



# UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

## Maestría y Doctorado en Ciencias Bioquímicas

Caracterización de microRNAs circulantes maternos asociados al desarrollo  
fetoplacentario

TESIS

QUE PARA OPTAR POR EL GRADO DE:  
Maestro en Ciencias

PRESENTA:

Alma Lilia Hernández Olvera

TUTOR PRINCIPAL

Luis Antonio Mendoza Sierra

[Instituto de Investigaciones Biomédicas](#)

MIEMBROS DEL COMITÉ TUTOR

Erika Chavira Suárez

[Unidad de Vinculación Científica de la Facultad de Medicina, UNAM en el  
INMEGEN](#)

ILEANA PATRICIA CANTO CETINA

[Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán](#)

Ciudad de México. Mayo, 2023



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.

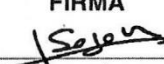
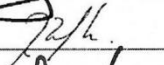
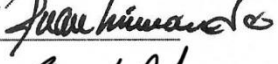


**SINODALES DESIGNADOS**

**Presente**

Estimado académico:

Los miembros del Subcomité Académico en reunión extraordinaria del **27 de marzo de 2023**, conocieron la solicitud de asignación de **JURADO DE EXAMEN** para optar por el grado de **Maestría en Ciencias** del/la estudiante **HERNÁNDEZ OLVERA ALMA LILIA**, con la tesis "**Caracterización de microRNAs circulantes maternos asociados al desarrollo fetoplacentario**", dirigida por el/la Dr(a). **MENDOZA SIERRA LUIS ANTONIO**.

De su análisis se acordó nombrar el siguiente jurado en el que se encuentra usted incluido:

			ACEPTA	FECHA	FIRMA
SEGOVIA FORCELLA LORENZO PATRICK	PMDCBQ	PRESIDENTE	SI <input type="checkbox"/> NO <input type="checkbox"/>	10 / 04 / 2023	
GUTIÉRREZ RÍOS ROSA MARÍA	PMDCBQ	SECRETARIO	SI <input checked="" type="checkbox"/> NO <input type="checkbox"/>	10 / 04 / 2023	
MIRANDA RÍOS JUAN	PMDCBQ	VOCAL	SI <input checked="" type="checkbox"/> NO <input type="checkbox"/>	10 / 04 / 2023	
RESENDIS ANTONIO OSBALDO	PMDCBQ	VOCAL	SI <input checked="" type="checkbox"/> NO <input type="checkbox"/>	11 / 04 / 2023	
ZAMBRANO GONZÁLEZ ELENA	PMDCBQ	VOCAL	SI <input checked="" type="checkbox"/> NO <input type="checkbox"/>	11 / 04 / 2023	

Sin otro particular por el momento, aprovecho la ocasión para enviarle un cordial saludo.

**Atentamente**  
**"POR MI RAZA HABLARÁ EL ESPÍRITU"**  
 Cd. Universitaria, Cd. Mx., a 28 de marzo de 2023



**Coordinadora**  
**Dra. Claudia Lydí Treviño Santa Cruz**

## **Agradecimientos**

Agradezco el financiamiento recibido por el fondo sectorial de investigación para la educación de CONACyT (Ciencias Básicas 2017-2018: A1-S-35245) apoyo con el cual se solventaron los gastos pertinentes para servicios, reactivos e insumos que permitieron el desarrollo experimental de este trabajo.

Agradezco al Instituto de Nacional de Medicina Genómica y al laboratorio de la Unidad de Vinculación Científica – Facultad de Medicina y a sus integrantes por permitir el uso de sus equipos e instalaciones para el desarrollo experimental precedente a este proyecto.

# Índice

Siglas y Abreviaturas .....	5
Resumen .....	7
Introducción .....	8
Marco teórico .....	10
1. Generalidades de los ncRNAs .....	10
2. Generalidades de los miRNAs .....	11
2.1. Biogénesis y función de los miRNAs .....	13
2.2. Los miRNAs como reguladores de la expresión génica.....	15
2.3. Aplicación de los miRNAs como biomarcadores .....	17
3.1. Desarrollo embrionario y fetal.....	18
3.2. Desarrollo placentario .....	22
Planteamiento del problema .....	24
Objetivos.....	25
Objetivos particulares.....	25
Hipótesis.....	25
Metodología.....	25
Recolección de muestras .....	26
Extracción de RNA y datos de secuenciación .....	26
Recopilación de datos clínicos .....	27
Validación de la firma de c-miRNAs con indicadores del crecimiento fetal .....	27
Asociación entre miRNAs del C14 y C19 con firma de CF .....	27
Determinación de blancos génicos.....	28
Red de interacción proteína- proteína .....	28
Enriquecimiento Funcional .....	29
Resultados.....	30
Discusión .....	41
Conclusiones .....	50
Referencias .....	51
Apéndice A: Características clínicas de las participantes .....	58
Apéndice B: Código en Python modelado de random forest de regresión para predecir la longitud del fémur.....	59
Apéndice C: Código en R para determinar enriquecimiento .....	61
Apéndice D: Código en R predicción de blancos y construcción de redes de IPP .....	63
Apéndice E: Código en R para validar significancia de redes de IPP por proporción de interacciones máximas.....	65

# Siglas y Abreviaturas

## A

AGO: Argonaut

A: Adenina

ARE: AU-rich elements (Elementos ricos en AU)

## C

c-miRNAs: circulating miRNA (miRNAs circulantes)

CCR4-NOT: The carbon catabolite repression 4 (CCR4)–negative on TATA-less (NOT) complex (complejo represor de catabolito de carbono 4 negativo) citotrofoblastos (CTB).

CA: Circunferencia abdominal

C14: Cromosoma 14

C19: Cromosoma 19

CEV: Citotrofoblasto extravelloso

## D

DGCR8: Proteína de la región crítica del gen 8 del Síndrome de DiGeorge

DMSG: Diámetro medio del saco gestacional

## E

ECNT: Enfermedades crónico-degenerativas no transmisibles

E: embarazadas

## F

FXR1: Fragile X Mental Retardation, Autosomal Homolog 1 (Proteína 1 relacionada con el síndrome de retraso mental del X frágil)

FcRn: Neonatal Fc receptor (Receptores Fc neonatal)

FC: Fold Change (Cambio de pliegue)

## G

GE: Grande para la edad gestacional

GFP: Green Fluorescent Protein (Proteína verde fluorescente)

GO: Gene ontology (Ontología génica)

## I

INMEGEN: Instituto Nacional de Medicina Genómica

IMC-pg: índice de masa corporal pregestacional

IPP: Interacciones proteína-proteína

IPP<sub>MAX</sub>: Interacciones proteína-proteína máximas

lncRNAs: Long non-coding RNA (RNAs largos no codificantes)

## L

LCR: Longitud cráneo-rabadilla

LF: Longitud del fémur

## M

miRNAs: microRNAs

mRNAs: RNA mensajeros

miRISC: miRNA-induced silencing complex (Complejo silenciador inducido por miRNAs)

MCI: Masa compacta de células internas

MMP: Metalloproteinase (metaloproteínasa)

## N

nt: nucleótidos

## O

OMS: Organización mundial de la salud

## P

piRNAs: RNAs asociados a piwi

pol II: Polimerasa II

pol III: polimerasa III

pri-miRNA: miRNA primario

pre-miRNA: precursor de miRNA

PACT: Protein Activator of the interferon-induced protein kinase (proteína activadora de interferón inducida por quinasas)

PABPC: Cytoplasmic poly(A) binding protein (proteína de unión a poliadenilato citoplasmática)

Peso-pg: Peso pregestacional

PP: Posterior al embarazo

PEG: Pequeño para la edad gestacional

PAPP-A: Pregnancy associated plasma protein-A (Proteína placentaria A asociada al embarazo)

PIGF: Placental growth factor (factor de crecimiento placentario)

## R

RanGTP: GTP-bound Ran (Ran enlazado a GTP)

RCI: Restricción del crecimiento intrauterino

RCIU: Restricción del crecimiento intrauterino

## S

snoRNAs: Small nucleolar RNAs (pequeños nucleolares)

snRNAs: Small nuclear RNAs (Pequeños nucleares)

siRNAs: Small interfering RNAs (RNAs de interferencia)

SDG: Semanas de gestación

STB: Sincitiotrofoblasto

SNC: Sistema nervioso central

sFLT1: Circulating soluble fms-like tyrosine kinase-1 (receptor 1 del factor de crecimiento vascular)

