Drosha(Kwon et al., 2019; Roden et al., 2017). Por lo anterior, la estructura tallo-asa del transcrito primario, dicta eficacia del procesamiento mediado por DROSHA, la cual se compone de dos tallos, superior e inferior y un tallo apical, dentro de los cuales, existen motivos de unión que permiten la correcta unión del microprocesador con el pri-miRNA, potenciando o disminuyendo la tasa de procesamiento del transcrito primario.

Por lo tanto, todos los motivos de unión encontrados en la estructura tallo-asa, así como la estructura misma del pri-miRNA aumentan la eficiencia y la precisión del procesamiento del pri-miRNA mediado por el complejo heterotrimérico de Drosha y

DGCR8, específicamente, en donde, Drosha reconoce los elementos basales y DGCR8 estabiliza a Drosha e interactúa con los elementos apicales con el fin de facilitar el procesamiento del pri-miRNA-21 para obtener al precursor de miRNA.

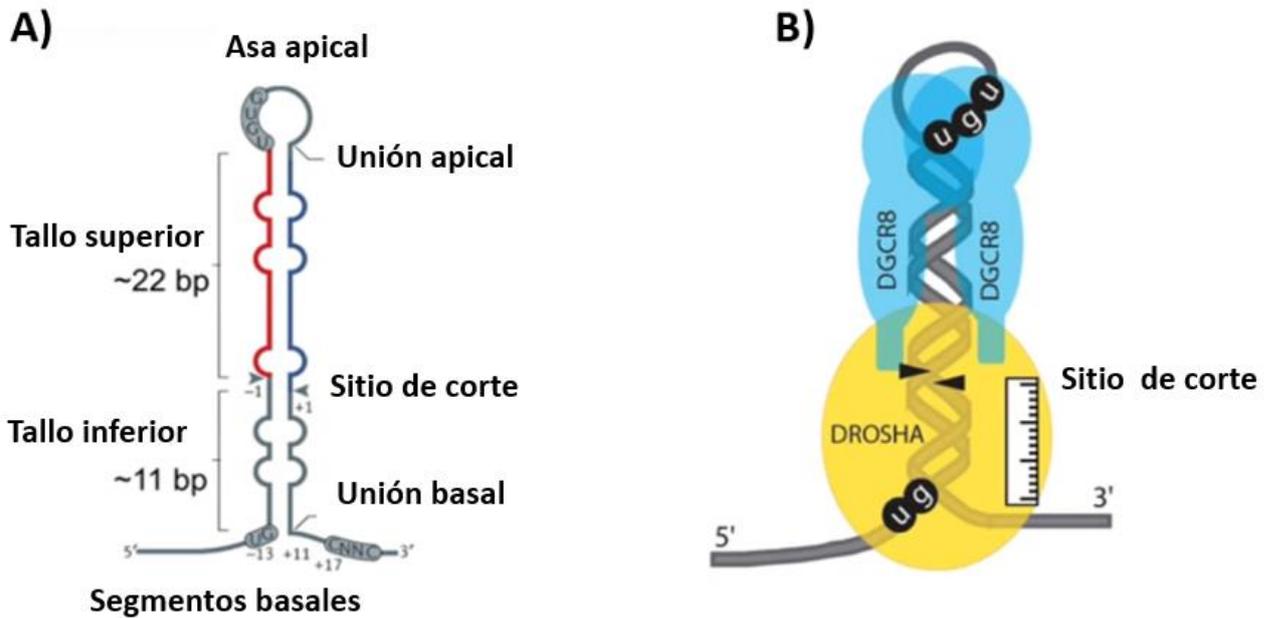


Figura 2. Estructura tallo-asa del transcrito primario de miRNA y su reconocimiento por el complejo microprocesador. A) Estructura general tallo-asa formada por el pri-miRNA, en la cual se muestran elementos en común aun teniendo diversas estructuras a nivel de secuencia. Se compone de un tallo largo de aproximadamente 35 nucleótidos, el cual se divide en el tallo superior (aprox. 22 nt), un sitio de corte y el tallo inferior (aprox. 11 nt). Asimismo, en los extremos del tallo, se encuentra un asa apical (parte superior) y un segmento basal de una sola hebra (parte inferior). Dentro de la secuencia del tallo-asa del pri-miRNA, se encuentran motivos de unión como el motivo en la unión basal "UG", en la unión apical "UGU/GUG" y el motivo CNNC en el extremo 3' del segmento basal. B) Los motivos de unión dentro de la secuencia del pri-miRNA son reconocidos por el complejo microprocesador DROSHA/DGCR8. El microprocesador es un complejo trimérico de 360 kDa, compuesto por dos copias de DGCR8 y una copia de DROSHA. En conjunto, reconocen la estructura general del pri-miRNA, DROSHA reconoce los elementos basales (motivo "UGU") y se posiciona en el tallo inferior, midiendo 11 pares de bases desde la secuencia basal. Por otro lado, las dos moléculas de DGCR8 se posicionan en la parte superior del tallo, reconociendo los elementos apicales (motivo "UGU/GUG" y el asa apical). (Imágenes obtenidas de V. Narry Kim y traducidas al español)

Una vez procesado el pri-miRNA, se obtiene como resultado una estructura de tallo-asa, que lleva por nombre pre-miRNA (precursor de miRNA). El pre-miRNA, es una estructura de tallo-asa de ~70 nt, es transportado del núcleo al citoplasma por el complejo multi-peptídico Ran-GTP²⁵, y por la Exportina-5 (XPO5, receptor de exportación)²⁶, para posteriormente ser procesado por la ribonucleasa miembro de la familia de las RNAasas tipo III conocida como DICER, la cual se une al pre-miRNA mediante su dominio PAZ (dominio de unión a RNA de doble cadena), el cual interacciona con el extremo 3' del pre-miRNA, permitiendo a los dominios catalíticos de DICER realizar un corte asimétrico del extremo 5' de la hebra 3p y del extremo 3' de la hebra 5p, de la estructura tallo-asa del pre-miRNA (Meijer et al., 2014). La fidelidad del corte, la selección de la hebra guía y la unión entre la ribonucleasa DICER, así como las proteínas Argonauta (AGO) son mediadas a través de la proteína TRBP, la cual es un cofactor para el procesamiento eficiente de DICER. Como producto final de este corte se obtiene una hebra dúplex de miRNA maduro de aproximadamente 22 nt de largo, las cuales son conocidas como la hebra 5p y la hebra 3p, que corresponden a las cadenas sentido y antisentido del miRNA maduro (Chendrimada et al., 2005; Koscianska et al., 2011; Wilson et al., 2015).

Por último, el dúplex de miRNAs maduros sufre un proceso conocido como, selección de hebras, dicho proceso tiene como finalidad determinar cuál de las dos hebras en el dúplex de miRNA será considerada la hebra funcional (hebra guía o miRNA), y cual será considerada la hebra pasajera o miRNA*, la cual es generalmente degradada. Esta selección se basa en parte, en un mecanismo en donde participa la fuerza de enlaces de hidrógeno dependiente de proteínas AGO (Schwarz et al., 2003), el cual se somete a cambios conformacionales con el fin de permitir la unión del dúplex de RNA para formar el complejo miRISC, en donde la proteína AGO2 interactúa con DICER para cargar al dúplex de miRNAs. Específicamente, el extremo 5' de la cadena guía se une al dominio Ago2 MID, asistido por el dominio PIWI (por sus siglas en inglés, *P-element Induced Wimpy testis*), mientras que el extremo 3' interactúa con el dominio PAZ (por sus siglas en inglés, PIWI / Argonaute / Zwillig) (Meijer et al., 2014; Yoda et al., 2010). En particular, la formación del complejo miRISC se compone de dos etapas, en primer lugar, el proceso de carga del dúplex de miRNAs a AGO, en donde intervienen dos proteínas chaperonas, HSP70 y HSP90, que inducen cambios conformacionales de AGO permitiendo una carga asimétrica del dúplex de miRNAs en AGO, de manera que el extremo menos estable del dúplex se una al dominio MID. La segunda etapa consta de la separación de una hebra

del dúplex, en esta etapa, a diferencia de la carga del dúplex al complejo miRISC la cual es un proceso activo que requiere ATP, el proceso de separación de hebras es un proceso pasivo independiente de ATP, en donde se propone que el dominio N-terminal de AGO sufre cambios conformacionales al unirse al dúplex(Koscianska et al., 2011; Medley et al., 2021; Schwarz et al., 2003; Yoda et al., 2010). Finalmente, estos dos procesos llevan a la activación del complejo multipéptido miRISC el cual reconocerá su mRNA blanco a nivel post-transcripcional y dará como resultado un arresto de la traducción de la proteína regulada por un miRNA.

Las estrategias experimentales han establecido que una vez activado el complejo miRISC, este permite la presentación directa de la secuencia semilla del miRNA, la cual corresponde a una secuencia de reconocimiento específica de 2-8 nt en longitud, localizada hacia el extremo 3' del miRNA guía, que establece dónde se va a ocurrir la interacción por complementariedad de bases entre la secuencia del miRNA y la secuencia complementaria de su mRNA blanco(Klein et al., 2017), llevando al mRNA a un silenciamiento génico postranscripcional(Granados-López et al., 2017; Meijer et al., 2014).

En la actualidad se ha demostrado que este silenciamiento ocurre dentro de los cuerpos de procesamiento conocidos como cuerpos P (del inglés, *Processing-bodies*) dentro del citoplasma, donde los mRNAs pueden ser acumulados y almacenados, con dos finalidades: servir como reservorios de polisomas y dada una señal intracelular volver a entrar en la vía de traducción de mRNAs, o entrar en una vía de degradación de mRNAs(Rana, 2007). Los cuerpos P, son gránulos de ribonucleoproteínas citoplasmáticos compuestos de mRNAs arrestados traduccionalmente y proteínas que funcionan en el metabolismo de mRNAs, las cuales son requeridas para inhibir la traducción, como las proteínas AGO, RCK/P54, LSM1 y GW182(Chu & Rana, 2006, p. 7, 2006; J. Liu et al., 2005), y para la degradación de mRNA, como es la proteína CCR4 con actividad de des-adenilasa, el complejo multipéptido DCP1–DCP2 que tiene la función de eliminar la estructura CAP que se encuentra en el extremo 5' (7mGDP) de todos los mRNAs y la exonucleasa XRN1(Eulalio et al., 2007; Shyu et al., 2008). En conjunto, las evidencias experimentales sugieren que los cuerpos P son subcompartimentos citoplásmicos que sirven como reservorios de mRNAs que necesitan

estar en pausa, impidiendo su traducción, hasta que haya una señal intracelular que dicte su liberación para que puedan iniciar la traducción y dar origen a una nueva proteína, o en su caso, que las proteínas almacenadas sean degradadas.

De manera específica, el mecanismo de silenciamiento génico mediado por la degradación de los mRNA es desencadenada por un acortamiento gradual de la cola poly (A) del mRNA, como resultado, los mRNA desadenilados son sustratos para la degradación catalizada por exosomas o la eliminación de la modificación en el extremo 5' de la 7-metilguanosina trifosfato por el complejo de *decapping* mediado por Dcp1/2. Como producto de degradación, se genera un intermediario de RNA con un monofosfato en su extremo 5' que es reconocido e hidrolizado rápidamente por la exonucleasa citoplasmática 5'-3' Xrn1 (Garneau et al., 2007; Jinek et al., 2011). Por lo tanto, el mecanismo de silenciamiento génico post-transcripcional mediado por miRNAs mediante la vía de degradación del mRNA se lleva a cabo mediante una desadenilación o un *deccapping*, con la finalidad de facilitar el reconocimiento de los mRNAs blanco por exonucleasas, ya sean en el citoplasma, o dentro de los cuerpos P.

Por otro lado, dentro del mecanismo de silenciamiento génico post-transcripcional mediado por el arresto de la traducción, se han propuesto dos mecanismos: Los miRNAs inhiben el inicio de la traducción. Este proceso se lleva a cabo mediante un impedimento estérico que se da como consecuencia de la unión del complejo miRISC con el mRNA, evitando que las subunidades ribosómicas se unan al mRNA y puedan iniciar la traducción del mismo. Por otro lado, el segundo mecanismo se da en un paso "posterior a la iniciación" en la traducción, que también provoca la degradación co-traducciona del péptido naciente (Shyu et al., 2008).

Además, otro mecanismo que ocasionalmente ocurre, pero es poco común, se da cuando la región semilla del miRNA y el extremo 3'UTR del mRNA son perfectamente complementarios, lo cual desencadena una vía de degradación mediada por el dominio PIWI de AGO2 (única proteína de esta familia que puede inducir una ruptura endonucleótica con actividad de tipo RNasa-H, el cual realiza un corte catalítico en el centro del dúplex de mRNA-miRNA (Bartel, 2009; Filipowicz et al., 2008). Por lo tanto, las diferentes vías de silenciamiento post-transcripcional mediadas por la inhibición de la traducción son consecuencia de los diversos mecanismos de regulación mediados por los miRNAs, así como, estas vías son evidencia de demostrar la multifuncionalidad de

un mismo miRNA en diferentes vías, interaccionando con diversas proteínas y complejos mapeptídicos para llevar a cabo el silenciamiento génico post-transcripcional.

En resumen, los diferentes tipos de mecanismos de silenciamiento post-transcripcional utilizados por los miRNAs: el arresto de la traducción, la degradación y/o almacenamiento en cuerpos P, son mecanismos de regulación indispensables en varios procesos biológicos finamente orquestados, indicando, que al existir una alteración, ya sea desde etapas tempranas de la biogénesis (factores epigenéticos, o alteración de proteínas importantes en la biogénesis) o una alteración post-transcripcional de las hebras maduras, esta desregulación, impacta directamente en los niveles de expresión de los mRNA que son regulados por miRNAs, y por lo tanto, tienen efectos sobre las vías metabólicas intracelulares en las que estos genes se encuentran involucrados, derivando en el desarrollo de enfermedades y/o patologías, reiterando la importancia de la función de los miRNAs en la regulación de sus genes blanco.

1.2.1 Biogénesis: Selección de hebras

La selección de hebras es el proceso por el cual se determina cuál de las dos hebras maduras será incorporada en el complejo miRISC (hebra guía, miRNA) y cuál será degradada (hebra pasajera, miRNA*). La expresión miRNA/miRNA* varía dependiendo de la especie, el tipo de tejido, el estado de desarrollo, el estado de progresión de la enfermedad (cáncer), entre otras, implicando que el repertorio de miRNAs sea regulado a nivel transcripcional y post-transcripcional a través de factores de transcripción, a nivel procesamiento del dúplex de miRNAs, a nivel de la formación y activación del complejo miRISC y a nivel de la estabilidad de los miRNAs(Krol et al., 2010; Li et al., 2012), sugiriendo que la regulación de blancos génicos mediada por miRNAs, no se basa únicamente por la teoría de la fuerza de unión de enlaces de hidrógeno, si no también se encuentra influenciada por diferentes factores moleculares. En particular, a nivel post-transcripcional, la selección de hebras está mediada por las proteínas AGO, y se basa en gran medida en la estabilidad de los extremos del dúplex de RNA, es decir, la hebra del dúplex con el extremo 5' con una interacción más débil (miRNA-miRNA) es incorporada a RISC (hebra guía), mientras que la hebra con la interacción más fuerte en su extremo 5' es generalmente degradada (hebra pasajera)(Schwarz et al., 2003).

Entonces, se propone que esta selección preferencial mediada por la termodinámica del dúplex de RNA se debe principalmente a que una interacción más débil del extremo 5' inducirá la separación de las hebras en el dúplex de RNA, facilitando a las proteínas AGO el extremo del dúplex RNA expuesto e iniciar el proceso de selección de hebras (Kwak & Tomari, 2012).

En este mecanismo se han propuesto varios mecanismos, como el mecanismo de separación de hebras pasivo, independiente de ATP desencadenado por los cambios conformacionales que sufre el proceso de carga del dúplex de miRNAs en AGO. Por otro lado, se sugiere que el proceso de separación de hebras radica en la función de una enzima con actividad de helicasa (aún por identificar) (Schwarz et al., 2003).

Las evidencias experimentales han descrito y sugerido una predisposición en la estructura de cada hebra, es decir, la hebra guía tiene tendencia a tener uracilos en su extremo 5', así como un exceso de purinas a lo largo de su estructura; y la hebra pasajera, presenta citosinas en su extremo 5' y un mayor número de pirimidinas a lo largo de su estructura (H. Y. Hu et al., 2009). Sin embargo, aquí se sugiere que otro factor está involucrado en la preferencia de carga de la selección de hebra al complejo multipéptido RISC, donde la orientación en la cual la hebra guía es cargada al complejo, mediada por los dominios del complejo miRISC; en donde las purinas de la futura hebra guía se sabe interaccionan de forma hidrofóbica con los residuos aromáticos del dominio AGO: PAZ (H. Y. Hu et al., 2009), orientando a la hebra del dúplex de miRNAs complementaria a una posición en la cual esta sea separada con mayor facilidad y sea degradada. En conjunto, la estructura dictada por el contenido de nucleótidos de los miRNAs, así como las diversas proteínas y complejos que actúan desde nivel transcripcional hasta el nivel post-transcripcional, afectan al conjunto de miRNAs, llevando a estas RNA pequeños a una regulación específica y controlada.

Por otro lado, los antecedentes experimentales han establecido durante mucho tiempo que solamente una hebra del miRNA es funcional y la otra hebra es degradada, sin embargo, con el avance de las técnicas de biología molecular se ha establecido que lo anterior no es del todo cierto. Los estudios más recientes han demostrado que, en el

caso de algunos miRNAs, es posible identificar ambas hebras del dúplex cargadas en RISC, sugiriendo que tanto la hebra 5p cómo la hebra 3p son incorporadas y cargadas en el complejo multipéptido RISC, sugiriendo que las hebras tienen una función biológica cuando ambas son seleccionadas y cargadas en el complejo proteínico RISC(Medley et al., 2021).

Es importante mencionar que en los modelos biológicos donde ambas hebras se encuentran presentes, se ha establecido que ambas hebras pueden actuar regulando distintos blancos génicos o el mismo mRNA, regulando diferentes funciones biológicas, de manera sinérgica o antagónica, a pesar de tener secuencias casi idénticas entre ellas. Dentro de esta evidencia experimental limitada se encuentran estudios específicos en cáncer. Por ejemplo, tanto miR-31-5p y 3p en el carcinoma oral de células escamosas (Chang et al., 2013), así como miR-574-5p y 3p en cáncer gástrico(Z. Zhang et al., 2019), tienen papeles antagónicos, contrarrestando la función de la otra hebra. En específico, ambas hebras de miR-31 se encuentran sobreexpresadas en carcinoma oral de células escamosas, en donde, la hebra miR-31-5p se había reportado anteriormente como un miRNA oncogénico, a diferencia de la hebra 3p, la cual no había sido descrita anteriormente. El grupo de Chang y colaboradores, describieron por primera vez la expresión de miR-31-3p en cáncer de boca, y su función; ellos reportan que miR-31-3p antes considerada como pasajera y no funcional, se encontraba sobreexpresada al igual que miR-31-5p, aunque en menor proporción, y reportaron la función supresora de tumor de miR-31-3p, mediante experimentos de ganancia y pérdida de función, reportando una disminución en la proliferación y migración celular de dos líneas celulares neoplásicas de cáncer de boca (SAS y FADU), cuando se tenían condiciones de sobreexpresión de miR-21-3p, teniendo como consecuencia la regulación negativa del gen *RhoA*, un gen con funciones oncogénicas reportadas(Chang et al., 2013, p. 31).

Asimismo, miR-574-5p y miR-574-3p en cáncer gástrico presentan funciones antagónicas, pero a diferencia de miR-31, miR-574-5p y miR-574-3p presentan un patrón de expresión diferencial entre sus dos hebras, la hebra 5p se encuentra sobreexpresada en tejido tumoral mientras la hebra 3p se encuentra subexpresada en tejido tumoral; al realizar experimentos de sobreexpresión, Zhang y colaboradores demostraron el efecto supresor de tumor al sobreexpresar miR-574-3p, la cual potenciaba un fenotipo normal,

a diferencia del efecto oncogénico observado de la miR-574-5p, la cual potenciaba un fenotipo oncogénico(Z. Zhang et al., 2019). Las evidencias experimentales anteriores dieron pauta a la posibilidad de tener dos patrones de expresión entre las cuales se sugería que ambas hebras podrían estar tanto sobreexpresadas así como subexpresadas. En el segundo modelo biológico se muestra una expresión diferencial de hebras en condiciones neoplásicas, es decir, una hebra se encuentra sobreexpresada y otra hebra se encuentra subexpresada, independientemente de ser la hebra guía o hebra pasajera, teniendo como consecuencia, en este caso, que ambas hebras de un mismo miRNA tienen funciones diferentes en una misma patología, una hebra presenta funciones oncogénicas y la otra hebra presenta funciones como gen supresor de tumor.

Por otro lado, se encuentran los miRNAs que presentan un comportamiento sinérgico entre sus hebras, como lo demostraron Uchino y colaboradores, este grupo de investigación reportó que ambas hebras de miR-582 en cáncer de vejiga, se encuentran subexpresadas, y al realizar experimentos de sobreexpresión tanto de la hebra 5p como de la hebra 3p, la proliferación celular y el crecimiento tumoral disminuía en ambos casos, adjudicándole tanto a la hebra 5p como a la hebra 3p una función sinérgica como genes supresores de tumor(Uchino et al., 2013). En cambio, el miRNA, miR-17 se encuentra sobreexpresado en carcinoma hepatocelular, y al ser sobreexpresados de manera coordinada ambas hebras de miR-17 el fenotipo celular encontrado se asoció a un aumento en el tamaño e invasión tumoral(Shan et al., 2013). Por lo tanto, en un modelo de función sinérgica, las dos hebras de un mismo miRNA, pueden estar sobreexpresadas y/o subexpresadas, y su co-expresión aumenta el efecto biológico, ya sea potenciando o disminuyendo un fenotipo oncogénico.

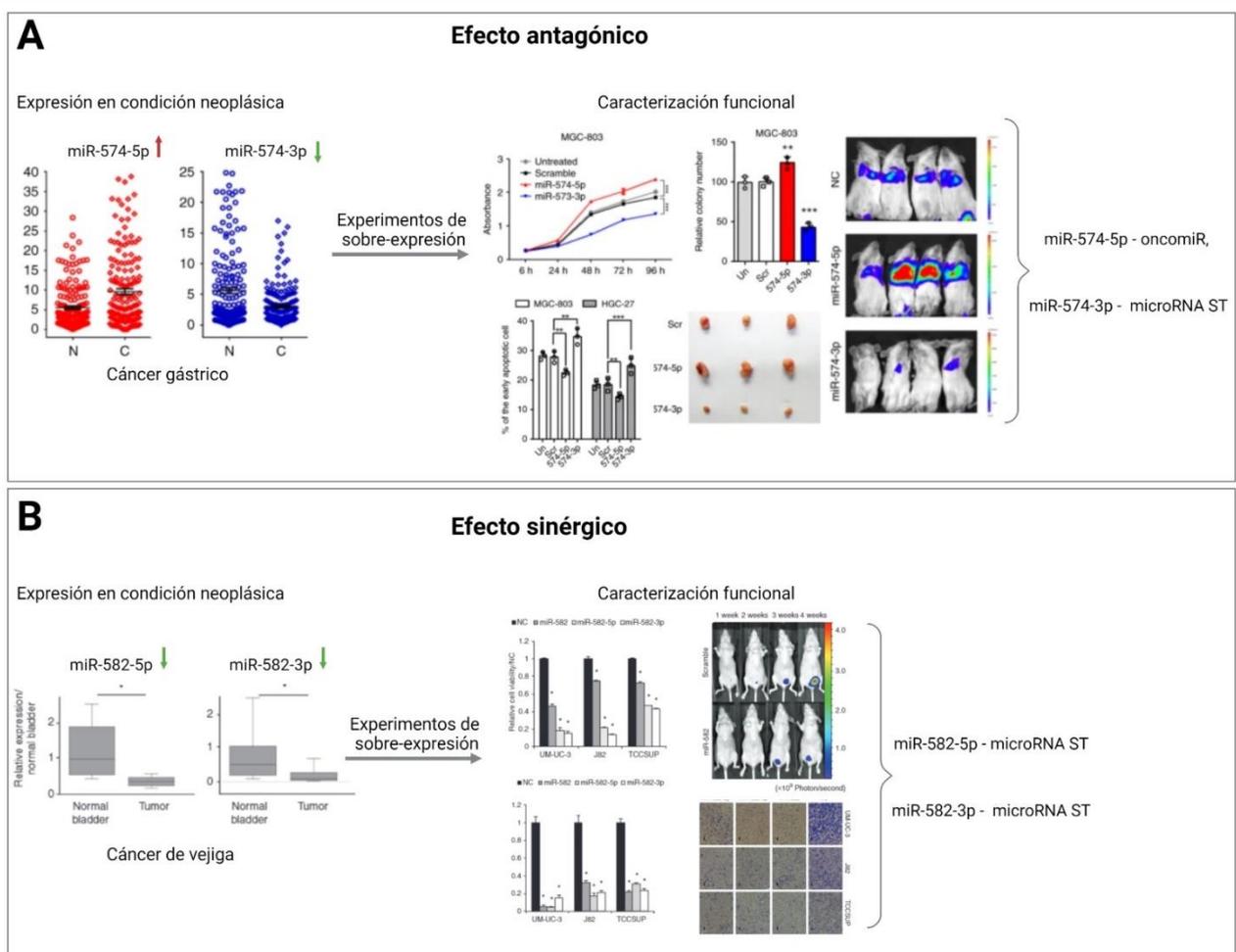


Figura 3. Figura representativa del efecto biológico, antagónico y sinérgico de ambas hebras de un mismo miRNA que se presentan en condiciones neoplásicas. A) En condiciones neoplásicas, Zhang y colaboradores (2019), observaron un patrón de expresión diferencial entre las dos hebras del miRNA-574, reportando una sobreexpresión de la hebra 5p, caracterizada como un oncomiR debido a su potencial para favorecer la proliferación e invasión celular, metástasis, volumen y peso tumoral, así como disminuir la apoptosis celular, favoreciendo un fenotipo oncogénico en cáncer gástrico. A diferencia de la hebra miR-574-3p, que se encontró subexpresada en cáncer gástrico, y al ser sobreexpresada, se caracterizaba como un miRNA supresor de tumor (ST) debido a que contrarrestó los efectos oncogénicos de miR-574-5p, disminuyendo la proliferación, invasión, metástasis, peso y volumen tumoral, así como aumentando la apoptosis, favoreciendo un fenotipo normal. Ejemplificando la forma en que dos hebras de un mismo miRNA son capaces de tener efectos biológicos antagónicos. B) En cáncer de vejiga, Uchino y colaboradores (2013) reportaron un ejemplo de función sinérgica de dos hebras en una misma patología, en donde, tanto la expresión de ambas hebras y su efecto biológico se encontraron de manera coordinada en miR-582, en donde ambas hebras, 5p y 3p se encuentran subexpresados en cáncer de vejiga, y al ser sobreexpresadas, se caracterizan como miRNAs ST, debido a que disminuyen la proliferación e invasión celular, así como en la metástasis, disminuyendo el fenotipo oncogénico de manera sinérgica.

En resumen, las evidencias experimentales dan un nuevo enfoque hacia la importancia de la manera en que los miRNAs presentan sus mecanismos de regulación génica, presentando funciones bimodales independientemente del patrón de expresión de las hebras, indicando que existen más mecanismos de regulación aún por ser caracterizados. La evidencia experimental de que existe una interacción entre las hebras 5p y 3p de cada miRNA que juegan un papel importante en el desarrollo de enfermedades y patologías, son mecanismos poco estudiados y caracterizados los cuales necesitan ser mejor comprendidos para entender cómo es que los miRNAs se encuentran asociados al desarrollo de un fenotipo oncogénico y tratar de entender nuevos mecanismos de regulación y su asociación de sus funciones en el cáncer.

1.3 miRNAs y su asociación con cáncer

A partir de la publicación del estudio científico de Calin y colaboradores en el año 2002, la cual dio a conocer la primera evidencia directa de un miRNA involucrado en el desarrollo de cáncer (Calin et al., 2002), el estudio del papel de los miRNAs en cáncer y sus efectos biológicos, han cobrado gran interés. En donde cada vez se reportan nuevas funciones de miRNAs asociados a vías metabólicas oncogénicas, como proliferación, metástasis (Lou et al., 2017), apoptosis (Jovanovic & Hengartner, 2006), mantenimiento y progresión del cáncer (B. Zhang et al., 2007), entre otros.

El mecanismo propuesto por el cual los miRNAs favorecen o disminuyen un fenotipo neoplásico es con base en el potencial de regulación que poseen los miRNAs, al provocar un silenciamiento génico de diversos mRNAs mediante complementariedad de bases entre la región semilla del miRNA y su región complementaria en el mRNA, llevando a este mRNA a un silenciamiento génico post-transcripcional, ya sea por inhibir su traducción, así como la acumulación y/o degradación del mRNA blanco. Específicamente, la función de miRNA supresor de tumor o miRNA oncogénico depende totalmente en los mRNAs que puedan ser potenciales blancos de un miRNA, y a qué vías metabólicas se encuentren asociados a estos mRNAs. Dependiendo de sus blancos génicos, los miRNAs se clasifican en oncogénicos (oncomiRs), cuando regulan

negativamente a genes supresores de tumor y/o a genes asociados a vías metabólicas oncogénicas como son, la diferenciación celular, migración e invasión, apoptosis, entre otras; en donde, la regulación negativa del mRNA blanco tiene como consecuencia, el desarrollo de un fenotipo oncogénico(B. Zhang et al., 2007). De igual manera, los oncomiRs, al ser moléculas reguladoras que promueven el desarrollo tumoral, se encuentran sobreexpresados en cáncer, como la familia de oncomiRs miR-17-92, miR-21, miR-27a(Bautista-Sánchez et al., 2020; Olive et al., 2013; Sharma et al., 2016). Por otro lado, los miRNAs que regulan negativamente a oncogenes y/o genes asociados a vías oncogénicas, que disminuyen el fenotipo tumoral, inhibiendo el desarrollo tumoral, se les conoce como miRNAs supresores de tumor, y a diferencia de los oncomiRs, los miRNAs supresores de tumor se encuentran subexpresados en condiciones tumorales cuando se comparan contra las condiciones normales. miR-145, la familia de let-7y miR-34a son ejemplos de miRNAs supresores de tumor, que disminuyen el desarrollo tumoral al regular negativamente sus blancos génicos(Kwon et al., 2019; Mizuno et al., 2018; B. Zhang et al., 2007).

De manera conceptual, un miRNA puede fungir como oncomiR así como un miRNA supresor de tumor, presentando funciones duales en escenarios específicos(Svoronos et al., 2016). El potencial que tiene un mismo miRNA de potenciar o disminuir un fenotipo oncogénico, se debe a la capacidad de regulación por medio de unión complementaria a un amplio repertorio de mRNAs blanco, en donde un solo miRNA es capaz de regular una gran cantidad de genes, puesto que la región semilla de un miRNA, al tener una estructura de aproximadamente 7 nt tiene la facultad de identificar y unirse por complementariedad de bases a diversos sitios de unión dentro de los transcritos, así como a regiones promotoras y extremos 5' y 3'UTR. Adicionalmente de la capacidad de unión de la región semilla a diversos transcritos, recientemente se ha demostrado que las hebras maduras 5p y 3p de un mismo pre-miRNA son capaces de compartir mRNAs blanco, y de igual manera cada hebra madura es capaz de regular a mRNAs específicos, ampliando el universo de mRNAs blanco y vías metabólicas que son regulados por un mismo miRNA.

En resumen, los miRNAs en cáncer son clasificados como oncomiRs y miRNAs supresores de tumor, dependiendo los blancos génicos que regulen y las vías metabólicas que se encuentren asociadas a los genes blanco. Un mismo pre-miRNA

tiene la capacidad de generar dos hebras maduras que, en una misma patología regulen genes y vías metabólicas que potencien o disminuyan un fenotipo oncogénico, ya sea de manera sinérgica o antagónica. Entonces, es importante resaltar el complejo mecanismo de acción y la interacción que existe entre ambas hebras de un mismo miRNA, en donde es necesario ampliar el campo de estudio de los miRNAs hacía ambas hebras, con el fin de poder elucidar funciones de las hebras desconocidas anteriormente que han sido consideradas en su momento como las “pasajeras”.

1.4 miR-21

El RNA pequeño miR-21, fue uno de los primeros oncomiRs descritos, teniendo como blanco génico a numerosos genes supresores de tumores asociados con la proliferación, la apoptosis y la invasión, en donde su regulación y papel en la carcinogénesis ha sido ampliamente investigado(Bautista-Sánchez et al., 2020).

Conforme a la base de datos de acceso público miRBase(Kozomara et al., 2019), el pri-miRNA-21 generado por la RNA pol II es de 3,433 nt de largo, una vez procesado por DROSHA, genera un pre-miR-21 de 72 nt de largo, el cual al ser procesado por DICER genera dos hebras como transcrito maduro: miR-21-5p (21 nt) y miR-21-3p (20 nt). El análisis en la estabilidad del dúplex pre-miR-21, reveló que el extremo 5´ de miR-21*, el cual se conoce como la hebra pasajera es ligeramente menos estable que miR-21, y, por lo tanto, miR-21* es generalmente degradada(Kumarswamy et al., 2011).

En particular, el gen no codificante de miR-21, se encuentra ubicado en el cromosoma 17 (17q23.2)(Calin et al., 2004), en el exón 11, el cual está codificado dentro del gen codificante *TMEM49* (proteína transmembranal 49, por sus siglas en inglés), también conocido como el gen *VMP1*. Asimismo, miR-21 se transcribe independientemente del gen *TMEM49*, mediante sus propias regiones promotoras(Krichevsky & Gabriely, 2009). El grupo de Fujita y colaboradores, propusieron una región promotora para miR-21, en la cual, se identificaron secuencias consenso, las cuales contenían cajas TATA, GC, y CCAAT, así como sitios de unión a factores transcripcionales positivos y represores, como lo son, la proteína activadora 1 (AP-1), E-26 transformación específica/PU.1

(Ets/PU.1), CCAAT/proteína de unión a *enhancer* α (C/EBP α), factor nuclear I (NFI), factor de respuesta a suero (SRF), p53, y el transductor de señal y activador de la transcripción 3 (STAT3) (Figura 4). En el cual demostraron que existe una regulación de retroalimentación doble negativa, en donde el mRNA de NFIB es un blanco directo de miR-21 y la proteína NFIB es un regulador negativo de miR-21, al unirse al promotor de miR-21 en células HL-60, operando como un mecanismo de represión para mantener la expresión de miR-21 (Fujita et al., 2008).

En resumen, las evidencias experimentales sugieren que miR-21, es un miRNA endógeno intragénico, que mediante la biogénesis se obtienen dos hebras maduras, miR-21-5p y miR-21-3p, siendo la hebra miR-21-5p menos estable, influyendo en el proceso de selección de hebras, por lo cual, la hebra miR-21-5p es cargada en la proteína AGO con mayor preferencia que la hebra miR-21-3p que es considerada la hebra pasajera, y se cree que generalmente es degradada mediante exonucleasas después del proceso de selección de hebras. Por otro lado, miR-21 se encuentra altamente regulado a lo largo de su biogénesis, tanto transcripcionalmente, por factores de transcripción como AP-1 y el factor nuclear NF-1 (Bautista-Sánchez et al., 2020), así como post-transcripcionalmente mediante proteínas, mRNAs, circRNAs y lncRNAs. Los factores anteriormente mencionados, demuestran la importancia de una regulación conservada de miR-21 en condiciones normales, mismos factores que pueden influenciar en una alteración desde niveles transcripcionales hasta post-transcripcionales, teniendo como consecuencia una desregulación de genes y vías de señalización oncogénicas en condiciones neoplásicas.

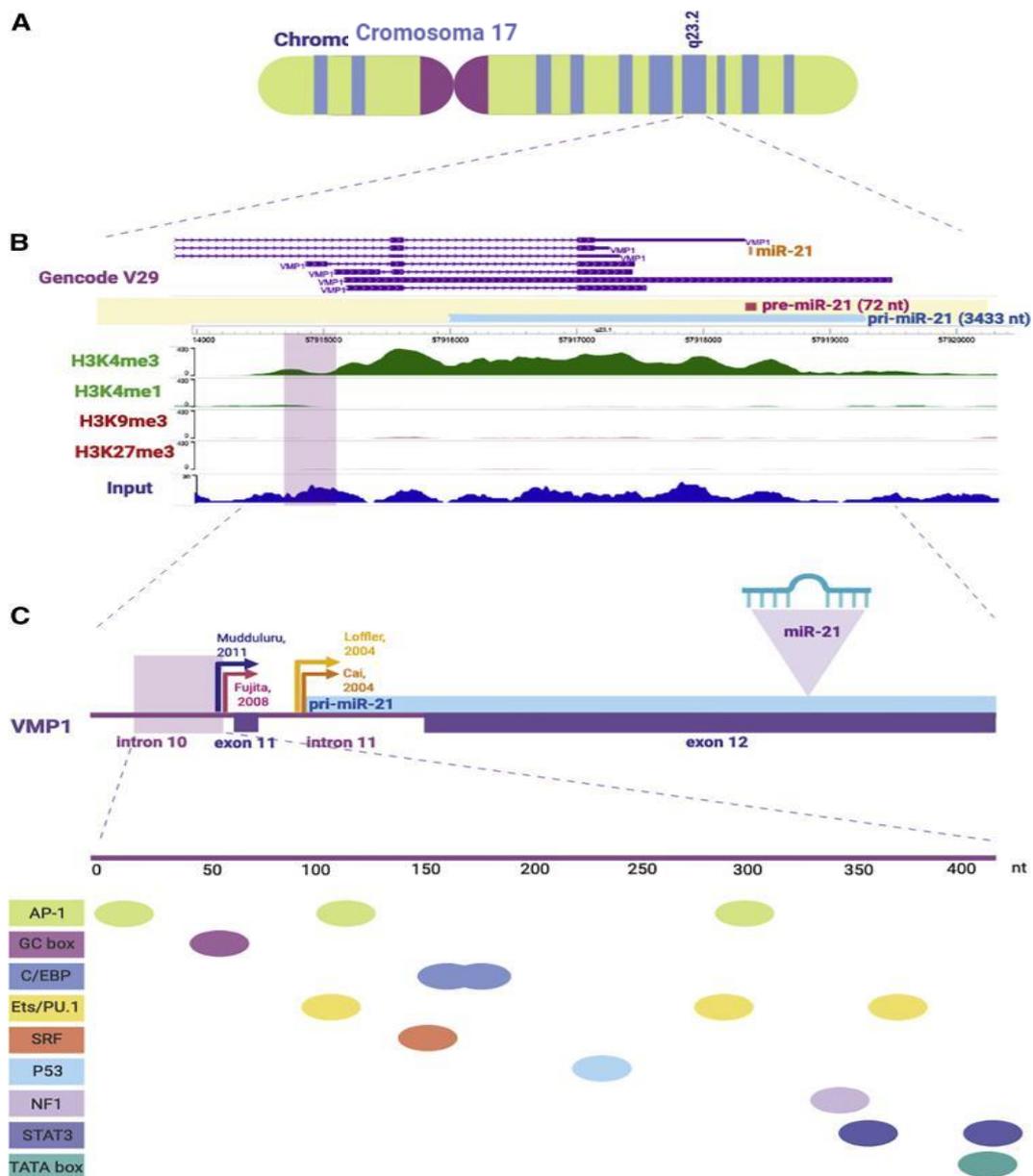


Figura 4. Localización y caracterización genómica de miR-21.

(A) Localización genómica, miR-21 se localiza en el brazo largo del cromosoma 17 (17q23.2), siendo un transcrito intragénico, dentro del 11^{vo} intrón del gen codificante *TMEM49*. **(B)** Mapa de cromatina de miR-21. Las modificaciones post-traduccionales de histonas muestran que existe actividad transcripcional activa cerca de la región promotora y a lo largo del transcrito de miR-21, marcada por H3K4me3 y H3K4m1 (en verde), a diferencia de las marcas de represión H3K9me3 y H3K27me3, las cuales se encuentran auscentes. Y por otro lado se observan los histogramas de secuenciación de inmunoprecipitación de cromatina (ChiP-seq) (en azul) **(C)** Región promotora putativa de miR-21 (Cai et al., 2004; *Curcumin Regulates miR-21 Expression and Inhibits Invasion and Metastasis in Colorectal Cancer - PubMed*, s. f.; Fujita et al., 2008; Löffler et al., 2007), en las cuales se encuentran diversos sitios de unión a factores de transcripción en la región promotora putativa <Modificado de Bautista, et al, 2020²⁰

1.4.1 Vías metabólicas intracelulares reguladas por miR-21-5p en cáncer.

Mediante evidencia experimental (RT-qPCR, microarreglos, RNA-Seq), se ha demostrado que miR-21-5p se encuentra sobreexpresado en varios tipos de cáncer como pulmón, estómago, cabeza y cuello, así como en cáncer de mama⁵(H. Wang et al., 2019), entre otros(Feng & Tsao, 2016; Pfeffer et al., 2015; Selcuklu et al., 2009). La expresión diferencial de miR-21-5p en un contexto anormal tiene un efecto directo en los genes que son regulados negativamente a nivel post-transcripcional por miR-21-5p, específicamente, los genes supresores de tumor mejor caracterizados como blancos génicos de miR-21-5p son *PTEN*, *SPRY2*, *PDCD4*, *TIMP3* y *RECK*, los cuales están asociados a vías metabólicas oncogénicas como la proliferación celular, la migración, la invasión y la apoptosis (Figura 5), y se ha demostrado que una alteración en la expresión de miR-21-5p, tiene un efecto directo en la expresión génica de sus mRNAs blanco, y en la acumulación de la proteína. Por lo anterior, miR-21-5p es un oncomiR fundamental en los procesos oncogénicos cuando su expresión se encuentra alterada, debido a que regula genes importantes en procesos metabólicos asociados al cáncer.

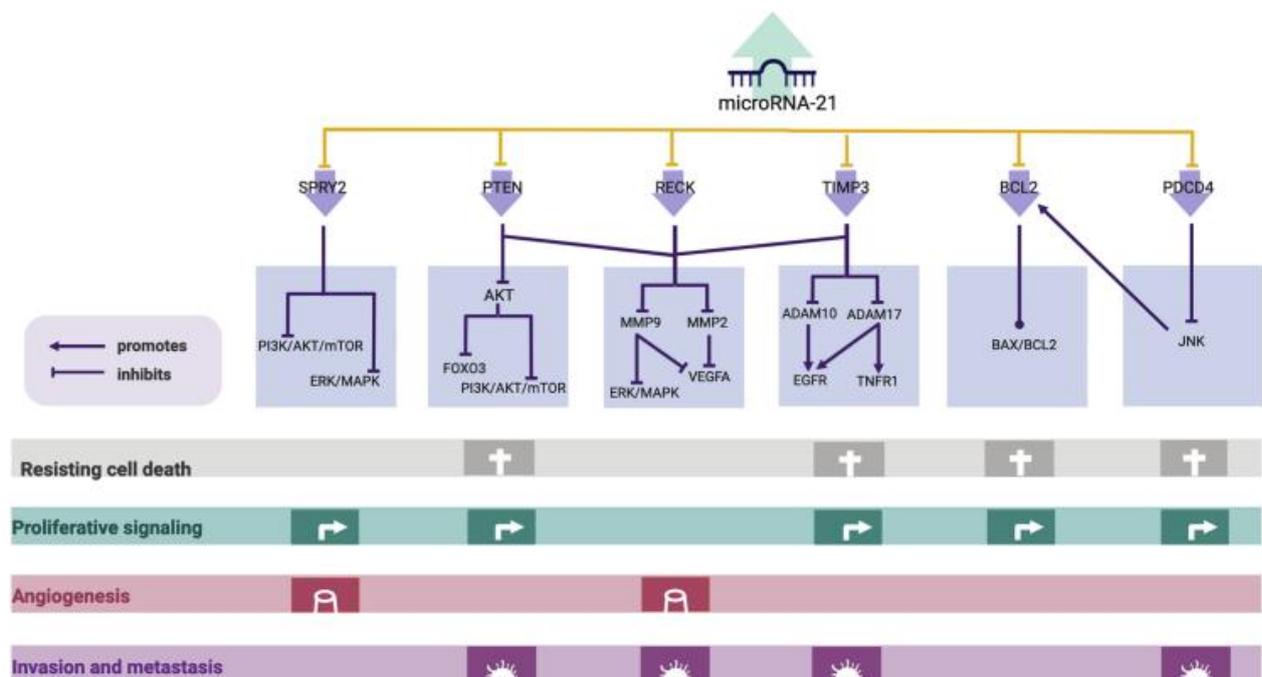


Figura 5. Blancos génicos que se encuentran asociados a procesos oncogénicos.

MiR-21 se encuentra sobre-expresado en varios tipos de cáncer, como cáncer de pulmón, gliomas, próstata, colorectal, entre otros. Al encontrarse esta sobre-expresión de miR-21, la expresión de los genes blanco involucrados en procesos oncogénicos se encuentra alterada, demostrando una sub-expresión de SPY2, PTEN, RECK, TIMP3, BCL2 y PDCD4. Esta alteración en la expresión de los genes regulados por miR-21, conduce a una regulación negativa de vías de señalización asociadas a procesos oncogénicos como resistencia a muerte celular, proliferación, angiogénesis, invasión y metástasis. Específicamente, SPY2 regula negativamente la angiogénesis(Wietecha et al., 2011) y la proliferación celular, mediante las vías PI3K/AKT/mTOR potenciando el transporte de Erb2 y HER2, y la vía ERK/MAPK, inhibiendo la expresión de Raf1(Gao et al., 2012). Por otro lado, PTEN regula la vía PI3K/AKT/mTOR inhibiendo la activación de AKT, y por lo tanto potenciando la expresión de FOXO3, oncogen que regula sobrevivencia celular, proliferación y diferenciación(X. Zhang et al., 2011). Asimismo, AKT activa la vía mTORC1(Memmott & Dennis, 2009), alterando crecimiento celular, proliferación y motilidad celular. De igual manera, RECK controla negativamente a MMP-2 y MMP-9 (metaloproteinasas de la matriz extracelular), alterando los procesos de angiogénesis e invasión celular(Oh et al., 2001; Takahashi et al., 1998). TIMP3 regula a ADAM10 y ADAM17 (dos MMP), así como a sustratos de receptores: TGF-B, HER2, HER4 y VEGFR2, potenciando la resistencia a muerte celular, proliferación e invasión y metástasis(Miller et al., 2017). De manera contraria, miR-21 induce la expresión de BCL2, teniendo como consecuencia una alteración del radio BAX/BCL2, disminuyendo apoptosis y aumentando la proliferación celular(Dong et al., 2011). Por último, PDCD4, se encuentra asociado a la activación de la vía JNK, potenciando una sobre-expresión de proteínas tipo BLC2, desregulando apoptosis, proliferación y migración(Shen et al., 2013)¹¹⁵. <Modificado de Bautista, *et al*, 2020(Bautista-Sánchez et al., 2020)

En particular, miR-21-5p ha sido reportado en análisis experimentales tanto *in vitro* como *in vivo*, como un regulador de genes supresores de tumor fundamentales como *PTEN* (homólogo de fosfatasa y tensina), *SPYR3* (antagonista de señalización RTK 2), *PDCD4* (factor de muerte celular programada 4), etc. Se ha demostrado mediante experimentos de sobreexpresión de miR-21-5p, que tanto el nivel de expresión y la proteína de *PTEN*, *SPYR3*, *PDCD4*, disminuyen de manera que las vías metabólicas oncogénicas asociadas son alteradas, éstas incluyen la proliferación celular, la apoptosis, la migración y la invasión, entre otras, potenciando un fenotipo oncogénico. Por otro lado, al realizar experimentos de inhibición mediante oligonucleótidos antisentido (LNAs, *Locked Nucleic Acid*), e inhibidores, se demostró que al inhibir la expresión de miR-21-5p, el nivel de expresión y de la proteína de *PTEN*, *SPYR3* y *PDCD4* aumentaban, disminuyendo el desarrollo tumoral, y, en consecuencia, el fenotipo oncogénico. Asimismo, otros genes

que se encuentran regulados por miR-21-5p son *RECK* (la proteína rica en cisteína inductora de reversión con motivos Kazal), *TPM1* (tropomiosina-1), *TIMP3* (el inhibidor de metaloproteínasa 3), *SERPINB5* (el inhibidor de la peptidasa de serpina B), *CASP3* (caspasa 3), *FASLG* (ligando de Fas), *RHOB* (miembro B de la familia de homólogos de Ras). En conjunto, miR-21-5p ha demostrado tener un papel importante en la regulación de la proliferación celular, la metástasis y la apoptosis, mediante el silenciamiento post-transcripcional de sus mRNAs, los cuales son genes asociados a vías metabólicas oncogénicas, teniendo como consecuencia, un efecto directo en el desarrollo tumoral.

No obstante, las investigaciones se han enfocado en el estudio de la hebra 5p, considerada la hebra guía de miR-21, por lo tanto, y a pesar de ser uno de los primeros oncomiRs descritos, miR-21-3p, a la fecha sus funciones han sido reportadas en limitadas investigaciones. En cáncer de ovario, cáncer colorrectal y en líneas celulares de cáncer de próstata, se demostró mediante experimentos de inhibición utilizando inhibidores específicos para miR-21-3p, que, la inhibición en la expresión de miR-21-3p, potencia el desarrollo tumoral (Báez-Vega et al., 2016; Zhao et al., 2015).

En conclusión, miR-21-5p es un oncomiR muy bien caracterizado, el cual regula a genes asociados a las vías metabólicas oncogénicas que incluyen los genes *PTEN*, *PDCD4*, *SPRY2*, *RECK*, entre otros. Los cuales al ser silenciados post-transcripcionalmente por miR-21-5p potencian un fenotipo oncogénico mediante la proliferación celular, la migración, la invasión y la metástasis. Por otro lado, a diferencia de miR-21-5p, la hebra de miR-21-3p no ha sido ampliamente estudiada, debido a que se ha considerado la hebra pasajera, y, por lo tanto, se daba por hecho que esta hebra no era funcional porque se creía que ésta era degradada, en consecuencia, existen pocos estudios acerca de la expresión y la función de miR-21-3p en cáncer. Por lo tanto, es necesario establecer cuál es la función biológica de miR-21-3p en el desarrollo de cáncer de mama, así como establecer los mRNAs que se encuentran regulados por miR-21-3p con el objetivo de determinar cuáles serían las vías intracelulares metabólicas reguladas por los blancos génicos miR-21-3p y establecer como la alteración de estas vías metabólicas y de los genes regulados por de cómo miR-21-3p potenciarían el desarrollo de un proceso carcinogénico en los tumores de la mama.

Resumen Gráfico del Proyecto

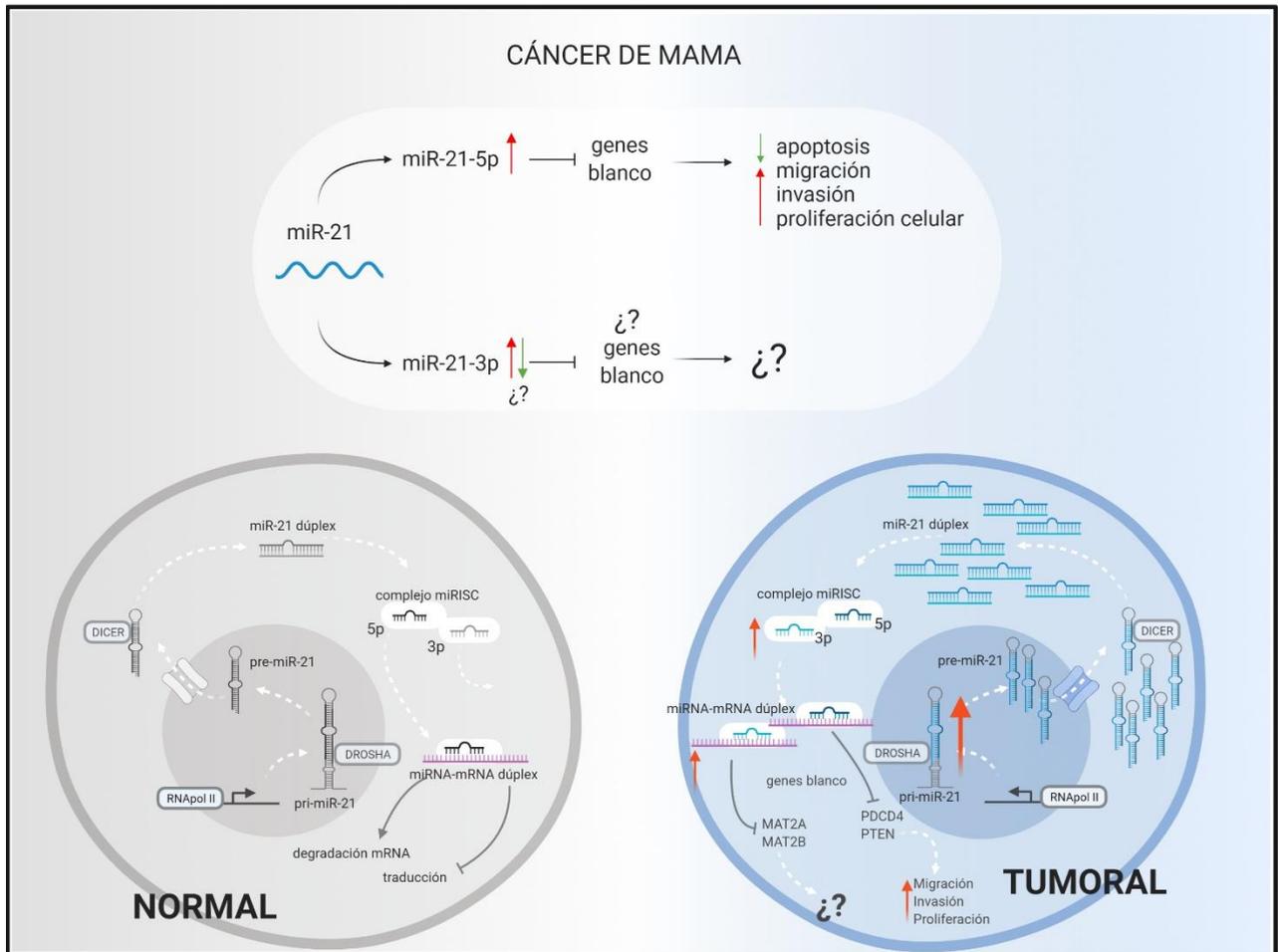


Figura 6: Importancia del proyecto de investigación

La hebra de miR-21-5p se encuentra sobreexpresada en cáncer de mama, lo cual tiene una relación con el proceso de transformación e inicio del proceso carcinogénico en cáncer de mama a través de la regulación de genes como *PTEN*, *SPRY2*, *PDCD4*, *RECK*, *TIMP3*, entre otros. Dichos genes se encuentran asociados a la regulación de vías metabólicas intracelulares como la proliferación celular, la migración y la adhesión celular, lo cual se encuentra asociado con un fenotipo neoplásico en mama. Sin embargo, en cáncer de mama se desconoce si la expresión de miR-21-3p se encuentra alterada en cáncer de mama, y no se conocen los blancos génicos que son regulados por miR-21-3p por lo que en este proyecto nos planteamos establecer el papel funcional de miR-21-3p y qué blancos génicos regula en líneas celulares de cáncer de mama.

2 Justificación

Las evidencias experimentales han establecido que miR-21-5p funciona como un oncomiR el cual regula genes asociados a vías intracelulares metabólicas oncogénicas como *PTEN*, *PDCD4*, *SPRY2*, *RECK*. Dichos genes se encuentran asociados a la regulación de vías de señalización oncogénicas como la proliferación celular, la migración y la adhesión celular, lo cual se encuentra asociado con un fenotipo neoplásico en cáncer de mama. Sin embargo, se desconoce si la expresión de la hebra miR-21-3p se encuentra alterada en cáncer de mama, y por lo tanto se desconocen los blancos génicos que son regulados por miR-21-3p en cáncer de mama. Por lo que en este proyecto nos planteamos establecer el papel funcional de miR-21-3p y proponer genes blancos, los cuales son candidatos a ser regulados en líneas celulares de cáncer de mama. Además, la expresión miR-21-3p, así como sus genes blanco, no han sido descritos en cáncer de mama, ignorando los potenciales efectos biológicos que la hebra 3p pudiera tener en este tipo tumoral. Por lo tanto, los resultados de este trabajo podrían ayudar a comprender la función biológica de la hebra 3p de miRNA-21 en cáncer de mama, así como el efecto biológico que causa el aumento o disminución de su expresión para establecer el efecto sobre sus genes blanco. Específicamente, lo anterior, permitirá establecer el efecto biológico de miR-21-3p en cáncer de mama, elucidando si miR-21-5p y miR-21-3p trabajan de manera sinérgica potenciando un fenotipo oncogénico como oncomiRs, o trabajan de manera antagónica, miR-21-5p como un oncomiR y miR-21-3p como un miRNA supresor de tumor.

3 Planteamiento del problema

Los estudios moleculares se han enfocado principalmente al estudio funcional de la hebra 5p de miR-21. Por el contrario, la expresión, así como la caracterización funcional de la hebra miR-21-3p en cáncer de mama no ha sido explorado. Por lo tanto, en este proyecto se pretende determinar si existe una alteración en la expresión de miR-21-3p en líneas celulares de cáncer de mama, así como, establecer a mRNAs blanco candidatos y vías metabólicas oncogénicas, que se encuentren reguladas por miR-21-3p en cáncer de mama.

4 Pregunta de investigación

¿Existirá un efecto sinérgico entre miR-21-5p y miR-21-3p al encontrarse sobre-expresados en líneas celulares de cáncer que promueva el desarrollo de un fenotipo neoplásico?

5 Hipótesis

La sobreexpresión de miR-21-5p y miR-21-3p promoverá de forma sinérgica el desarrollo del fenotipo neoplásico en líneas celulares de cáncer de mama.

6 Objetivo general

Determinar el efecto sinérgico de la sobreexpresión de miR-21-5p y miR-21-3p en el desarrollo del fenotipo neoplásico en líneas celulares de cáncer de mama.

6.1 Objetivos particulares

- Establecer los niveles de expresión de pri-miRNA-21, miR-21-5p y miR-21-3p en bases de datos de acceso público.
- Identificar los blancos génicos a estudiar de miR-21-5p y miR-21-3p a través de un análisis *in silico*.
- Validar la sobreexpresión de pri-miRNA-21, miR-21-5p y miR-21-3p en líneas celulares de cáncer de mama.

7 Metodología

7.1 Etapa A. Análisis bioinformático

7.1.1 Análisis de expresión diferencial de pri-miRNA-21 mediante RNA-Seq

Con el objetivo de determinar la expresión diferencial de pri-miRNA-21 se utilizaron transcriptomas de líneas celulares de cáncer de mama depositadas en la base de datos pública de Gene Expresión Omnibus (GEO, <https://www.ncbi.nlm.nih.gov/gds>, consultado en enero 2020)(Barrett et al., 2012). GEO es un repositorio público de datos de genómica funcional basados en arreglos y secuencias, dentro de la cual se proporcionan herramientas para ayudar a los usuarios a consultar y descargar experimentos y perfiles de expresión genética específicos.

Se realizó la búsqueda de resultados de secuenciación por RNA-Seq de las líneas celulares de mama de MCF-10A, MCF-7, BT474 y MDA-MB-231 utilizando la base de datos GEO Datasets para posteriormente descargar en el servidor Galaxy(Afgan et al., 2018) los archivos del Archivo Europeo de Nucleótidos (*ENA*, por sus siglas en inglés) (Tabla 1). De los cuales se consultaron los datos de expresión de pri-miRNA-21, específico para cada línea celular. Se realizó la selección de archivos, considerando que cada resultado fuera el RNA-Seq de las líneas celulares utilizadas, además de la línea celular control, se buscó que no hubiesen sido administrados con ningún tipo de vehículo, disolvente o control de vector con el objetivo de analizar la expresión de pri-miRNA-21 en su contexto natural o silvestre de cada línea celular analizada.

Tabla 1. Claves de acceso en GEO Datasets para los archivos de RNA-Seq de líneas celulares de cáncer de mama.

Línea Celular	Clave de GEO	Clave de SRA	Tipo	Plataforma de Secuenciación
MCF-10A	GSM1172882	SRX317727	Paired-end	Illumina Genome Analyzer IIx
	GSM1897320	SRX1293333		Illumina HiSeq 2000
	GSM1915044	SRX1361306		Illumina Genome Analyzer IIx
MCF-7	GSM2072527	SRX1603568		Illumina HiSeq 2000
	GSM2072571	SRX1603615		
	GSM2072572	SRX1603616		
BT474	GSM1172853	SRX317702		Illumina Genome Analyzer IIx
	GSM1466928	SRX671583		Illumina HiSeq 2000
	GSM1897280	SRX1293293		
MDA-MB-231	GSM2242132	SRX1960593		Illumina HiSeq 2000
	GSM2791584	SRX3210867		
	GSM2791576	SRX3210859		

7.1.2 Análisis de expresión diferencial de miR-21-5p y miR-21-3p por RNA-Seq.

Se consultó en la base de datos TACCO (*Transcriptome Alterations in CanCer Omnibus*, <http://tacco.life.nctu.edu.tw/>, consultado mayo 2020)(Chou et al., 2019). TACCO es una base de datos diseñada específicamente para conectar alteraciones del transcriptoma, de vías y resultados clínicos en diferentes tipos de cánceres. Los usuarios pueden

seleccionar genes expresados diferencialmente con diferentes criterios dentro de la base de datos TACCO o cargar su propia lista de genes para ser analizados.

Los datos de expresión de ambas hebras de RNA que corresponden a miR-21-5p y miR-21-3p, las cuales pertenecen a una cohorte de pacientes de cáncer de mama en donde se compara tejido normal contra tejido tumoral. Se graficó, utilizando la tasa de cambio (2-FoldChange) obtenida en transcritos por millón (TPM), de la expresión de cada hebra entre tejido normal y tejido tumoral, además se incluyeron gráficos de barras comparativos en diferentes tipos de cáncer.

7.1.3 Búsqueda de genes blanco de miR-21-3p

Se realizó una búsqueda de blancos génicos candidatos de miR-21-3p, utilizando durante todo el proceso de búsqueda, a miR-21-5p como control interno de los análisis *in silico*.

Esta búsqueda se realizó utilizando servidores web que mediante algoritmos predicen los blancos biológicos de miRNAs mediante la búsqueda de interacciones entre miRNA-blancos (sitios de unión miRNA-mRNA, minería de texto (*text mining*), evidencia experimental (Western Blot, microarreglos, sobreexpresión e inhibición de miRNAs). Los servidores utilizados fueron miRTarBase (https://mirtarbase.cuhk.edu.cn/~miRTarBase/miRTarBase_2019/php/index.php, consultado Junio 2020) , TargetScan V7.2 (http://www.targetscan.org/vert_72/, consultado Junio 2020), TarBase V8 (https://carolina.imis.athena-innovation.gr/diana_tools/web/index.php?r=tarbasev8%2Findex, consultado Junio 2020) y RNAinteractome (<https://www.rna-society.org/rnainter/>, consultado junio 2020), en los cuales se seleccionó un punto de corte determinado en 100 genes blanco para la lista de genes obtenidos para cada servidor, considerando los genes con mayor valor de predicción, mayor número de publicaciones, y experimentos realizados que confirmen a estos genes como blancos directos.

Una vez obtenida la lista de genes de cada servidor, se utilizó como herramienta, un diagrama de Venn(Bardou et al., 2014) para representar visualmente qué genes blanco obtenidos de los 4 servidores, coinciden entre sí, con el fin de reducir nuestro universo de genes blanco. Posteriormente, de los genes obtenidos en la intersección de los 4

servidores, se realizó una búsqueda mediante minería de textos, de evidencia experimental que los validara como blancos directos, con el fin de seleccionar a los mejores blancos candidatos para cada hebra. En donde se seleccionaron 2 genes blanco para miR-21-5p y 2 genes blanco para miR-21-3p.

7.1.4 Expresión diferencial de mRNAs que son blancos génicos de miR-21-5p y miR-21-3p.

Con el objetivo de analizar la expresión de los genes blanco candidatos de miR-21-3p en un contexto neoplásico, a partir de los datos obtenidos de la búsqueda de resultados de secuenciación por RNA-Seq de las líneas celulares MCF-10A, MCF-7, BT474 y MDA-MB-231 utilizando la base de datos GEO Datasets (Barrett et al., 2012), se consultaron los datos de expresión de los mRNAs de los genes *PTEN*, *PDCD4*, *MAT2A* y *MAT2B*, con las cuales se realizaron gráficos de caja (del inglés, *boxplot*), utilizando como unidad de medida estandarizada los transcritos por millón (TPM) para comparar la expresión que presenta cada línea celular, así como establecer si existía alguna diferencia significativa en expresión mediante análisis ANOVA, seguido de una prueba de Tukey, considerando un intervalo de confianza de 95%, por medio de la herramienta *RStudio*.

Para la expresión en la cohorte de pacientes de cáncer de mama, se consultó la base de datos de acceso público conocida como *GEPIA* (<http://gepia.cancer-pku.cn/>, consultado enero 2020). *GEPIA* es una aplicación web interactiva para el análisis de la expresión génica basada en muestras de tumores y muestras normales de las bases de datos TCGA y GTEx, que utiliza un procesamiento estándar para datos de secuenciación de RNA, la cual permite el análisis de expresión diferencial, trazado de perfiles, análisis de correlación, análisis de supervivencia del paciente, detección de genes similares y análisis de reducción de dimensionalidad.

La expresión de los genes *PTEN*, *PDCD4*, *MAT2A* y *MAT2B* fue analizada mediante TPM. Asimismo, con la herramienta de *GEPIA* se realizaron gráficos de caja de cada gen, comparando la expresión presentada en tejido normal versus tejido tumoral, utilizando *t-student* no pareada de una cola, con intervalo de confianza del 95%, para verificar si existían diferencias significativas entre la expresión de genes en las condiciones analizadas.

7.1.5 Elaboración de gráficos de caja de expresión *in silico* e *in vitro*

Los gráficos de caja se elaboraron con el paquete para la interfaz RStudio *ggplot2*, mediante el uso de la función *boxplot*, utilizando los parámetros preestablecidos por la interfaz de RStudio, *ggplot*.

7.1.6 Determinación de vías metabólicas intracelulares asociadas a cáncer de los genes *PTEN*, *PDCD4*, *MAT2A* y *MAT2B*.

Con el objetivo de establecer cuáles son vías metabólicas intracelulares asociadas a cáncer de los genes *PTEN*, *PDCD4*, *MAT2A* y *MAT2B*, se utilizó la base de datos de acceso público RAIN V 1.0 (*RNA-protein Association and Interaction Networks*, <https://rth.dk/resources/rain/>, consultado marzo 2021)(Junge et al., 2017). La plataforma RAIN facilita la comprensión de redes moleculares mediante la integración de varias interacciones de RNA y redes de asociación proteína-proteína, en la cual se observa la red de interacción del gen a estudiar, con 5 blancos no codificantes y 5 codificantes, en donde estos 10 blancos, resultan ser los de mayor puntaje de predicción. La aplicación de esta herramienta, permitió el desarrollo de una escala común utilizando diferentes puntajes de confianza heterogéneos, con el fin de obtener puntajes de probabilidad a través de la evaluación comparativa. Mediante el uso del puntaje de confianza a interacciones (máxima confianza 0.9 en STRING), a las asociaciones con apoyo experimental se les asignaron puntuaciones de confianza basadas en el número de experimentos / publicaciones de apoyo, a las asociaciones derivadas de la minería de texto se les asigna una puntuación en función de la coincidencia de nombres de genes, y para las predicciones de blancos de miRNA.

Asimismo, utilizando la herramienta de análisis del servidor *RAIN*. La cual es un recurso que abarca ncRNA y sus asociaciones con otros ncRNA y proteínas. RAIN integra interacciones de ncRNA de un conjunto diverso de fuentes, utilizando 4 organismos: humano (*Homo sapiens*), ratón (*Mus musculus*), rata (*Rattus norvegicus*) y levadura de panadería (*Saccharomyces cerevisiae*), con la finalidad de calificar la confiabilidad de cada interacción mediante un esquema de calificación basado en la comparación con un conjunto seleccionado de interacciones. Finalmente, integra las asociaciones ncRNA-RNA y ncRNA-proteína con asociaciones proteína-proteína

contenidas en la base de datos STRING. Esto permite a los usuarios explorar redes de interacción compleja.

Se muestran las vías metabólicas en las cuales el gen seleccionado está involucrado mediante análisis de vías KEGG (*Kyoto Encyclopedia of Genes and Genomes*, <https://www.genome.jp/kegg/> consultado marzo 2021) y GO (*Gene ontology*: procesos biológicos, <http://geneontology.org/>, consultado marzo 2021), una vez obtenidas las vías metabólicas asociadas a *PTEN*, *PDCD4*, *MAT2A* y *MAT2B*, se realizó una búsqueda de literatura que validara las vías metabólicas afectadas, en donde se haya reportado que *PTEN*, *PDCD4*, *MAT2A* y *MAT2B* sean genes reguladores de dichas vías metabólicas.

7.2 Etapa B. Validación experimental

7.2.1 Purificación de RNA total de líneas celulares de cáncer de mama.

Para estandarizar y validar el método de cuantificación de la expresión de pri-miRNA-21, miR-21-5p y miR-21-3p se cultivaron las líneas celulares MCF-10A, MCF-7, BT474 y MDA-MB-231, y se obtuvo el RNA total de cada una de ellas, comprobando la calidad de este y determinando su integridad, así como su concentración.

7.2.1.1 Cultivo de las líneas celulares de cáncer de mama.

Las líneas celulares que se usaron en este trabajo son derivadas de tejido mamario, y se cultivaron en condiciones libres de antibiótico. Las características y condiciones específicas para cada línea celular se encuentran en la siguiente tabla:

Línea celular	Características		Medio de cultivo	Suplementos	Temperatura	CO ₂
MCF-10A (CRL-10317)	Proveniente de tejido de glándula mamaria de <i>Homo sapiens</i> (humano), femenino caucásico de 36 años, cuyo origen es un paciente con fibrosis quística.	Línea celular epitelial no tumorigénica	MEBM	- SFB al 10%	37 °C	5%
MCF-7 (HTB-22)	Proveniente del tejido de glándula mamaria de <i>H. sapiens</i> , femenino caucásico de 69 años, derivada de adenocarcinoma en un sitio de metástasis.	Línea celular tumorigénica, de morfología epitelial. Subtipo molecular luminal A	DMEM/F-12	- SFB al 10% - insulina recombinante (0.01 mg/mL)		5%
BT474 (HTB-20)	Proveniente de tejido de los ductos de la glándula mamaria de <i>H. sapiens</i> , femenino caucásico de 60 años, derivada de un carcinoma	Línea celular tumorigénica, de morfología epitelial. Subtipo molecular luminal B	46-X Hybri Care	- SFB al 10% - Bicarbonato de sodio (1.5 g/L)		5%
MDA-MB-231 (HTB-26)	Proveniente de tejido de la glándula mamaria de <i>H. sapiens</i> , femenino caucásico de 51 años, derivada de adenocarcinoma en un sitio de metástasis.	Línea celular tumorigénica, de morfología epitelial. Subtipo molecular Triple negativo	Leibovitz's L-15	- SFB al 10%		--

7.2.1.2 Purificación de RNA total

Se extrajo RNA total de los cultivos de líneas celulares MCF10-A, MCF7, BT474 y MDA-MB-231 realizados en botellas de cultivo de 25 cm², utilizando *TRIzol Reagent /Ambion, Life Technologies, de Thermo Fisher Scientific*).