## T

TARBP: Trans-Activation-Responsive RNA-Binding Protein (proteína de unión a RNA TAR)

TE: Trofoectodermo

TA: Tensión arterial

## U

UTR: Untranslated region (Región no traducida)

U: Uracilo

UVA-FM: Unidad de Vinculación Científica-Facultad de medicina

UAT: Unidad de Alta Tecnología

## X

XPO5: Exportin-5 (exportina 5)

XRN1: 5'-3' Exoribonuclease 1 (exoribonucleasa 1)

## Resumen

El desarrollo fetal marca un punto crucial en la salud actual y futura de un individuo. Actualmente, a nivel clínico el desarrollo fetal se ve reducido únicamente a mediciones antropométricas estimadas por ultrasonografía (USG). Las cuales, nos brinda poca información para determinar un desarrollo fetal adecuado. En este sentido, estudios recientes han revelado que diversos microRNAs (miRNAs) circulantes se ven alterados en diferentes complicaciones del embarazo implicadas en el crecimiento fetal, lo cual, sugiere que estos transcritos circulantes en la madre podrían reflejar alteraciones asociadas al desarrollo fetal. Nosotros buscamos caracterizar una firma de miRNAs circulantes correlacionados con indicadores de crecimiento fetal. Para ello, obtuvimos el perfil de expresión de miRNAs circulantes en 8 mujeres muestreadas en el primero, segundo y tercer trimestre de gestación. Con los datos obtenidos se identificaron seis miRNAs correlacionados positivamente con longitud de fémur (LF) y peso fetal estimado (PFE). A partir de esta firma validamos por random forest la asociación de estos miRNAs con el crecimiento fetal. Por otro lado, consultamos bases de datos validadas experimentalmente para identificar los blancos génicos de la firma, con base en ello construimos una red de interacciones proteína-proteína (IPP) para determinar si los miRNAs de la firma regulan potenciales blancos génicos claves en la regulación de procesos biológicos asociados con el desarrollo fetal. Nuestro estudio verifico que la firma de crecimiento fetal presenta un alto valor de importancia para predecir el crecimiento fetal a través de una regresión por random foreste, además, identificamos que estos miRNAs pertenecen al grupo de miRNAs del cromosoma 19 (C19) y que los miRNAs de este clúster tienden a presentar una mayor correlación con indicadores del crecimiento fetal. Además, determinamos que los blancos génicos regulados por estos miRNAs presentan un mayor número de IPP entre sí, esta evidencia nos sugiere que pueden participar en la regulación de vías implicadas en el crecimiento fetal que además podemos observar en circulación, por lo que, en un futuro con un mayor entendimiento de su función pueden ser trasladados a la clínica como biomarcadores tempranos poco invasivos del desarrollo fetal.



## Introducción

El embarazo es un periodo de cambios anatómicos, bioquímicos, hormonales e inmunitarios en la mujer para lograr el desarrollo de órganos fetales y estructuras especializadas como lo es la placenta. En este periodo transcurren procesos biológicos materno y fetales críticos para la evolución del embarazo.

La exposición a componentes tanto intrínsecos, como extrínsecos, modulan el desarrollo fetal durante la vida intrauterina. Factores adversos, tales como: el cigarrillo, alcohol, fármacos e incluso el estrés materno puede desencadenar el desarrollo de enfermedades crónico-degenerativas no transmisibles (ECNT) en la vida adulta (Gluckman et. al., 2008; Neiger, 2017). No obstante, la investigación en dicho campo ha avanzado lentamente. Estudios recientes se han sumado para determinar los mecanismos biológicos diversos que pueden estar interviniendo en el desarrollo normal y patológico del embarazo, así como sus implicaciones en la salud materna y fetal.

Actualmente, los investigadores y científicos buscan identificar biomarcadores potenciales que sean más precisos y eficientes para el monitoreo del embarazo y la detección temprana tanto de complicaciones y/o alteraciones en el desarrollo y crecimiento fetal, como de la propia evolución del embarazo. Todo ello es con la finalidad de permitir una intervención clínica oportuna y especializada a la mujer embarazada, y así, reducir los riesgos de complicaciones prenatales.

La detección de ácidos nucleicos en biofluidos, como el plasma sanguíneo, se ha convertido en un campo de investigación ampliamente explorado para determinar futuros biomarcadores, que nos permitan esbozar los eventos moleculares activos en un organismo. La evaluación de transcritos circulantes en el plasma materno es uno de los campos más atractivos en este sentido. Esto se debe a que representa una forma menos invasiva de evaluar indirectamente sutiles alteraciones en los diversos procesos biológicos presentes en la evolución del embarazo (Karin-Kujundzic et. al., 2019).

Dentro de los ácidos nucleicos circulantes más estudiado hasta el momento se encuentran los microRNAs (miRNAs), los cuales son secuencias cortas de RNAs no codificantes (ncRNAs) de ~22 nucleótidos (nt), que regulan la expresión genética

a nivel postranscripcional (Gebert & MacRae, 2019). Por un lado, debido a su función como reguladores postranscripcionales, son un blanco importante de estudio en la regulación de la expresión génica en la célula (Goodall & Wickramasinghe, 2021). Además, de poseer un gran potencial de biomarcación en diferentes procesos biológicos y patológicos (Taft, 2010) en la circulación sanguínea, así como, en otros biofluidos.

Estudios previos, realizados por nuestro grupo de investigación (Chavira-Suárez et al., 2023), así como, otras investigaciones identificaron una gran variedad de transcritos en el plasma de las mujeres embarazadas, siendo los miRNAs la clase más abundante en muestras de plasma (Vashukova et al., 2021). Además, hemos identificado la existencia de conjuntos diferentes o firmas de miRNAs circulantes (c-miRNAs) específicamente asociados a la evolución normal del embarazo, entre ellas un conjunto de miRNAs correlacionado con indicadores de crecimiento fetal (Chavira-Suárez et al., 2023). Por lo cual, este estudio caracteriza conjuntos de c-miRNAs altamente correlacionados con eventos relacionados con el desarrollo placentario y crecimiento fetal.

# Marco teórico

## 1. Generalidades de los ncRNAs

La expresión génica promueve la síntesis de proteínas necesarias para la función, mantenimiento y supervivencia de las células que componen los tejidos y órganos del cuerpo. Los factores de transcripción son proteínas con funciones reguladoras de la expresión génica. No obstante, estas proteínas no son las únicas moléculas capaces de regular la expresión de los genes. Existen diferentes tipos o clases de RNAs, los cuales cumplen con funciones celulares específicas, entre ellas la regulación de la expresión génica. Es bien sabido que dentro de los diversos tipos de RNAs están los transcritos codificantes o RNA mensajeros (mRNA) que codifican para proteínas y los RNAs de transferencia y ribosomales que participan en el proceso de la traducción del mRNA a proteína (Slaby & Calin Editors, 2016).

Hasta la fecha, los avances en el conocimiento sobre la regulación de la expresión genética por medio de los factores de transcripción no han sido suficientes para descifrar los efectos de inhibición o activación en la expresión de los genes. Los desarrollos tecnológicos recientes sobre la secuenciación masiva y el análisis en las ómicas han permitido identificar recientemente un número mayor de moléculas de diversos tipos y características entre las cuales destacan los RNAs.

La célula sintetiza diversos tipos de ncRNAs que no se encuentran directamente asociados con el proceso de traducción. Tales como los RNAs miRNAs, pequeños nucleares (snRNAs), pequeños nucleolares (snoRNAs), RNAs asociados a piwi (piRNAs) y RNAs de interferencia (siRNAs) que cumplen funciones críticas en la regulación de la expresión génica (Esteller, 2011). La clasificación de los diferentes tipos de ncRNAs está principalmente basada en la longitud de su secuencia nucleotídica. Como por ejemplo RNAs largos no codificantes (lncRNAs) cuando la secuencia está conformada por más de 200 nt y sncRNAs cuando la secuencia está conformada por menos de 200 nt (Panni et. al., 2020).

### 1.1. RNAs circulantes

Dennis Lo y sus colaboradores fueron los primeros investigadores en reportar la presencia de ácidos nucleicos periféricos en el suero y plasma de mujeres

embarazadas (Dennis Lo et. al., 1997). Este hallazgo despertó un gran interés en el estudio de los ácidos nucleicos como potenciales biomarcadores de procesos biológicos y patológicos (Condrat et al., 2020).