Se retiró el medio de cultivo de cada caja con lavado posterior con amortiguador fosfato salino (PBS, por siglas en inglés) agregando 1 mL de TRIzol de manera homogénea por toda la superficie de la caja. Esto fue almacenado en un tubo eppendorf de 1.5 mL. Por cada mL de TRIzol utilizado se agregó a la mezcla 0.2 mL de cloroformo, agitando en vortex y centrifugando a 14,000 revoluciones por minuto (RPM), durante 15 minutos a

4°C, en otro tubo de centrifuga de 1.5 ml se le agregó el sobrenadante y a este 1 mL de isopropanol, incubándolo 40 minutos en hielo, posteriormente se centrifugó a 14,000 RPM, 15 minutos a 4°C. Después, se retiró el sobrenadante, y la pastilla que pertenecía al RNA total extraído se lavó con etanol al 70% y se centrifugó a 14,000 RPM por 5 minutos a 4°C. Por último, se decantó el sobrenadante y se dejó secar la pastilla a temperatura ambiente por 15 minutos. Finalmente, se hizo una elución de la pastilla en 20 μ L con agua libre de RNasas, incubando a 55°C durante 10 minutos. El RNA total se conservó a -80°C.

7.2.1.3 Cuantificación de RNA

Se llevó a cabo la cuantificación de RNA total con el dispositivo NanoDRop (AccesoLab, de *Thermo Fisher Scientific*), determinando la señal de absorbancia de ácidos nucleicos (RNA). Donde se utilizó como solución blanco, agua libre de RNasas. Una vez determinado el blanco, se cuantificó 1 μ L de RNA total extraído de cada línea celular, verificando los ratios de absorbancia 260/280 (con el fin de determinar contaminación con DNA o proteínas) y 230/260 (para contaminación con fenoles). El ratio 260/280 se considera ideal cuando su valor es cercano a 2.0, mientras que el rango del ratio 260/230 es de 2.0 a 2.2.

7.2.1.4 Determinación de la calidad e integridad del RNA

7.2.1.4.1 Gel Agarosa

Se cargaron 500 ng de cada RNA total extraído en un gel de Agarosa al 1.2% preparado de la siguiente manera. Se pesó 0.36 g de Agarosa y se disolvió en 27 mL de amortiguador MOPS 1X y 3 mL de agua con dietil pirocarbonato (DEPC) libre de RNasas, agregando 1 μ L de la solución para tinción de ácidos nucleicos *Gel Red* (concentración final 1X). A través del uso de una fuente de poder, se realizó la corrida del gel de agarosa en una cámara de electroforesis (*Bio-Rad*) a 90 Volts, durante 40 minutos y posteriormente se utilizó un transiluminador (*UVP*) para visualizar el gel.

7.2.1.4.2 Bioanalizador

Se utilizó el bioanalizador *TapeStation 2200* (Agilent Technologies) para determinar la integridad y calidad del RNA total extraído de cada línea celular. Se preparó para cada muestra 1 μL de RNA total (100 ng/ μL) en 5 μL de solución amortiguadora desnaturalizante de RNA (*RNA Screen Tape Sample Buffer y Screen Tape*) para RNA, de Agilent Technologies). La mezcla fue incubada a 72°C por 3 minutos, posteriormente fue colocada en hielo 2 minutos, y posteriormente, la mezcla de RNA y solución desnaturalizante fue cargada en la placa *ScreenTape* e introducida al bioanalizador. Se obtuvo tanto la concentración de RNA, como el número de integridad del RNA (RIN, por sus siglas en inglés). Se considera que el RNA tienen buena calidad e integridad cuando el valor del RIN es igual o mayor a 7.0.

7.2.2 Expresión de los transcritos pri-miRNA-21, miR-21-5p, miR-21-3p, y sus genes blanco mediante PCR en tiempo real en líneas celulares de cáncer de mama.

Para determinar la expresión diferencial del transcrito pri-miRNA-21 por RT-qPCR fue necesario obtener cDNA a partir del RNA total extraído de las líneas celulares MCF-10A, MCF-7, BT474 y MDA-MB-231, para lo cual se siguieron los siguientes pasos:

7.2.2.1.1 Síntesis de cDNA para ensayo de expresión de pri-miRNA-21

1. **Tratamiento con DNasa I.** Se utilizó la enzima DNasa I (*Thermo Fisher Scientific*), para eliminar contaminación con DNA genómico,

Para cada línea celular y con un volumen final de 10 μL , se preparó una alícuota de 1 μg de RNA total, 1 μL de solución amortiguadora 10X para DNasa I con MgCl₂, 1 μL de enzima DNasa I (1U/ μL) y se completó el volumen final con agua libre de RNasas. Se incubó la mezcla a 37°C por 40 minutos. Posteriormente, se agregó 1 μL de ácido etilaminotetraacético (EDTA) 50 mM, incubando la muestra a 65°C por 10 minutos. Finalmente, se dejó incubando la mezcla en hielo y se almacenó a -80°C (se utilizó el kit *DNase I, RNase-free, molecular biology*, de *Thermo Fisher Scientific*).

Una vez realizado el tratamiento de DNasa I a las muestras de RNA total. Se utilizó el kit *High Capacity cDNA Reverse Transcription, applied biosystems*, de *Thermo Fisher Scientific* para sintetizar cDNA a utilizar en el ensayo de cuantificación de mRNAs, pri-miRNA, *PTEN*, *PDCD4*, *MAT2A* y *MAT2B*.

2. **Reacción de transcripción reversa.** Para la síntesis de cDNA , se mezclaron 10 μL de RNA total tratado con DNasa I, 4.2 μL de agua bi-destilada, desionizada y estéril, 2 μL de solución amortiguadora PCR 10X, 2 μL de *Random Primers*, 0.8 μL de dNTPs 10 mM y 1 μL de enzima transcriptasa reversa de alta capacidad (RT) a 50 U/ μL . La reacción se incubó de la siguiente manera en un termociclador:

Paso	1	2 Reacción de transcriptasa reversa.	3 Stop	4
Temperatura (°C)	25	37	85	4
Tiempo (min)	10	120	5	Indefinido

*La reacción RT contiene todo lo antes mencionado con excepción de la enzima RT, en su lugar, se agrega 1 μL de agua libre de RNasas.

7.2.2.1.2 Determinación de la calidad de cDNA por PCR punto final

Para verificar la calidad del cDNA sintetizado mediante el kit *High Capacity cDNA Reverse Transcription, applied biosystems*, de *Thermo Fisher Scientific* se realizó un PCR de punto final, utilizando el gen constitutivo *RPS28* como control de amplificación. Donde los oligonucleótidos sentido y antisentido (*RPS28FW* y *RPS28RV*, respectivamente) generan un amplicón de la secuencia de aproximadamente 100 pares de bases. Cuando se genera el amplicón de 100 pb se considera que el CDNA sintetizado tiene la calidad necesaria para ser utilizado en ensayos de qPCR.

Se preparó para cada reacción obtenida de la Transcriptasa reversa la siguiente mezcla: 7.06 μL de agua libre de RNasas, 1 μL de amortiguador PCR 10X -Mg, 0.3 μL de Magnesio (Mg^{2+}) 50 mM, 0.2 μL de mezcla de dNTPs 10 mM, 0.04 μL de DNA Polimerasa Taq Platinum (5U/ μL), mezcla de oligonucleótidos (forward y reverse) RPS28 10 μM y 1 μL de cDNA (dilución 1:2), para obtener un volumen final de reacción de 10 μL (se utilizó el kit *Platinum™ Taq DNA Polymerase*, invitrogen, de *Thermo Fisher Scientific*).

La reacción se procesó en un termociclador con la siguiente programación de temperaturas:

Paso	Temperatura	Tiempo
Desnaturalización inicial	94°C	2 minutos
Desnaturalización*	94°C	20 segundos*
Alineación*	60°C	20 segundos*
Extensión*	72°C	20 segundos*
Extensión final	72°C	7 minutos
<i>Stop</i>	4°C	Indefinido

*36 ciclos de amplificación de la PCR.

Una vez terminada la amplificación, se cargaron 10 μL de reacción de amplificación junto con 2 μL de amortiguador de corrida de DNA 6X en un gel de agarosa 1% con amortiguador TBE al 0.5%, a 90 Volts por 40 minutos. Se visualizó el gel en un transiluminador (UVP).

7.2.2.1.3 Cuantificación de expresión diferencial de pri-miRNA-21.

Con el cDNA que fue validado por medio de PCR punto final, y que se consideró como cDNA de calidad para ser utilizado en ensayos de qPCR. Se realizó la cuantificación de la expresión relativa del transcrito pri-miRNA-21 por PCR en tiempo real (qPCR), utilizando cuadruplicados técnicos y quintuplicados biológicos para cada línea celular, tanto para la sonda del gen constitutivo (hsa-GAPDH) como para el gen problema (hsa-pri-miRNA-21).

La cuantificación se realizó con el equipo QuantStudio 3 (applied biosystems, de Thermo Fisher Scientific), preparando la siguiente reacción: 5 μL de TaqMan PCR Universal Master Mix, 0.5 μL de cada sonda 20X TaqMan pri-miRNA Assays, 2.5 μL de agua libre de RNasas y 2 μL de cDNA (dilución 1:5) o 2 μL de agua libre de RNasas para los controles negativos (NTC).

7.2.2.2 Determinación de la expresión relativa de miR-21-5p y miR-21-3p mediante PCR tiempo real

7.2.2.2.1 Síntesis de cDNA para ensayo de expresión de miRNAs

Para la síntesis de cDNA a ser utilizado para ensayos de amplificación de miRNAs mediante sondas TaqMan, se realizó un tratamiento de DNasas I al RNA total extraído de las líneas celulares MCF-10A, MCF-7, BT474 y MDA-MB-231, siguiendo las mismas características experimentales mencionadas en el apartado **7.2.2.1.1. (1. Tratamiento con DNasa I)**. Una vez realizado el tratamiento de DNasa I a las muestras de RNA total. Se utilizó el kit: *TaqMan™ Advanced miRNA cDNA Synthesis, applied biosystems*, de *Thermo Fisher Scientific* para sintetizar el cDNA que sería utilizado para cuantificar los transcritos de miRNAs maduros, miR-21-5p y miR-21-3p.

Este kit de síntesis de cDNA se divide en 5 etapas principales: la reacción de poliadenilación, la ligación de un adaptador, la retrotranscripción, una amplificación universal, y, por último, la detección de los transcritos por medio de qPCR con sondas TaqMan. En la primera etapa de poliadenilación, a partir de RNA total, por medio de la

polimerasa poly(A) se sintetiza una cola de adenina en el extremo 3' de todos los miRNAs, secuencia que servirá como secuencia molde para el reconocimiento del oligonucleótido *reverse* en la reacción de amplificación de miRNAs. Después, se liga una secuencia adaptadora en el extremo 5' de las moléculas. Esta secuencia en el extremo 5', funcionará como oligonucleótido *forward* en la reacción de amplificación de miRNAs. Posteriormente, la reacción de transcripción reversa se lleva a cabo al unir un oligonucleótido RT universal al extremo poly(A) de las moléculas. El cDNA resultante se somete a una amplificación universal de miRNAs, haciendo posible la detección más sensible de transcritos poco expresados.

1. **Reacción de poliadenilación:** En esta reacción, a partir de RNA total, por medio de la polimerasa poly(A) se sintetizó una cola de adenina en el extremo 3' de todos los miRNAs

Para cada línea celular se prepararon alícuotas de la siguiente reacción, en donde se utilizó un volumen final de 5 μL , y se añadieron: 0.5 μL de amortiguador 10X Poly(A), 0.5 de ATP 10 mM, 0.3 μL de enzima Poly(A), 1.7 μL de agua libre de RNasas y 2 μL de RNA (100 ng/L) tratado con DNasa I. Se colocó la reacción en un termociclador con el siguiente programa:

Paso	Temperatura	Tiempo
Poliadenilación	37°C	45 minutos
Inactivación	65°C	10 minutos
Stop	4°C	Indefinido

Al término, se procedió inmediatamente con la siguiente reacción de ligación del adaptador.

2. **Reacción de ligación del adaptador:** Para cada línea celular se prepararon dos alícuotas (RT+ y RT-) con un volumen final de 15 μL , de la siguiente mezcla: 3 μL de amortiguador 10X DNA ligasa, 4.5 μL de PEG 8000 al 50%, 0.6 μL de Adaptador de ligación 25X, 1.5 μL de RNA ligasa, 0.4 μL de agua libre de RNasas y los 5 μL resultantes de la reacción de poliadenilación. Se colocó en un termociclador con la siguiente configuración:

Paso	Temperatura	Tiempo
Ligación	16°C	60 minutos
Espera	4°C	Indefinido

Se procedió inmediatamente a la reacción de transcripción reversa.

3. **Reacción de transcripción reversa:** Para cada línea celular se prepararon dos alícuotas (RT+ y RT-) de la siguiente mezcla: 6 μL de amortiguador RT 10X, 1.2 μL de dNTP mix (25mM), 1.5 μL oligonucleótido RT 20X Universal, 3 μL de mix Enzima RT 10X (únicamente a las muestras RT+), 15 μL de la reacción de ligación y se completó con agua libre de RNasas hasta obtener un volumen final de 30 μL . Se colocaron las muestras en el termociclador con la siguiente programación:

Paso	Temperatura	Tiempo
Transcripción reversa	42°C	15 minutos
Inactivación	85°C	5 minutos
<i>Stop</i>	4°C	Indefinido

Una vez finalizada la reacción se guardaron los productos a -20°C, o se procedió a la amplificación de miRNAs.

4. Amplificación de miRNAs

El cDNA resultante de la reacción de transcriptasa reversa (RT) se sometió a una amplificación universal de miRNAs.

Para cada reacción RT, se preparó la siguiente mezcla, para obtener un volumen final de 50 μL : 25 μL de 2X miR-Amp Master Mix, 2.5 μL de 2X miR-Amp Primer Mix, 17.5 μL de agua libre de RNasas y 5 μL de la reacción resultante de la RT.

Se colocaron las muestras en un termociclador y se incubaron con la siguiente configuración de temperaturas:

Paso	Temperatura	Tiempo
------	-------------	--------

Desnaturalización inicial	95°C	2 minutos
Desnaturalización*	95°C	3 segundos*
Alineación/ Extensión *	60°C	30 segundos*
Inactivación	99°C	10 segundos
<i>Stop</i>	4°C	Indefinido

*14 ciclos de amplificación de RT.

Una vez finalizada la reacción se guardaron los productos de PCR a -20°C.

7.2.2.2.2 Cuantificación de expresión diferencial de miR-21-5p y miR-21-3p por RT-qPCR.

Se realizó la cuantificación de la expresión relativa de miR-21-5p y miR-21-3p, por qPCR tiempo real, utilizando cuadruplicados técnicos y quintuplicados biológicos para cada línea celular, tanto para la sonda del miRNA constitutivo hsa-miR-423-3p así como para las sondas de miRNAs hsa-miR-21-5p y hsa-miR-21-3p.

La cuantificación se realizó utilizando el termociclador de tiempo real QuantStudio 3 (Applied Biosystems, de Thermo Fisher Scientific), preparando la siguiente reacción: 5 μ L de TaqMan PCR Universal Master Mix (II) con UNG, 0.5 μ L de cada sonda 20X TaqMan Advanced Assays, 2 μ L de agua libre de RNasas y 2.5 μ L de cDNA (dilución 1:10 del producto de PCR resultante de la amplificación de miRNAs) o 2.5 μ L de agua libre de RNasas para los controles negativos (NTC, son muestras que no contienen cDNA, y por lo tanto no deben mostrar amplificación).

7.2.2.3 Determinación de la expresión relativa de los transcritos PTEN, PDCD4, MAT2A y MAT2B, mediante PCR en tiempo real

7.2.2.3.1 Diseño de oligonucleótidos

Utilizando la secuencia FASTA de cada gen candidato (proveniente de *NCBI nucleotide*), se diseñaron dos pares de oligonucleótidos para PCR en tiempo real con el uso de la

herramienta *Primer-BLAST* (NCBI)(*Clustal Omega < Multiple Sequence Alignment < EMBL-EBI*, s. f.; *UCSC In-Silico PCR*, s. f.; Ye et al., 2012). Donde los parámetros de diseño fueron los pre-establecidos por la herramienta, seleccionando los pares de oligonucleótidos con mayor contenido de GC (%GC > 50\%), con amplicones de longitud entre 80 a 150 bases, localizados en los exones, y que cumplieran con las siguientes características:

- Presencia de las bases G o C en el extremo 3' de cada oligonucleótido.
- Valor de auto complementariedad menor o igual a 3.00.
- Valor de Tm cercano a 60°C (58°C-62°C).
- Longitud de cada oligonucleótido no mayor a 24 bases.

Se corroboró *in silico* la existencia de un amplicón por cada par de oligonucleótidos con la herramienta en línea *In-Silico PCR* del servidor *UCSC Genome Browser*⁸⁶, y se solicitó su síntesis (Tabla 1, Anexo.).

7.2.2.3.2 Determinación de la temperatura de desnaturalización (TM) de cada par de oligonucleótidos

Con las TMs reportadas por el proveedor y las calculadas ($TM = 2^{\circ}C (A + T) + 4^{\circ}C (G + C)$) se establecieron gradientes de temperatura específicos para cada par de oligonucleótidos. Posteriormente se realizó una PCR punto final utilizando el mismo procedimiento experimental mencionado en el apartado **7.2.2.1.3**, preparando una reacción por cada temperatura (por cada par de oligonucleótidos se prepararon 6 reacciones idénticas), y programando el gradiente de temperatura en el paso de desnaturalización, en el termociclador. Por último, la temperatura a la cual se visualice mejor el amplicón (mayor concentración y definición de la banda de amplificación observada en el gel de agarosa), se escogió como la TM a utilizar en los siguientes experimentos.

7.2.2.3.3 Determinación de la eficiencia de amplificación de los oligonucleótidos mediante PCR tiempo real

La eficiencia de amplificación de los oligonucleótidos se llevó a cabo realizando PCR en tiempo real con el uso del termociclador *QuantStudio 3 (applied biosystems, de Thermo Fisher Scientific)*, preparando la siguiente reacción: 5 μ L de SYBR Green/ROX qPCR Master Mix (*molecular biology, de Thermo Scientific*), 2.2 μ L de agua calidad Biología molecular, 0.3 μ L de oligonucleótidos 10 μ M sentido y antisentido, 2.5 μ L de cDNA (en el caso del control negativo [NTC], se añadió la misma cantidad de agua). Los oligonucleótidos sentido y antisentido generan un amplicón de la secuencia del gen constitutivo *RPS28* con lo que se corrobora la adecuada amplificación de los productos de reacción.

Para realizar una curva de correlación, se realizaron diluciones seriadas 1:10, 1:100, 1:1000 y 1:10000 de cDNA de la línea celular MCF-7 a partir de una dilución primaria 1:4. Para cada dilución se preparó un cuadruplicado técnico, tanto para los oligonucleótidos del gen constitutivo como para los de la secuencia problema.

Se estableció el programa de reacción en el termociclador de la siguiente manera:

Etapa de desnaturalización		Etapa de Amplificación			Etapa de disociación		
50°C	95°C	95°C	60°C	72°C	95°C	60°C	95°C
2 min	10 min	15 s	30 s	30 s	15 s	1 min	1 s

Los resultados de C_q se relacionaron linealmente con la dilución correspondiente, y se llevó a cabo una regresión lineal para la determinación del valor de la pendiente, que se relaciona directamente con el porcentaje de eficiencia de amplificación. Para ello, se utilizó la herramienta disponible en Thermo Fisher Web Tools \rightarrow qPCR Efficiency Calculator. Se considera que un par de oligonucleótidos capaz de ser utilizado de manera óptima en ensayo de PCR en tiempo real cuando, es eficiente, cuando la eficiencia de amplificación de los oligonucleótidos se encuentra entre el 95% y el 105%.

7.2.2.3.4 Cuantificación de expresión diferencial de miR-21-5p y miR-21-3p por RT-qPCR.

La determinación de la expresión diferencial de los genes blancos propuestos: PTEN, PDCD4, MAT2A y MAT2B se realizaron con las mismas características experimentales mencionadas en el apartado **7.2.2.2.1**, en un contexto de condiciones normales, utilizando la línea celular MCF-10A, y en condiciones neoplásicas, utilizando las líneas celulares MCF-7, BT474 y MDA-MB-231.

7.2.2.4 Análisis de resultados de expresión obtenidos mediante qPCR, por medio de $\Delta\Delta CT$

1. **Método de expresión relativa $\Delta\Delta CT$:** Tanto para el ensayo de transcrito primario pri-miRNA-21 y los genes blancos, como para el ensayo de miRNAs (miR-21-5p/3p) se realizó el análisis tanto por el método delta CT (ΔCT , ecuación 1) así como por la ecuación delta delta CT ($\Delta\Delta CT$, ecuación 2),
 - Ecuación 1: $\Delta CT = CT(\text{gen problema}) - CT(\text{gen constitutivo})$
 - Ecuación 2: $\Delta\Delta CT = \Delta CT(\text{muestra problema}) - \Delta CT(\text{muestra control}) = (CT(\text{gen problema}) - CT(\text{gen constitutivo}))_{\text{muestra problema}} - (CT(\text{gen problema}) - CT(\text{gen constitutivo}))_{\text{referencia}}$

En donde se determinó el promedio de la cuantificación relativa para los quintuplicados biológicos y cuadruplicados técnicos y se calculó el error estándar de los valores obtenidos.

2. Análisis estadístico

Para analizar tanto los datos de expresión diferencial de RNA-Seq así como los obtenidos experimentalmente en líneas celulares por RT-qPCR, como primer paso, se realizó un análisis de distribución de datos, con la finalidad de comprobar si los datos obtenidos se comportaban de una manera normal. La prueba de normalidad se realizó utilizando Rstudio, empleando la prueba de Shapiro Wilt.

Para los datos que se comportaron con una distribución normal se llevó a cabo un análisis *ANOVA* de una vía, seguido de una prueba de *Tukey*, considerando un intervalo de confianza del 95%, utilizando el paquete *RStudio*.

Por otro lado, los datos que se comportaron con una distribución no paramétrica, se llevó a cabo un análisis Krustal Wallis y prueba de Wilcoxon, considerando un intervalo de confianza del 95%, utilizando el paquete *RStudio*.

8 Resultados

8.1 Análisis bioinformáticos

8.1.1 Determinación de la expresión diferencial de pri-miRNA-21, miR-21-5p y miR-21-3p a partir de datos de RNA-Seq de bases públicas de secuenciación.

Las evidencias experimentales han establecido que miR-21 se encuentra diferencialmente expresado en distintos tipos de cáncer, incluyendo el cáncer de mama (Gong et al., 2011; Loh et al., 2019; Najjary et al., 2020). Con el objetivo de determinar la expresión diferencial de miR-21, se evaluó la expresión del pri-miRNA-21 en líneas celulares que corresponden a diferentes subtipos moleculares de CaMa. La línea celular MCF-10A corresponde a un tejido epitelial normal, por lo que fue utilizada como un control de expresión de tejido normal; mientras que MCF-7, BT474 y MDA-MB-231, corresponden a los subtipos Luminal A, Luminal B y Triple negativo, respectivamente. Se obtuvieron datos de RNA-Seq de las líneas celulares mencionadas anteriormente, de la base de datos pública de GEO (*Gene Expression Omnibus*). De los cuales se realizó un análisis de expresión diferencial siguiendo las características mencionadas en el apartado 7.1.1 (metodología), los cuales fueron cuantificados mediante TPMs. Asimismo, se comprobó si existía diferencia significativa en la expresión del pri-miRNA-21, entre las líneas celulares mediante un análisis estadístico de ANOVA, seguido de una prueba estadística de Tukey, considerando un intervalo de confianza de 95%.

Los resultados mostraron que pri-miRNA-21 en la línea celular MCF10A muestra una expresión basal lo cual corresponde a una expresión en condiciones normales, sin embargo, pri-miRNA-21, se vio sobreexpresado de manera estadísticamente significativa en la línea celular MCF7, cuando se comparó contra las condiciones de expresión normal en la línea celular MCF10A. Por otro lado, la línea celular BT474 y MDA-MB-231 que corresponden a los subtipos moleculares de Luminal B y triple negativo mostraron una expresión basal del pri-miRNA-21 al igual que la línea celular normal de MCF10A, sin mostrar diferencia estadística significativa entre estos fenotipos moleculares (Figura 7). Entonces, los resultados muestran que pri-miRNA-21 se sobreexpresa de manera

estadísticamente significativa en la línea celular neoplásica Luminal A que corresponde a MCF7.

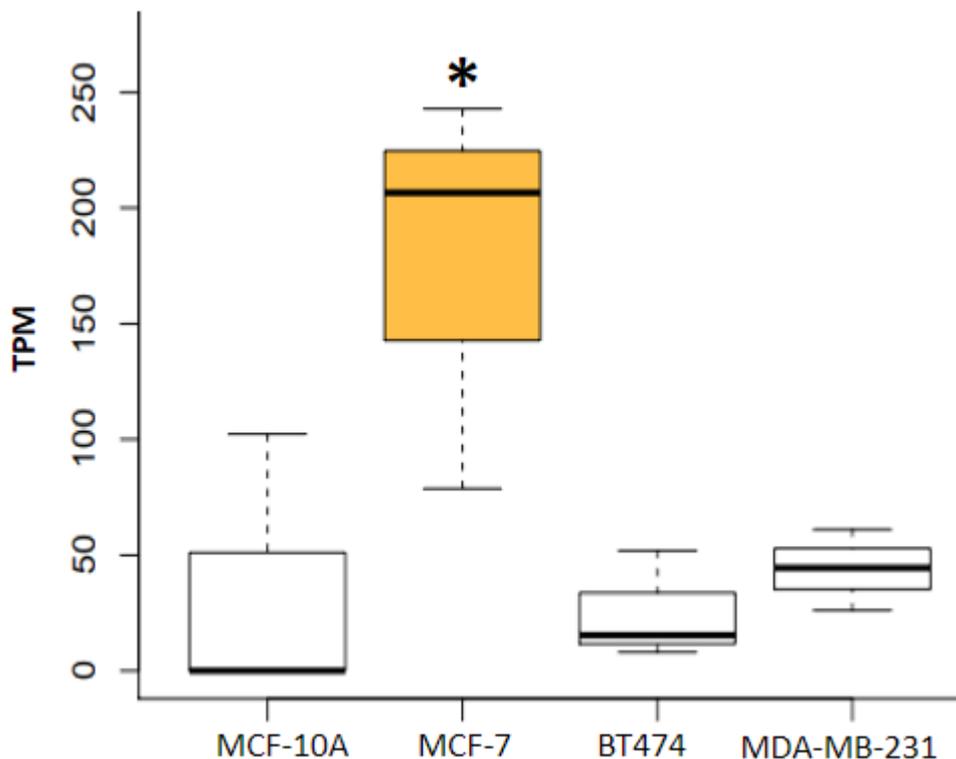


Figura 7. Expresión diferencial de pri-miRNA-21 en líneas celulares de cáncer de mama en distintos subtipos moleculares. A partir de datos obtenidos de RNA-Seq, se observa que el transcrito de naturaleza no-codificante pri-miRNA-21, presenta mayor expresión en la línea celular neoplásica subtipo molecular Luminal A que la línea celular control MCF-10A (no neoplásica) y las otras líneas celulares neoplásicas, luminal B y Triple Negativo. Presentando a pri-miRNA-21 en una frecuencia 200 veces mayor en la línea celular neoplásica MCF-7 (ANOVA con prueba de Tukey, $p=0.05$).

Una vez corroborado la expresión diferencial de pri-miRNA-21 en diferentes líneas celulares de cáncer mama, se realizó una búsqueda de la expresión de las hebras maduras de miR-21 (miR-21-5p y miR-21-3p), comparando 22 diferentes tipos de cáncer entre los cuales se incluyeron cáncer de mama, cáncer de pulmón, cáncer de próstata, entre otros, con el objetivo de comprobar si la sobreexpresión observada en pri-miRNA-21 también se presentaba en las formas procesadas así como maduras de miR-21-5p y miR-21-3p. Con la ayuda de la base de datos de acceso público TACCO la cual tiene depositada alrededor de 1085 pacientes de cáncer de mama de

ascendencia caucásica, se realizó la búsqueda de expresión de los transcritos maduros de miR-21-5p y miR-21-3p, en donde se muestra la tasa de cambio (*Fold Change*) del nivel de expresión al comparar una cohorte de tejidos normales y tumorales (Figura 8).

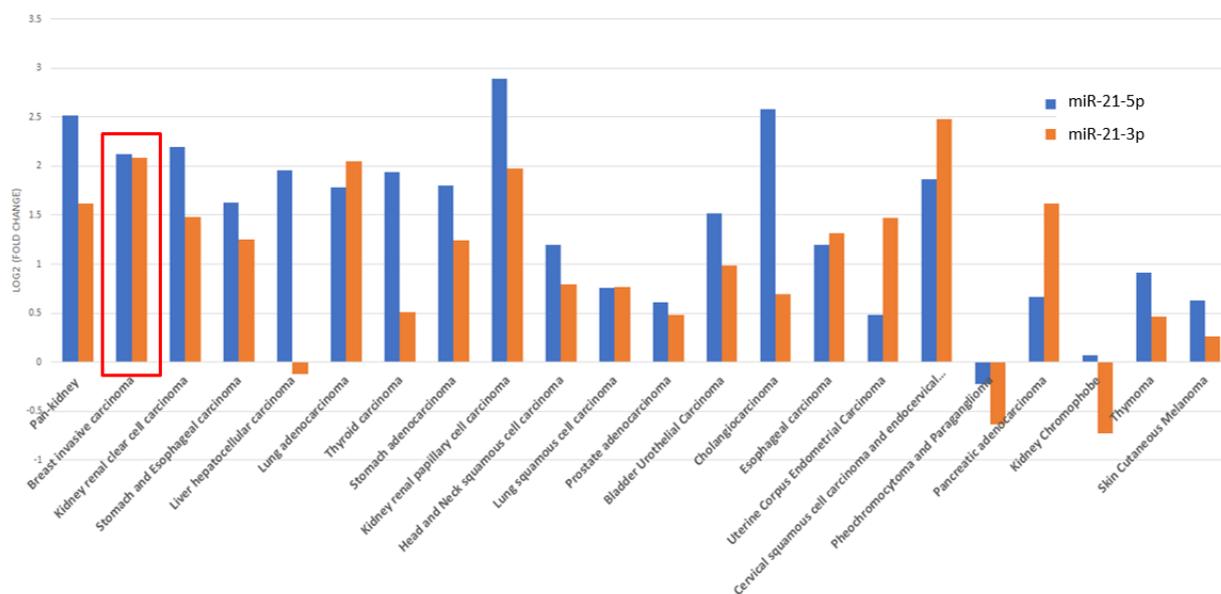


Figura 8. Expresión diferencial de las hebras maduras de miR-21-5p y miR-21-3p en distintos tipos de cáncer en la base de acceso público TACCO. Gráfico de barras donde se muestra el análisis de expresión diferencial de miR-21-5p (gráficas azules) y miR-21-3p (gráficas naranjas) para diferentes tipos de cáncer como cáncer de mama, cáncer de pulmón, cáncer de próstata (los datos fueron consultados de TACCO). Se observó que existe una sobreexpresión en nuestro modelo de estudio cáncer de mama (marcado con recuadro rojo). En donde, se ha reportado que una sobre-expresión de miR-21 en cáncer de mama, se encuentra asociado a proliferación y metástasis, resistencia a tratamientos quimioterapéuticos, así como la etapa clínica y el pronóstico(Gong et al., 2011; Najjary et al., 2020; Yan et al., 2008; Zhao et al., 2015).

Los resultados muestran un fenómeno en donde una de las hebras presenta un patrón de degradación evidente al comparar su expresión con la otra hebra 5p, esto se observa en 20 de los 22 tipos de cáncer, apegándose al modelo canónico de selección de hebras, en donde una hebra es generalmente degradada (se expresa

menos). Por otro lado, en cáncer de mama (recuadro rojo) y en cáncer de células escamosas de pulmón, se muestra un nivel de expresión similar entre sus dos hebras, presentando un fenómeno de selección de hebras no canónico. Por lo tanto, los resultados sugieren que las hebras miR-21-5p y miR-21-3p se encuentran sobreexpresadas en pacientes caucásicas con cáncer de mama y presentan un incremento (*Fold Change*) muy similar entre ambas hebras, fenómeno que es único entre las neoplasias comparadas, presentando un incremento de miR-21-5p y miR-21-3p de hasta seis veces cuando se compara tejido tumoral contra tejido normal y, por lo tanto, ambas hebras se expresan con un comportamiento similar. Comprobando que, en cáncer de mama, existe una posibilidad de que no únicamente miR-21-5p, si no también miR-21-3p se exprese de tal manera que pueda ser capaz de regular genes asociados a vías metabólicas oncogénicas.

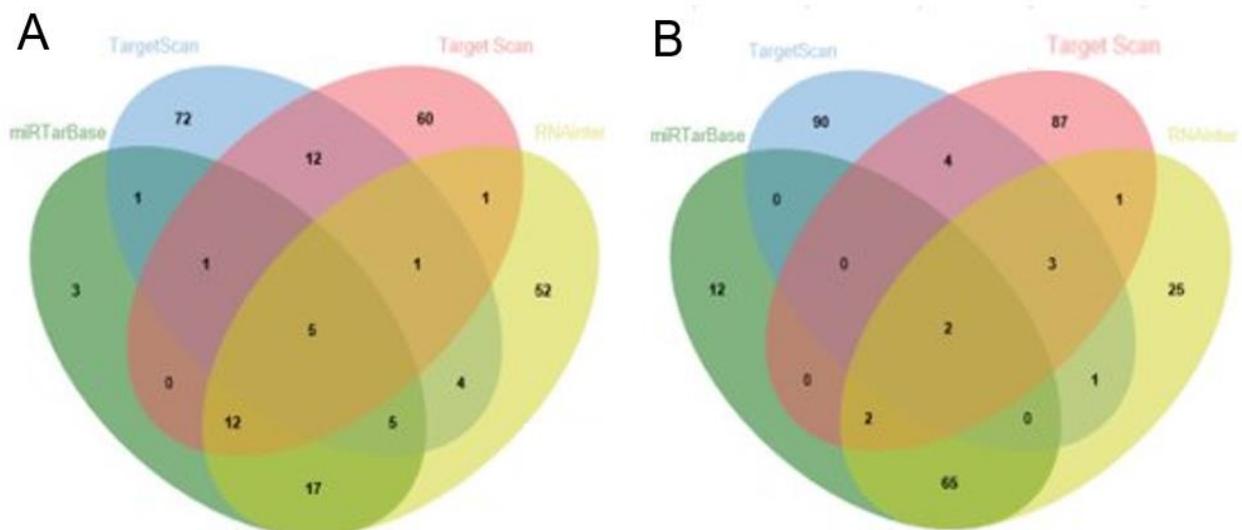
8.1.2 Determinación de los mRNAs blancos para miR-21-5p y miR-21-3p

Los miRNAs son pequeños RNAs no codificantes que regulan la expresión génica mediante complementariedad de bases con sus mRNAs blanco, llevando a estos mRNAs a vías de degradación y/o arresto de la traducción a través del reconocimiento por complementariedad de bases del extremo 3' UTR del mRNA. Al observar que existe una alteración en la expresión de miR-21-5p y miR-21-3p, se hipotetiza que los mRNAs blanco que sean complementarios a ambas hebras de miR-21 reconocerán estos sitios en su extremo 3'UTR. Por lo anterior, se realizó la búsqueda los mRNAs blanco para cada hebra por medio de distintos servidores web de acceso público, los cuales, mediante algoritmos establecidos, predicen blancos biológicos de miRNAs. Una vez analizado y consultado las cuatro bases de datos de acceso público se elaboró un Diagrama de Venn con los mRNAs que tuvieran los mejores valores de interacción con las hebras maduras miR-21-5p y miR-21-3p.

El diagrama de Venn mostró una intersección de las cuatro bases de datos en donde se mostró 5 mRNAs blanco que tenían la capacidad de interactuar por complementariedad con miR-21-5p, los genes encontrados fueron *PTEN*, *PDCD4*, *TIMP3*, *RECK* y *MSH2* (Figura 9A). Además, se hicieron los mismos análisis con la hebra de miR-21-3p y se encontró a dos genes candidatos dentro de la intersección de los 4 servidores, los cuales

se predice interactúan por complementariedad con miR-21-3p, *MAT2A* y *MAT2B* (Figura 9B).

Por otro lado, mediante la herramienta *RNAinteractome*, se obtuvo el mapa de interacción, el cual nos permite visualizar la red de interacciones de miR-21-5p y miR-21-3p con los 100 genes blanco, tomando en cuenta el puntaje dado por la misma herramienta, en donde se presentan tanto interacciones RNA-RNA como RNA-proteína, RNA-compuestos (tratamientos con fármacos) así como factores de transcripción. En la cual fueron señalados los genes candidatos de cada miRNA, los cuales coincidían con los genes candidatos encontrados en la intersección del diagrama de Venn, corroborando que *PTEN* y *PDCD4* interactúan con miR-21-5p (Figura 9C), así como *MAT2A* y *MAT2B* se encuentran dentro de los 100 genes blanco con miR-21-3p, con mejor puntaje (Figura 9D). Por lo tanto, se seleccionaron a los genes *PTEN*, *PDCD4*, *MAT2A* y *MAT2B* como mRNAs candidatos, los cuales pueden estar asociados a procesos oncogénicos en cáncer de mama, siendo *PTEN* y *PDCD4* blancos para miR-21-5p y *MAT2A* y *MAT2B* blancos para miR-21-3p.



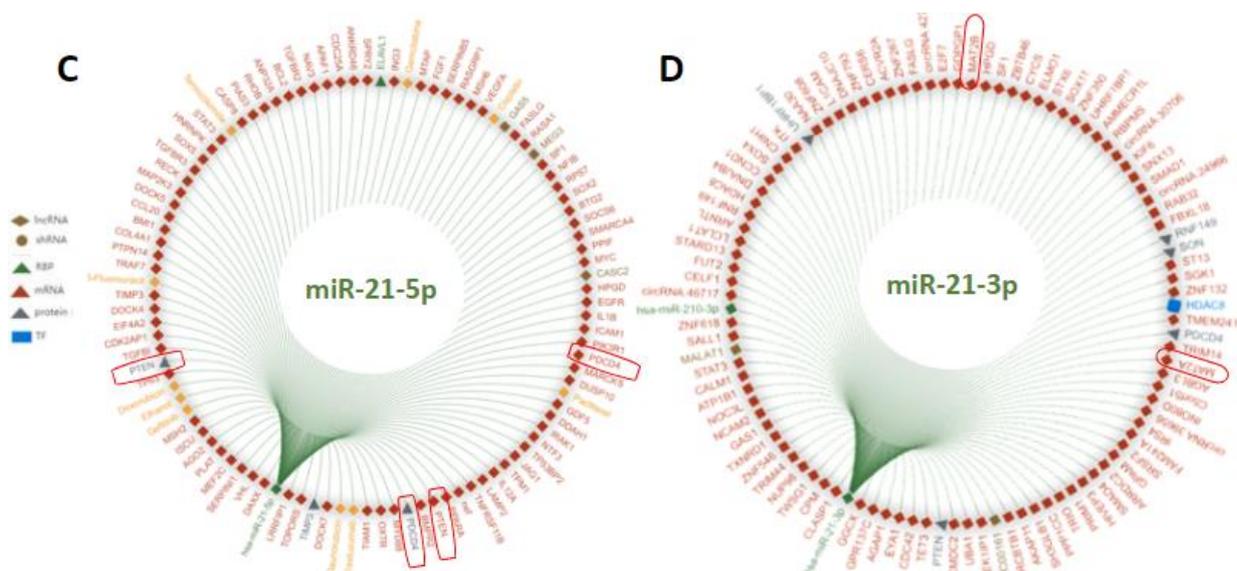


Figura 9. Determinación de blancos génicos mediante algoritmos de predicción de interacciones entre miRNA-mRNAs. Diagrama comparativo de Venn en el cual se muestra en la intersección, los genes que coinciden entre los cuatro servidores web de predicción utilizados. A) En el diagrama de Venn realizado para miR-21-5p, se observan cinco blancos génicos obtenidos en la intersección: PTEN, PDCD4, RECK, MSH2 y TIMP3. Los cuales, son blancos génicos que ya han sido reportados como mRNAs blanco directos mediante la unión de miR-21-5p a sus extremos 3'UTR. B) En el diagrama de Venn realizado para miR-21-3p, se obtuvieron dos únicos blancos génicos en la intersección: MAT2A y MAT2B. Ambos mRNAs fueron reportados anteriormente como mRNAs blanco de miR-21-3p en cáncer de hígado. C) Mapa de interacción entre miR-21-5p y sus 100 blancos con mayor valor de predicción, en donde se presentan interacciones RNA-RNA, RNA-proteínas, RNA-compuestos. En el cual se observa una interacción entre miR-21-5p con *PTEN* y *PDCD4*, a nivel RNA-RNA y RNA-proteína. D) Mapa de interacción entre miR-21-3p y sus 100 blancos con mayor valor de predicción. En el cual se observa una interacción entre miR-21-3p con los genes encontrados en la intersección, *MAT2A* y *MAT2B*, únicamente a nivel RNA-RNA. *RNAInteracome* es una herramienta de predicción web, la cual integra datos de interactomas de RNA validados experimentalmente y predichos computacionalmente de la literatura, además de información de blancos como lo son los sitios de modificaciones de RNA, la localización de RNA, la estructura de interacción miRNA-mRNA mediante las secuencias de transcritos y proteínas, sitios de unión experimentalmente validados, así como el grupo de genes ontólogos y parálogos.

Por último, mediante la herramienta *Diana MicroT-CDS* se realizó un análisis de los posibles sitios de unión miRNA-mRNA de cada gen candidato. En donde se obtuvo que miR-21-5p reconoce sitios de unión al extremo 3'UTR de los mRNAs *PTEN* y *PDCD4*

(Figura 10A). Asimismo, miR-21-3p reconoce sitios de unión al extremo 3'UTR de los mRNAs *MAT2A* y *MAT2B* (Figura 10B), en donde esta unión es mediada por una complementariedad de bases entre el miRNA y el RNA mensajero, mediante la región semilla; observando que miR-21-3p tiene la capacidad de unión por complementariedad mediante 2 sitios de unión a *MAT2A* y 1 sitio de unión a *MAT2B*.

Los resultados establecieron que tanto miR-21-5p como miR-21-3p contienen sitios de unión al extremo 3'UTR de los mRNA blanco encontrados. Los análisis realizados *in silico* sugieren que los mRNAs blanco encontrados para miR-21-5p y miR-21-3p, corresponden a genes candidatos que podrían ser regulados por ambas hebras sugiriendo tener un efecto biológico en la regulación miRNA-mRNA de cada uno de ellos.

A

Gen	Clave ENSEMBL	miTG score	Binding sites number	Region	Binding Type	score
PDCD4	ENSG00000150593	0.9996	2	UTR3	8mer	0.13
				UTR3	6mer	0.01
PTEN	ENSG000000171862	0.405	2	UTR3	6mer	0.02
				UTR3	7mer	0.02

PTEN



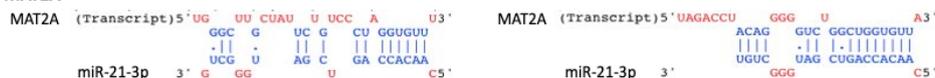
PDCD4



B

Gen	Clave ENSEMBL	miTG score	Binding sites number	Region	Binding Type	score
MAT2A	ENSG00000168906	0.892	2	UTR3	6mer	0.010
				UTR3	8mer	0.035
MAT2B	ENSG0000003827	0.672	1	UTR3	8mer	0.022

MAT2A



MAT2B



Figura 10. Sitios de unión predichos por análisis in silico entre miRNA y sus mRNAs blanco. A) Sitios establecidos como regiones de unión de miR-21-5p con los extremos 3'UTR de los genes *PTEN* y *PDCD*, mostrando dos uniones canónicas entre la región semilla de miR-21-5p y los mRNAs blanco, sitios que habían sido reportados con anterioridad como sitios de unión. B) Sitios establecidos como regiones de unión putativos de miR-21-3p con los extremos 3'UTR de los genes *MAT2A* y *MATB*, en donde se observan uniones canónicas entre la región semilla de miR-21-3p y los mRNAs blanco propuestos. El algoritmo microT utiliza elementos de reconocimiento de miRNA (MRE) para el emparejamiento de bases RNA-RNA (reconocimiento de RNA-MRE). Los MRE se buscan en los extremos 3' UTR de mRNAs humanos extraídos de la base de datos RefSeq. En donde, se incluyen dos parámetros en el algoritmo de búsqueda, el primero es el cálculo de la energía libre del emparejamiento canónico Watson-Crick y, por otro lado, los pares de bases de dinucleótidos oscilantes G-U para la identificación putativa de MRE¹⁰⁵.

Al confirmar que tanto *MAT2A*, como *MAT2B* son genes candidatos a regulación por medio de complementariedad en sus extremos 3'UTR por medio de miR-21-3p. En este sentido, se realizó una búsqueda de los niveles de expresión de *PTEN*, *PDCD4*, *MAT2A* y *MAT2B*, comparando tejido normal contra tejido tumoral, con la ayuda de la plataforma de acceso público TACCO, utilizando los datos de RNA-Seq de una cohorte de pacientes de cáncer de mama de 1085 muestras neoplásicas. Adicionalmente, con el objetivo de comprobar si la expresión de los mRNAs, *PTEN*, *PDCD4*, *MAT2A* y *MAT2B* se ve afectada por la sobreexpresión de miR-21-5p y de miR-21-3p en líneas celulares de cáncer de mama. Se realizó un análisis de expresión diferencial a partir de datos de RNA-Seq de las cuatro líneas celulares de CaMa a estudiar, siguiendo las características mencionadas en el apartado 7.1.4 (metodología).

En particular, en la cohorte de pacientes con cáncer de mama, en la cual previamente se había observado una sobreexpresión de miR-21-3p, el mRNA de *MAT2A* se encuentra subexpresado en tejido tumoral (Figura 11A), es decir, cuando miR-21-3p se encuentra sobreexpresado, el mRNA de *MAT2A* se encuentra subexpresado, sugiriendo una regulación negativa mediada por miR-21-3p. Por el contrario, en tejido tumoral, *MAT2B* muestra una expresión basal cuando se compara con los tejidos normales (Figura 11B). Asimismo, la expresión del mRNA, *PTEN*, en tejido tumoral, muestra un patrón de

regulación negativa como el observado en *MAT2A*, en el cual, miR-21-5p se encuentra sobreexpresado (anteriormente observada), y se observa una subexpresión de *PTEN* en esta cohorte de tejidos tumorales (Figura 11C), y, por otra parte, la expresión de *PDCD4*, se observa como una expresión basal cuando se compara con tejido normal.

Por lo anterior, con lo obtenido en los datos de expresión de RNA-Seq, mediante la herramienta TACCO, se observa un patrón de regulación concordante entre la expresión de miR-21-3p y *MAT2A*, en la cual miR-21-3p regula negativamente en tejido tumoral al gen *MAT2A*, fenómeno observado también en miR-21-5p y su gen blanco, *PTEN*, regulación que ha sido validada experimentalmente con anterioridad, por lo tanto al ver este mismo patrón de regulación, en miR-21-3p y *MAT2A*, en CaMa, se sugiere que, en efecto, *MAT2A* es un posible blanco de miR-21-3p en cáncer de mama. Por otro lado, *MAT2B* muestra una expresión basal, al igual que *PDCD4*. No obstante, aún sin observar una sobreexpresión en esta cohorte, no se puede descartar a *MAT2B* como posible blanco génico de miR-21-3p, esto debido a que *PDCD4* muestra el mismo patrón de expresión, y de igual manera, *PDCD4* ha sido validado experimentalmente como blanco génico de miR-21-5p en cáncer de mama, al igual de evidencia en la que se reporta que no únicamente los la expresión de los mRNAs blanco es afectado, también los niveles de proteína acumulados son regulados, por lo que aún sin poder observar un efecto directo en *MAT2B*, la regulación de miR-21-5p, puede presentarse como una regulación miRNA-proteína. Estos datos sugieren que, tanto *MAT2A* como *MAT2B* son posibles genes blanco de miR-21-3p en cáncer de mama.

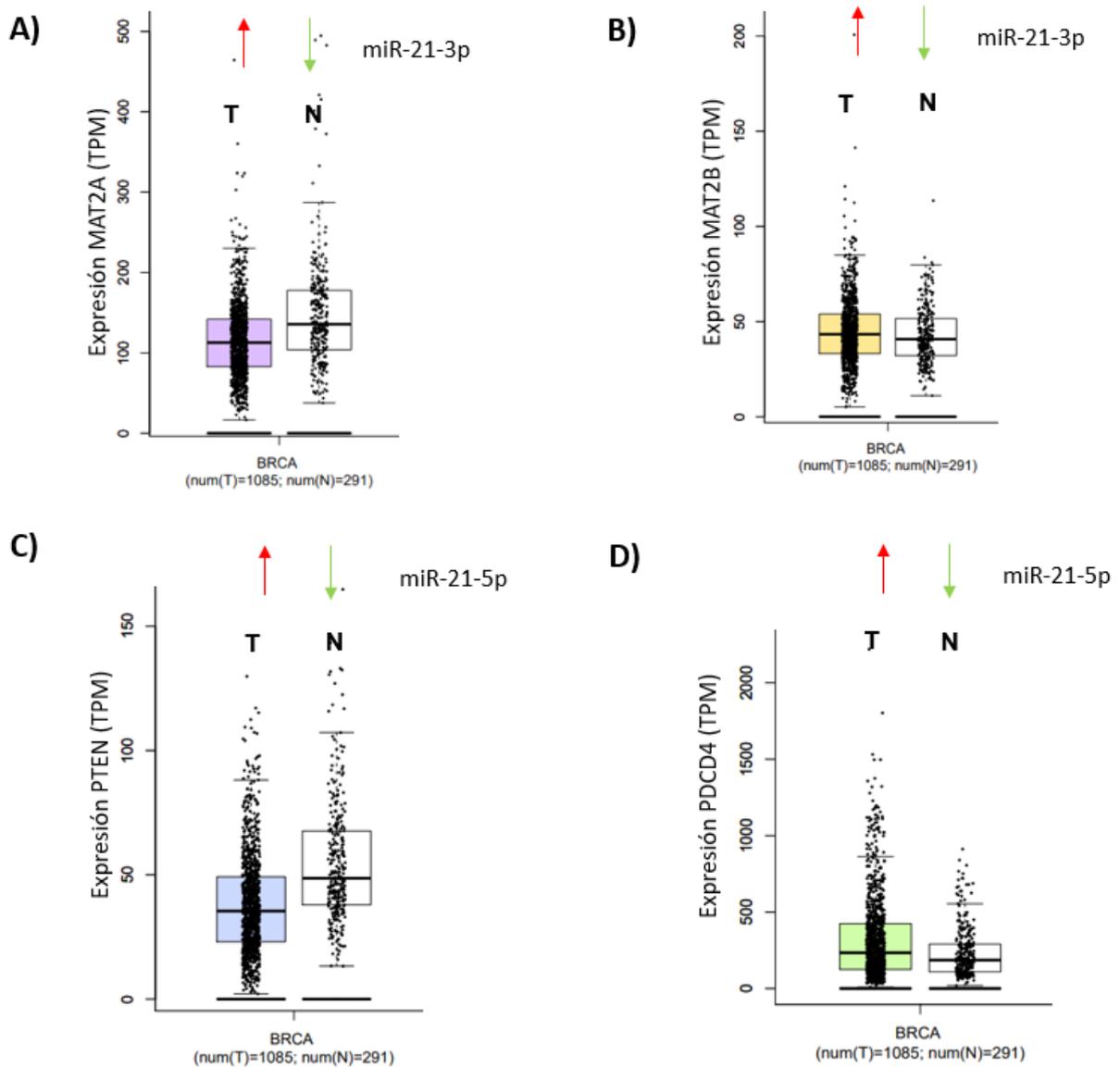


Figura 11. Análisis in silico de expresión miRNA-mRNA blanco en una cohorte de tejidos de cáncer de mama. A) Análisis de expresión de *MAT2A* en cáncer de mama mediante datos de expresión obtenidos por RNAsec utilizando la herramienta TACCO, en donde se graficó mediante *boxplots*, el nivel de expresión dado en TPMs en una cohorte de tejido tumoral, comparado con tejido normal. B) Análisis de expresión de *MAT2B*, C) Análisis de expresión de *PTEN* y D) Análisis de expresión de *PDCD4* (*Boxplots* fueron obtenidos mediante la misma herramienta de TACCO, cohorte de tejido normal (N)= 291; cohorte de tejido tumoral (T)=1085)

Además de analizar la expresión de *MAT2A*, *MAT2B*, *PTEN* y *PDCD4* en la cohorte de pacientes de cáncer de mama, se realizó el análisis de expresión en las líneas celulares neoplásicas MCF-7 (subtipo molecular Luminal A), BT474 (subtipo molecular Luminal B) y MDA-MB-231 (subtipo molecular Triple Negativo). Se realizaron mediante la herramienta *ggplot* de *RStudio* diagramas de cajas para comparar el nivel de expresión de cada subtipo molecular contra la línea celular normal, MCF-10A. En las cuales, al encontrarse miR-21-3p sobreexpresado, *MAT2A* y *MAT2B* muestran un patrón de expresión similar, presentando una expresión basal en la línea celular neoplásica MCF-7, seguido por un aumento en su expresión en BT474 y de igual manera otro aumento en la expresión en MDA-MB-231 (Figura 12A y 12B). Por el contrario, al sobreexpresarse miR-21-5p, *PTEN* muestra un aumento de la expresión en las líneas celulares MCF-7 y BT474 cuando se compara con MCF-10A, y MDA-MB-231 presenta un nivel de expresión de *PTEN* basal (Figura 12C). Finalmente, *PDCD4* presenta niveles basales de expresión en MCF-7 y MDA-MB-231, a diferencia de la línea celular BT474, en la cual se observa un aumento en la expresión cuando se compara con la línea celular normal, MCF-10A (Figura 12D).

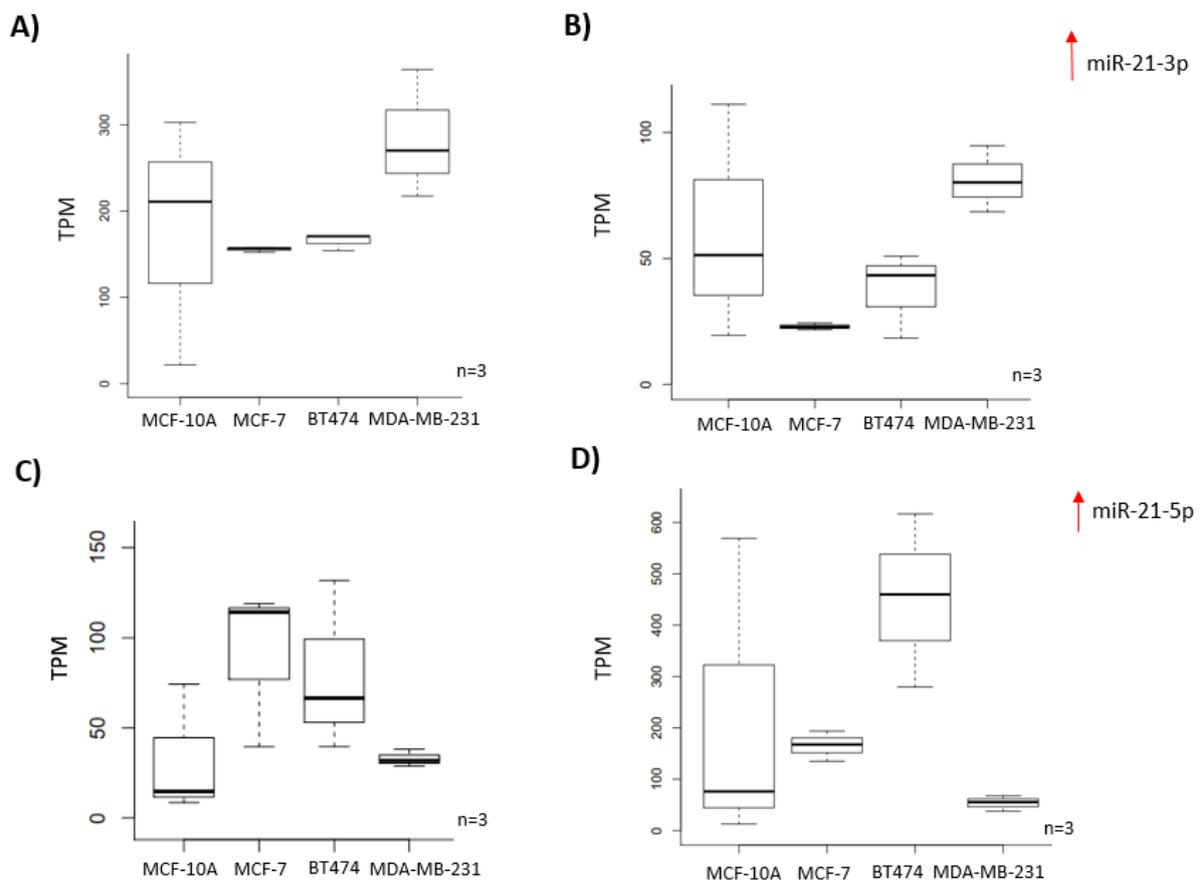


Figura 12. Análisis *in silico* de expresión de *MAT2A*, *MAT2B*, *PTEN* y *PDCD4* en líneas celulares neoplásicas de diferentes subtipos moleculares.

A) Análisis de expresión de *MAT2A* en cuatro diferentes líneas celulares de cáncer de mama y una línea celular normal, utilizando datos de expresión obtenidos por RNAseq, en donde se graficó mediante *boxplots*, el nivel de expresión dado en TPMs. B) Análisis de expresión de *MAT2B*, C) Análisis de expresión de *PTEN* y D) Análisis de expresión de *PDCD4* (MCF10-A (línea celular normal), MCF-7 (línea celular subtipo Luminal A), BT474 (línea celular Luminal B) y MDA-MB-231 (línea celular Triple Negativo); *Boxplots* fueron realizados con el paquete *ggplot RStudio*; n=3; se utilizó un análisis ANOVA con prueba de Tukey para los genes *PTEN*, *MAT2A* y *MAT2B*; para el gen *PDCD4* se utilizó un análisis Kruskal-Wallis y Wilcoxon, en todas las pruebas estadísticas se utilizaron valores de confianza de 95%.

8.1.3 Determinación de la asociación de *PTEN*, *PDCD4*, *MAT2A* y *MAT2B* en vías oncogénicas.

Las evidencias experimentales han sugerido que los genes que corresponden *PTEN*, *PDCD4*, *MAT2A* y *MAT2B* son regulados por miR-21-5p y miR-21-3p, respectivamente. Con el objetivo de establecer qué moléculas interaccionan con miR-21-5p y miR-21-3p se utilizó la plataforma de RAIN (*RAIN: RNA-protein Association and Interaction Networks*), la cual la cual facilita la comprensión de redes moleculares mediante la integración de varias interacciones de RNA y redes de asociación proteína-proteína. Esta herramienta a diferencia del mapa de interacción presentado que fue realizado con *RNAinteractome*, resulta ser una herramienta con mayores ventajas como lo son, la visualización de la red de interacción entre el miRNA, el gen propuesto, y los genes y/o proteínas que se encuentren interaccionando con estos, con el objetivo de poder utilizar la herramienta de RAIN para poder asociar estas redes, a procesos o vías metabólicas de nuestra importancia, las cuales mediante la misma herramienta de RAIN, son obtenidas de GSEA (*Gene Set Enrichment Analysis*), KEGG (*Kyoto Encyclopedia of Genes and Genomes*). Por lo anterior, se utilizó la herramienta de RAIN, para obtener un mapa de interacción, utilizando 5 *blancos* codificantes y 5 *blancos* no codificantes, en donde estos 10 *blancos*, resultan ser los de mayor puntaje de predicción, con el fin de reducir nuestro universo de blancos génicos.

Se observó que, *PTEN* y *PDCD4* se encuentran dentro de la red de interacción dada por el programa RAIN, en donde ambos mRNAs tienen una interacción RNA-RNA con miR-

21-5p (Figura 13A y 13B). Asimismo, se encontró, que dentro de las redes de interacción de *MAT2A* y *MAT2B*, ambos mRNAs interaccionaban entre ellos, así como con miR-21-3p, mediante interacciones RNA-RNA. (Figura 13C y 13D). Por lo tanto, estos resultados validan mediante otra plataforma de análisis *in silico*, que los genes *MAT2A* y *MAT2B*, son mRNAs blancos de miR-21-3p, los cuales se encuentran dentro de los 10 blancos de predicción principales para mir-21-3p, y por lo tanto, su expresión y/o genes así como sus vías metabólicas asociados a *MAT2A* y *MAT2B* sugieren que estas pueden encontrarse reguladas por miR-21-3p.

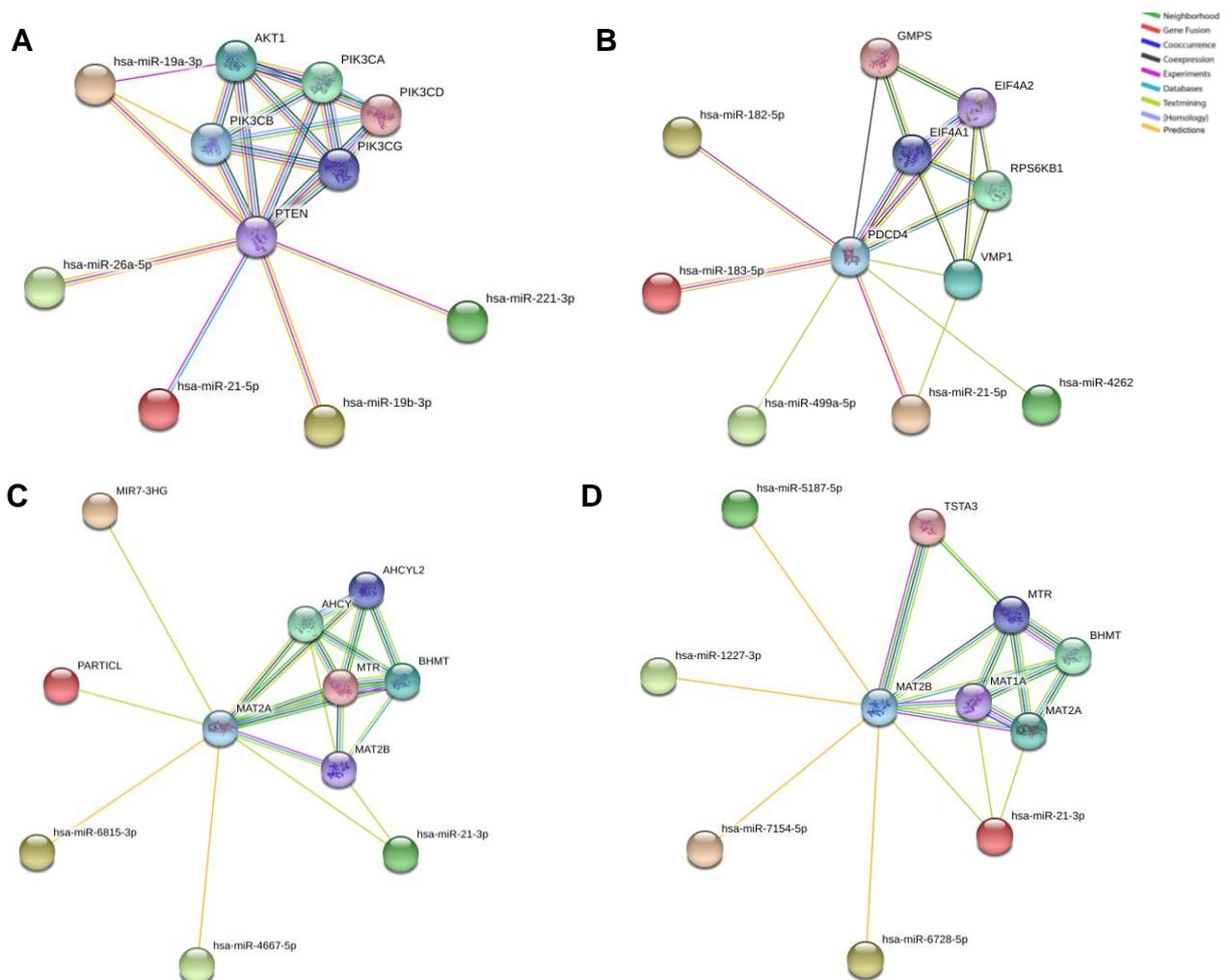


Figura 13. Redes moleculares de asociación e interacción de miR-21-5p y miR-21-3p. Red de interacción, obtenida mediante la herramienta RAIN, de las 5 moléculas codificantes y 5 moléculas no codificantes con mayor puntuación de predicción. **(A)** Red molecular de interacción entre *PTEN* y sus 10 *blancos*, dentro de los cuales se encuentra *AKT1*, *PIK3CA/B/D/G* y miR-21-

5p. **(B)** Red molecular de interacción entre *PDCD4* y los *blancos* principales, como lo son *EIF4A1/2* y miR-21-5p. **(C)** Red molecular de interacción entre *MAT2A* y sus 10 *blancos* de mayor predicción, en donde se observó una interacción entre *MAT2A* con *MAT2B* y miR-21-3p. **(D)** Asimismo, la red de interacciones de *MAT2B* mostró el mismo patrón de interacción, en donde se encontró a *MAT2A* y miR-21-3p como *blancos*. Se representa el tipo de interacción mediante un código de colores: coexpresión (negro), evidencia experimental (rosado), base de datos (azul), literatura (verde), predicción (naranja).

Al encontrar mediante diferentes herramientas que los genes *MAT2A* y *MAT2B* son blancos génicos regulados por miR-21-3p. Mediante recopilación de información obtenida de RAIN (interacciones génicas que se han visto asociados por GSEA y por Ontología de Genes, las cuales forman parte de una vía intracelular metabólica asociada a cáncer), se realizó una búsqueda en la literatura de evidencia experimental que validara a estos mRNAs como blancos génicos de miR-21-3p, así como de vías metabólicas que estuvieran reguladas por los genes *MAT2A* y *MAT2B* reportadas en cáncer. Con la información recopilada, se realizó una tabla (Tabla 2). Se obtuvo que el gen *PTEN*, ya conocido como un gen supresor de tumor, está asociado a vías oncogénicas en CaMa como la vía PI3K-AKT-mTOR y a la vía de FOXO3; el gen *PDCD4*, esta reportado como un gen supresor de tumor está involucrado en las vías JKN y MEK/ERK. Por otro lado, los genes *MAT2A* y *MAT2B* podrían considerarse como oncogenes, los cuales en vías oncogénicas están asociados a vías como RAS-RAF-MEK-ERK y/o metilación aberrante de regiones promotoras que corresponden a genes supresores de tumor.

miRNA	Símbolo	Función	Vías oncogénicas asociadas
miR-21-5p	PTEN	Gen supresor de tumor	Inhíbe la proliferación, la migración, la invasión y la supervivencia celular, mediante las vías PI3K-AKT-mTOR y FOXO3
	PDCD4	Gen supresor de tumor	Inhíbe el crecimiento celular, invasión tumoral, metástasis e inducción de la apoptosis mediante las vías: JKN y (MEK)/ERK.
miR-21-3p	MAT2A	Catalizan la síntesis de S-adenosilmetionina (SAM), el principal donante biológico de metilo. Las dos proteínas interactúan y se estabilizan entre sí.	Su sobre-expresión conduce al crecimiento, migración e invasión en cánceres de hígado, colon, cáncer gástrico humano y en células de cáncer de mama resistentes al tamoxifeno, mediante vías como RAS-RAF-MEK-ERK y/o metilación aberrante de regiones promotoras de oncogenes.
	MAT2B		

Tabla 2. Vías metabólicas oncogénicas asociadas a PTEN, PDCD4, MAT2A y MATB.

Resumen de las vías metabólicas que se encuentran reportadas en la literatura. Mediante búsqueda en la literatura, miR-21-5p regula a *PTEN* y *PDCD4*, genes involucrados en procesos celulares oncogénicos como la proliferación, la migración y la invasión, mediante su papel en vías como PI3K-AKT-mTOR y FOXO3. Asimismo, se proponen como genes blanco de miR-21-3p a *MAT2A* y *MAT2B*, los cuales han sido reportados como genes asociados a vías como RAS-RAF-MEK-ERK, además de su papel en la metilación, en donde se encuentran asociados en cáncer a procesos oncogénicos como el crecimiento celular, la migración y la invasión.

En resumen, mediante el análisis realizado a partir de los datos de expresión obtenidos mediante los datos de RNA-Seq y la plataforma de acceso público de TACCO para las diferentes líneas celulares de CaMa y tejidos neoplásicos de cáncer de mama, se mostró que el pri-miRNA-21 se sobreexpresa en la línea celular neoplásica MCF-7 en comparación con la línea celular de tejido mamario normal MCF-10A. De igual manera, la sobreexpresión de miR-21-5p y miR-21-3p se encuentran sobreexpresados en CaMa. Se encontró que ambas hebras de miR-21 se sobreexpresan en tejido tumoral en comparación con el tejido normal, en donde la hebra de miR-21-5p presenta una tasa de cambio (*Fold Change*) en una frecuencia 100 veces mayor en tejido normal. Por otro lado, al observar que la expresión de ambas hebras de miR-21 se encuentran sobreexpresadas, se sugiere que los

miRNAs regulan a los mRNAs analizados. Se propusieron como mRNAs blancos para estudiar en este proyecto a *PTEN* y *PDCD4* para miR-21-5p y a los mRNAs de *MAT2A* y *MAT2B* para miR-21-3p. Además, se estudió la expresión diferencial de cada gen en dos condiciones que incluye el análisis de comparar tejido normal contra tejido neoplásico, así como utilizar una línea celular epitelial de mama que asemeje un contexto normal y que sirva como línea celular control en los análisis *in silico*. En general, los resultados mostraron que, así como *PTEN* y *PDCD4* son genes blanco, regulados por miR-21-5p que han sido descritos en cáncer de mama, *MAT2A* y *MAT2B* son dos blancos génicos potenciales a ser regulados por miR-21-3p, ya sea a nivel miRNA-mRNA y/o miRNA-Proteína. Asimismo, mediante el análisis de interacción con la plataforma de acceso público de RAIN, se obtuvieron las interacciones con genes que se han visto asociados a vías intracelulares metabólicas con un papel en cáncer.

8.2 Validación experimental

8.2.1 Calidad e integridad de RNA total de líneas celulares de cáncer de mama.

Los resultados de los análisis de expresión de pri-miRNA-21, miR-21-5p y miR-3p a partir de datos de RNA-Seq provenientes de líneas celulares de cáncer de mama confirmaron que el pri-miRNA-21 y sus formas maduras se sobreexpresan en CaMa, específicamente en la línea celular neoplásica MCF-7. Con el objetivo de corroborar de manera experimental y validar la sobreexpresión de pri-miRNA-21, miR-21-5p y miR-3p, se llevó a cabo el método de cuantificación relativa de expresión de los tres RNAs por RT-qPCR en la línea celular MCF-7. Además, se incluyeron en los análisis las líneas celulares MCF-10A, BT474 y MDA-MB-231. Para esto, se extrajo RNA total de cada línea celular de mama y se verificó que tuviera la calidad e integridad adecuada para ser utilizado como templado para la reacción RT y así poder asegurar que se obtuvo un cDNA de calidad suficiente para su uso en PCR en tiempo real.