Las evidencias ómicas actuales indican que, diferentes biofluidos están enriquecidos con una gran variedad de moléculas que previamente no habían sido detectadas. De esta manera, el conocimiento sobre la transcriptómica en fluidos corporales en hoy base de estudios para el entendimiento sobre el origen y desarrollo de enfermedades multifactoriales (Godoy et. al., 2018).

A raíz de esto, se han presentado diversas teorías que plantean desde los posibles mecanismos de secreción hasta las posibles funciones de los transcritos circulantes, como un método de comunicación celular. Sin embargo, más allá de determinar la función de los ácidos nucleicos circulantes, se ha planteado la posibilidad de evaluarlos para ser utilizados como biomarcadores de diferentes condiciones fisiológicas y patológicas con miras de uso clínico (O'Brien et. al., 2018).

## 2. Generalidades de los miRNAs

El descubrimiento de los miRNAs surge como un hallazgo inesperado en la búsqueda de genes involucrados en el desarrollo larvario de *Caenorhabditis elegans*. Víctor Ambros y sus colaboradores en 1993, durante sus investigaciones lograron identificar que la subexpresión de la proteína lin-14 era esencial para la progresión del desarrollo larvario. Simultáneamente, observaron que la baja expresión de lin-14 era dependiente de la transcripción del gen lin-4. Sin embargo, no lograron identificar un producto proteico derivado de este transcrito. En su lugar identificaron dos transcritos pequeños de diferentes longitudes de 21 y 61 nucleótidos aproximadamente (Lee et. al., 1993).

Este grupo concluyó que uno de estos transcritos servía como precursor del más corto. Más tarde, notaron que el transcrito de menor longitud presentaba complementariedad antisentido a múltiples sitios en la región no traducida 3' (UTR) del gen lin-14. La complementariedad de ambos transcritos disminuye el producto proteico de lin-14 (Bartel, 2004). Años más tarde, un creciente número de estudios identificaron características particulares de este tipo de RNAs en diferentes

organismos, desde nemátodos hasta humanos, dando origen a la categorización de una nueva clase de RNAs reguladores a los que hoy conocemos como miRNA.

La rápida evolución en el descubrimiento de los miRNAs da pie al surgimiento de la base de datos miRBase (Griffiths-Jones, 2006). Este es el principal repositorio público para dar a conocer las secuencias y anotaciones de los miRNAs que se han ido identificando con el paso del tiempo, así, como para el registro de nuevas secuencias e isoformas identificadas. En la actualidad, se han descrito 38,589 miRNAs en 271 organismos diferentes. A la fecha, han reportado 1,917 precursores de miRNAs que dan origen a 2654 secuencias maduras anotadas para el humano (Griffiths-Jones, 2006).

miRBase cuenta con una serie de reglas de nomenclatura que determina las pautas para la anotación de estas secuencias (Griffiths-Jones, 2006). De acuerdo con ello, cada secuencia es identificada con un número, el cual se asigna de acuerdo con el orden de descubrimiento de la secuencia. Además, se agrega un prefijo, que consiste en una abreviación de 3 o 4 letras para designar la especie. En el caso de las anotaciones para el humano corresponden a 'hsa' (*Homo sapiens*) seguido de la etiqueta 'miR' para la secuencia madura. Cabe destacar, que los miRNAs se originan a partir de un transcrito precursor, que posteriormente es modificado para generar su forma madura, por lo que, miRBase hace una distinción entre estos transcritos con la letra R de la etiqueta 'miR' por lo que utiliza la letra en minúscula como 'mir' para referirse al transcrito precursor. Finalmente, se le adiciona la terminación 3p o 5p, en función de la direccionalidad de la cadena que le da origen al miRNA maduro. A la secuencia resultante del extremo 3' se le asigna la terminación 3p mientras que a la secuencia del extremo 5' tiene la terminación de 5p (Griffiths-Jones, 2006).

A pesar de que se han identificado una gran cantidad de nuevos miRNAs, siguen presentándose las incógnitas sobre su participación reguladora en las funciones biológicas. Hasta el momento, se sabe que los miRNAs están involucrados en la regulación génica postranscripcional y que tienen un papel fundamental en los procesos biológicos esenciales de la célula incluyendo

proliferación, diferenciación, metabolismo y crecimiento como los más reportados (Gebert & MacRae, 2019).

## 2.1. Biogénesis y función de los miRNAs

Los miRNAs inician su biogénesis en el núcleo, donde la RNA polimerasa II (pol II) comienza con la transcripción del DNA para sintetizar un transcrito primario que posteriormente se transformarán en el miRNA activo. Estos pueden originarse a través de intrones y/o exones de regiones codificantes (intragénicas) o de regiones no codificantes (intergénicas), por lo que, su ubicación en el genoma es muy diversa. Sin embargo, se sabe que la mayor parte de los miRNAs identificados hasta el momento se encuentran en regiones intergénicas.

Aunque se han descrito diferentes mecanismos por los cuales se pueden sintetizar los miRNAs. La vía canónica consta de múltiples fases en diferentes compartimentos celulares como se ilustra en el **Esquema 1**. comienza con la transcripción en el núcleo principalmente por la pol II, y en casos particulares la transcripción es mediada por la RNA polimerasa III (pol III). Derivado de este proceso se obtiene un primer transcrito que consiste en una secuencia larga con una estructura de horquilla. El transcrito es denominado como miRNA primario (pri-miRNA) (O'Brien et. al., 2018).

Una vez sintetizado el pri-miRNA es escindido en el núcleo por un complejo multi-proteico denominado como complejo microprocesador que está conformado por la ribonucleasa III (Drosha) y la proteína de la región crítica del gen 8 del Síndrome de DiGeorge (DGCR8). El pri-miRNA es reconocido por una molécula de DGCR8, mientras que una segunda proteína DGCR8 permite mantener fija la estructura sujeta por el tallo de doble cadena del pri-miRNA. Una vez ensamblado el complejo la proteína efectora Drosha puede cortar con precisión los extremos 5' y 3', reconociendo el punto de corte por estar constituido de pequeñas secuencias de cadena sencilla. De este proceso deriva una nueva estructura transitoria denominada precursor de miRNA (pre-miRNA). Se caracteriza por tener un grupo fosfato en la terminación 3' y 2 nucleótidos sobresalientes en el extremo 5'. Cabe mencionar, que Drosha pueden generar variaciones en la secuencia del miRNA

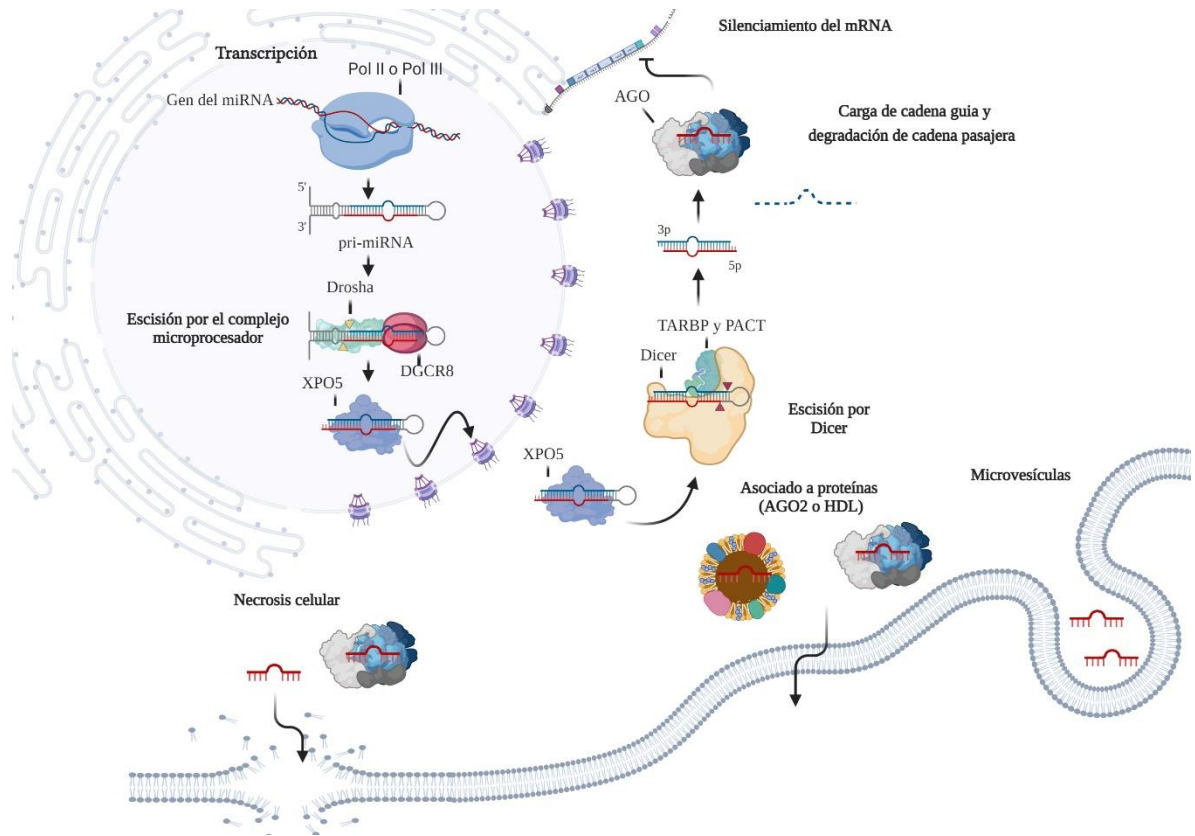
resultante por cortes alternativos variando el espaciado entre el corte realizado, dando lugar a formas isoméricas conocidas isomiRs (Gebert & MacRae, 2019).