Con el objetivo de evaluar la calidad e integridad de los RNA de las líneas celulares, se analizó el RNA total mediante un bionalizador de lo cual se obtuvieron valores

asociados de RIN mayores a 9.5, de cada una de las líneas celulares de mama, por lo que se considera que el RNA total extraído tiene la calidad e integridad suficiente para poder ser utilizado en ensayos de biología molecular por RT-qPCR (Figura 14). Por otro lado, mediante gel de agarosa al 1% se visualizó el RNAs total que se extrajo de todas las réplicas biológicas obtenidas de cada línea celular de mama (Anexo, Figura1). Como puede observarse, las bandas que corresponden al RNA ribosomal 28S y 18S, mostraron una relación 2:1, es decir, la banda que corresponde al 28S es dos veces más intensa en señal que la banda del RNA 18S, esto se vio en tres réplicas biológicas de cada línea celular analizada, lo cual se considera como que el RNA total extraído se encuentra íntegro y no muestra estar degradado. Por lo anterior, los resultados muestran que el RNA obtenido tiene la calidad e integridad apropiada y cumple con los criterios necesarios para realizar PCR en tiempo real.

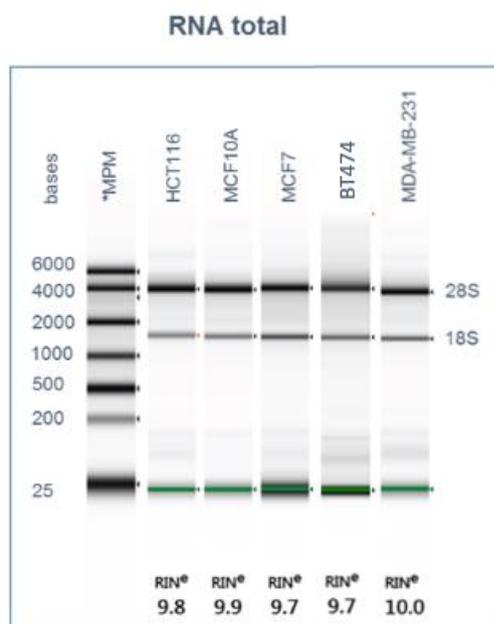


Figura 14. Determinación de la Calidad e Integridad de RNA total obtenido de líneas celulares de mama. La calidad del RNA total fue evaluada en bionalizador. La obtención de RNA total se realizó mediante el método de Trizol. Gel de electroforesis capilar que muestra la integridad del RNA (RIN: valor de integridad del RNA). El RIN establece la calidad e integridad de los ácidos nucleicos. La escala va del 1 al 10, donde 1 indica degradación y 10 indica alta calidad e integridad. En donde se observa la integridad de los RNA ribosomales 28s y 18s del RNA extraído de las líneas celulares MCF-10A, MCF7, BT474 y MDA-MB-231 respectivamente, en donde se incluyó un control de RNA alta calidad (HCT116).

Una vez comprobada la calidad e integridad del RNA total extraído de las líneas celulares de mama, se sintetizó el cDNA mediante RT, utilizando el RNA total extraído de cada réplica biológica y que corresponde a cada una de las líneas celulares MCF-10A, MCF7, BT474 y MDA-MB-231. Con el fin de verificar la calidad del cDNA sintetizado, se llevó a cabo una reacción de PCR punto final para amplificar el gen constitutivo *RPS28*.

Los resultados de PCR punto final para las réplicas biológicas de las cuatro líneas celulares de mama analizadas cumplieron con los criterios de calidad, ya que en todas las muestras se observa la presencia del amplicón del gen *RPS28* que corresponde a una banda de 100 pb (Anexo, Figura 2). En conclusión, el cDNA sintetizado de todas las muestras extraídas de RNA total cuentan con la calidad apropiada para el análisis de expresión por cuantificación de expresión relativa mediante PCR.

8.2.2 Determinación de la expresión de los RNA maduros de miR-21-5p, miR-21-3p, mediante PCR en tiempo real con sondas de hibridación en líneas celulares de CaMa.

Una vez determinada la calidad del cDNA se procedió a cuantificar los niveles de expresión relativa tanto del transcrito primario, así como de los transcritos maduros de ambas hebras de miR-21 por PCR tiempo real a través del uso del método de análisis de expresión relativa $\Delta\Delta Cq$. Lo anterior, fue con la finalidad de comparar y validar los resultados de sobreexpresión obtenidos en los análisis bioinformáticos para las diferentes líneas celulares de CaMa.

Los resultados mostraron que, con respecto a los niveles de expresión de pri-miRNA-21 en MCF-10A la cual corresponde a la línea celular calibradora en los ensayos de $\Delta\Delta Cq$, es decir, la línea celular de mama que corresponde a un tejido epitelial normal y que es donde se encontraría la expresión de miR21 en un contexto normal, nosotros encontramos que la línea celular MCF-7 que corresponde al subtipo molecular Luminal A de cáncer de mama, presenta una sobreexpresión de pri-miRNA-21

(Figura 15). Se observó un incremento de expresión de más de 15 veces de pri-miRNA-21 en la línea neoplásica MCF-7 cuando se comparó la expresión con la línea celular calibradora, MCF-10A. Asimismo, en las líneas celulares BT474 y MDA-MB-231 se observó un aumento en la expresión de 5 veces los niveles del pri-miRNA-21, comparando con la expresión observada en MCF-10A. En cuanto lo observado, se validó experimentalmente la sobreexpresión observada de pri-miRNA-21 en los datos de RNA-Seq, previamente analizados por análisis bioinformáticos.

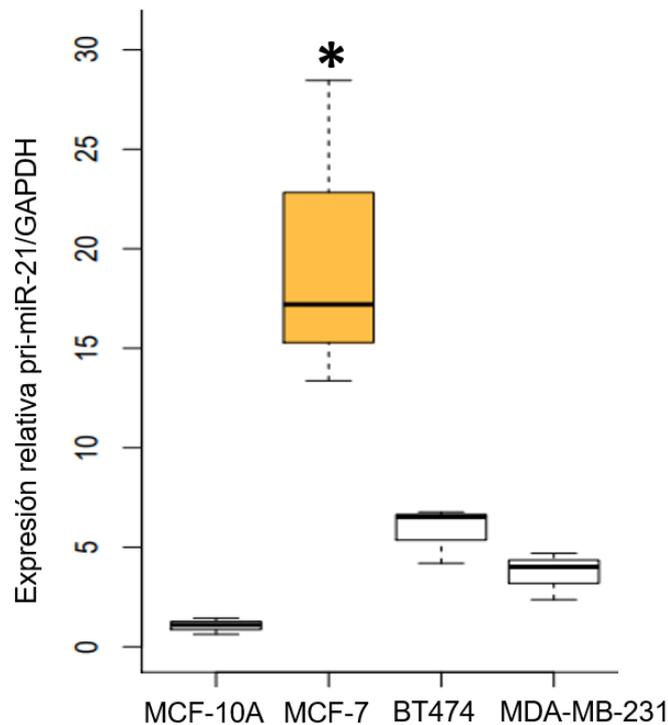


Figura 15. Expresión diferencial de pri-miRNA-21 en líneas celulares de Cáncer de mama. Cuantificación relativa de la expresión de pri-miRNA-21 mediante RT-qPCR, en tres líneas celulares neoplásicas y una línea celular no neoplásica. Los resultados de la cuantificación relativa de pri-miRNA-21, mostraron que, al comparar el nivel de expresión de las líneas celulares neoplásicas con la expresión de pri-miRNA-21 en la línea celular no neoplásica, MCF-10A, la línea celular neoplásica Luminal A, MCF7 sobreexpresa de manera significativa a pri-miRNA-21, mostrando una expresión 20 veces mayor que MCF10-A, a diferencia de las líneas celulares neoplásicas BT474, la cual representa al subtipo Luminal B y la línea celular neoplásica triple negativo, MDA-MB-231, las cuales muestran una expresión de 5 veces el pri-miRNA-21 comparado con MCF-10A. No expresan de manera significativa a pri-miRNA-21, con respecto al calibrador MCF-10A. Gen constitutivo utilizado, gen GAPDH. (n= 3, p=0.00005 ANOVA con prueba de Tukey).

Posteriormente, para realizar el análisis de expresión de los transcritos maduros de miR-21-5p y miR-21-3p, se realizó RT-qPCR con sondas de hibridación en las líneas neoplásicas de estudio, MCF-7, BT474, MDA-MB-231 comparando los valores de expresión contra la línea celular calibradora, MCF-10A, de la misma forma que con pri-miRNA-21, utilizando DDCT. En la línea celular MCF-7, se obtuvo una expresión de miR-21-5p de más de 5 veces la expresión obtenida para MCF-10A, a diferencia de la expresión de miR-21-5p en BT474 y MDA-MB-231, líneas celulares en las cuales esta hebra presenta valores basales de expresión (Figura 16A). Además, en MCF-7, se obtuvo una expresión de miR-21-3p que fue 3 veces mayor en comparación con la expresión de miR-21-3p en MCF-10A. Por otro lado, la línea celular neoplásica BT474 presenta niveles por debajo de los niveles de expresión de miR-21-3p en la línea calibradora MCF-10A, a diferencia de la línea MDA-MB-231, la cual presenta niveles basales de expresión de miR-21-3p. En resumen, el patrón de expresión de pri-miRNA-21, miR-21-5p y miR-21-3p coincidió con lo observado en los análisis *in silico*, validando los análisis realizados en datos públicos de RNA-Seq. Por lo anterior, la línea celular de cáncer de mama de subtipo molecular Luminal A, MCF-7, mostró sobreexpresión de pri-miRNA-21 y de ambas hebras maduras, miR-21-5p y miR-21-3p cuando éstas se comparan con una línea celular no neoplásica como es el caso de MCF-10A.

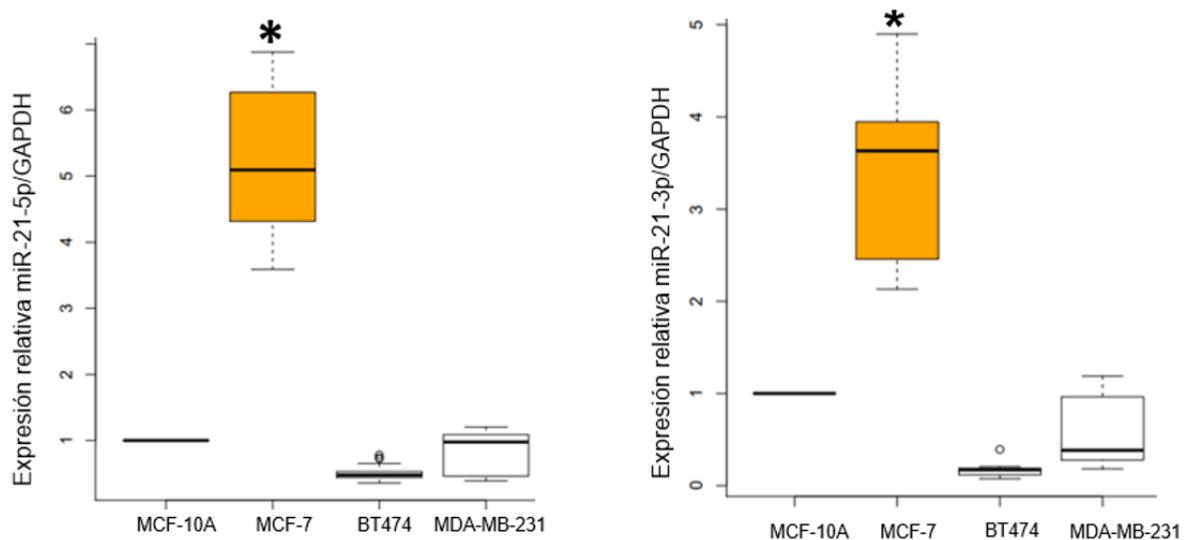


Figura 16. Expresión diferencial de la hebra guía y pasajera de miR-21 en líneas celulares de Cáncer de mama. A) Cuantificación relativa de miR-21-5p mediante RT-qPCR en tiempo real con sondas de hibridación, mostró que la línea celular neoplásica Luminal A

expresa 5 veces más a miR-21-5p, respecto a la línea celular calibradora, MCF-10A, y que las líneas celulares neoplásicas Luminal B, BT474 y Triple Negativo, MDA-MB-231 muestran una expresión basal de miR-21-5p con respecto a MCF-10A. B) De igual manera, en la cuantificación relativa de miR-21-3p mediante RT-qPCR en tiempo real con sondas de hibridación, se observó el mismo patrón de expresión, en donde la línea celular neoplásica Luminal A, sobreexpresa 4 veces más a miR-21-3p cuando se compara con MCF-10A, y las líneas celulares BT474 y MDA-MB-231 presentan expresión basal de miR-21-3p. Gen constitutivo, gen GAPDH. (n=6, p=0.00005 ANOVA con prueba de Tukey).

8.2.3 Determinación de la expresión diferencial de los mRNA de *PTEN*, *PDCD4*, *MAT2A* y *MAT2B* mediante RT-PCR en tiempo real en líneas celulares de Cáncer de Mama.

Los resultados obtenidos de la expresión diferencial de miR-21-5p y miR-21-3p en las líneas celulares de cáncer de mama, en conjunto con lo hallado en los análisis por RNA-Seq, nos permitió validar que ambas hebras se encuentran sobreexpresadas en CaMa, por lo tanto, se quiso determinar si los mRNAs blanco de cada hebra de miR-21 se expresan de manera diferencial, es decir, quisimos determinar si la sobreexpresión de miR-21-5p se asociaba directamente con la subexpresión del mRNA de PTEN para establecer si habría una correlación directa y, por lo tanto, saber si este gen pudiera ser blanco de miR-21-5p cuando este se sobreexpresa como ha sido reportado en la literatura (Shan et al., 2013). Por lo anterior, se llevó a cabo la determinación de la expresión diferencial de cada uno de los mRNA blanco por RT-PCR en tiempo real en la línea celular MCF-7; para esto, y como parte del proceso de validación, se procedió a evaluar la eficiencia de amplificación de los oligonucleótidos diseñados para cada gen blanco. Los resultados mostraron que el porcentaje de eficiencia de amplificación para los oligonucleótidos de *PTEN*, *PDCD4*, *MAT2A* y para *MAT2B* (Anexo, Tabla 1) se encuentran en el intervalo de 95% a 110% de eficiencia de amplificación (Anexo, Figura 3), por lo que se consideran oligonucleótidos funcionales en un procedimiento de RT-PCR en tiempo real para evaluar expresión génica.

Una vez comprobada la eficiencia de amplificación de los oligonucleótidos para los genes blanco de *PTEN*, *PDCD4*, *MAT2A* y *MAT2B*, se procedió a cuantificar los niveles de expresión relativa de estos mRNAs a través de PCR en tiempo real, con el uso del

método $\Delta\Delta Cq$, con el objetivo de validar los datos de expresión diferencial encontrados en los análisis del RNA-Seq para las diferentes líneas celulares de CaMa, utilizando como línea celular de referencia a MCF-10A y como gen constitutivo a *GAPDH* (Figura 17).

Se observó, en los resultados obtenidos mediante RT-qPCR que los mRNAs de *PTEN* y *PDCD4* son blancos génicos de miR-21-5p, debido a que presentan en las líneas celulares neoplásicas MCF-7 y MDA-MB-231 una expresión basal, concordando con lo encontrado en los análisis de RNA-Seq, en donde *PTEN* y *PDCD4* en tejido tumoral se encuentran subexpresados, y de manera contraria, en los resultados de RT-qPCR, *PTEN* y *PDCD4* se encuentran sobreexpresados en la línea celular BT474. Específicamente, miR-21-5p se encuentra sobreexpresado, por lo que se esperaría un efecto de regulación negativo de sus blancos génicos, efecto que es observado en MCF-7 para *PTEN* y *PDCD4*. Asimismo, miR-21-5p en la línea celular BT474, se encuentra subexpresado, línea celular en la cual, *PTEN* y *PDCD4* se encuentran sobreexpresados. Por lo tanto, se encontró un patrón de expresión, que sugiere la existencia de una regulación negativa en la expresión de *PTEN* y *PDCD4*, mediada por miR-21-5p.. (Figura 17A y 17B).

Con respecto a los blancos génicos propuestos para miR-21-3p, se observó que el transcrito de naturaleza codificante *MAT2A* presenta una mayor expresión en la línea celular MCF-7, cuando se compara con la línea calibradora, MCF10-A, sin embargo, esta expresión no fue significativa. Asimismo, *MAT2A* se encontró sobreexpresado de manera significativa en la línea celular neoplásica BT474, al igual que se observa una sobreexpresión en MDA-MB-231 (Figura 17C). Al analizar la expresión de miR-21-3p y la expresión de sus genes blanco, específicamente *MAT2A*. En MCF-7, se observó un patrón de regulación negativa, en donde miR-21-3p se encuentra sobreexpresado y *MA2TA* se observa subexpresado. Asimismo, *MAT2A* las líneas celulares en las cuales miR-21-3p se encuentra sub-expresado (BT474 y MDA-MB-231) presenta una sobreexpresión.

Adicionalmente, *MAT2B*, el segundo gen candidato a ser blanco de regulación de miR-21-3p, presentó un nivel de expresión en las líneas celulares luminales, MCF-7 y BT474 una expresión basal, al compararse con la línea calibradora MCF-10A. Al igual que con *MAT2A*, *MAT2B* presentó un patrón que sugiere un tipo de regulación negativa a nivel

mRNA en la línea celular MDA-MB-231, en la cual, miR-21-3p se encuentra subexpresado (Figura 17D).

En resumen, los análisis de expresión diferencial mediante RT-qPCR sugieren que, al igual que *PTEN* y *PDCD4* con miR-21-5p, Los mRNAs de *MAT2A* y *MAT2B* son dos blancos candidatos a posiblemente regulados por miR-21-3p a nivel post-transcripcional, en donde una sobreexpresión de miR-21-3p, resulta en la disminución de los niveles de *MAT2A* y *MAT2B*.

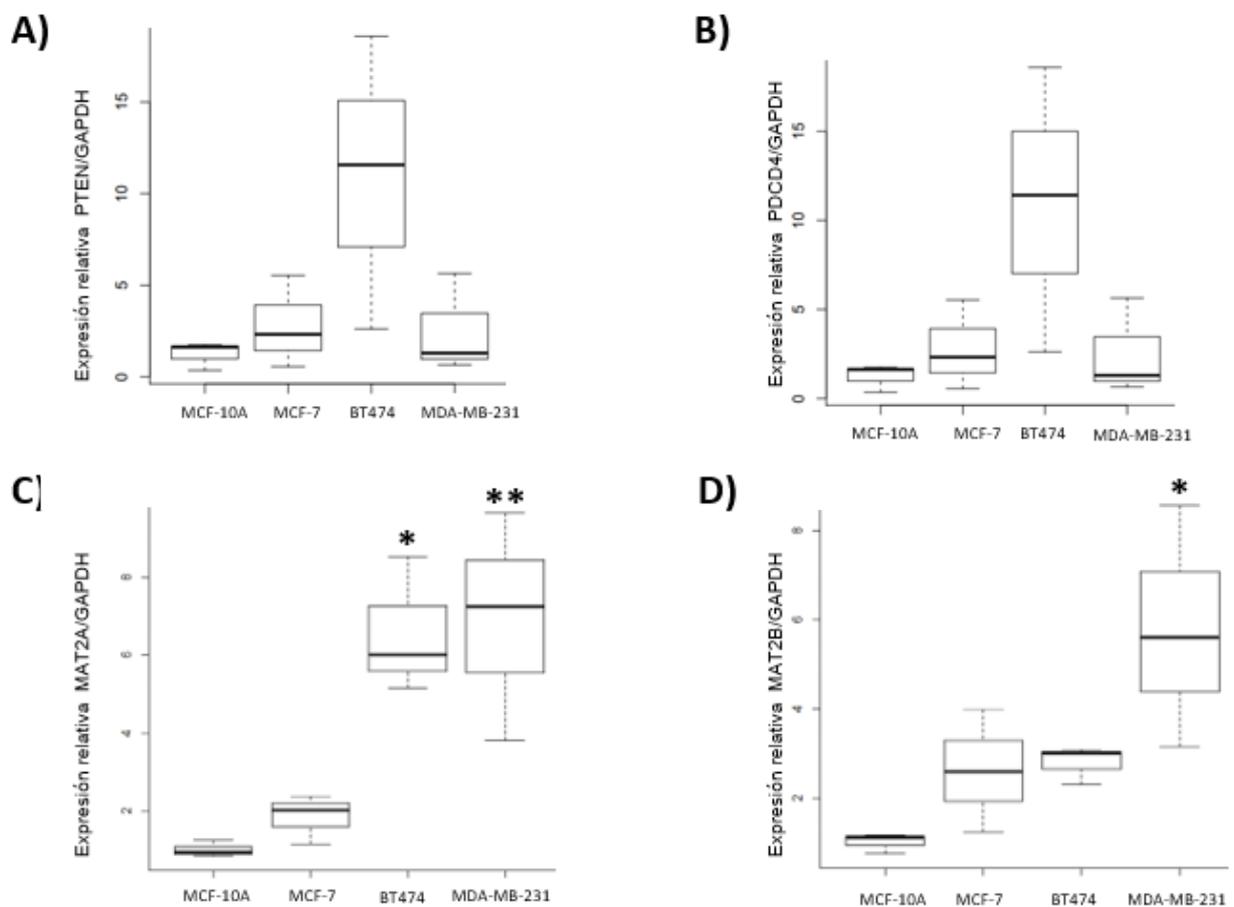


Figura 17. Expresión diferencial de los mRNAs de *PTEN*, *PDCD4*, *MAT2A* y *MAT2B* en líneas celulares de cáncer de mama. A partir de los experimentos de qPCR, se observa que los blancos génicos de miR-21-5p, *PTEN* (A) y *PDCD4* (B), en las líneas neoplásicas MCF-7 y MDA-MB-231 una expresión basal de los transcritos codificantes, a diferencia de BT474, la línea Luminal-B que muestra una mayor expresión de *PTEN* y *PDCD4*. Expresión diferencial de los transcritos codificantes *MAT2A* y *MAT2B* mediante RT-qPCR en tiempo real. En donde se observa que *MAT2A* presenta una sobreexpresión en la línea celular Luminal-B, BT474 y MDA-MB-231, a diferencia de la línea MCF-7 la cual presenta niveles basales de *MAT2A* (C). De la misma forma, *MAT2B* se expresa de forma basal en las líneas neoplásicas Luminales MCF-7 y

BT474, mientras que en la línea Triple Negativo, MDA-MB-231 se encuentra sobreexpresado (**D**). (n=3, p=0.05, *ANOVA con prueba de Tukey).

En conclusión, los resultados de este estudio mostraron que existe una alteración en la expresión del oncomiR miR-21. Se observó que el pri-miRNA-21 se encontró sobreexpresado tanto por los análisis de RNA-Seq, así como en los datos experimentales obtenidos de RT-qPCR. Por otro lado, se demostró que existe una sobreexpresión de la hebra miR-21-3p, la cual nunca antes había sido reportada en líneas celulares de cáncer de Mama. La sobreexpresión observada de miR-21-3p, se encontró específicamente en la línea celular MCF-7, la cual es una línea celular neoplásica de subtipo molecular Luminal A, en la cual, al igual que miR-21-3p que se sobreexpresó más de 4 veces, la hebra miR-21-5p también se encontró sobreexpresada, validando los datos analizados en la cohorte de pacientes con cáncer de mama, en donde miR-21-5p y miR-21-3p se sobreexpresaban en tejido tumoral con una frecuencia similar, patrón que únicamente se observó en dos de veintidós tipos de cáncer analizados, en particular se vio sobreexpresión de ambas hebras en cáncer de mama y pulmón. Asimismo, mediante información obtenida de diferentes herramientas, los genes *MAT2A* y *MAT2B* se propusieron como blancos génicos directos de miR-21-3p, mostrando un patrón de regulación en las líneas celulares neoplásicas, sugiriendo una regulación RNA-RNA por medio de miR-21-3p, y proponiendo una regulación RNA-Proteína. Por lo tanto, los resultados encontrados sugieren que la hebra que antes se creía que era hebra pasajera, en este estudio se encontró sobreexpresada con una frecuencia similar a miR-21-5p. Por lo tanto, los resultados presentados, sugieren que existen mecanismos de regulación mediados por miR-21-3p lo cual manifiesta que existan rutas metabólicas asociadas a procesos oncogénicos que no habían sido asociadas a esta hebra con anterioridad, ampliando el universo de blancos génicos y posibles vías metabólicas asociadas en las cuales miR-21-3p pueda ser caracterizado funcionalmente.

9 Discusión

El presente estudio de investigación es un enfoque hacia el estudio de la expresión y posible función biológica de la hebra miR-21-3p en cáncer de mama, a través del uso de herramientas bioinformáticas, así como el uso de líneas celulares de cáncer de mama para entender la función de miR-21-3p en el desarrollo de cáncer de mama.

Los miRNAs son pequeños RNAs no codificantes endógenos de aproximadamente 23 nt que desempeñan importantes funciones de regulación génica en animales y plantas al reconocer por complementariedad la 3'UTR de los mRNAs de genes que codifican proteínas para dirigir su represión postranscripcional (Bartel, 2009). En particular, se sabe que los miRNAs están involucrados en diferentes procesos celulares relacionados con la patogénesis del CaMa (Loh et al., 2019). Por otro lado, los análisis de transcriptoma a través de la secuenciación masiva en paralelo de miRNAs muestran que como producto final de la biogénesis se producen cantidades significativas de miRNA maduro de ambas hebras (5p y 3p) del dúplex de miRNAs (Baskerville, 2005; Kozomara et al., 2019). Pocos estudios han reportado que ambas hebras de un mismo miRNA son funcionales, y se encuentran altamente reguladas de manera transcripcional y post-transcripcional, siguiendo un modelo de selección de hebras no canónico, en donde ambas hebras son cargadas al complejo miRISC con una frecuencia similar, incluso ambas hebras son capaces de silenciar post-transcripcionalmente a mRNAs blanco (Mitra et al., 2020). A pesar de que recientemente varios grupos de trabajo se han enfocado en elucidar las funciones en diferentes patologías de ambas hebras de un mismo miRNA, debido al modelo canónico de selección de hebras, el papel de la mitad de los miRNAs aún no ha sido descrita, por lo que es necesario profundizar el estudio de este tipo de mecanismos no canónicos de regulación mediados por ambas hebras en distintas patologías, específicamente en cáncer, donde ambas hebras de un mismo miRNA se co-expresen de manera que puedan considerarse como miRNAs que siguen un modelo no canónico de selección de hebras, y por lo tanto, ambas hebras podrían ser funcionales, llevando al silenciando post-transcripcional de diferentes mRNAs, teniendo un efecto biológico sobre la regulación de distintas vías metabólicas asociadas a procesos oncogénicos, destacando la importancia del estudio de miRNAs con un comportamiento no canónico en el desarrollo de neoplasias.

Actualmente, miR-21 es uno de los oncomiRs mejor caracterizados y estudiados en investigación básica, debido a la enorme red de interacción que tiene con genes asociados a vías metabólicas intracelulares involucradas en procesos oncogénicos (Bautista-Sánchez et al., 2020), sin embargo, los estudios científicos se han enfocado principalmente en la hebra considerada “guía” la cual corresponde a la hebra 5p, sin tomar en consideración a la hebra pasajera de miR-21 que se conoce como 3p. Lo anterior ha llevado a preguntarnos si estas hebras funcionan de manera coordinada para regular procesos y vías oncogénicas durante el proceso de carcinogénesis, o en su caso, determinar si cada hebra tiene un efecto biológico antagónico, inhibiendo sus funciones individuales. Específicamente, se sabe que, en cáncer de mama, miR-21-5p se encuentra sobreexpresado y actúa como un oncomiR, regulando varios genes supresores de tumor asociados a mecanismos oncogénicos como la apoptosis (*PTEN* (Meng et al., 2007), *TIMP3* (X. Zhang et al., 2011, p. 3), *BCL2* (L. Xu et al., 2014, p. 2) y *PDCD4* (Asangani et al., 2008, p. 4)), la proliferación (*SPRY2* (Gao et al., 2012; Shen et al., 2013), *PTEN* (Z.-L. Liu et al., 2013; Meng et al., 2007), *TIMP3* (J. Hu et al., 2016, p. 3), *BCL2* (Dong et al., 2011, p. 2) y *PDCD4* (Asangani et al., 2008, p. 4; Shen et al., 2013, p. 4)), así como la metástasis y la invasión (*PTEN*, *RECK* (Takahashi et al., 1998), *TIMP3* y *PDCD4*), incluyendo también a la angiogénesis (*SPRY2* (Gao et al., 2012, p. 2) y *RECK* (Oh et al., 2001)), entre otras (Bautista-Sánchez et al., 2020). Además, las evidencias experimentales han determinado que la expresión miR-21-5p está relacionada con el desarrollo de resistencia a múltiples fármacos (MDR) utilizados como tratamientos neo-adyuvantes y adyuvantes, como lo son el Topotecán con el cual, al inhibir a miR-21 resulta en sensibilización de la línea celular MCF-7 (línea celular de cáncer de mama, subtipo molecular luminal A) a este tratamiento. También, en cáncer de pulmón, actúa como un factor de MDR con los fármacos Citarabina y Nogamicina, por mencionar algunos de los MDRs asociados a la expresión de miR-21-5p en cáncer (Najjary et al., 2020). Por lo anterior, el estudio del oncomiR, miR-21-5p, ha tomado gran importancia, específicamente en el estudio de su expresión y su papel en diferentes tipos de cáncer, sin embargo, el estudio de miR-21-3p en cáncer, ha sido reportada limitadas veces, por lo que es necesario tratar de establecer cuales la función de miR-21-3p en cáncer.

En este estudio, reportamos por primera vez en cáncer de mama, mediante análisis de datos de RNA-Seq y su posterior validación por experimentos de biología molecular, la importancia de la sobreexpresión de miR-21-3p. Los resultados sugirieron una posible

función biológica que no había sido estudiada anteriormente para miR-21-3p. Los análisis a partir de datos de RNA-Seq depositados en bases de acceso público como TACCO mostraron que miR-21-3p se encuentra sobreexpresado en una cohorte de pacientes caucásicos de CaMa, presentando un incremento de expresión de hasta 6 veces más cuando se compara la expresión en tejido normal contra tejido tumoral. Así mismo, en 2 de 22 tipos de cáncer analizados, específicamente en cáncer de mama, se observó un patrón de expresión que concuerda con un tipo de selección de hebras no canónico, en donde ambas hebras (5p y 3p) se expresan de manera similar en estas patologías.

Por otro lado, los resultados experimentales por RT-qPCR de líneas celulares de cáncer de mama corroboraron lo observado en los análisis bioinformáticos de RNA-Seq sugiriendo una expresión diferencial de miR-21-3p en líneas celulares de cáncer de mama. Los resultados muestran que tanto el pri-miRNA-21, así como ambas hebras maduras de miR-21 se sobre-expresan de manera significativa en la línea celular neoplásica, de morfología epitelial, subtipo molecular Luminal A que corresponde a MCF-7. Lo anterior sugiere que, en cáncer de mama, ambos miRNAs maduros de miR-21 se encuentran sobreexpresados, indicando que existe una posible regulación post-transcripcional mediada no únicamente por miR-21-5p como ha sido reportado anteriormente (H. Wang et al., 2019; Yan et al., 2008), si no que miR-21-3p también parece tener un función ya que se encuentra sobreexpresada, lo cual sugiere que ambas hebras de miR-21 podrían tener un papel funcional en cáncer de mama.

Por otro lado, mediante el uso de 4 servidores web de predicción de blancos génicos para miRNAs (*TargetScan*, *miRTarBase*, *TarBase* y *RNAinteractome*), se obtuvieron 5 blancos génicos para miR-21-5p, de los cuales se escogieron a *PTEN* y *PDCD4*, los cuales han sido reportados experimentalmente con anterioridad como blancos génicos de miR-21-5p (Asangani et al., 2008, p. 4; Meng et al., 2007) y nuestros resultados han apoyado lo anterior ya que en nuestros análisis *in silico* pudimos validar estos mismos blancos génicos validados experimentalmente, demostrando que con el uso de las herramientas bioinformáticas y plataformas de acceso público uno puede llegar a establecer los posibles blancos génicos a través del uso de este tipo de herramientas para hacer predicciones de blancos génicos *in silico*. Además, los genes *PTEN* y *PDCD4* han sido extensamente estudiados dentro del contexto de regulación génica mediada por miR-21-5p, en donde *PTEN* y *PDCD4* se encuentran asociados a vías metabólicas intracelulares oncogénicas como la proliferación celular, la apoptosis, la migración y la

invasión celular. En particular, el gen *PTEN* se encuentra asociado a las vías de señalización PI3K/AKT/mTOR, FOXO3 y MAPK/ERK, VEGFA (Miller et al., 2017). Asimismo, el gen *PDCD4* se encuentra asociada a la vía de señalización JKN (Shen et al., 2013). En conclusión, los resultados sugieren que miR-21-5p tiene una función importante en procesos oncogénicos como la proliferación celular, la apoptosis, así como la migración y la invasión celular. En donde miR-21-5p, a través de la regulación negativa de sus genes blanco como *PTEN* y *PDCD4*, desencadena el desarrollo de un fenotipo tumoral.

Por otro lado, en el caso particular de miR-21-3p, éste no ha sido ampliamente estudiado, y, por lo tanto, se desconoce si esta hebra de miR-21 tiene la capacidad de expresarse de manera que desempeñe una función biológica en cáncer de mama. Sin embargo, miR-21-3p se ha estudiado en otras neoplasias como en cáncer de ovario (Báez-Vega et al., 2016), en el cual, una disminución de miR-21-3p tiene como efecto una disminución en la proliferación celular y en la invasión, además se reportó como blanco directo de miR-21-3p, a la proteína de unión a RNA con empalme múltiple (*RBPMS*), la cual muestra asociación con la sensibilidad de las células de cáncer de ovario al tratamiento con Cisplatino (*The passenger strand, miR-21-3p, plays a role in mediating cisplatin resistance in ovarian cancer cells - ScienceDirect*, s. f.). De igual manera, en cáncer colorrectal, mediante un estudio basado en datos de TCGA (*The Cancer Genome Atlas*), se propone a miR-21-3p como un oncomiR, debido a que promueve la movilidad celular mediante la transición epitelio-mesénquima (Jiao et al., 2017). Por lo tanto, los limitados estudios de la función de miR-21-3p en cáncer proponen a esta hebra como un oncomiR, que al igual que miR-21-5p, promueve el desarrollo tumoral, mediante la regulación negativa de sus blancos génicos, los cuales se encuentran asociados a vías metabólicas intracelulares oncogénicas.

En este estudio mediante herramientas de búsqueda informática de predicción se encontró que *MAT2A* y *MAT2B* son genes blanco candidatos de miR-21-3p, en donde *MAT2A* y *MAT2B* contienen regiones de unión en sus extremos 3'UTR, que son reconocidos por miR-21-3p, sugiriendo que existe interacción entre miRNA-mRNA mediante una unión complementaria de bases en la región semilla de miR-21-3p y las regiones de unión encontradas en los extremos 3'UTR del mRNA de *MAT2A* y *MAT2B*, teniendo como consecuencia, una posible regulación negativa a nivel post-transcripcional de estos mRNAs, resultando en la alteración de vías metabólicas en las

cuales *MAT2A* y *MAT2B* se encuentren asociadas. Por lo anterior, se estudiaron los niveles de expresión de *MAT2A* y *MAT2B* mediante datos obtenidos de TACCO de una cohorte de 1085 pacientes de cáncer de mama, de los cuales se obtuvo que en tejido tumoral (tejido en el cual miR-21-3p se encuentra sobreexpresado), *MAT2A* se encuentra subexpresado cuando se compara con los niveles de expresión de tejido normal. En el caso del gen *MAT2B* éste presenta niveles basales de expresión, sugiriendo que una sobreexpresión de miR-21-3p, da como resultado una disminución en la expresión de *MAT2A*, por lo cual se sugiere que podría haber una regulación a nivel post-transcripcional del mRNA de *MAT2A* a través de miR-21-3p mediante una interacción miRNA-mRNA en pacientes con cáncer de mama, sin embargo, esta hipótesis aún está por confirmarse experimentalmente. Una vez observado este patrón de regulación, se estudiaron los niveles de expresión de *MAT2A* y *MAT2B*, utilizando los datos de RNA-Seq de las líneas celulares de cáncer de mama utilizadas en este estudio, MCF-7, BT474 y MDA-MB-231, utilizando a la línea celular MCF-10A como línea celular no neoplásica. Específicamente, en la línea celular MCF-7 (línea celular que sobreexpresó a miR-21-3p) se encontró que *MAT2A* y *MAT2B* presentan niveles basales de expresión cuando son comparados con MCF-10A. Además, mediante cuantificación de la expresión por RT-qPCR de los genes *MAT2A* y *MAT2B* en las líneas celulares de cáncer de mama, el nivel basal de expresión de *MAT2A* y *MAT2B* se volvió a encontrar en MCF-7, y por el contrario, se observó una sobreexpresión de los genes *MAT2A* y *MAT2B* en las líneas celulares BT474 y MDA-MB-231, líneas celulares en las cuales miR-21-3p se encuentra subexpresado. Los resultados anteriores sugieren que posiblemente los genes *MAT2A* y *MAT2B* se encuentran regulados directamente por miR-21-3p, mediante la unión de la región semilla de miR-21-3p y los extremos 3'UTR de *MAT2A* y *MAT2B*, sin embargo, esta es aún una hipótesis por confirmar experimentalmente en cáncer de mama.

Con respecto a la función de los genes *MAT2A* y *MAT2B*, el grupo de Ting- Fang Lo y colaboradores en el 2013 reportaron en cáncer de hígado, la primera evidencia de *MAT2A* y *MAT2B* como genes blanco directos de miR-21-3p, en el cual miR-21-3p disminuye directamente los niveles de expresión de *MAT2* y *MAT2B* mediante un emparejamiento por complementariedad de bases en la región semilla de miR-21-3p y los extremos 3'UTR de *MAT2* y *MAT2B*.(Lo et al., 2013) Además, las evidencias experimentales han demostrado que el efecto de miR-21-3p sobre la regulación de los niveles de expresión de *MAT2* y *MAT2B*, tiene efecto directo en la regulación de la cantidad de S-Adenosil metionina (SAM) que es el principal donador de grupos metilo en

la célula (Frau et al., 2012; Pañeda et al., 2002). Se encontró que la sobreexpresión de miR-21-3p incrementa los niveles intracelulares de SAM, el cual, específicamente en cáncer de hígado, ha sido demostrado que altos niveles de SAM disminuyen el crecimiento celular e induce apoptosis en líneas celulares de cáncer de hígado (Lo et al., 2013). En resumen, el grupo de Ting-Fang Lo, reportó por primera vez a miR-21-3p como un miRNA supresor de tumor mediante la regulación de sus blancos génicos *MAT2* y *MAT2B*, específicamente en la regulación indirecta de la proteína que es catalizada por estas enzimas, SAM en cáncer de hígado. Adicionalmente a lo reportado en cáncer de hígado, se han reportado estudios en otros tipos de cáncer como cáncer de colon, gástrico, mama, páncreas y próstata, en donde se relaciona el papel de los genes *MAT* con el desarrollo del cáncer, más allá de su función en hígado. En este subconjunto de cánceres extrahepáticos, *MAT2A* y *MAT2B* se inducen a través de diferentes mecanismos reguladores. Específicamente en cáncer de mama, en células resistentes a tamoxifeno (TAMR-MCF-7), se reportó una sobreexpresión de *MAT2A* cuando se comparó la expresión con las células control (Phuong et al., 2016). Además, fue reportado en otro estudio, que los niveles de SAM endógenos (dependientes de *MAT2A*) encontrados en las células TAMR-MCF-7 se encontraba correlacionado con la hipermetilación del gen supresor de tumor *PTEN*, disminuyendo su expresión y por lo tanto facilitando el crecimiento tumoral (Phuong et al., 2011).

Por otro lado, *MAT2B* se ha correlacionado con mal pronóstico en el subtipo molecular triple negativo de cáncer de mama, en donde con experimentos de pérdida de función mediante inhibidores (RNAi), demostraron que la inhibición de *MAT2B* en las células subtipo molecular triple negativo (MDA-MB-231) redujo el crecimiento celular y migración e indujo apoptosis celular, en donde ellos atribuyen el mecanismo de regulación a la vía AKT y ERK. (J. Xu et al., 2019). Por lo tanto, teniendo como antecedente, las funciones oncogénicas reportadas con anterioridad de *MAT2* y *MAT2B*, se proponen a estas dos moléculas, como oncogenes. En donde *MAT2* y *MAT2B* fueron reportados previamente como blancos directos de de miR-21-3p en cáncer hepático.

Finalmente, lo reportado en este estudio, adicional a las funciones de *MAT2A* y *MAT2B* en cáncer de mama, y la validación experimental de una regulación directa por complementariedad de bases mediada por miR-21-3p Ting-Fang lo y colaboradores, sugiere que en cáncer de mama, específicamente en la línea celular neoplásica subtipo molecular Luminal A que corresponde a MCF-7, miR-21-3p se encuentra

sobreexpresado, donde los genes *MAT2A* y *MAT2B* parecen ser potenciales blancos génicos a ser regulados directamente por miR-21-3p, presentando una regulación negativa, en la cual una sobreexpresión de miR-21-3p, tenga como consecuencia biológica una disminución en los niveles de expresión de *MAT2A* y *MAT2B*, y por lo tanto, adicionalmente exista una regulación indirecta de los niveles intracelulares de SAM mediados por miR-21-3p. Por lo tanto, es necesario elucidar si en cáncer de mama, *MAT2A* y *MAT2B* son blancos directos de miR-21-3p y, en consecuencia, los niveles intracelulares de SAM puedan ser regulados indirectamente por los niveles de expresión de miR-21-3p, y qué vías metabólicas oncogénicas se encuentran asociadas a la regulación de *MAT2A* y *MAT2B* mediada por miR-21-3p, con el fin de caracterizar de una manera funcional, a miR-21-3p como un oncomiR, potenciando un fenotipo oncogénico, y por lo tanto funcionando de manera sinérgica con miR-21-5p en cáncer de mama, o por otro lado, caracterizando a miR-21-3p como un miRNA supresor de tumor, que disminuya un fenotipo, presentando una función antagónica a miR-21-5p.

10 Conclusiones

A continuación, se concluye lo siguiente de los resultados obtenidos en este trabajo:

- Se encontró una sobreexpresión de pri-miRNA-21, miR-21-5p y miR-21-3p en MCF7 y en pacientes con cáncer de mama.
- Los análisis *in silico* sugieren que los blancos génicos de miR-21-3p son los genes *MAT2A* y *MAT2B*.
- Los análisis de interacción realizados en la plataforma de RAIN, sugieren que los genes *MAT2A* y *MAT2B* se encuentran asociados a la vía de señalización RAS-RAF-MERK-ERK, la metilación de DNA, así como la regulación de los niveles intracelulares de SAM.
- Se encontró en los análisis de RT-qPCR, que miR-21-3p y miR-21-5p se sobreexpresan en la línea celular neoplásica MCF7.

La conclusión general de esta tesis es, que las evidencias experimentales sugieren una sobreexpresión de pri-miRNA-21, miR-21-5p y miR-21-3p en el cáncer de mama a través del estudio de datos de expresión de pacientes y de líneas celulares de cáncer de mama. Específicamente, se encontró una sobreexpresión miR-21-3p que no había sido reportada previamente en cáncer de mama, sugiriendo la existencia de un mecanismo de regulación post-transcripcional mediada por sobreexpresión de ambas hebras de miR-21. Asimismo, mediante un análisis *in silico*, se propusieron a *MAT2A* y *MAT2B* como blancos génicos de miR-21-3p. Por lo tanto, la sobreexpresión de miR-21-3p observada en la línea celular neoplásica Luminal A que corresponde a MCF-7, sugiere que existe una regulación post-transcripcional de los genes *MAT2A* y *MAT2B* mediante la unión de sus extremos 3'UTR con miR-21-3p, en donde esta regulación miRNA-mRNA se encuentra alterada, lo cual sugiere un posible efecto biológico en vías metabólicas oncogénicas, que pudieran tener un efecto coordinado (sinérgico) con miR-21-5p, potenciando un fenotipo oncogénico en cáncer de mama.

11 Perspectivas

El presente proyecto se dirigió específicamente a determinar el papel de miR-21-3p en cáncer de mama, con el fin de establecer si existe una regulación post-transcripcional mediada por la alteración en la expresión de ambas hebras de miR-21, y así establecer cuál es la función biológica de la hebra guía y pasajera de miR-21 en el desarrollo y progresión del cáncer de mama. Asimismo, debido a la antigua concepción del mecanismo canónico en la biogénesis de los miRNAs, la cual dicta que una hebra es funcional y la otra hebra del dúplex de miRNAs es degradada, no existe evidencia que demuestre la función biológica de miR-21-3p en cáncer de mama, por lo que queda como propuesta para continuar el proyecto los siguientes abordajes experimentales:

Propuesta de ganancia y pérdida de expresión mediante el uso de inhibidores y mimics específicos para miR-21-5p y miR-21-3p

- Realizar ensayos de Luciferasa, en los cuales sean utilizados los sitios de unión 3'UTR de los genes blanco en condiciones silvestres, así como mutaciones puntuales dirigidas en los sitios de complementariedad de bases miRNA-mRNA. Validar que los genes *PTEN*, *PDCD4*, *MAT2A* y *MAT2B* son blancos directos, y que se unen mediante sus regiones 3'UTR a las hebras 5p y 3p respectivamente.
- Realizar mediante ensayos de ELISA, la cuantificación de SAM, en condiciones normales, en MCF-10A y en condiciones neoplásicas, en MCF-7. Además, en condiciones donde miR-21-5p o miR-21-3p se encuentre inhibido o sobreexpresado, mediante el uso de mimics e inhibidores de ambas hebras de miR-21. Con el fin de justificar si la regulación que se propone de miR-21-3p a *MAT2A* y *MAT2B*, afecta a los niveles intracelulares de SAM.
- Realizar ensayos de inhibición y sobreexpresión de miR-21-5p y miR-21-3p, en donde se realizarán ensayos de inhibición y sobreexpresión individual de cada miRNA en la línea celular MCF-7. Con el objetivo de observar si el abatimiento de la expresión del miRNA es suficiente para que exista una alteración en la

expresión de los genes blanco candidatos, así como en los niveles intracelulares de SAM.

- Evaluar el efecto biológico que tendrá el abatir y/o sobreexpresar los niveles de expresión de cada hebra de miR-21 mediante ensayos de apoptosis (marcadores, TUNEL, citometría de flujo), proliferación celular (curva de crecimiento celular) y migración e invasión (ensayo de herida y *Transwell*). Con el fin de elucidar si miR-21-3p y miR-21-5p trabajan de manera sinérgica en cáncer de mama, potenciando un fenotipo oncogénico.

12 Bibliografía

- Afgan, E., Baker, D., Batut, B., van den Beek, M., Bouvier, D., Čech, M., Chilton, J., Clements, D., Coraor, N., Grüning, B. A., Guerler, A., Hillman-Jackson, J., Hiltemann, S., Jalili, V., Rasche, H., Soranzo, N., Goecks, J., Taylor, J., Nekrutenko, A., & Blankenberg, D. (2018). The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Research*, *46*(W1), W537-W544. <https://doi.org/10.1093/nar/gky379>
- Asangani, I. A., Rasheed, S. A. K., Nikolova, D. A., Leupold, J. H., Colburn, N. H., Post, S., & Allgayer, H. (2008). MicroRNA-21 (miR-21) post-transcriptionally downregulates tumor suppressor Pcd4 and stimulates invasion, intravasation and metastasis in colorectal cancer. *Oncogene*, *27*(15), 2128-2136. <https://doi.org/10.1038/sj.onc.1210856>
- Báez-Vega, P. M., Vargas, I. M. E., Valiyeva, F., Encarnación-Rosado, J., Roman, A., Flores, J., Marcos-Martínez, M. J., & Vivas-Mejía, P. E. (2016). Targeting miR-21-3p inhibits proliferation and invasion of ovarian cancer cells. *Oncotarget*, *7*(24), 36321-36337. <https://doi.org/10.18632/oncotarget.9216>
- Bagga, S., Bracht, J., Hunter, S., Massirer, K., Holtz, J., Eachus, R., & Pasquinelli, A. E. (2005). Regulation by let-7 and lin-4 miRNAs results in target mRNA degradation. *Cell*, *122*(4), 553-563. <https://doi.org/10.1016/j.cell.2005.07.031>
- Bardou, P., Mariette, J., Escudié, F., Djemiel, C., & Klopp, C. (2014). jvenn: An interactive Venn diagram viewer. *BMC Bioinformatics*, *15*(1), 293. <https://doi.org/10.1186/1471-2105-15-293>
- Barrett, T., Wilhite, S. E., Ledoux, P., Evangelista, C., Kim, I. F., Tomashevsky, M., Marshall, K. A., Phillippy, K. H., Sherman, P. M., Holko, M., Yefanov, A., Lee, H., Zhang, N., Robertson, C. L., Serova, N., Davis, S., & Soboleva, A. (2012). NCBI GEO: Archive for functional genomics data sets—update. *Nucleic Acids Research*, *41*(D1), D991-D995. <https://doi.org/10.1093/nar/gks1193>
- Bartel, D. P. (2004). MicroRNAs. *Cell*, *116*(2), 281-297. [https://doi.org/10.1016/S0092-8674\(04\)00045-5](https://doi.org/10.1016/S0092-8674(04)00045-5)
- Bartel, D. P. (2009). MicroRNAs: Target Recognition and Regulatory Functions. *Cell*, *136*(2), 215-233. <https://doi.org/10.1016/j.cell.2009.01.002>

- Baskerville, S. (2005). Microarray profiling of microRNAs reveals frequent coexpression with neighboring miRNAs and host genes. *RNA*, *11*(3), 241-247.
<https://doi.org/10.1261/rna.7240905>
- Bautista-Sánchez, D., Arriaga-Canon, C., Pedroza-Torres, A., De La Rosa-Velázquez, I. A., González-Barrios, R., Contreras-Espinosa, L., Montiel-Manríquez, R., Castro-Hernández, C., Fragoso-Ontiveros, V., Alvarez-Gomez, R. M., & Herrera, L. A. (2020). The promising role of miR-21 as a cancer biomarker and its importance in RNA-based therapeutics. *Molecular Therapy - Nucleic Acids*, S2162253120300949. <https://doi.org/10.1016/j.omtn.2020.03.003>
- Breving, K., & Esquela-Kerscher, A. (2010). The complexities of microRNA regulation: Mirandering around the rules. *The International Journal of Biochemistry & Cell Biology*, *42*(8), 1316-1329. <https://doi.org/10.1016/j.biocel.2009.09.016>
- Cai, X., Hagedorn, C. H., & Cullen, B. R. (2004). Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs. *RNA (New York, N. Y.)*, *10*(12), 1957-1966. <https://doi.org/10.1261/rna.7135204>
- Calin, G. A., Dumitru, C. D., Shimizu, M., Bichi, R., Zupo, S., Noch, E., Aldler, H., Rattan, S., Keating, M., Rai, K., Rassenti, L., Kipps, T., Negrini, M., Bullrich, F., & Croce, C. M. (2002). Frequent deletions and down-regulation of micro- RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(24), 15524-15529. <https://doi.org/10.1073/pnas.242606799>
- Calin, G. A., Sevignani, C., Dumitru, C. D., Hyslop, T., Noch, E., Yendamuri, S., Shimizu, M., Rattan, S., Bullrich, F., Negrini, M., & Croce, C. M. (2004). Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(9), 2999-3004.
<https://doi.org/10.1073/pnas.0307323101>
- Chang, K.-W., Kao, S.-Y., Wu, Y.-H., Tsai, M.-M., Tu, H.-F., Liu, C.-J., Lui, M.-T., & Lin, S.-C. (2013). Passenger strand miRNA miR-31* regulates the phenotypes of oral cancer cells by targeting RhoA. *Oral Oncology*, *49*(1), 27-33.
<https://doi.org/10.1016/j.oraloncology.2012.07.003>

- Chendrimada, T. P., Gregory, R. I., Kumaraswamy, E., Norman, J., Cooch, N., Nishikura, K., & Shiekhattar, R. (2005). TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing. *Nature*, *436*(7051), 740-744. <https://doi.org/10.1038/nature03868>
- Chou, P.-H., Liao, W.-C., Tsai, K.-W., Chen, K.-C., Yu, J.-S., & Chen, T.-W. (2019). TACCO, a Database Connecting Transcriptome Alterations, Pathway Alterations and Clinical Outcomes in Cancers. *Scientific Reports*, *9*(1), 3877. <https://doi.org/10.1038/s41598-019-40629-z>
- Chu, C., & Rana, T. M. (2006). Translation Repression in Human Cells by MicroRNA-Induced Gene Silencing Requires RCK/p54. *PLoS Biology*, *4*(7). <https://doi.org/10.1371/journal.pbio.0040210>
- Clustal Omega < Multiple Sequence Alignment < EMBL-EBI.* (s. f.). Recuperado 25 de mayo de 2020, de <https://www.ebi.ac.uk/Tools/msa/clustalo/>
- Curcumin Regulates miR-21 Expression and Inhibits Invasion and Metastasis in Colorectal Cancer—PubMed.* (s. f.). Recuperado 1 de julio de 2020, de <https://pubmed.ncbi.nlm.nih.gov/20815812/>
- Denli, A. M., Tops, B. B. J., Plasterk, R. H. A., Ketting, R. F., & Hannon, G. J. (2004). Processing of primary microRNAs by the Microprocessor complex. *Nature*, *432*(7014), 231-235. <https://doi.org/10.1038/nature03049>
- Department of Medicine and Division of Nephrology, UT Southwestern Medical Center, Dallas, TX, USA, Hajarnis, S., Lakhia, R., Department of Medicine and Division of Nephrology, UT Southwestern Medical Center, Dallas, TX, USA, Patel, V., & Department of Medicine and Division of Nephrology, UT Southwestern Medical Center, Dallas, TX, USA. (2015). MicroRNAs and Polycystic Kidney Disease. En Department of Internal Medicine, Department of Biochemistry and Molecular Biology, University of Kansas Medical Center, Kansas City, KS 66160, USA & X. Li (Eds.), *Polycystic Kidney Disease* (pp. 313-334). Codon Publications. <https://doi.org/10.15586/codon.pkd.2015.ch13>
- Dong, J., Zhao, Y.-P., Zhou, L., Zhang, T.-P., & Chen, G. (2011). Bcl-2 Upregulation Induced by miR-21 Via a Direct Interaction Is Associated with Apoptosis and Chemoresistance in MIA PaCa-2 Pancreatic Cancer Cells. *Archives of Medical Research*, *42*(1), 8-14. <https://doi.org/10.1016/j.arcmed.2011.01.006>

- Eulalio, A., Behm-Ansmant, I., Schweizer, D., & Izaurralde, E. (2007). P-body formation is a consequence, not the cause, of RNA-mediated gene silencing. *Molecular and Cellular Biology*, 27(11), 3970-3981. <https://doi.org/10.1128/MCB.00128-07>
- Feng, Y.-H., & Tsao, C.-J. (2016). Emerging role of microRNA-21 in cancer. *Biomedical Reports*, 5(4), 395-402. <https://doi.org/10.3892/br.2016.747>
- Filipowicz, W., Bhattacharyya, S. N., & Sonenberg, N. (2008). Mechanisms of post-transcriptional regulation by microRNAs: Are the answers in sight? *Nature Reviews Genetics*, 9(2), 102-114. <https://doi.org/10.1038/nrg2290>
- Frau, M., Tomasi, M. L., Simile, M. M., Demartis, M. I., Salis, F., Latte, G., Calvisi, D. F., Seddaiu, M. A., Daino, L., Feo, C. F., Brozzetti, S., Solinas, G., Yamashita, S., Ushijima, T., Feo, F., & Pascale, R. M. (2012). Role of transcriptional and posttranscriptional regulation of methionine adenosyltransferases in liver cancer progression. *Hepatology*, 56(1), 165-175. <https://doi.org/10.1002/hep.25643>
- Fujita, S., Ito, T., Mizutani, T., Minoguchi, S., Yamamichi, N., Sakurai, K., & Iba, H. (2008). MiR-21 Gene expression triggered by AP-1 is sustained through a double-negative feedback mechanism. *Journal of Molecular Biology*, 378(3), 492-504. <https://doi.org/10.1016/j.jmb.2008.03.015>
- Gao, M., Patel, R., Ahmad, I., Fleming, J., Edwards, J., McCracken, S., Sahadevan, K., Seywright, M., Norman, J., Sansom, O., & Leung, H. Y. (2012). SPRY2 loss enhances ErbB trafficking and PI3K/AKT signalling to drive human and mouse prostate carcinogenesis. *EMBO Molecular Medicine*, 4(8), 776-790. <https://doi.org/10.1002/emmm.201100944>
- Garneau, N. L., Wilusz, J., & Wilusz, C. J. (2007). The highways and byways of mRNA decay. *Nature Reviews Molecular Cell Biology*, 8(2), 113-126. <https://doi.org/10.1038/nrm2104>
- Gong, C., Yao, Y., Wang, Y., Liu, B., Wu, W., Chen, J., Su, F., Yao, H., & Song, E. (2011). Up-regulation of miR-21 mediates resistance to trastuzumab therapy for breast cancer. *The Journal of Biological Chemistry*, 286(21), 19127-19137. <https://doi.org/10.1074/jbc.M110.216887>
- Granados-López, A., Ruiz-Carrillo, J., Servín-González, L., Martínez-Rodríguez, J., Reyes-Estrada, C., Gutiérrez-Hernández, R., & López, J. (2017). Use of Mature miRNA Strand

- Selection in miRNAs Families in Cervical Cancer Development. *International Journal of Molecular Sciences*, 18(2), 407. <https://doi.org/10.3390/ijms18020407>
- Hu, H. Y., Yan, Z., Xu, Y., Hu, H., Menzel, C., Zhou, Y. H., Chen, W., & Khaitovich, P. (2009). Sequence features associated with microRNA strand selection in humans and flies. *BMC Genomics*, 10, 413. <https://doi.org/10.1186/1471-2164-10-413>
- Hu, J., Ni, S., Cao, Y., Zhang, T., Wu, T., Yin, X., Lang, Y., & Lu, H. (2016). The Angiogenic Effect of microRNA-21 Targeting TIMP3 through the Regulation of MMP2 and MMP9. *PLOS ONE*, 11(2), e0149537. <https://doi.org/10.1371/journal.pone.0149537>
- Jiao, W., Leng, X., Zhou, Q., Wu, Y., Sun, L., Tan, Y., Ni, H., Dong, X., Shen, T., Liu, Y., & Li, J. (2017). Different miR-21-3p isoforms and their different features in colorectal cancer. *International Journal of Cancer*, 141(10), 2103-2111. <https://doi.org/10.1002/ijc.30902>
- Jinek, M., Coyle, S. M., & Doudna, J. A. (2011). Coupled 5' Nucleotide Recognition and Processivity in Xrn1-Mediated mRNA Decay. *Molecular Cell*, 41(5), 600-608. <https://doi.org/10.1016/j.molcel.2011.02.004>
- Jovanovic, M., & Hengartner, M. O. (2006). miRNAs and apoptosis: RNAs to die for. *Oncogene*, 25(46), 6176-6187. <https://doi.org/10.1038/sj.onc.1209912>
- Junge, A., Refsgaard, J. C., Garde, C., Pan, X., Santos, A., Alkan, F., Anthon, C., von Mering, C., Workman, C. T., Jensen, L. J., & Gorodkin, J. (2017). RAIN: RNA–protein Association and Interaction Networks. *Database*, 2017(baw167). <https://doi.org/10.1093/database/baw167>
- Klein, M., Chandradoss, S. D., Depken, M., & Joo, C. (2017). Why Argonaute is needed to make microRNA target search fast and reliable. *Seminars in Cell & Developmental Biology*, 65, 20-28. <https://doi.org/10.1016/j.semcdb.2016.05.017>
- Koscianska, E., Starega-Roslan, J., & Krzyzosiak, W. J. (2011). The Role of Dicer Protein Partners in the Processing of MicroRNA Precursors. *PLoS ONE*, 6(12), e28548. <https://doi.org/10.1371/journal.pone.0028548>
- Kozomara, A., Birgaoanu, M., & Griffiths-Jones, S. (2019). miRBase: From microRNA sequences to function. *Nucleic Acids Research*, 47(D1), D155-D162. <https://doi.org/10.1093/nar/gky1141>

- Krichevsky, A. M., & Gabriely, G. (2009). miR-21: A small multi-faceted RNA. *Journal of Cellular and Molecular Medicine*, 13(1), 39-53. <https://doi.org/10.1111/j.1582-4934.2008.00556.x>
- Krol, J., Loedige, I., & Filipowicz, W. (2010). The widespread regulation of microRNA biogenesis, function and decay. *Nature Reviews Genetics*, 11(9), 597-610. <https://doi.org/10.1038/nrg2843>
- Kumarswamy, R., Volkmann, I., & Thum, T. (2011). Regulation and function of miRNA-21 in health and disease. *RNA Biology*, 8(5), 706-713. <https://doi.org/10.4161/rna.8.5.16154>
- Kwak, P. B., & Tomari, Y. (2012). The N domain of Argonaute drives duplex unwinding during RISC assembly. *Nature Structural & Molecular Biology*, 19(2), 145-151. <https://doi.org/10.1038/nsmb.2232>
- Kwon, S. C., Baek, S. C., Choi, Y.-G., Yang, J., Lee, Y., Woo, J.-S., & Kim, V. N. (2019). Molecular Basis for the Single-Nucleotide Precision of Primary microRNA Processing. *Molecular Cell*, 73(3), 505-518.e5. <https://doi.org/10.1016/j.molcel.2018.11.005>
- Lai, E. C., Tomancak, P., Williams, R. W., & Rubin, G. M. (2003). [No title found]. *Genome Biology*, 4(7), R42. <https://doi.org/10.1186/gb-2003-4-7-r42>
- Lee, R. C., Feinbaum, R. L., & Ambros, V. (1993). The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell*, 75(5), 843-854. [https://doi.org/10.1016/0092-8674\(93\)90529-Y](https://doi.org/10.1016/0092-8674(93)90529-Y)
- Li, S.-C., Liao, Y.-L., Ho, M.-R., Tsai, K.-W., Lai, C.-H., & Lin, W. (2012). MiRNA arm selection and isomiR distribution in gastric cancer. *BMC Genomics*, 13(Suppl 1), S13. <https://doi.org/10.1186/1471-2164-13-S1-S13>
- Lim, L. P. (2003a). Vertebrate MicroRNA Genes. *Science*, 299(5612), 1540-1540. <https://doi.org/10.1126/science.1080372>
- Lim, L. P. (2003b). The microRNAs of *Caenorhabditis elegans*. *Genes & Development*, 17(8), 991-1008. <https://doi.org/10.1101/gad.1074403>
- Liu, J., Rivas, F. V., Wohlschlegel, J., Yates, J. R., Parker, R., & Hannon, G. J. (2005). A role for the P-body component GW182 in microRNA function. *Nature Cell Biology*, 7(12), 1261-1266. <https://doi.org/10.1038/ncb1333>

- Liu, Z.-L., Wang, H., Liu, J., & Wang, Z.-X. (2013). MicroRNA-21 (miR-21) expression promotes growth, metastasis, and chemo- or radioresistance in non-small cell lung cancer cells by targeting PTEN. *Molecular and Cellular Biochemistry*, 372(1-2), 35-45.
<https://doi.org/10.1007/s11010-012-1443-3>
- Lo, T.-F., Tsai, W.-C., & Chen, S.-T. (2013). MicroRNA-21-3p, a berberine-induced miRNA, directly down-regulates human methionine adenosyltransferases 2A and 2B and inhibits hepatoma cell growth. *PloS One*, 8(9), e75628. <https://doi.org/10.1371/journal.pone.0075628>
- Löffler, D., Brocke-Heidrich, K., Pfeifer, G., Stocsits, C., Hackermüller, J., Kretzschmar, A. K., Burger, R., Gramatzki, M., Blumert, C., Bauer, K., Cvijic, H., Ullmann, A. K., Stadler, P. F., & Horn, F. (2007). Interleukin-6 dependent survival of multiple myeloma cells involves the Stat3-mediated induction of microRNA-21 through a highly conserved enhancer. *Blood*, 110(4), 1330-1333. <https://doi.org/10.1182/blood-2007-03-081133>
- Loh, H.-Y., Norman, B. P., Lai, K.-S., Rahman, N. M. A. N. Abd., Alitheen, N. B. M., & Osman, M. A. (2019). The Regulatory Role of MicroRNAs in Breast Cancer. *International Journal of Molecular Sciences*, 20(19), 4940. <https://doi.org/10.3390/ijms20194940>
- Lou, W., Liu, J., Gao, Y., Zhong, G., Chen, D., Shen, J., Bao, C., Xu, L., Pan, J., Cheng, J., Ding, B., & Fan, W. (2017). MicroRNAs in cancer metastasis and angiogenesis. *Oncotarget*, 8(70), 115787-115802. <https://doi.org/10.18632/oncotarget.23115>
- Medley, J. C., Panzade, G., & Zinovyeva, A. Y. (2021). microRNA strand selection: Unwinding the rules. *WIREs RNA*, 12(3). <https://doi.org/10.1002/wrna.1627>
- Meijer, H. A., Smith, E. M., & Bushell, M. (2014). Regulation of miRNA strand selection: Follow the leader? *Biochemical Society Transactions*, 42(4), 1135-1140.
<https://doi.org/10.1042/BST20140142>
- Memmott, R. M., & Dennis, P. A. (2009). Akt-dependent and -independent mechanisms of mTOR regulation in cancer. *Cellular Signalling*, 21(5), 656-664.
<https://doi.org/10.1016/j.cellsig.2009.01.004>
- Meng, F., Henson, R., Wehbe-Jane, H., Ghoshal, K., Jacob, S. T., & Patel, T. (2007). MicroRNA-21 Regulates Expression of the PTEN Tumor Suppressor Gene in Human

- Hepatocellular Cancer. *Gastroenterology*, 133(2), 647-658.
<https://doi.org/10.1053/j.gastro.2007.05.022>
- Miller, M. A., Sullivan, R. J., & Lauffenburger, D. A. (2017). Molecular Pathways: Receptor Ectodomain Shedding in Treatment, Resistance, and Monitoring of Cancer. *Clinical Cancer Research*, 23(3), 623-629. <https://doi.org/10.1158/1078-0432.CCR-16-0869>
- Mitra, R., Adams, C. M., Jiang, W., Greenawalt, E., & Eischen, C. M. (2020). Pan-cancer analysis reveals cooperativity of both strands of microRNA that regulate tumorigenesis and patient survival. *Nature Communications*, 11(1), 968. <https://doi.org/10.1038/s41467-020-14713-2>
- Mizuno, R., Kawada, K., & Sakai, Y. (2018). The Molecular Basis and Therapeutic Potential of Let-7 MicroRNAs against Colorectal Cancer. *Canadian Journal of Gastroenterology and Hepatology*, 2018, 5769591. <https://doi.org/10.1155/2018/5769591>
- Najjary, S., Mohammadzadeh, R., Mokhtarzadeh, A., Mohammadi, A., Kojabad, A. B., & Baradaran, B. (2020). Role of miR-21 as an authentic oncogene in mediating drug resistance in breast cancer. *Gene*, 738, 144453. <https://doi.org/10.1016/j.gene.2020.144453>
- Oh, J., Takahashi, R., Kondo, S., Mizoguchi, A., Adachi, E., Sasahara, R. M., Nishimura, S., Imamura, Y., Kitayama, H., Alexander, D. B., Ide, C., Horan, T. P., Arakawa, T., Yoshida, H., Nishikawa, S., Itoh, Y., Seiki, M., Itohara, S., Takahashi, C., & Noda, M. (2001). The Membrane-Anchored MMP Inhibitor RECK Is a Key Regulator of Extracellular Matrix Integrity and Angiogenesis. *Cell*, 107(6), 789-800. [https://doi.org/10.1016/S0092-8674\(01\)00597-9](https://doi.org/10.1016/S0092-8674(01)00597-9)
- Olive, V., Li, Q., & He, L. (2013). mir-17-92: A polycistronic oncomir with pleiotropic functions. *Immunological Reviews*, 253(1), 158-166. <https://doi.org/10.1111/imr.12054>
- Pañeda, C., Gorospe, I., Herrera, B., Nakamura, T., Fabregat, I., & Varela-Nieto, I. (2002). Liver cell proliferation requires methionine adenosyltransferase 2A mRNA up-regulation: Liver cell proliferation requires methionine adenosyltransferase 2A mRNA up-regulation. *Hepatology*, 35(6), 1381-1391. <https://doi.org/10.1053/jhep.2002.32538>
- Pfeffer, S. R., Yang, C. H., & Pfeffer, L. M. (2015). The Role of miR-21 in Cancer. *Drug Development Research*, 76(6), 270-277. <https://doi.org/10.1002/ddr.21257>

- Phuong, N. T. T., Kim, S. K., Im, J. H., Yang, J. W., Choi, M. C., Lim, S. C., Lee, K. Y., Kim, Y.-M., Yoon, J. H., & Kang, K. W. (2016). Induction of methionine adenosyltransferase 2A in tamoxifen-resistant breast cancer cells. *Oncotarget*, *7*(12), 13902-13916.
<https://doi.org/10.18632/oncotarget.5298>
- Phuong, N. T. T., Kim, S. K., Lim, S. C., Kim, H. S., Kim, T. H., Lee, K. Y., Ahn, S.-G., Yoon, J.-H., & Kang, K. W. (2011). Role of PTEN promoter methylation in tamoxifen-resistant breast cancer cells. *Breast Cancer Research and Treatment*, *130*(1), 73-83.
<https://doi.org/10.1007/s10549-010-1304-2>
- Rana, T. M. (2007). Illuminating the silence: Understanding the structure and function of small RNAs. *Nature Reviews. Molecular Cell Biology*, *8*(1), 23-36. <https://doi.org/10.1038/nrm2085>
- Reinhart, B. J. (2002). MicroRNAs in plants. *Genes & Development*, *16*(13), 1616-1626.
<https://doi.org/10.1101/gad.1004402>
- Reinhart, B. J., Slack, F. J., Basson, M., Pasquinelli, A. E., Bettinger, J. C., Rougvie, A. E., Horvitz, H. R., & Ruvkun, G. (2000). The 21-nucleotide let-7 RNA regulates developmental timing in *Caenorhabditis elegans*. *Nature*, *403*(6772), 901-906.
<https://doi.org/10.1038/35002607>
- Roden, C., Gaillard, J., Kanoria, S., Rennie, W., Barish, S., Cheng, J., Pan, W., Liu, J., Cotsapas, C., Ding, Y., & Lu, J. (2017). Novel determinants of mammalian primary microRNA processing revealed by systematic evaluation of hairpin-containing transcripts and human genetic variation. *Genome Research*, *27*(3), 374-384. <https://doi.org/10.1101/gr.208900.116>
- Schwarz, D. S., Hutvagner, G., Du, T., Xu, Z., Aronin, N., & Zamore, P. D. (2003). Asymmetry in the Assembly of the RNAi Enzyme Complex. *Cell*, *115*(2), 199-208.
[https://doi.org/10.1016/S0092-8674\(03\)00759-1](https://doi.org/10.1016/S0092-8674(03)00759-1)
- Selcuklu, S. D., Donoghue, M. T. A., & Spillane, C. (2009). MiR-21 as a key regulator of oncogenic processes. *Biochemical Society Transactions*, *37*(4), 918-925.
<https://doi.org/10.1042/BST0370918>
- Shan, S. W., Fang, L., Shatseva, T., Rutnam, Z. J., Yang, X., Du, W., Lu, W.-Y., Xuan, J. W., Deng, Z., & Yang, B. B. (2013). Mature miR-17-5p and passenger miR-17-3p induce

hepatocellular carcinoma by targeting PTEN, GalNT7 and vimentin in different signal pathways. *Journal of Cell Science*, 126(Pt 6), 1517-1530. <https://doi.org/10.1242/jcs.122895>