La estructura madura del miRNA se lleva a cabo en el citoplasma, por lo que el pre-miRNA es transportado fuera del núcleo. Para ello, se forma un complejo transportador integrado por la exportina 5 (XPO5) y Ran enlazado a GTP (RanGTP). Fuera del núcleo el pre-miRNA es reconocido por la ribonucleasa III Dicer donde se realiza una segunda escisión para eliminar la estructura de bucle. La actividad de Dicer es regulada por la proteína de unión a RNA TAR (TARBP) y por la proteína activadora de interferón inducida por quinasas (PACT). El producto resultante son dos secuencias complementarias cortas, cuya longitud es propia de cada especie (Daly & Driscoll, 2017).

Ambas secuencias son miRNAs maduros, no obstante, solo una de ellas es seleccionada por la proteína efectora, mientras que la otra secuencia es degradada. La proteína efectora que selecciona al miRNA maduro es una de las proteínas de la familia Argonauta (AGO). En los humanos existen cuatro proteínas argonautas (AGO 1-4) y el ensamblaje es mediado por la maquinaria de proteínas chaperonas Hsp70 / Hsp90 (Kobayashi & Tomari, 2016). La elección de la hebra 3p o 5p cargada por AGO2 está en función de la estabilidad termodinámica de la secuencia. Además, presenta preferencia por la secuencia cuya terminación 5' sea Adenina (A) o Uracilo (U) (Gebert & MacRae, 2019). Sin embargo, también se ha observado que la proporción de hebras 5p o 3p cargadas por AGO puede presentar una gran variación según el tipo de célula o el entorno celular (O'Brien et. al., 2018).

Los miRNAs presentan una región comprendida entre los primeros 2-8 nt de la secuencia del extremo 5', que se definen como región semilla. Los miRNAs y su mRNA mensajero blanco no necesariamente requieren presentar una complementariedad perfecta entre ambos, sin embargo, una complementariedad perfecta o muy alta de la región semilla permite un mejor reconocimiento del mRNA. Además, estudios *in silico* han demostrado que la presencia de una adenina flanqueando ambos lados de la región semilla aumenta la estabilidad de la unión miRNA-mRNA (McGeary et al., 2019) es por ello por lo que muchos autores agrupan

a los miRNAs de acuerdo con la similitud que existe entre sus secuencias semillas (Wahid et. al., 2010).



### Esquema 1. Biogénesis y secreción de los miRNAs

La biogénesis de los miRNAs comienza con la transcripción en el núcleo del gen del miRNA seguido de una serie de incisiones del transcrito primario para formar un precursor. Posteriormente la proteína AGO realiza una selección de la cadena guía y la transporta a su sitio de acción; generalmente el mRNA asociado a ese miRNA. Además, se puede liberar a la matriz extracelular a partir de diferentes mecanismos de secreción como necrosis, apoptosis, asociado a lipoproteínas y AGO2 o bien a través de microvesículas.

### 2.2. Los miRNAs como reguladores de la expresión génica

La principal función de los miRNAs se centra en el silenciamiento genético a través de la represión postranscripcional del mRNA. Este proceso, se desarrolla al inhibir la traducción, por un mecanismo efectuado por el complejo silenciador inducido por miRNAs (miRISC) que está integrado por el miRNA maduro y su proteína argonauta (O'Brien et. al., 2018).

El miRISC reconoce su blanco por complementariedad de bases entre la región semilla del miRNA y la región UTR 3' del mRNA. No obstante, se ha



observado que el miRNA también puede reconocer secuencias en diferentes regiones del cuerpo del mRNA (Xu et. al., 2014) tales como la región codificante (Plotnikova et. al., 2019).

Como se ha mencionado la región de reconocimiento o semilla que consta de los nucleótidos 2-8 de la secuencia del miRNA, es decir, es una región muy pequeña, por lo tanto, existe una gran variedad de emparejamientos de miRNA-mRNA. De este modo, tanto un miRNA puede regular una gran cantidad de mRNAs como un mismo mRNA puede ser regulado por múltiples miRNAs. Además, estos pueden reconocer a su blanco por complementariedad de bases de los nucleótidos 13 a 16, conocida como región suplementaria (Gebert & MacRae, 2019).

Una vez reconocido el mRNA blanco se produce el silenciamiento génico por dos mecanismos de acción. Por un lado, se reprime el proceso de traducción del mRNA al impedirse el acceso de la maquinaria traduccional en la región UTR 3'. Por otra parte, se lleva a cabo la degradación del mRNA. La degradación se puede realizar directamente por la proteína AGO, o bien, la AGO puede iniciar el reclutamiento de proteínas de degradación. En el humano únicamente AGO2 presenta actividad de endonucleasa (Jo et. al., 2015).

Los otros miembros de la familia AGO igualmente pueden inducir la degradación del mRNA, a través del reclutamiento y ensamblaje de la maquinaria de degradación. Este reclutamiento proteico comienza con la familia de proteínas GW182; TNRC6A, TNRC6B y TNRC6C (Niaz & Hussain, 2018) y la proteína de unión a poliadenilato citoplasmática (PABPC) (Schäfer et. al., 2019). Las cuales servirán como marcaje del blanco y se encargarán de reunir proteínas efectoras tales como: el complejo de deadenilasa PAN2-PAN3 y el complejo represor de catabolito de carbono 4 negativo (CCR4-NOT) (Behm-Ansmant et. al., 2006) que actuarán en la deadenilación de la cola de poli A. Simultáneamente, se lleva a cabo la decapitación del mRNA por el complejo de enzimas decapitadoras de mRNA DCP1–DCP2 (Mugridge et. al., 2016) produciendo la desestabilización molecular del mRNA. Con ello, se produce una degradación rápida de la secuencia por la exoribonucleasa 1 (XRN1) (Gebert & MacRae, 2019).

El 66 al 90% de los mRNAs que son sintetizados en la célula se regulan por RISC (Gebert & MacRae, 2019). Principal razón por la cual se asume que los miRNAs actúan únicamente como silenciadores de la expresión génica. Estudios experimentales recientes han determinado que AGO2 y la proteína 1 relacionada con el síndrome de retraso mental X frágil (FXR1), están involucrados en la regulación génica positiva inducida por miRNAs dependiente de la fase del ciclo celular. En el citoplasma se han identificado elementos conocidos como cuerpos p, que están compuestos de una gran variedad de proteínas reguladoras. Entre ellas AGO2 y FXR1, quienes se asocian con elementos ricos en AU (ARE) altamente conservados en la región UTR 3' para activar la traducción. No obstante, éste es un campo de investigación que se encuentra aún en desarrollo (Vasudevan & Steitz, 2007). Es de vital importancia en la interpretación de resultados de los estudios con miRNAs que estos no solo participan en la regulación postranscripcional de los genes, si no, que también pueden estar activando su expresión.

### 2.3. Aplicación de los miRNAs como biomarcadores

Los miRNAs que se conocen actualmente se encuentra altamente caracterizada en el entorno intracelular. Sin embargo, se ha visto que además de desempeñar un importante papel como regulador de la expresión génica dentro de las células, los miRNAs se han encontrado presentes en una gran diversidad de fluidos biológicos, tales como: suero, plasma, lágrimas, orina, leche materna, lavado broncoalveolar, líquido cefalorraquídeo, semen, líquido amniótico, entre otros. Su presencia en estos fluidos corporales sugiere que los c-miRNAs son secretados y pueden ser un reflejo de los eventos celulares acontecidos en los tejidos colindantes (Godoy et. al., 2018).