Sharma, A., Abdelfatah, E., Al Eissa, M., & Ahuja, N. (2016). Chapter 11—Prognostic Epigenetics. En T. O. Tollefsbol (Ed.), *Medical Epigenetics* (pp. 177-195). Academic Press. <https://doi.org/10.1016/B978-0-12-803239-8.00011-9>

Shen, L., Ling, M., Li, Y., Xu, Y., Zhou, Y., Ye, J., Pang, Y., Zhao, Y., Jiang, R., Zhang, J., & Liu, Q. (2013). Feedback Regulations of miR-21 and MAPKs via Pcd4 and Spry1 Are Involved in Arsenite-Induced Cell Malignant Transformation. *PLoS ONE*, 8(3), e57652. <https://doi.org/10.1371/journal.pone.0057652>

Shyu, A.-B., Wilkinson, M. F., & van Hoof, A. (2008). Messenger RNA regulation: To translate or to degrade. *The EMBO Journal*, 27(3), 471-481. <https://doi.org/10.1038/sj.emboj.7601977>

Svoronos, A. A., Engelman, D. M., & Slack, F. J. (2016). OncomiR or Tumor Suppressor? The Duplicity of MicroRNAs in Cancer. *Cancer Research*, 76(13), 3666-3670. <https://doi.org/10.1158/0008-5472.CAN-16-0359>

Takahashi, C., Sheng, Z., Horan, T. P., Kitayama, H., Maki, M., Hitomi, K., Kitaura, Y., Takai, S., Sasahara, R. M., Horimoto, A., Ikawa, Y., Ratzkin, B. J., Arakawa, T., & Noda, M. (1998). Regulation of matrix metalloproteinase-9 and inhibition of tumor invasion by the membrane-anchored glycoprotein RECK. *Proceedings of the National Academy of Sciences*, 95(22), 13221-13226. <https://doi.org/10.1073/pnas.95.22.13221>

The passenger strand, miR-21-3p, plays a role in mediating cisplatin resistance in ovarian cancer cells—ScienceDirect. (s. f.). Recuperado 30 de junio de 2020, de <https://www.sciencedirect.com/science/article/abs/pii/S0090825815000049>

Uchino, K., Takeshita, F., Takahashi, R., Kosaka, N., Fujiwara, K., Naruoka, H., Sonoike, S., Yano, J., Sasaki, H., Nozawa, S., Yoshiike, M., Kitajima, K., Chikaraishi, T., & Ochiya, T. (2013). Therapeutic Effects of MicroRNA-582-5p and -3p on the Inhibition of Bladder Cancer Progression. *Molecular Therapy*, 21(3), 610-619. <https://doi.org/10.1038/mt.2012.269>

UCSC In-Silico PCR. (s. f.). Recuperado 25 de mayo de 2020, de <https://genome.ucsc.edu/cgi-bin/hgPcr>

- Wang, B. (2013). Base Composition Characteristics of Mammalian miRNAs. *Journal of Nucleic Acids*, 2013, 1-6. <https://doi.org/10.1155/2013/951570>
- Wang, H., Tan, Z., Hu, H., Liu, H., Wu, T., Zheng, C., Wang, X., Luo, Z., Wang, J., Liu, S., Lu, Z., & Tu, J. (2019). MicroRNA-21 promotes breast cancer proliferation and metastasis by targeting LZTFL1. *BMC Cancer*, 19. <https://doi.org/10.1186/s12885-019-5951-3>
- Wietecha, M. S., Chen, L., Ranzer, M. J., Anderson, K., Ying, C., Patel, T. B., & DiPietro, L. A. (2011). Sprouty2 downregulates angiogenesis during mouse skin wound healing. *American Journal of Physiology-Heart and Circulatory Physiology*, 300(2), H459-H467. <https://doi.org/10.1152/ajpheart.00244.2010>
- Wilson, R. C., Tambe, A., Kidwell, M. A., Noland, C. L., Schneider, C. P., & Doudna, J. A. (2015). Dicer-TRBP Complex Formation Ensures Accurate Mammalian MicroRNA Biogenesis. *Molecular Cell*, 57(3), 397-407. <https://doi.org/10.1016/j.molcel.2014.11.030>
- Xu, J., Wu, D., Wang, S., & Wang, Z. (2019). MAT2B expression correlates with poor prognosis in triple-negative breast cancer. *Cancer Management and Research*, Volume 11, 5501-5511. <https://doi.org/10.2147/CMAR.S200716>
- Xu, L., Wu, Z., Chen, Y., Zhu, Q., Hamidi, S., & Navab, R. (2014). MicroRNA-21 (miR-21) Regulates Cellular Proliferation, Invasion, Migration, and Apoptosis by Targeting PTEN, RECK and Bcl-2 in Lung Squamous Carcinoma, Gejiu City, China. *PLoS ONE*, 9(8), e103698. <https://doi.org/10.1371/journal.pone.0103698>
- Yan, L.-X., Huang, X.-F., Shao, Q., Huang, M.-Y., Deng, L., Wu, Q.-L., Zeng, Y.-X., & Shao, J.-Y. (2008). MicroRNA miR-21 overexpression in human breast cancer is associated with advanced clinical stage, lymph node metastasis and patient poor prognosis. *RNA*, 14(11), 2348-2360. <https://doi.org/10.1261/rna.1034808>
- Ye, J., Coulouris, G., Zaretskaya, I., Cutcutache, I., Rozen, S., & Madden, T. L. (2012). Primer-BLAST: A tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics*, 13, 134. <https://doi.org/10.1186/1471-2105-13-134>
- Yoda, M., Kawamata, T., Paroo, Z., Ye, X., Iwasaki, S., Liu, Q., & Tomari, Y. (2010). ATP-dependent human RISC assembly pathways. *Nature structural & molecular biology*, 17(1), 17-23. <https://doi.org/10.1038/nsmb.1733>

- Zeng, Y., Yi, R., & Cullen, B. R. (2003). MicroRNAs and small interfering RNAs can inhibit mRNA expression by similar mechanisms. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(17), 9779-9784.
<https://doi.org/10.1073/pnas.1630797100>
- Zhang, B., Pan, X., Cobb, G. P., & Anderson, T. A. (2007). MicroRNAs as oncogenes and tumor suppressors. *Developmental Biology*, *302*(1), 1-12.
<https://doi.org/10.1016/j.ydbio.2006.08.028>
- Zhang, X., Tang, N., Hadden, T. J., & Rishi, A. K. (2011). Akt, FoxO and regulation of apoptosis. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, *1813*(11), 1978-1986.
<https://doi.org/10.1016/j.bbamcr.2011.03.010>
- Zhang, Z., Pi, J., Zou, D., Wang, X., Xu, J., Yu, S., Zhang, T., Li, F., Zhang, X., Zhao, H., Wang, F., Wang, D., Ma, Y., & Yu, J. (2019). MicroRNA arm-imbalance in part from complementary targets mediated decay promotes gastric cancer progression. *Nature Communications*, *10*(1), 4397. <https://doi.org/10.1038/s41467-019-12292-5>
- Zhao, Z., Cai, Y., Wang, Y., Xia, C., Li, C., Chen, S., Yang, Q., & Chen, C. (2015). [Effects of miRNA-21 on paclitaxel-resistance in human breast cancer cells]. *Zhejiang Da Xue Xue Bao. Yi Xue Ban = Journal of Zhejiang University. Medical Sciences*, *44*(4), 400-409.

13 Anexo

Tabla A1. Información de los oligonucleótidos diseñados para los experimentos de PCR en tiempo real.

Nombre	Secuencia 5´- 3´	Longitud (bases)	TM (°C)	Longitud de amplicón (bases)
PTEN_132 FW PTEN_132 RV	GCT GTG GTT GCC ACA AAG TG GGT AGA AGG CAA CTC TGC CA	20 20	57.5 57.1	132
PTEN_87 FW PTEN_87 RV	CTG TGG TTG CCA CAA AGT GC TTA GTG CAC AGT TCC ACC CC	20 20	57.5 57.3	87
PDCD4_138 FW PDCD4_138 RV	GGC AAA AAG GCG ACT AAG GAA GAT CGC CTA TCC AGC AAC CT	21 20	55.8 57	138
PDCD4_143 FW PDCD4_143 RV	CCA AGG CAA AAA GGC GAC TAA GGA TCG CCT ATC CAG CAA CC	21 20	56 58.1	143
MAT2A_101 FW MAT2A_101 RV	ACG AGT AGA ACG CTG TCC G GTT CAT GTT GGT GTC GGT GTG	19 21	56.8 56.6	101
MAT2A_110 FW MAT2A_110 RV	TAC GAG TAG AAC GCT GTC CG AGC TGT CCG TTC ATG TTG GT	20 20	56.2 57	110
MAT2B_130 FW MAT2B_130 RV	AGG AGC TGC TGT TTT GAG GAT CTG CTG CCA GTG ATC CAT GT	21 20	57 56.5	130
MAT2B_79 FW MAT2B_79 RV	GCA CAA CGT CCG AGA AAT GC TTC GAA ATG GTG TTC GTT GGC	20 21	57.9 57	79

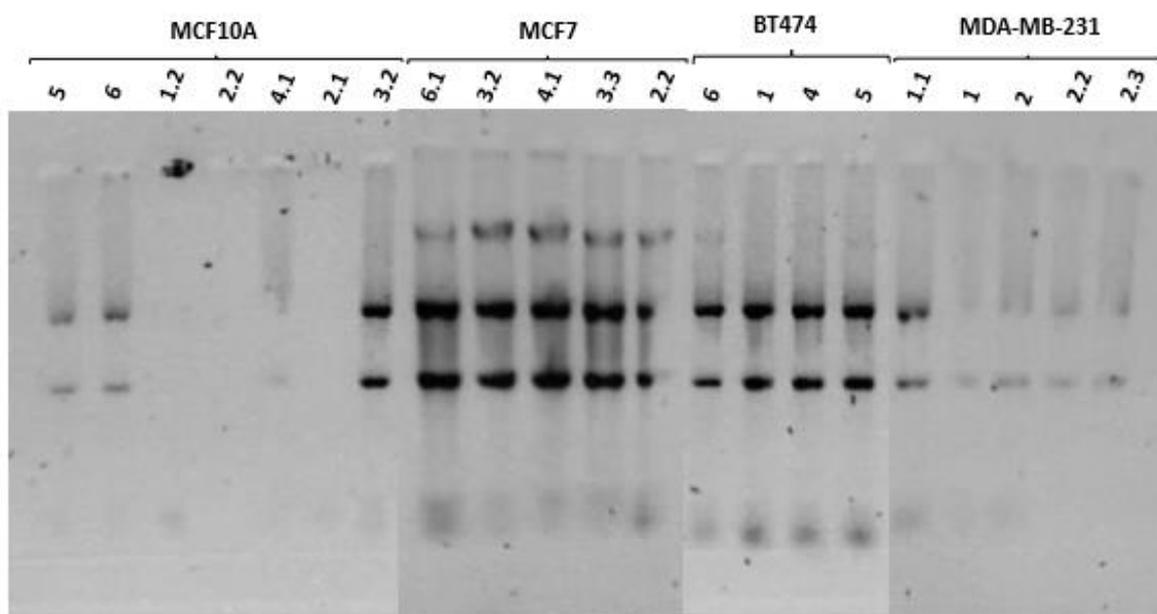


Figura A1. Integridad del RNA total de líneas celulares de mama evaluada mediante un gel agarosa. Gel de agarosa al 1.2% teñido con *Gel-Red*, donde se muestra la integridad del RNA total. Se observan los RNA ribosomales 28s (4 Kb) y 18s (2Kb) del RNA total extraído de las líneas celulares MCF-10A, MCF7, BT474 y MDA-MB-231. El RNA total que se consideró íntegro muestra bandas de RNA ribosomal 28S y 18S, con una relación 2: 1 (28S: 18S).

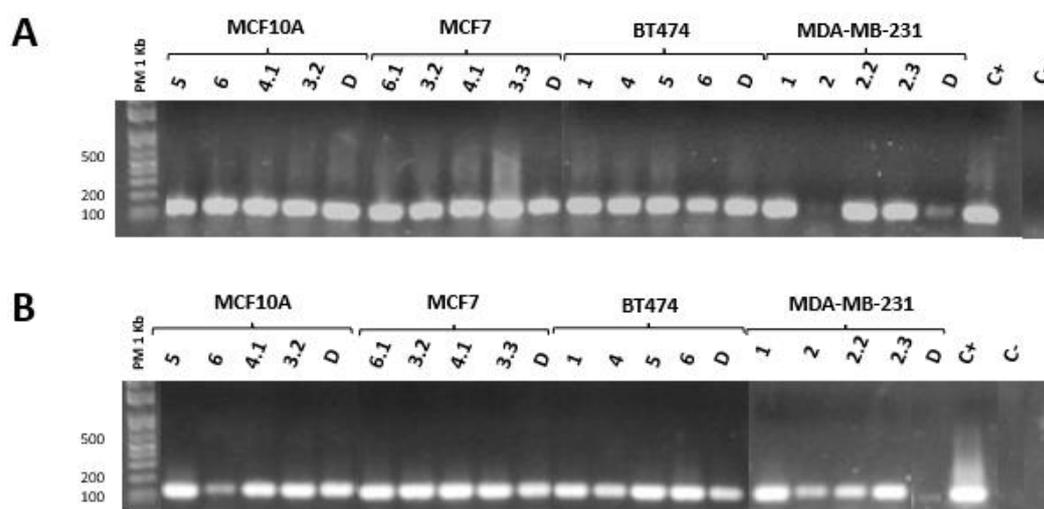


Figura A2. Determinación de la calidad del cDNA sintetizado a partir del RNA extraído de líneas celulares de mama. Resultados de la amplificación del gen constitutivo *RPS28* por medio de reacción en cadena de la polimerasa punto final. El producto de PCR se corrió en un gel de Agarosa al 1% teñido con Gel Red 1X, a 90V por 40 minutos. **A)** cDNA sintetizado con kit cDNA High Capacity para pri-

miRNA TaqMan Assay **B)** cDNA sintetizado con kit de cDNA synthesis miRNA TaqMan Assay. Se considera que el cDNA sintetizado tiene buena calidad cuando es posible visualizar el amplicón de aproximadamente 100 pb. Se muestra la calidad del cDNA en donde se observa el amplicón de aproximadamente 100 pb. Para la amplificación del control positivo (C+) se utilizó cDNA de una línea celular previamente comprobada su calidad (LnCAP) y el control negativo (C-) con agua libre de RNAsas.

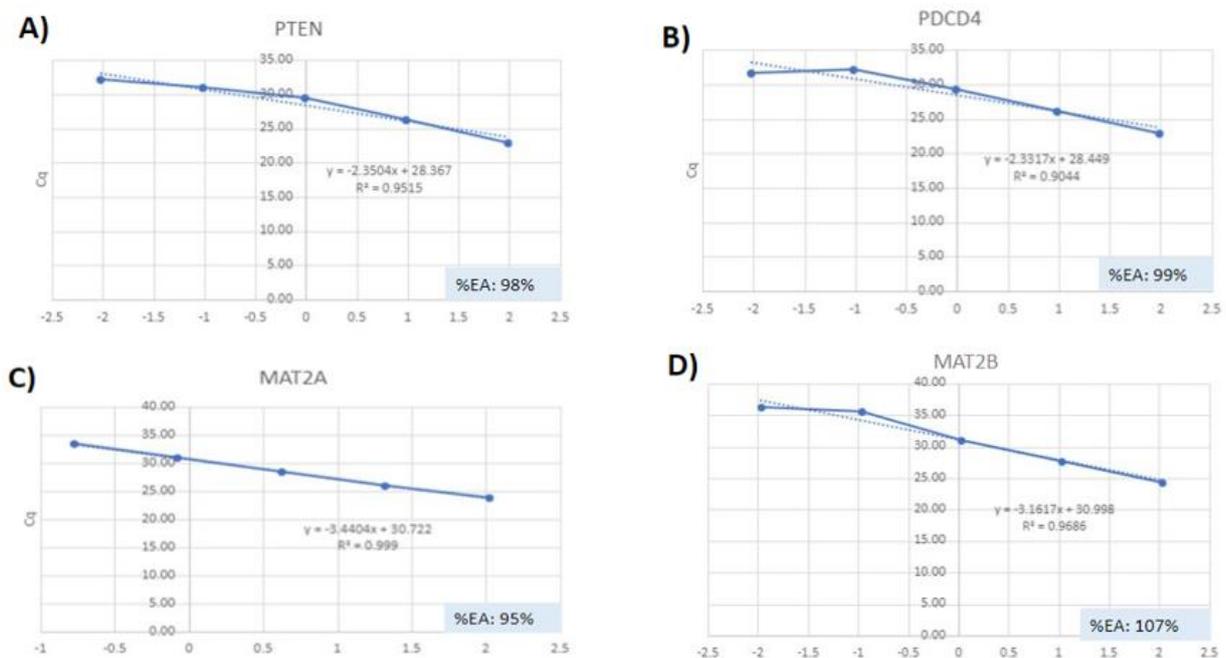


Figura A3. Eficiencia de amplificación de los oligonucleótidos de los genes PTEN, MAT2A, MAT2B y PDCD4 blanco. A) Curva de eficiencia de amplificación del par de oligonucleótidos PTEN_87_FW y PTEN_87_RV, para el gen *PTEN*, amplificando un amplicón de 87 pares de bases, presentando una eficiencia del 98%. B) Curva de eficiencia de amplificación de los oligonucleótidos para el gen *PDCD4*, PDCD4_143_FW y PDCD4_143_RV, obteniendo un amplicón de 143 pares de bases, presentando una eficiencia del 99%. C) Curva de eficiencia de amplificación de los oligonucleótidos para el gen *MAT2A*, MAT2A_110_FW y MAT2A_110_RV, mediante los cuales se obtiene un amplicón de 110 pares de bases, y una eficiencia del 95%. D) Curva de eficiencia de amplificación del par de oligonucleótidos MAT2B_79_FW y MAT2B_79_RV, para el gen *MAT2B* amplificando un amplicón de 79 pares de bases, presentando una eficiencia de 107%. La determinación de la eficiencia de amplificación se realizó con cDNA de la línea

celular MCF-7, utilizando 4 diluciones seriadas: 1:10, 1:100, 1:1000 y 1:10000. El análisis de la eficiencia de amplificación se realizó con la herramienta *qPCR Efficiency Calculator* disponible en la página de internet de *Thermo Fisher Scientific* (<https://www.thermofisher.com/mx/es/home/brands/thermo-scientific/molecular-biology/molecular-biology-learning-center/molecular-biology-resource-library/thermo-scientific-web-tools/qpcr-efficiency-calculator.html>)

Apéndice A: Habilidades aprendidas durante la realización de este proyecto de tesis

Durante la realización de este proyecto, las habilidades adquiridas fueron:

- **Análisis Bioinformático:**
 - Manejo de herramientas bioinformáticas en el servidor Galaxy y en el lenguaje de programación R para realizar el alineamiento de secuencias al genoma de referencia, cuantificación de transcritos, análisis de expresión diferencial.
 - Manejo de lenguaje R para análisis estadístico.
 - Diseño de oligonucleótidos
 - Manejo de plataformas de acceso público como TACCO, RNAinteractome, RAIN
 - Búsqueda de genes blanco específicos para miRNAs
 - Cuantificación relativa mediante PCR en tiempo real y por los métodos ΔCq y $\Delta\Delta Cq$.

- **Habilidades Experimentales**
 - Cultivo celular
 - Extracción de RNA total de líneas celulares
 - Cuantificación de RNA y análisis de integridad del RNA
 - Síntesis de cDNA mediante RT-PCR punto final y análisis de calidad del cDNA sintetizado.
 - PCR en tiempo real para la cuantificación relativa de expresión génica.

Apéndice B: Artículo en el que se participó como co-autora: Saliva is a reliable and accessible source for the detection of SARS-CoV-2 publicado en 2021.



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



Saliva is a reliable and accessible source for the detection of SARS-CoV-2

Luis A. Herrera^{a,b,*}, Alfredo Hidalgo-Miranda^{c,*}, Nancy Reynoso-Noverón^b, Abelardo A. Meneses-García^d, Alfredo Mendoza-Vargas^e, Juan P. Reyes-Grajeda^e, Felipe Vadillo-Ortega^f, Alberto Cedro-Tanda^a, Fernando Peñaloza^b, Emmanuel Frías-Jimenez^a, Cristian Arriaga-Canon^b, Rosaura Ruiz^g, Ofelia Angulo^g, Imelda López-Villaseñor^h, Carlos Amador-Bedollaⁱ, Diana Vilar-Compte^d, Patricia Cornejo^d, Mireya Cisneros-Villanueva^c, Eduardo Hurtado-Cordova^c, Mariana Cendejas-Orozco^c, José S. Hernández-Morales^e, Bernardo Moreno^b, Irwin A. Hernández-Cruz^b, César A. Herrera^b, Francisco García^b, Miguel A. González-Woge^b, Paulina Munguía-Garza^b, Fernando Luna-Maldonado^b, Antonia Sánchez-Vizcarra^b, Vincent G. Osnaya^b, Nelly Medina-Molotla^e, Yair Alfaro-Mora^b, Rodrigo E. Cáceres-Gutiérrez^b, Laura Tolentino-García^b, Patricia Rosas-Escobar^e, Sergio A. Román-González^a, Marco A. Escobar-Arrazola^b, Julio C. Canseco-Méndez^e, Diana R. Ortiz-Soriano^a, Julieta Domínguez-Ortiz^b, Ana D. González-Barrera^a, Diana I. Aparicio-Bautista^a, Armando Cruz-Rangel^a, Ana Paula Alarcón-Zendejas^b, Laura Contreras-Espinosa^b, Rodrigo González^b, Lissania Guerra-Calderas^b, Marco A. Meraz-Rodríguez^b, Michel Montalvo-Casimiro^b, Rogelio Montiel-Manríquez^b, Karla Torres-Arciga^b, Daniela Venegas^b, Vasti Juárez-González^b, Xiadani Guajardo-Barreto^b, Verónica Monroy-Martínez^h, Daniel Guillén^h, Jacquelinna Fernández^h, Juliana Herrera^h, Renato León-Rodríguez^h, Israel Canela-Pérez^h, Blanca H. Ruíz-Ordaz^h, Rafael Valdez-Vazquez^j, Jennifer Bertin-Montoya^j, María Niembro-Ortega^j, Liudmila Villegas-Acosta^j, Daniela López-Castillo^j, Andrea Soriano-Ríos^j, Michael Gastelum-Ramos^j, Tonatiuh Zamora-Barandas^j, Jorge Morales-Baez^j, María García-Rodríguez^k, Mariano García-Martínez^k, Erik Nieto-Patlán^k, Maricarmen Quirasco-Baruch^l, Irma López-Martínez^m, Ernesto Ramírez-Gonzalez^m, Hiram Olivera-Díaz^m, Noe Escobar-Escamilla^m

^a Instituto Nacional de Medicina Genómica, INMEGEN, Mexico City, Mexico

^b Unidad de Investigación Biomédica en Cáncer, Instituto Nacional de Cancerología-Instituto de Investigaciones Biomédicas, UNAM, Mexico City, Mexico

^c Laboratorio de Genómica del Cáncer, Instituto Nacional de Medicina Genómica, INMEGEN, Mexico City, Mexico

^d Instituto Nacional de Cancerología, INCAN, Mexico City, Mexico

^e Unidad de Secuenciación, Instituto Nacional de Medicina Genómica, INMEGEN, Mexico City, Mexico

^f Unidad de Vinculación Científica Facultad de Medicina-INMEGEN, Mexico City, Mexico

^g Secretaría de Educación, Ciencia, Tecnología e Innovación, Mexico City, Mexico

^h Instituto de Investigaciones Biomédicas, Universidad Nacional Autónoma de México, Mexico City, Mexico

ⁱ Departamento de Física y Química Teórica, Facultad de Química, Universidad Nacional Autónoma de México, Mexico City, Mexico

^j Unidad Temporal COVID-19, Mexico City, Mexico

^k Unidad de Investigación Preclínica, Facultad de Química, Mexico City, Mexico

^l Departamento de Alimentos y Biotecnología, Facultad de Química, Universidad Nacional Autónoma de México, Mexico City, Mexico

^m Instituto de Diagnóstico y Referencia Epidemiológicos, InDRE, Mexico City, Mexico

* Corresponding authors at: National Institute of Genomic Medicine, Periferico Sur 4809, Arenal Tepepan, 14610 Mexico City, Mexico.
E-mail addresses: lherrera@inmegen.gob.mx (L.A. Herrera), ahidalgo@inmegen.gob.mx (A. Hidalgo-Miranda).

ARTICLE INFO

Article history:

Received 22 October 2020

Received in revised form 2 February 2021

Accepted 4 February 2021

Keywords:

COVID-19

SARS-CoV-2

Diagnostic test

Saliva testing

Pooling strategy

ABSTRACT

Objectives: The aim of this study was to investigate the feasibility of saliva sampling as a non-invasive and safer tool to detect severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and to compare its reproducibility and sensitivity with nasopharyngeal swab samples (NPS). The use of sample pools was also investigated.

Methods: A total of 2107 paired samples were collected from asymptomatic healthcare and office workers in Mexico City. Sixty of these samples were also analyzed in two other independent laboratories for concordance analysis. Sample processing and analysis of virus genetic material were performed according to standard protocols described elsewhere. A pooling analysis was performed by analyzing the saliva pool and the individual pool components.

Results: The concordance between NPS and saliva results was 95.2% ($\kappa = 0.727$, $p = 0.0001$) and 97.9% without considering inconclusive results ($\kappa = 0.852$, $p = 0.0001$). Saliva had a lower number of inconclusive results than NPS (0.9% vs 1.9%). Furthermore, saliva showed a significantly higher concentration of both total RNA and viral copies than NPS. Comparison of our results with those of the other two laboratories showed 100% and 97% concordance. Saliva samples are stable without the use of any preservative, and a positive SARS-CoV-2 sample can be detected 5, 10, and 15 days after collection when the sample is stored at 4 °C.

Conclusions: The study results indicate that saliva is as effective as NPS for the identification of SARS-CoV-2-infected asymptomatic patients. Sample pooling facilitates the analysis of a larger number of samples, with the benefit of cost reduction.

© 2021 The Author(s). Published by Elsevier Ltd on behalf of International Society for Infectious Diseases. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Introduction

The rapid spread of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) worldwide has generated considerable demand for medical supplies for use in fighting the pandemic. Among other problems, this has resulted in a shortage of nasopharyngeal swabs (NPS) and tests for the detection of SARS-CoV-2. Scarce consumables and invasive sample collection, which can expose medical personnel to biohazards, are obstacles to effective mass screening of the population to identify infected individuals. Mass screening is essential to identify and isolate infected individuals during reopening. Additionally, fast massive effective screening is essential in the event of a coronavirus disease 2019 (COVID-19) resurgence and for the safe return to productive activities, an approach that has been implemented by several governments around the globe. Although this situation has been addressed using different innovative approaches, such as three-dimensional printing of NPS (Callahan et al., 2020), additional solutions for sample collection that are easier and less invasive, with minimal risk to health professionals, together with strategies aiming to maximize the number of samples analyzed, must be explored.

The gold standard test for the diagnosis of SARS-CoV-2 infection involves sample collection via NPS, followed by viral RNA extraction and detection by real-time polymerase chain reaction (RT-qPCR). Recent reports have indicated that saliva is a viable option for testing with several potential advantages over NPS, including that it is a less invasive procedure, making it more viable for repeated testing. Furthermore, saliva can be self-collected by the patient with minimal guidance and intervention from healthcare personnel (Azzi et al., 2020). SARS-CoV-2 can be detected in more than 95% of saliva samples, and the virus can be cultured from saliva samples (To et al., 2020b). Detection of the virus in saliva has also been used to monitor viral load dynamics over time, indicating that the highest viral load in saliva presents during the first week after symptom onset and then declines over time (To et al., 2020a). Recently, the Food and Drug Administration (FDA) in the United States approved the first diagnostic test with the option for saliva sampling for SARS-CoV-2 detection (U.S. Food and Drug Administration, 2020a). Another study found that the home-based

collection method of saliva, supervised by a clinician, performed similarly to or even better than NPS for the detection of infection (Noah et al., 2020). These findings were confirmed by recent studies, which found that saliva is more sensitive for SARS-CoV-2 detection than NPS in patients with COVID-19 (Wyllie et al., 2020). In another report, 229 paired samples from 95 patients also showed a high concordance and no significant temporal variation in viral load between the two sample types (Cheuk et al., 2020).

The combined advantages offered by saliva sampling and sample pooling result in an inexpensive diagnostic procedure suitable for assaying large numbers of samples, as has been required during the current pandemic (Abdalhamid et al., 2020; Yelin et al., 2020). Sample pooling has proven its efficacy in different applications, including retrospective testing (Hogan et al., 2020) and, more importantly, in large-scale screening of asymptomatic populations (Ben-Ami et al., 2020; Lohse et al., 2020; U.S. Food and Drug Administration, 2020b). There is work showing that pooling saliva samples for the detection of SARS-CoV-2 provides a mechanism to support testing for a greater number of individuals with substantial cost savings, especially at lower prevalence levels (Pasomsub et al., 2020a, 2020b; Watkins et al., 2020). Mirimus Clinical Labs in their SalivaClear test already use the pooling strategy to monitor and detect infections in groups of symptomatic and asymptomatic individuals (SalivaClear by Mirimus Clinical, 2020).

This study was performed to compare the reproducibility, accuracy, and feasibility of saliva sampling using NPS followed by RT-qPCR for the detection of SARS-CoV-2 in paired samples from asymptomatic clinical and laboratory personnel working in two Mexico National Institutes of Health laboratories and from asymptomatic office workers ($N = 2107$ individuals). This study presents evidence that saliva sample pooling is a reliable and inexpensive method that allows for the screening of a large number of samples.

Materials and methods

Participants

A cross-sectional study design was used to collect samples from personnel engaged in clinical and laboratory activities at Mexico's

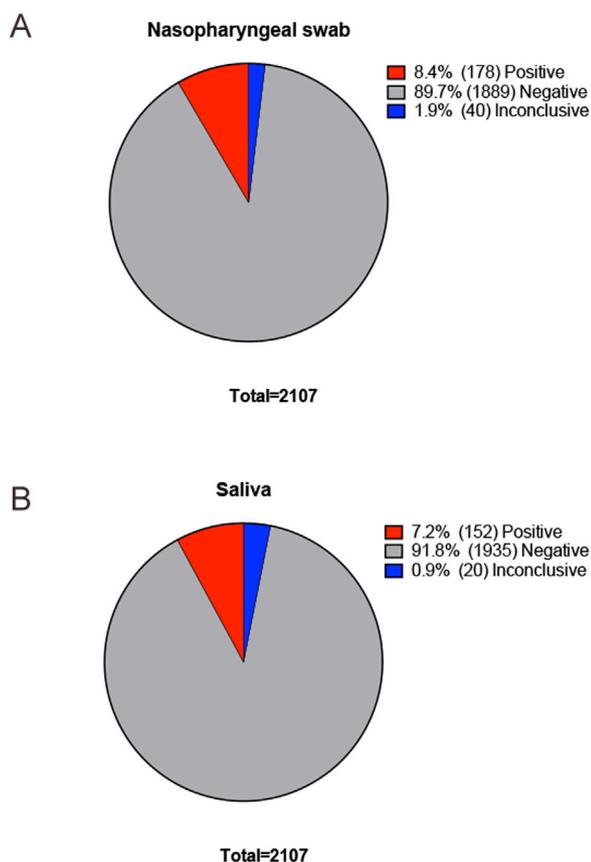


Figure 1. Frequencies and percentages of positive, negative, and inconclusive samples in 2107 paired nasopharyngeal swab and saliva samples: (A) nasopharyngeal swabs; (B) saliva samples.

National Cancer Institute and National Institute of Genomic Medicine. Consecutive asymptomatic subjects were sampled after signing an informed consent form. The study was approved by the ethics and research committees of both institutes (CEI/1479/20 and CEI 2020/21). Paired saliva and NPS samples were collected from 2107 asymptomatic healthcare and office workers to compare the two sample sources for SARS-CoV-2 detection. Additionally, saliva samples were collected from 3983 asymptomatic office workers, 2126 asymptomatic healthcare personnel, and 846 symptomatic office workers to detect SARS-CoV-2.

Table 1

Detection of SARS-CoV-2 by RT-qPCR between nasopharyngeal swab and saliva samples: (A) positive, negative, and inconclusive samples; (B) only positive and negative samples.

A		Nasopharyngeal swab			Total
		Positive	Negative	Inconclusive	
Saliva	Positive	139 (6.6%)	10 (0.5%)	3 (0.1%)	152 (7.2%)
	Negative	34 (1.6%)	1867 (88.6%)	34 (1.6%)	1935 (91.8%)
	Inconclusive	5 (0.2%)	12 (0.6%)	3 (0.1%)	20 (0.9%)
	Total	178 (8.4%)	1889 (89.7%)	40 (1.9%)	2107 (100%)
B		Nasopharyngeal swab		Total	
		Positive	Negative		
Saliva	Positive	139 (6.8%)	10 (0.5%)	149 (7.3%)	
	Negative	34 (1.7%)	1867 (91.1%)	1901 (92.7%)	
	Total	173 (8.4%)	1877 (91.6%)	2050 (100%)	

Prevalence positive test = 8.44% (95% CI 7.27–9.73%), sensitivity = 80.35% (95% CI 73.63–85.99%), specificity = 99.47% (95% CI 99.02–99.74%), positive predictive value = 93.29% (95% CI 88.18–96.28%), negative predictive value = 98.21% (95% CI 97.60–98.67%), positive likelihood ratio = 150.81 (95% CI 80.92–281.06), negative likelihood ratio = 0.20 (95% CI 0.15–0.27).

Sample collection

NPS were collected by a trained clinician with a flexible nylon swab that was inserted through the patient’s nostrils to reach the posterior nasopharynx. It was left in place for several seconds and slowly removed while rotating. The swab was then placed in 3 mL of sterile viral transport medium. Swabs from both nostrils were deposited in a single viral transport tube. Saliva samples were self-collected by the individuals without any stimulation and without rinsing the mouth before sample collection. Five milliliters of saliva was collected in a 50-ml sterile conical centrifuge tube without preservatives. Sample collection was done within the same facilities where the viral diagnosis laboratory is located. They were also collected from nearby hospitals. As a result, the swabs and the saliva samples were processed for viral RNA extraction within 5 h after collection.

SARS-CoV-2 RNA extraction and detection

Total nucleic acid was extracted from 300 µL of viral transport medium from the NPS or 300 µL of whole saliva using the MagMAX Viral/Pathogen Nucleic Acid Isolation Kit (Thermo Fisher Scientific) and eluted into 75 µL of elution buffer. For SARS-CoV-2 RNA detection, 5 µL of RNA template was tested using the US CDC real-time RT-qPCR primer/probe sets for 2019-nCoV_N1 and 2019-nCoV_N2 and human RNase P (RP) as an extraction control. Samples were classified as positive for SARS-CoV-2 when both the N1 and N2 primer/probe sets were detected with a Ct value of less than 40 (Centers for Disease Control and Prevention, 2020). If only one of these genes was detected, the sample was labeled inconclusive. All tests were run on Thermo Fisher ABI QuantStudio 5 or QuantStudio 7 real-time thermal cyclers.

Validation of saliva performance in independent laboratories and different extraction and detection methods

For validation purposes, 60 samples that were analyzed in our laboratory were also processed in two independent authorized laboratories (30 samples in each laboratory: Instituto de Investigaciones Biomédicas and Facultad de Química, Universidad Nacional Autónoma de México) using two additional RNA extraction methods and detection systems. The additional extraction methods consisted of spin-column-based RNA extraction (Total RNA Purification Kit, Jena Biosciences) and the use of a quick extraction solution from Lucigen. The two additional methods for SARS-CoV-2 detection were conducted using the GoTaq Probe 1-

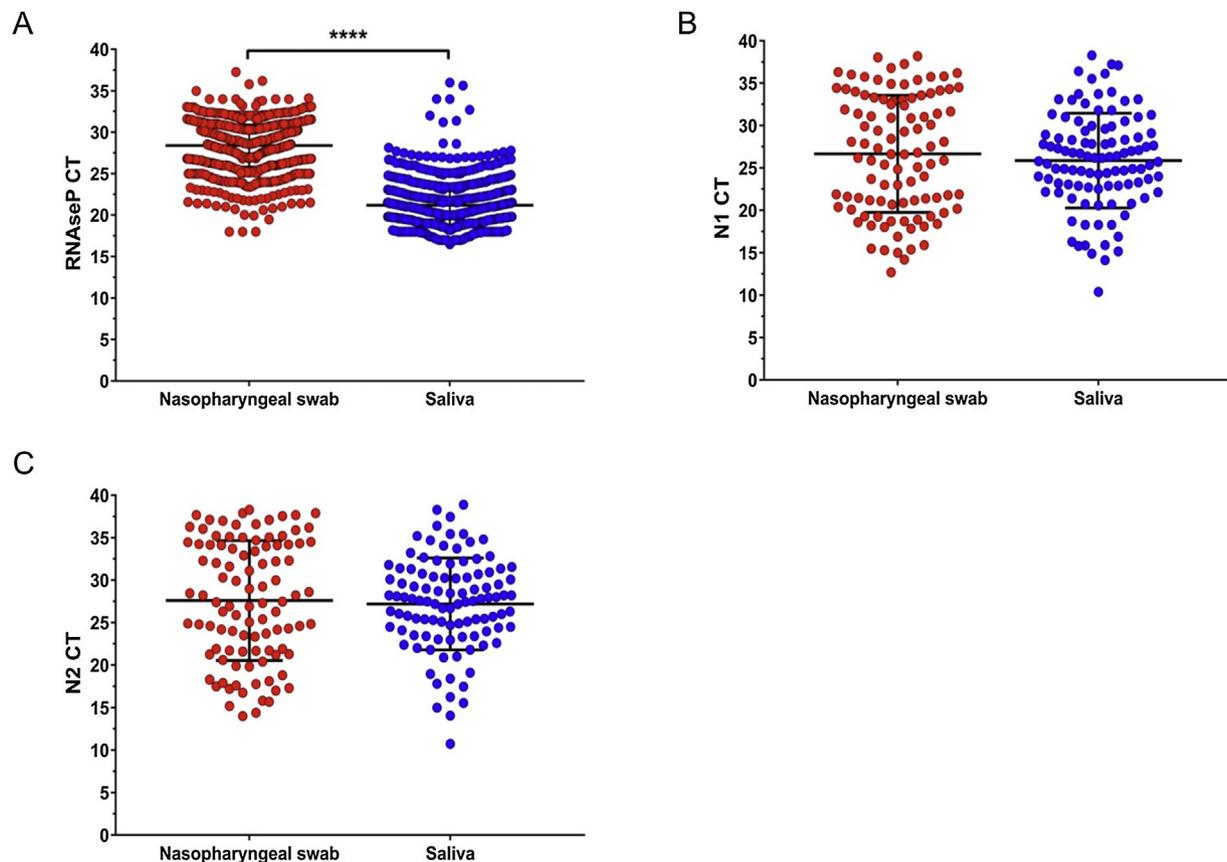


Figure 2. Cycle threshold values (Ct) in nasopharyngeal swab versus saliva. (A) RNase P gene in all samples; (B) N1 in SARS-CoV-2-positive samples; (C) N2 in SARS-CoV-2-positive samples. RNase P had a significantly higher concentration of total RNA in saliva compared to nasopharyngeal swab; $p < 0.00001$ (t -test).

Step RT-qPCR System from Promega on a 7500 ABI system and the Star Q One-step RT-qPCR from Genes2Life.

Viral copy number analysis

Copies of the SARS-CoV-2 virus were quantified using a standard curve with serial dilutions (10-fold) using the 2019-nCoV_N and Hs_RPP30 positive controls synthesized by Integrated DNA Technologies (IDT, Coralville, IA, USA), with the same detection protocol as the clinical samples. The Ct values obtained from each dilution were used to interpolate the Ct value of each gene from the samples and to calculate viral copy numbers.

Stability assay

The stability of viral RNA in saliva for the detection of SARS-CoV-2 over time after sampling was assessed. A second RNA and an RT-qPCR extraction were performed from 150 SARS-CoV-2-positive saliva samples (stored at 4 °C) at 5, 10, and 15 days after the first positive result.

Sample pooling

SARS-CoV-2 detection in the pooling strategy was performed using the DAAN-Gene Kit following the manufacturer's instructions. Briefly, the kit detects the ORF1ab and N genes of the virus. Five microliters of total RNA were used in the RT-qPCR reaction, and Ct values less than 40 were considered positive.

The evaluation of sample pools was conducted for both viral transport medium with NPS and saliva samples collected in

parallel. To test the sensitivity of the pooling strategy, several pools were prepared from saliva and NPS. In the first pooling approach, one positive saliva sample (Ct values of 19.6 and 28.0 for the N viral gene) and its paired positive NPS (Ct values of 18.0 and 36.3 for the N viral gene) were mixed with five and nine known negative samples, respectively. Five hundred microliters of saliva was used from each sample to obtain the pool. This volume was necessary to obtain a homogeneous mixture in the saliva pool, given the differences in viscosity between different samples.

Based on the results obtained from the 10 sample pools and with the premise that asymptomatic individuals might have lower Ct values, which might result in false-negatives in the 10-sample pool, we generated five-sample pools from NPS and saliva from asymptomatic non health-related workers. For NPS, 51 pools made out of 255 individuals were evaluated. For saliva, 26 pools made out of 130 individuals were evaluated.

Statistical analysis

The accuracy of SARS-CoV-2 saliva detection, including sensitivity, specificity, predictive values, and likelihood ratios, was determined using RT-qPCR in NPS as the 'gold standard'. Other statistical analyses were performed using GraphPad Prism 7.0 and IBM SPSS Statistics version 24 software. One-tailed parametric (Student t -test) and non-parametric (Mann-Whitney U -test) statistical tests were used to determine the significance of the data, considering a statistically significant value of $p \leq 0.05$. The kappa coefficient was used to estimate the concordance between saliva and NPS results (McHugh, 2012).

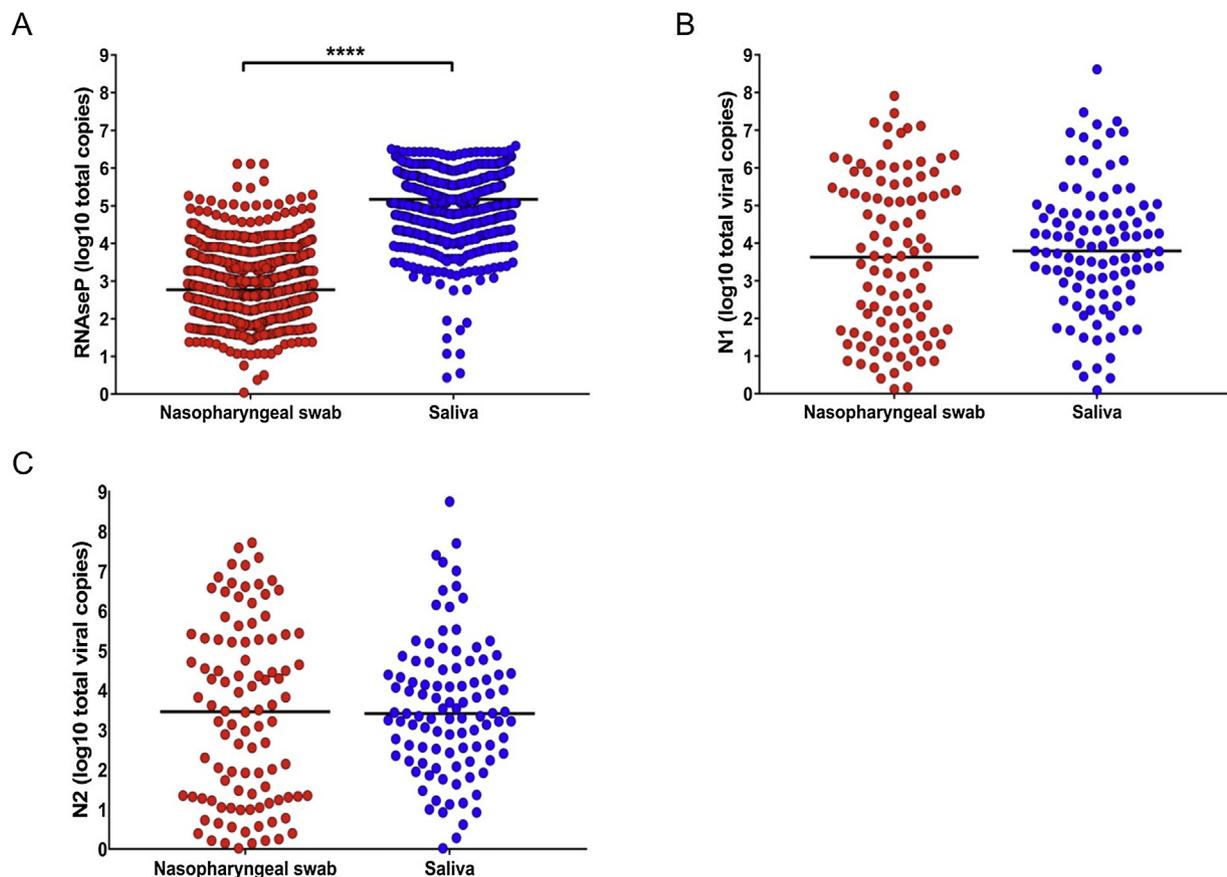


Figure 3. Total viral copies in nasopharyngeal swab versus saliva. (A) RNAse P gene in all samples; (B) N1 in SARS-CoV-2-positive samples; (C) N2 in SARS-CoV-2-positive samples. RNAse P had significantly higher copies in saliva than nasopharyngeal swab; $p < 0.00001$ (Mann-Whitney U -test).

Results and discussion

The design of this study was intended to compare the reproducibility, accuracy, and feasibility of saliva sampling followed by RT-qPCR to identify SARS-CoV-2 and to evaluate the use of saliva in sample pooling strategies. It was a priori accepted that the use of NPS followed by RT-qPCR is the gold standard for identification of the virus, despite current studies showing marked variation in the accuracy of this test.

A total of 2107 paired NPS and saliva samples were included in the analysis. The distribution of the results is described in Figure 1. Concordance between saliva and NPS results was statistically significant (Cohen's kappa 0.727, standard error 0.025; $p = 0.0001$; Table 1A). Concordance improved when inconclusive samples were removed from the analysis (Cohen's kappa 0.852, standard error = 0.022; $p = 0.0001$). Overall, 2006 out of 2050 tests (98%) showed the same results in both saliva and NPS (Table 1B). Saliva had a lower number of inconclusive results than NPS (0.9% vs 1.9%) (Table 1 and Fig. 1).

The concordance between the RT-qPCR results from viral RNA obtained from saliva and NPS was statistically significant, indicating that saliva is at least as sensitive as NPS for SARS-CoV-2 detection. Comparison of both the Ct values and the transcript copies of RNAse P showed a significantly higher total RNA concentration in saliva than in NPS (Figures 2A and 3A). However, when the two viral genes in the positive samples were analyzed, saliva and NPS did not show significant differences in viral load (Figure 2B, C and Figure 3B, C). Spearman correlation analysis of viral copies confirmed that saliva and NPS are both reliable sources for SARS-CoV-2 detection (N1: $r = 0.4217$, $p = 0.0001$; N2: $r = 0.4261$, $p = 0.0001$).

Saliva and paired NPS, which were previously analyzed in our laboratory (60 paired samples), were sent to two independent laboratories for extraction and SARS-CoV-2 detection and processed using different extraction and detection kits. Each laboratory processed 30 paired samples. A 100% concordance was observed in the results between our laboratory and the Instituto de Investigaciones Biomédicas (27 negative and three positive both in saliva and NPS), while 96.7% of the samples sent to Facultad de Química had the same result as in our laboratory (28 negative, one positive, and one discordant). This independent validation is an initial and exploratory assessment.

The accuracy of the saliva test is useful for clinical purposes. The positive likelihood ratio strongly supports its use as a reliable clinical test. A statistically significant correlation and concordance of the RT-qPCR detection of the virus in the saliva samples compared to NPS was identified, and a high concordance between the two types of samples was observed (Table 2A). Given the high number of paired samples analyzed, the results clearly indicate that saliva is as good as NPS for viral detection in the diagnosis of COVID-19, as it has been shown in other studies in hospitalized patients (Table 2B). The data also demonstrated that saliva is stable even without the use of any preservative during sample collection and that a positive SARS-CoV-2 sample can be detected 5, 10, and 15 days after collection when the sample is stored at 4 °C: variation in Ct values in the viral N gene was 0.88 ± 1.92 at 5 days, -0.93 ± 3.01 at 10 days, and -0.76 ± 2.12 at 15 days. Other studies have also demonstrated the stability of saliva for the detection of SARS-CoV-2, with storage for 10–25 days at room temperature (Uwamino et al., 2021) without buffers or stabilizers (Ott et al., 2020).

Table 2

Detection of SARS-CoV-2 in samples of saliva and nasopharyngeal swabs: (A) paired samples; (B) saliva samples only; (C) saliva only in the present study.

A		Paired samples			
Country	Study population	Paired samples	Viral genes	Concordance %	Reference
Australia	Ambulatory patients	522	ORF1a, ORF8	84.6	(Williams et al., 2020)
Canada	Hospitalized patients	91	RdRp, E, N	61.0	(Jamal et al., 2020)
China	Ambulatory patients	229	E	76.0	(Cheuk et al., 2020)
China	Hospitalized patients	58	RdRp/Hel, E, N2	84.5	(Chen et al., 2020)
China	Patients from 12 independent cohorts	944	S, E, ORF1ab, N, RdRp, 5'UTR	92.1	(Zhu et al., 2020)
China	Hospitalized patients	95	E, RdRp	78.9	(Leung Chi-man et al., 2021)
Japan	Ambulatory patients	76	N	97.4	(Iwasaki et al., 2020)
Mexico	Ambulatory patients	253	E	78.6	(Moreno-Contreras et al., 2020)
Thailand	Hospitalized patients	200	ORF1ab, N	97.5	(Pasomsub et al., 2020a)
USA	Hospitalized patients and asymptomatic healthcare workers	29	N1, N2	79.0	(Wyllie et al., 2020)
USA	Ambulatory patients	91	N1, N2	94.0	(Miller et al., 2020)
Mexico	Asymptomatic healthcare and office workers	2107	N1, N2	97.9	Our study

B		Saliva samples only			
Country	Study population	Saliva samples	Viral genes	Positivity %	Reference
China	Hospitalized patients	12	S	91.7	(To et al., 2020a)
China	Hospitalized patients	18	E	84.0	(Hung et al., 2020)
Italy	Hospitalized patients	25	5'UTR	100.0	(Azzi et al., 2020)
Japan	Hospitalized patients	103	N1, N2, ORF1, E	93.4	(Ikeda et al., 2020)

C		Saliva only in the present study				
Setting	Total samples	Number of tests	Positive samples	Positivity (%)	Reduction in testing costs (%) ^a	Reduction in total sample collection direct costs (USD) ^{b,c}
Asymptomatic office workers	3983	1032	26	0.65	74	\$10 754.50
Asymptomatic healthcare personnel	2126	870	98	4.6	59	\$5740.20
Symptomatic office workers	846	846	67	7.9	0	\$2284.20

RdRp, RNA-dependent RNA polymerase; RdRp/Hel, RNA-dependent RNA polymerase/helicase; ORF1, open reading frame 1 (a,b); ORF8, open reading frame 8; E, envelope; N, nucleocapsid.

^a Cost reduction was calculated considering the number of tests necessary to identify the positive individuals in the positive pools.

^b Sample collection direct cost: 3 USD vs 0.3 USD, nasopharyngeal and saliva, respectively.

^c The sample cost includes both direct and indirect costs.

Sample pooling

Positive samples were selected according to their RT-qPCR results, representing low and high Ct values, to evaluate the effect of a 1:10 pooling with negative samples in the detection capacity of the test. In the first set of saliva samples, the initial Ct values for the positive sample were 22.3 for ORF1ab, 19.6 for N, and 19.8 for RNase P. As expected, after pooling with the additional nine negative samples, the Ct values increased to 23.8 for ORF1ab, 22.4 for N, and 21.6 for RNase P, showing that pooling did not affect the detection capacity of the test. A similar situation was observed in the NPS sample pool. In the second saliva pool, the positive sample had higher Ct values (31.9 ORF1ab, 28 for N, and 19.1 for RNase P). After pooling, an increase in four Ct values in both viral genes was observed. Even though this result is still within the acceptable range for detecting the positive sample in the pool (Figure 4), samples with a higher Ct value might become false-negatives if analyzed by pooling; for this reason, the subsequent experiments were focused on the analysis of five-sample pools.

A total of 130 individuals were tested in 26 NPS pools with five samples each, identifying 20 positive cases (15.4%). All positive cases identified in the pools were confirmed through the analysis of the individual samples used to generate the pool. In the case of saliva, 255 individuals were grouped into 26 pools with five samples each. In this case, two positive cases were identified (7.7%), which were also confirmed through analysis of the individual samples.

Additionally, asymptomatic office and healthcare personnel were tested, as well as office workers presenting mild symptoms

who were suspected of being SARS-CoV-2 carriers. Only saliva was used and five samples were pooled in the first two groups. Table 2C shows the positivity among the three groups, which was increased in healthcare personnel and symptomatic office workers. Substantial reductions in direct costs for sampling compared with NPS and in the costs by testing pools instead of individual samples were observed.

These results showed that it is feasible to apply pooling strategies using saliva. However, some considerations should be taken into account, including the use of 500 µL of saliva to generate the pool to obtain a homogeneous mixture. Dilution of one positive sample with nine negative samples showed that, even though positive results can still be obtained in the pool, samples with a low viral load might become difficult to detect. Therefore, we suggest pooling no more than five samples, even though other reports indicate that pooling strategies of 16 and 24 samples are useful in high prevalence populations ($\geq 10\%$) (Verwilt et al., 2020).

Concluding remarks

The study data indicate that saliva is a reliable source for the detection of SARS-CoV-2 infection. However, several aspects must be addressed to successfully use saliva testing: (1) Sample collection: even though saliva self-collection might be easier than NPS sampling, proper biosafety and risk evaluation protocols must be followed by medical personnel to minimize contagions due to the production of potential aerosols during saliva collection. (2) Sample handling: the application of proven and standardized methods for the inactivation and handling of a saliva sample

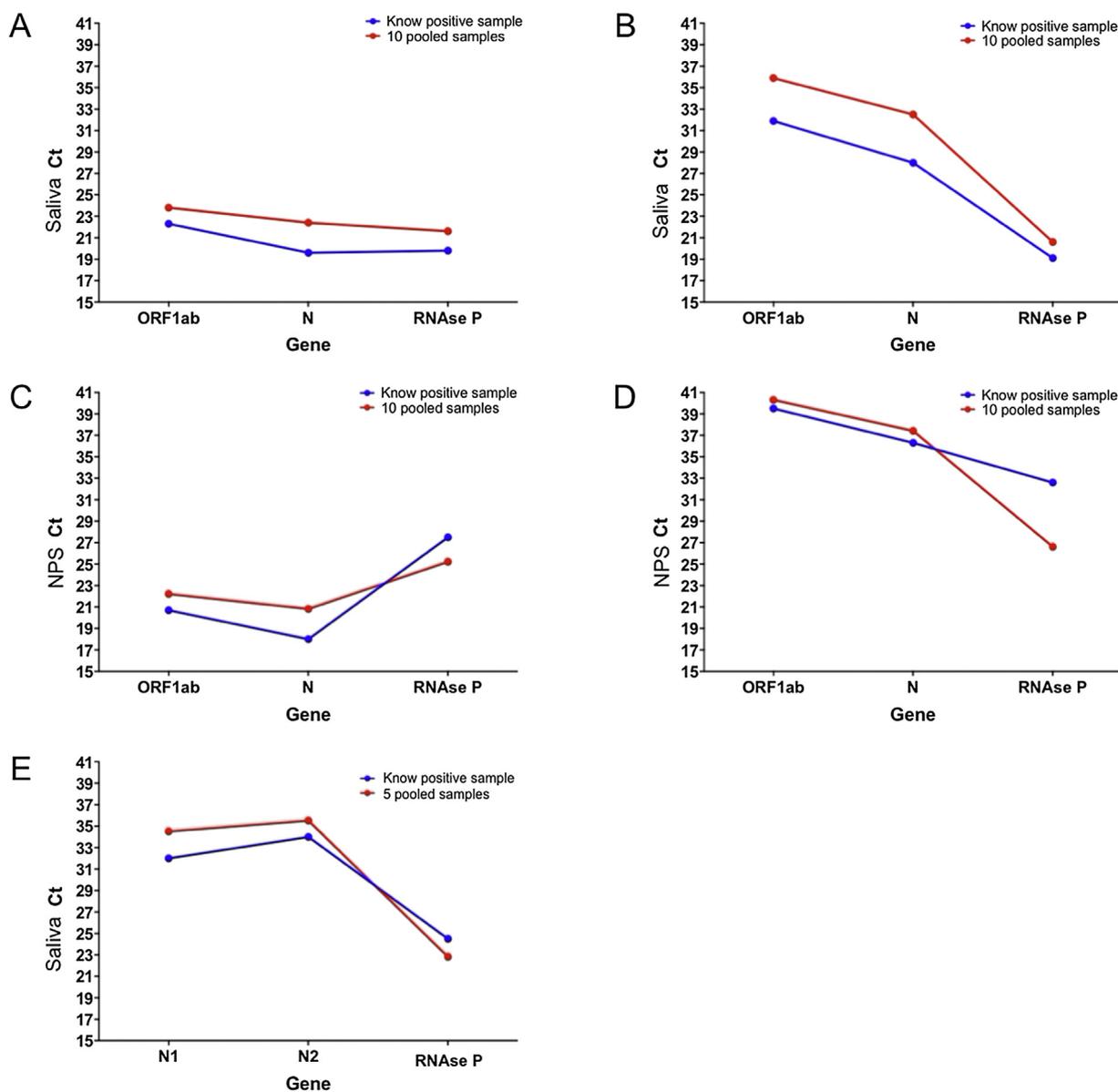


Figure 4. Analysis of pooled (1:5 and 1:10) saliva and swab paired samples. Saliva and nasopharyngeal swab pools were generated by mixing one positive sample with four/nine known negative samples. Positive samples with early and late Ct of the N gene were selected to evaluate the impact of dilution on its detection. (A) and (B) show saliva pooled 1:10; (C) and (D) show nasopharyngeal swab pooled 1:10; (E) shows saliva pooled 1:5.

should be considered, and saliva samples must always be regarded as potentially infected. The packaging and cold-chain protocols used for NPS samples must be followed. (3) RNA extraction and RT-qPCR: it has been well documented that several components of saliva can inhibit PCR, highlighting the importance of using viral RNA extraction systems that have been tested and approved by regulatory agencies that generate pure and high-quality RNA for RT-qPCR analysis. We did not use any preservative for saliva samples and suggest that samples should be stored at 4 °C after collection and processed within 4 days post collection.

Given the situations mentioned above, the use of saliva represents a viable option for SARS-CoV-2 detection. Thus, saliva and the pooling strategy presented here are effective options for the analysis of samples in well-controlled cohorts, which provide a cost-effective screening tool in asymptomatic populations. The cost reduction was calculated considering the number of tests

necessary to identify the positive individuals in the positive pool. This is particularly suitable, for example, in office workers, faculty, or other groups where testing is necessary on a periodic basis to identify and isolate infected individuals. The implementation of testing for SARS-CoV-2 infection by RT-qPCR using saliva as a source for viral RNA constitutes an easy, non-invasive, inexpensive, and less risky option compared to NPS, without compromising the accuracy of the test. The combination of saliva sampling and pooling represents a viable and useful method for population-based studies that will be necessary for a safe return to economic activities.

Funding

This work was funded by the Secretaría de Educación, Ciencia, Tecnología e Innovación de la Ciudad de México (SECTEI).

Conflict of interest

The authors do not have an association that might pose a conflict of interest.

Acknowledgments

The authors acknowledge M.C. Isabel Gracia Mora for providing BSL2 facilities at UNAM; Facultad de Química-UNAM, Patronato de la Facultad de Química-UNAM for providing all materials used in the RT-qPCR analyses; and CONACyT 314298 for project funding.

References

- Abdalahamid B, Bilder CR, McCutchen EL, Hinrichs SH, Koepsell SA, Iwen PC. Assessment of specimen pooling to conserve SARS-CoV-2 testing resources. *Am J Clin Pathol* 2020;153;. doi:<http://dx.doi.org/10.1093/ajcp/aqaa064>.
- Azzi L, Carcano G, Gianfagna F, Grossi P, Gasperina DD, Genoni A, et al. Saliva is a reliable tool to detect SARS-CoV-2. *J Infect* 2020;81;. doi:<http://dx.doi.org/10.1016/j.jinf.2020.04.005>.
- Ben-Ami R, Klochendler A, Seidel M, Sido T, Gurel-Gurevich O, Yassour M, et al. Large-scale implementation of pooled RNA extraction and RT-PCR for SARS-CoV-2 detection. *Clin Microbiol Infect* 2020;26;. doi:<http://dx.doi.org/10.1016/j.cmi.2020.06.009>.
- Callahan CJ, Lee R, Zulauf KE, Tamburello L, Smith KP, Previtiera J, et al. Open development and clinical validation of multiple 3D-printed sample-collection swabs: rapid resolution of a critical COVID-19 testing bottleneck. *MedRxiv* 2020;. doi:<http://dx.doi.org/10.1101/2020.04.14.20065094>.
- Centers for Disease Control and Prevention. (2019-nCoV), CDC 2019–Novel Coronavirus Panel, Real-Time RT-PCR Diagnostic. 2020. <https://www.fda.gov/media/134922/download>.
- Chen JHK, Yip CCY, Poon RWS, Chan KH, Cheng VCC, Hung IFN, et al. Evaluating the use of posterior oropharyngeal saliva in a point-of-care assay for the detection of SARS-CoV-2. *Emerg Microbes Infect* 2020;9;. doi:<http://dx.doi.org/10.1080/22221751.2020.1775133>.
- Cheuk S, Wong Y, Tse H, Siu HK, Kwong TS, Chu MY, et al. Posterior oropharyngeal saliva for the detection of SARS-CoV-2. *Clin Infect Dis* 2020;71:2939–46, doi:<http://dx.doi.org/10.1093/cid/ciaa797>.
- Hogan CA, Garamani N, Sahoo MK, Huang CH, Zehnder J, Pinsky BA. Retrospective screening for SARS-CoV-2 RNA in California, USA, late 2019. *Emerg Infect Dis* 2020;26;. doi:<http://dx.doi.org/10.3201/eid2610.202296>.
- Hung DL-L, Li X, Chiu KH-Y, Yip CC-Y, To KK-W, Chan JF-W, et al. Early-morning vs spot posterior oropharyngeal saliva for diagnosis of SARS-CoV-2 infection: implication of timing of specimen collection for community-wide screening. *Open Forum Infect Dis* 2020;7;. doi:<http://dx.doi.org/10.1093/ofid/ofaa210>.
- Ikeda M, Imai K, Tabata S, Miyoshi K, Murahara N, Mizuno T, et al. Clinical evaluation of self-collected saliva by RT-qPCR, direct RT-qPCR, RT-LAMP, and a rapid antigen test to diagnose COVID-19. *MedRxiv* 2020;58:e01438-20, doi:<http://dx.doi.org/10.1101/2020.06.06.20124123>.
- Iwasaki S, Fujisawa S, Nakakubo S, Kamada K, Yamashita Y, Fukumoto T, et al. Comparison of SARS-CoV-2 detection in nasopharyngeal swab and saliva. *J Infect* 2020;81;. doi:<http://dx.doi.org/10.1016/j.jinf.2020.05.071>.
- Jamal AJ, Mozafarihashjin M, Coomes E, Powis J, Liu AX, Paterson A, et al. Sensitivity of nasopharyngeal swabs and saliva for the detection of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). *MedRxiv* 2020;. doi:<http://dx.doi.org/10.1101/2020.05.01.20081026>.
- Leung Chi-man E, Chow Chi-ying V, Lee Kin-ping M, Lai Wai-man R. Deep throat saliva as an alternative diagnostic specimen type for the detection of SARS-CoV-2. *J Med Virol* 2021;93;. doi:<http://dx.doi.org/10.1002/jmv.26258>.
- Lohse S, Pfuhl T, Berkó-Göttel B, Rissland J, Geißler T, Gärtner B, et al. Pooling of samples for testing for SARS-CoV-2 in asymptomatic people. *Lancet Infect Dis* 2020;20;. doi:[http://dx.doi.org/10.1016/S1473-3099\(20\)30362-5](http://dx.doi.org/10.1016/S1473-3099(20)30362-5).
- McHugh ML. Interrater reliability: The kappa statistic. *Biochem Medica* 2012;22;. doi:<http://dx.doi.org/10.11613/bm.2012.031>.
- Miller M, Jansen M, Bisignano A, Mahoney S, Wechsberg C, Albanese N, et al. Validation of a Self-administrable, Saliva-based RT-qPCR Test Detecting SARS-CoV-2. *MedRxiv* 2020;. doi:<http://dx.doi.org/10.1101/2020.06.05.20122721>.
- Moreno-Contreras J, Espinoza MA, Sandoval-Jaime C, Cantú-Cuevas MA, Barón-Olivares H, Ortiz-Orozco OD, et al. Saliva sampling and its direct lysis, an excellent option to increase the number of SARS-CoV-2 diagnostic tests in settings with supply shortages. *J Clin Microbiol* 2020;58;. doi:<http://dx.doi.org/10.1128/JCM.01659-20>.
- Noah K, Fred T, Vlad S, Agatha B, Laura D, Siri K, et al. Self-collected oral fluid and nasal swabs demonstrate comparable sensitivity to clinician collected nasopharyngeal swabs for Covid-19 detection. *MedRxiv* 2020;. doi:<http://dx.doi.org/10.1101/2020.04.11.20062372>.
- Ott IM, Strine MS, Watkins AE, Boot M, Kalinich CC, Harden CA, et al. Simply saliva: Stability of SARS-CoV-2 detection negates the need for expensive collection devices. *MedRxiv* 2020;. doi:<http://dx.doi.org/10.1101/2020.08.03.20165233>.
- Pasomsub E, Watcharananan SP, Boonyawat K, Janchompoo P, Wongtabtim G, Suksuwan W, et al. Saliva sample as a non-invasive specimen for the diagnosis of coronavirus disease 2019: a cross-sectional study. *Clin Microbiol Infect* 2020a;. doi:<http://dx.doi.org/10.1016/j.cmi.2020.05.001>.
- Pasomsub Ekawat, Watcharananan SP, Watthanachokchai T, Rakmanee K, Tassaneethitip B, Kiertiburanakul S, et al. Saliva sample pooling for the detection of SARS-CoV-2. *J Med Virol* 2020b;. doi:<http://dx.doi.org/10.1002/jmv.26460>.
- SalivaClear by Mirimus Clinical. The SalivaClear Solution. 2020. <https://www.salivaclear.com/>.
- To KKW, Tsang OTY, Leung WS, Tam AR, Wu TC, Lung DC, et al. Temporal profiles of viral load in posterior oropharyngeal saliva samples and serum antibody responses during infection by SARS-CoV-2: an observational cohort study. *Lancet Infect Dis* 2020a;20;. doi:[http://dx.doi.org/10.1016/S1473-3099\(20\)30196-1](http://dx.doi.org/10.1016/S1473-3099(20)30196-1).
- To KKW, Tsang OTY, Yip CCY, Chan KH, Wu TC, Chan JMC, et al. Consistent detection of 2019 novel coronavirus in saliva. *Clin Infect Dis* 2020b;71;. doi:<http://dx.doi.org/10.1093/cid/ciaa149>.
- U.S. Food and Drug Administration. Coronavirus (COVID-19) update: FDA authorizes first diagnostic test using at-home collection of saliva specimens. 2020. <https://www.fda.gov/news-events/press-announcements/coronavirus-covid-19-update-fda-authorizes-first-diagnostic-test-using-home-collection-saliva>.
- U.S. Food and Drug Administration. Coronavirus (COVID-19) update: FDA authorizes first diagnostic test for screening of people without known or suspected COVID-19 infection authorization is also second to allow testing of pooled samples. 2020.
- Uwamino Y, Nagata M, Aoki W, Fujimori Y, Nakagawa T, Yokota H, et al. Accuracy and stability of saliva as a sample for reverse transcription PCR detection of SARS-CoV-2. *J Clin Pathol* 2021;74;. doi:<http://dx.doi.org/10.1136/jclinpath-2020-206972>.
- Verwilt J, Mestdagh P, Vandesompele J. Evaluation of efficiency and sensitivity of 1D and 2D sample pooling strategies for diagnostic screening purposes. *MedRxiv* 2020;. doi:<http://dx.doi.org/10.1101/2020.07.17.20152702>.
- Watkins AE, Fenichel EP, Weinberger DM, Vogels CBF, Brackney DE, Casanovas-Massana A, et al. Pooling saliva to increase SARS-CoV-2 testing capacity. *MedRxiv* 2020;. doi:<http://dx.doi.org/10.1101/2020.09.02.20183830>.
- Williams E, Bond K, Zhang B, Putland M, Williamson DA. Saliva as a noninvasive specimen for detection of sars-cov-2. *J Clin Microbiol* 2020;58;. doi:<http://dx.doi.org/10.1128/JCM.00776-20>.
- Wyllie AL, Fournier J, Casanovas-Massana A, Campbell M, Tokuyama M, Vijayakumar P, et al. Saliva is more sensitive for SARS-CoV-2 detection in COVID-19 patients than nasopharyngeal swabs. *MedRxiv* 2020;. doi:<http://dx.doi.org/10.1101/2020.04.16.20067835>.
- Yelin I, Aharonov N, Tamar ES, Argoetti A, Messer E, Berenbaum D, et al. Evaluation of COVID-19 RT-qPCR Test in Multi sample Pools. *Clin Infect Dis* 2020;71;. doi:<http://dx.doi.org/10.1093/cid/ciaa531>.
- Zhu J, Guo J, Xu Y, Chen X. Viral dynamics of SARS-CoV-2 in saliva from infected patients. *J Infect* 2020;81;. doi:<http://dx.doi.org/10.1016/j.jinf.2020.06.059>.