Bajo esta premisa, numerosos estudios genómicos se han dedicado a caracterizar el perfil de expresión de miRNAs en biofluidos en diferentes condiciones fisiopatológicas. De este modo, se han determinado irregularidades en el patrón de expresión de una gran variedad de miRNAs que se han asociado a diferentes patologías tales como: diabetes *Mellitus*, cáncer, enfermedades cardíacas y neurodegenerativas y en diferentes complicaciones perinatales (Wu et. al., 2021; Zhao et. al., 2013).

### 3. Fisiología del embarazo

El embarazo es un periodo en el que acontecen diversos eventos biológicos a nivel molecular que involucran variaciones sistémicas en la mujer para mantener el adecuado crecimiento y desarrollo del feto. No obstante, en la actualidad dichos eventos aún no se encuentran completamente caracterizados (Soma-pillay et. al., 2016).

Clínicamente se considera que la evolución a término del embarazo presenta una duración aproximada de 280 días. La manera de contabilizar el inicio del embarazo es a partir del primer día de la última fecha de menstruación y es el método primario por el cual se calcula la edad gestacional o semanas de gestación (SDG). No obstante, a nivel celular el desarrollo comienza el día de fecundación del ovocito por el espermatozoide (Posfai et. al., 2019).

En este momento la mujer genera una respuesta sistémica coordinada entre sus diferentes órganos para sostener el embarazo, ya sean, endocrinos, metabólicos, inmunológicos, o cardiovasculares (Hu & Pasca, 2016; Sanghavi & Rutherford, 2014). Dentro de los cuales destacan los cambios morfológicos en el útero para albergar al feto en desarrollo (Pijnenborg et. al., 2006), el aumento del gasto cardíaco, disminución de la resistencia vascular, aumento en el volumen globular y plasmático y disminución de la capacidad pulmonar total por la elevación del diafragma, entre otros cambios fisiológicos (Soma-Pillay et. al., 2016).

La aportación nutricional de la madre al feto exige una mayor ingesta tanto de macro como de micronutrientes, incluidos mayores requisitos de proteínas, hierro, calcio, ácido fólico y otras vitaminas y minerales. Asimismo, los niveles séricos de colesterol y triglicéridos aumentan, para elevar la biodisponibilidad de glucosa sérica para el feto, todo ello estrictamente modulado por la placenta (Sun et. al., 2020).

#### 3.1. Desarrollo embrionario y fetal

La gestación en los humanos comienza desde la fecundación del ovocito por el espermatozoide para dar lugar a un cigoto diploide. Comienza con un proceso proliferativo duplicando el número celular por múltiples segmentaciones del cigoto a los que denominamos blastómeros (Milewski & Ajduk, 2017). Una vez alcanzada la

segmentación de 32 blastómeros, estos se compactan entre sí sin cambiar el tamaño del embrión y reciben el nombre de mórula (Georgadaki et. al., 2016).

La mórula atraviesa un proceso de polarización de las células externas comenzando a dar origen a las primeras dos estructuras del blastocito el trofoectodermo (TE) compuesto por células polarizadas y la masa compacta de células internas (MCI) no polarizadas (Posfai et. al., 2019). Finalmente, ocurre un proceso morfológico llamado cavitación por el cual los blastómeros se compactan para formar una cavidad compuesta por líquido, el blastocele. En este momento el embrión es denominado como blastocito (Gerri et. al., 2020).

Una vez completada la etapa de fecundación, el blastocito comienza su migración para continuar su desarrollo en el útero, después de la implantación. Durante este periodo se desarrollan un gran número de eventos celulares y moleculares principalmente de proliferación y diferenciación para la formación de los tejidos propios del embrión y los extraembrionarios (Gerri et. al., 2020). No obstante, esta etapa crucial en el desarrollo embrionario ha sido poco explorada en el humano debido a cuestiones éticas. Por lo que, el entendimiento de estos eventos es muy limitado, basándonos únicamente en modelos de células madre y cultivos in vitro para simular la implantación y las etapas posteriores de la embriogénesis (Ojosnegros et. al., 2021; Semi & Takashima, 2020).

Debido a los diferentes niveles hormonales de estrógenos y progesterona durante la implantación, el endometrio sufre numerosos cambios moleculares, morfológicos e histológicos, para la recepción del ovocito fecundado (Kim & Kim, 2017). Una vez establecido el primer contacto entre el blastocito y el endometrio comienza el proceso de implantación el cual se ha clasificado en tres etapas principales: la aposición, adhesión e invasión (Ochoa-Bernal & Fazleabas, 2020).

Durante la aposición se establece el primer contacto físico, donde se establece el contacto con las membranas celulares receptoras de las células apicales del endometrio y el blastocito. De manera general, el blastocito se posiciona en contacto con el endometrio. No obstante, cuando este contacto no se dé en esta posición, la MCI se moverá dentro del blastómero para obtener la orientación adecuada. De esta manera, comienza el proceso de adhesión, entre las células trofoblásticas del

blastocisto y las células endoteliales del endometrio estableciendo así el primer puente de comunicación molecular entre la madre y el embrión (Ojosnegros et al., 2021b).

En la etapa de invasión las células trofoblásticas del blastocisto penetran en el endometrio uterino e invaden el estroma endometrial para alcanzar los vasos sanguíneos maternos. Para ello, se activan vías de degradación de la matriz extracelular, donde participan diferentes gelatinasas y metaloproteinasas (MMP), que permiten al trofoblasto invadir las diferentes capas celulares del endometrio (Jia-Yu Zhu et. al., 2012). Al mismo tiempo las células trofoblásticas van sufriendo un proceso de diferenciación dando origen al sincitiotrofoblasto (STB), mientras que las células trofoblásticas que rodean la MCI se denominan citotrofoblastos (CTB) (Ojosnegros et al., 2021b).

A medida que el proceso de implantación avanza, el CTB también sufre un proceso de diferenciación por el cual se da la formación de dos capas celulares conocidas como hipoblasto y epiblasto, que más adelante darán origen al ectodermo y al endodermo respectivamente. Esta estructura de disco plano es conocida como disco embrionario bilaminar, dentro de la capa epiblastica se genera un hueco que dará origen a la cavidad amniótica. Mientras que las células hipoblasticas que tapizan el blastocele dan origen al saco vitelino primario (West et. al., 2019).

Más adelante en el desarrollo embrionario alrededor de la tercera semana de gestación, se genera la diferenciación de tres tipos celulares distintos: el mesodermo, el endodermo y el ectodermo, en un proceso conocido como gastrulación (Muhr & Ackerman, 2022). Una vez se encuentra bien definido el disco embrionario trilaminar comienza un proceso de plegamiento que da pie al desarrollo de la organogénesis.

Durante el transcurso de la tercera semana de gestación el primer sistema en desarrollarse es el sistema circulatorio a partir del mesodermo (Rossi et al., 2021). Aparecen los primeros vasos sanguíneos fetales. Al finalizar la cuarta semana de gestación, las cuatro cavidades del corazón se han establecido. No obstante, es alrededor de la sexta semana cuando comienza a establecerse el flujo sanguíneo. Simultáneamente, se forma la placa neural que posteriormente se pliega para

formar el tubo neural, cerrándose al concluir la cuarta semana de gestación (Zhai et al., 2022).

En la octava semana de gestación se han desarrollado la mayor parte de los diferentes órganos y sistemas del feto (Amadei et al., 2022). La transición del periodo embrionario al periodo fetal ocurre en la novena semana de gestación. En este momento el feto entra en un periodo de crecimiento y maduración. No obstante, un número importante de órganos continúan con su desarrollo a lo largo del periodo fetal, como: el sistema nervioso central (SNC), el tracto urinario y los pulmones (Amadei et al., 2022).

Los estudios del desarrollo fetal en humanos se han visto limitado prácticamente a la medición de datos clínicos (Gaccioli et al., 2017). Dentro de los cuales los más relevantes son las antropometrías fetales obtenidas a través de la ultrasonografía. De esta manera, se puede evaluar el crecimiento del feto conforme avanza el embarazo de forma no invasiva.

Durante el primer trimestre de gestación se comprueba la viabilidad embrionaria y/o fetal, la cual, se refiere a la presencia de la actividad cardiaca al momento de la medición; se mide el diámetro medio del saco gestacional (DMSG), o bien, la longitud cráneo-rabadilla (LCR) solo en presencia del feto, para determinar la edad gestacional. Adicionalmente, de ser posible en el primer trimestre también se realizan mediciones del diámetro biparental y la circunferencia cefálica, además, de la evaluación de estructuras anatómicas tales como: la cabeza, el cuello, el rostro, la espina, el pecho, el corazón, las extremidades, la placenta y el cordón umbilical. No obstante, estas mediciones se realizan hasta después de la treceava semana de gestación. Además, de acuerdo con las políticas de salud vigentes puede realizarse el tamizaje de aneuploidías por la medición de translucencia nucal (Salomon et al., 2019).

Durante el embarazo medio y tardío se incluyen las mediciones de circunferencia abdominal (CA), la longitud del fémur (LF), estimación del peso fetal, volumen de líquido amniótico, y movimientos fetales. Del mismo modo, se realiza nuevamente las evaluaciones señaladas anteriormente para el primer trimestre, de

una manera más detalla para la evaluación de estructuras anatómicas y de la descripción de la apariencia y la localización de la placenta (Salomon et al., 2019).

No obstante, hasta el momento se ha observado que la medición de biometrías fetales y placentarias, a pesar de ser una poderosa herramienta para esbozar el desarrollo del embarazo, concretamente del feto, su funcionalidad queda restringido a evaluar el crecimiento fetal (Arenas et al., 2005; DeVilbiss et al., 2020). Además, esto presenta grandes limitaciones, debido a sesgos en las mediciones, imprecisiones para determinar rangos normales de crecimiento de acuerdo con diferentes poblaciones y aún más importante, las anomalías observadas por USG se ven limitadas a alteraciones morfológicas o del crecimiento que solo son perceptibles cuando el daño es demasiado grande (Kumar et al., 2017).

### 3.2. Desarrollo placentario

La placenta es un órgano derivado del desarrollo embrionario, su formación se ha categorizado en dos etapas; la etapa pre-vellositaria y vellositaria. A su vez, la etapa pre-vellositaria se divide en la fase prelacunar y lacunar. Durante la primera fase se desarrolla la implantación del blastocito, donde las células polares del TE adyacente a la MCI comienzan a invadir el epitelio endometrial, el cual a partir de este momento recibe el nombre de decidua. Las células trofoblásticas invasoras se fusionan entre sí para formar un sincitio primario.

En el momento en que el blastocito está completamente incrustado en la decidua y se encuentra cubierto por el tejido epitelial externo, comienza la fase lacunar. Durante esta fase aparecen espacios compuestos por líquido que se expanden y fusionan seccionando la masa sincitial por la formación de trabéculas. Por otra parte, durante la etapa vellositaria el CTB prolifera hasta atravesar el sincitio primario y establecer contacto con la decidua. A su vez, se comienzan a formar ramificaciones que darán lugar a las vellosidades coriónicas, de modo que las lagunas se convierten en el espacio intervelloso. Mientras que células individuales de CTB invaden la decidua convirtiéndose en el citotrofoblasto extravelloso (CEV).

El STB recubre la cara exterior de las vellosidades placentarias y establece contacto con los vasos sanguíneos y glándulas secretoras maternos. De esta manera, se establece el medio de comunicación entre la madre y el feto a través del

intercambio de nutrientes, gases, hormonas y diferentes productos de desecho. Con la finalidad de abastecer al feto y asimismo con los requerimientos necesarios para su desarrollo. Es por ello, que el STB cuenta con una extensa gama de proteínas transportadoras de glucosa, aminoácidos y otros nutrientes tanto en sus membranas apicales como basales. Además, cuentan con una gran cantidad de receptores de hormonas y estímulos de crecimiento, además de expresar receptores Fc neonatales (FcRn) que permiten el transporte de anticuerpos de la circulación materna a la circulación fetal.

El desarrollo de placenta en euterios es considerado un símbolo evolutivo que permite a estos animales diversos beneficios que procuren su supervivencia, tales como: la protección contra fluctuaciones ambientales, depredadores, el suministro constante de nutrientes, entre otros factores, aunque esto implica un alto cargo fisiológico en la salud materna.

En este sentido se han identificado grupos de miRNAs altamente asociados con el desarrollo placentario entre ellos los miRNAs del cromosoma 14 (C14), se encuentra ubicado en el loci 14q32. Su expresión se encuentra ligada a la impronta materna y se ha encontrado principalmente expresado en tejido placentario y embrionario. La placenta humana presenta una mayor expresión en el embarazo tardío de este grupo de miRNAs. Además, es altamente conservado a través del tiempo en mamíferos placentarios.

Por otra parte, los miRNAs del cromosoma 19 (C19) ubicados en el loci 19q13.41 del humano, su expresión está ligada a la impronta paterna y se expresa casi únicamente en el tejido placentario. Presenta una alta expresión en etapas tempranas de la gestación por lo que se ha asociado primordialmente al desarrollo temprano de la placenta. Estudios ortológicos han demostrado que su expresión se reduce únicamente a primates. Además, estudios de transfección en ratos han demostrado, que su alta expresión está altamente vinculada con el arreglo tisular que presenta la placenta en los primates y un tamaño más grande. Además, se ha demostrado como su inhibición en diferentes cultivos celulares de trofoblasto interfiere con el proceso de migración e invasión del trofoblasto.



## Planteamiento del problema

En la actualidad, no se sabe con certeza cuales son los diferentes eventos biológicos que se desarrollan a lo largo del embarazo, ni la forma en que estos se modulan entre sí para el mantenimiento del embarazo. Además, un número creciente de estudios ha demostrado que el desarrollo de complicaciones perinatal no está ligadas a un solo agente causal, es decir, derivan de eventos multifactoriales que conllevan el desarrollo de condiciones patológicas durante el embarazo que afectan tanto la salud de la madre como la del feto.

En general en los países de bajos ingresos prevalecen altos índices de mortalidad materna principalmente asociados a infección y hemorragias o a otras complicaciones relacionadas con la limitada atención prenatal que recibe la mujer durante la gestación y el parto (OMS, 2019). No obstante, aún en países desarrollados y en desarrollo siguen presentándose altos índices de mortalidad materna y de complicaciones perinatales (Say et. al., 2014).

Además, cerca de dos millones de bebés nacen muertos cada año en el mundo. A pesar de que las cifras estiman que al igual que la mortalidad materna, el 87% de las muertes fetales se producen principalmente en países de bajos ingresos, debido a los limitados servicios de salud disponibles (OMS, 2020a). Es evidente que un porcentaje importante de ellas no tienen una causa concreta que desencadene la muerte fetal. Sin mencionar, que además de las pérdidas fetales, prevalece un alto índice de muertes neonatales y estados de salud inadecuados para el feto tanto en el embarazo como en los primeros días de vida (OMS, 2020b).

Los datos anteriores reafirman lo que varios investigadores han postulado desde hace varias décadas sobre el desarrollo de complicaciones perinatales en la mujer y el desenlace adverso del feto al nacer como: cuestiones de etiología multifactorial, cuyo desarrollo depende de factores intrínsecos y extrínsecos, tales como el ambiente, la dieta, exposición a contaminantes, estrés y el propio genoma aunado a sus múltiples mecanismos de regulación.

Por lo tanto, en los últimos años ha aumentado el interés por determinar posibles biomarcadores poco invasivos capaces de monitorear el desarrollo

saludable del embarazo con precisión desde etapas tempranas de la gestación. El objetivo de este estudio es determinar la asociación de una firma de c-miRNAs con el desarrollo fetal y placentario. Al mismo tiempo, evaluaremos sus patrones de expresión a lo largo del embarazo, así como, una estimación de los diferentes procesos biológicos que están interviniendo en el desarrollo del embarazo normal y que podemos observar reflejados en los c-miRNAs.

## Objetivos

Determinar cómo los c-miRNAs maternos correlacionados con estimadores de crecimiento fetal están asociados con el desarrollo fetoplacentario

### Objetivos particulares

- Validar si la firma de crecimiento fetal permite estimar indicadores de crecimiento fetal (longitud de fémur).
- Determinar si la firma de crecimiento fetal se correlaciona con c-miRNAs de función y/o expresión en placenta.
- Identificar los blancos génicos de la firma de crecimiento, los procesos biológicos y vías de señalización principales de los blancos.

## Hipótesis

Los patrones de expresión de los c-miRNAs de la firma de crecimiento fetal están asociados a procesos del desarrollo fetoplacentario

## Metodología

Para los análisis presentados en este trabajo, se emplearon datos de expresión de c-miRNAs en el plasma de mujeres embarazadas. Estudios previos realizados por

la *Unidad de Vinculación Científica-Facultad de medicina (UVA-FM)* del Instituto Nacional de Medicina Genómica (INMEGEN), determinaron el transcriptoma de sncRNAs en diferentes puntos temporales de un mismo grupo de mujeres embarazadas por secuencia de nueva generación (Hernández, et. al., 2021.).

#### Recolección de muestras

Se seleccionaron 8 mujeres embarazadas de la Cohorte Perinatal de la CDMX construida en el Hospital General “Dr. Enrique Cabrera”. Para las mujeres incluidas en el estudio se realizó un muestreo por trimestre de gestación y se incluyeron únicamente a las mujeres con embarazos normo evolutivos, con una edad entre 19-35 años y parámetros bioquímicos dentro de los límites de referencia de acuerdo con la edad gestacional, que presentaran un producto único de gestación y con un seguimiento clínico desde la decimoctava semana de gestación. Se descartaron mujeres que desarrollaran complicaciones perinatales, que presentaran pruebas bioquímicas fuera de los límites de referencia, embarazo gemelar o que refirieron el consumo de agentes nocivos para la salud durante el periodo de estudio. Todos los embarazos fueron a término (> 37 SDG) y sus bebés presentaron peso y talla adecuados para la edad gestacional, de las 8 mujeres 5 tuvieron bebés de sexo femenino y de 3 de sexo masculino.

#### Extracción de RNA y datos de secuenciación

El RNA total se extrajo a partir de 2 mL de plasma materno por trimestre de gestación por la técnica de Trizol-cloroformo. Para la construcción de la biblioteca se empleó el kit NEXTflex Small RNA-Seq v3 (Bioo Scientific) y se corroboró su integridad por TapeStation de Agilent Technologies. La secuenciación se realizó en la Unidad de Alta Tecnología (UAT) de Secuenciación del INMEGEN en la plataforma de NextSeq500 Illumina durante 150 ciclos de lectura única (Illumina, 2019).

La alineación y el análisis de los datos de secuenciación se procesaron a partir de archivos FASTQ sin procesar. Para el control de calidad de las secuencias por FastQC, se empleó un criterio de calidad de 30 en la escala Phred score (Babraham Bioinformatics, 2020). Para el corte de adaptadores se utilizó el programa computacional de Cutadapt v.1.10 (Marcel, 2020). El alineamiento de las

secuencias se realizó en BWA; v.0.7.12 (H. Li & Durbin, 2010) empleando el genoma de referencia humano Hg38p12. Se determinó el perfil de anotación para miRNAs con miRDeep2 (Friedländer et. al., 2012) con miRBase como base de referencia. Los conteos de las lecturas para todas las muestras se normalizaron por cuantiles.

#### Recopilación de datos clínicos

La recopilación de datos clínicos consistió en las siguientes variables: edad, talla, peso pregestacional (Peso-pg), peso por trimestres de gestación, índice de masa corporal pregestacional (IMC-pg) y tensión arterial (TA). Por otra parte, dentro de los datos sociodemográficos se consideró la escolaridad, etnia, estado civil, toxicomanías reportadas, actividad económica e historial obstétrico. Finalmente se consideraron los niveles de glucosa y el perfil lipídico (triacilglicéridos, colesterol, HDL, LDL) dentro de los límites de referencia de acuerdo con la edad gestacional y grupo de estudio (Abbassi-Ghanavati et. al., 2009).

#### Validación de la firma de c-miRNAs con indicadores del crecimiento fetal

La firma está constituida por seis c-miRNAs correlacionados positivamente con indicadores de crecimiento fetal (LF y PFE). Para validar la firma de los c-miRNAs asociados con longitud de fémur y peso estimado fetal, se realizó un análisis de random forest. Para ello, se examinaron un total de 588 miRNAs identificados en el plasma de mujeres embarazadas y se evaluó la capacidad de cada uno de ellos en la estimación de valores de LF. Posteriormente se comparó si la firma de crecimiento fetal previamente establecida se encontraba significativamente enriquecida entre los miRNAs de importancia en el modelado de random forest para la estimación de LF.

#### Asociación entre miRNAs del C14 y C19 con firma de CF

Con el propósito de determinar si los miRNAs de la firma de CF se encontraban significativamente enriquecidos entre el grupo de miRNAs del C14 y C19. Se determinó la superposición de la firma de crecimiento fetal con cada uno de estos conjuntos, simultáneamente se realizó esta misma prueba un millón de veces con conjuntos de miRNAs aleatorizados y se calculó la probabilidad de que estos presentaran una superposición igual a la encontrada con la firma de crecimiento fetal.

Adicionalmente, se calculó el coeficiente de correlación de Spearman que presentan los miRNAs de C14 y C19 con datos de peso fetal estimado y LF. Posteriormente, se determinó su correlación media. Para determinar si esta correlación era significativa se calcularon coeficiente de correlación de Spearman con muestras de miRNAs aleatorias y determinó la correlación media de cada una de estas muestras randomizadas. El ensayo se realizó un millón de veces y se calculó el valor de significancia de la prueba.

#### Determinación de blancos génicos

Para determinar los blancos génicos de los miRNAs que integran la firma de crecimiento fetal, se consultó la base de miRTarBase 9.0 v22 (Huang et. al., 2020), este es un repositorio público de datos curada manualmente que integra los blancos génicos validados experimentalmente para los diferentes miRNAs identificados actualmente. De este modo, conjuntamente identificamos a todos los blancos génicos en donde exista evidencia en la que se demuestre la regulación de un blanco por al menos uno de los miRNAs de la firma de crecimiento fetal, C14 y C19. Adicionalmente, filtramos el total de blancos anotados, utilizando un criterio de fuerte evidencia experimental, que consistió en conservar aquellos blancos génicos que fueran validados por al menos una de las siguientes técnicas: Western-Blot, qRT-PCR y ensayos de reporte de luciferasa o GFP (Green Fluorescent Protein).

#### Red de interacción proteína- proteína

A partir de los blancos identificados se construyó una red de interacciones proteína-proteína (IPP), empleando interfaz gráfica disponibles en la base de datos de STRING 11.5 (Szklarczyk et. al., 2021). Para la construcción de la red se consultaron únicamente aquellas interacciones físicas que presentaran evidencia experimental.

Para determinar la significancia de las interacciones detectadas en la red IPP se realizó una prueba independiente a la proporcionada en la interfaz de STRING, esta consistió en determinar la proporción de interacciones que conforman la red entre el número máximo de interacciones posibles dado el número de proteínas que la conforman. Para ello, se consultaron aquellas interacciones físicas con evidencia experimental reportadas en la base de datos de STRING con un score mínimo de

interacción de 0.400. Simultáneamente se calculó el número máximo de interacciones posibles bajo la siguiente formula:  $IPP_{max} = \frac{nProteínas^2 - nProteínas}{2}$

Posteriormente, se determinó la relación de IPP presentes en la red ( $IPP/IPP_{max}$ ). Finalmente, se realizó una prueba numérica, en la cual se determinaron 10, 000 conjuntos aleatorios de miRNAs, para cada uno de ellos se determinaron sus blancos génicos y posteriormente sus productos proteicos. Una vez determinados se calculó la proporción de IPP presentes en cada red. Con base en esto se construyó una distribución de proporciones de IPP y se determinó la probabilidad de que la red construida con los productos regulados a partir de la firma de CF fuera más alta que las proporciones de IPP de redes aleatorias.

Para determinar el sesgo de la proporción de IPP entre redes construidas a partir de productos proteicos aleatorios contra productos proteicos regulados por uno o más miRNAs en común. Se empleo la estrategia implementada anteriormente, apoyándonos de los datos recopilados en la base de datos de miRTarBase y TargetScas. Esta consistió en determinar los blancos génicos para cada uno de los miRNAs anotados en estas bases, en el caso de miRTarBase se emplearon tanto el total de blancos anotados por miRNA, así como, aquellos blancos que presentaran fuerte evidencia experimental. En el caso de TargetScan, únicamente se evaluaron los blancos génicos que presentaran un score de contexto dentro del percentil 90. En todos los casos se construyó una distribución aleatoria basada en blancos génicos aleatorios y blancos génicos predichos por miRNAs, ya sean con evidencia experimental (miRTarBase) o por complementariedad de bases con su mRNA (TargetScan).

### Enriquecimiento Funcional

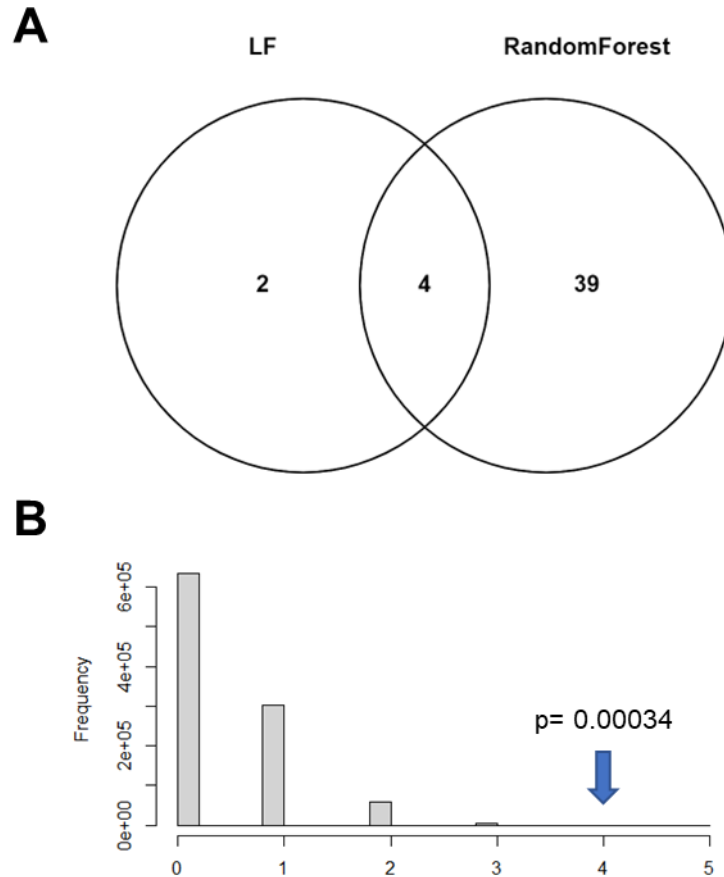
En primera instancia, determinamos los procesos biológicos enriquecidos por solapamientos, asociados a los blancos génicos de la firma de crecimiento fetal a través de PANTHER (Carbon et. al., 2021), empleando como lista de referencia los blancos génicos de los 1449 miRNAs identificados en circulación. Del mismo modo, identificamos los procesos biológicos enriquecidos dentro de los blancos génicos asociados a los miRNAs del C14, C19 y de la firma del CF.

Además, realizamos una prueba automatizada en la cual se determina para cada categoría enriquecida en circulación, el número total de c-miRNAs anotados a los genes de cada categoría y calculamos las veces de cambio entre el grupo de mujeres NE y el grupo de mujeres E por trimestre de embarazo y DE (E/NE) o Fold Change (FC) para cada categoría. Para determinar si el FC es significativo realizamos pruebas randomizadas, en donde se determinaron 10, 000 conjuntos aleatorios de genes del mismo tamaño que el del término GO evaluado y se determinó el FC para cada conjunto. Con estos datos se construyó una distribución de FC y se determinó la probabilidad de que el FC obtenido entre los c-miRNAs entre mujeres embarazadas y NE fuera mayor de lo esperado al azar.

## Resultados

Un número creciente de estudios han argumentado el potencial de los c-miRNAs como biomarcadores no invasivos en diferentes condiciones patológicas (Wang et. al., 2018; Wu et. al., 2021). Estudios previos realizados por la UVC del INMEGEN han observado que los miRNAs son la clase de ncRNA más abundante en el plasma de mujeres embarazadas cuyas características clínicas se describen en la Tabla 1 del Apéndice A. Además, de demostrar la existencia de conjuntos de miRNAs a los que denominamos firmas que se encuentran específicamente asociadas a la evolución normal del embarazo. Entre ellas, se encuentra una firma de crecimiento fetal que presenta una alta correlación positiva con estimadores de crecimiento fetal medidos por USG como la LF y el PFE.

En este trabajo se caracterizó a los miRNAs que integran esta firma de crecimiento fetal para estimar su potencial como biomarcadores del desarrollo fetoplacentario. A través de un modelado por random forest validamos la capacidad de los miRNAs identificados en el plasma de mujeres embarazadas, para estimar la LF fetal. Este modelo identificó un total de 43 miRNAs de importancia para la estimación de la longitud del fémur. Al comparar estos con la firma de CF encontramos que compartían 4 miRNAs en común. Además, de que este solapamiento se encontraba significativamente enriquecido.

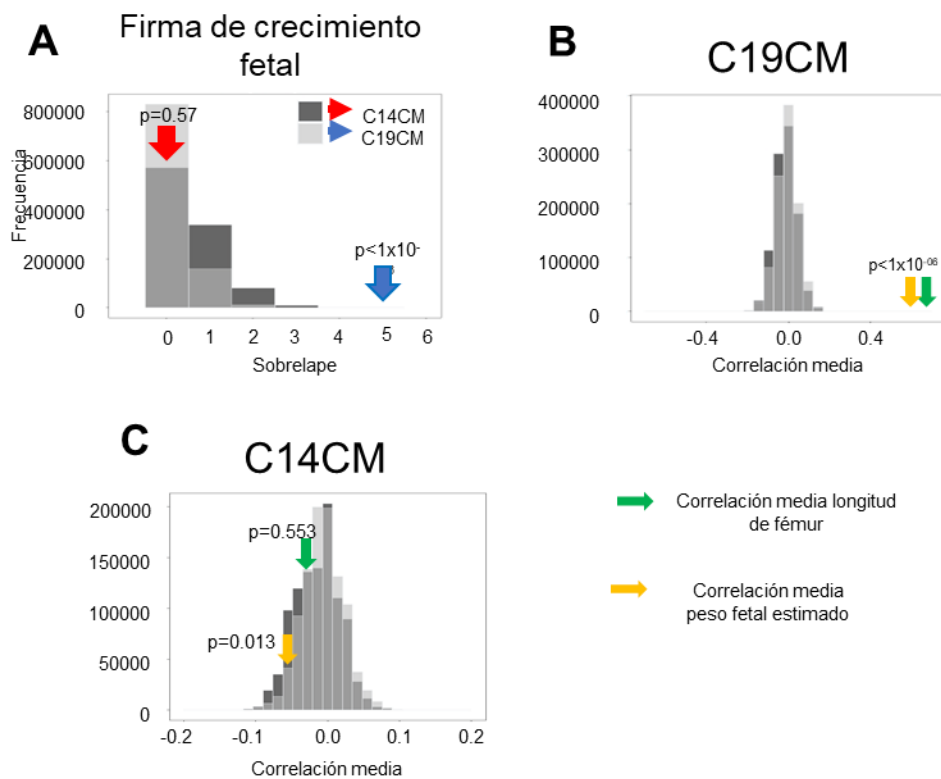


**Figura 1. miRNAs de la firma de CF son componentes de importancia en la estimación de la LF por random forest.** A) Diagrama de Venn que muestra los miRNAs compartidos entre la firma de CF obtenida por correlaciones de Spearman y los miRNAs de importancia en un modelo de regresión por random forest

Posteriormente, evaluamos su enriquecimiento entre los miRNAs pertenecientes al C14 y C19, estos consisten en dos conjuntos de miRNAs asociados a la función placentaria. Encontrando que los miRNAs correlacionados con crecimiento fetal se encuentran significativamente enriquecidos entre los miRNAs de C19, donde cinco de los seis miRNAs pertenecientes a la firma CF pertenecen a los miRNAs de C19, mientras que no se encontraron coincidencias



con los miRNAs de C14 (**Figura 2. A**). Adicionalmente, se determinó la correlación media que presentan los miRNAs del C14 y C19 con los datos de LF y peso fetal estimado. Del mismo modo, calculamos correlaciones medias aleatorias de diferentes conjuntos de miRNAs del mismo tamaño que las del C14 y C19, para determinar la probabilidad de obtener los mismos valores de correlación media obtenida previamente para C14 y C19. Encontrando, que la correlación media de los miRNAs del C19 presentan una correlación media positiva significativa con los indicadores de crecimiento fetal. Por el contrario, los miRNAs pertenecientes al C14 solo presentan una correlación media significancia con los datos de peso fetal estimado, cabe mencionar que la significancia que estos presentan es considerablemente menor a lo encontrado con los miRNAs de C19.

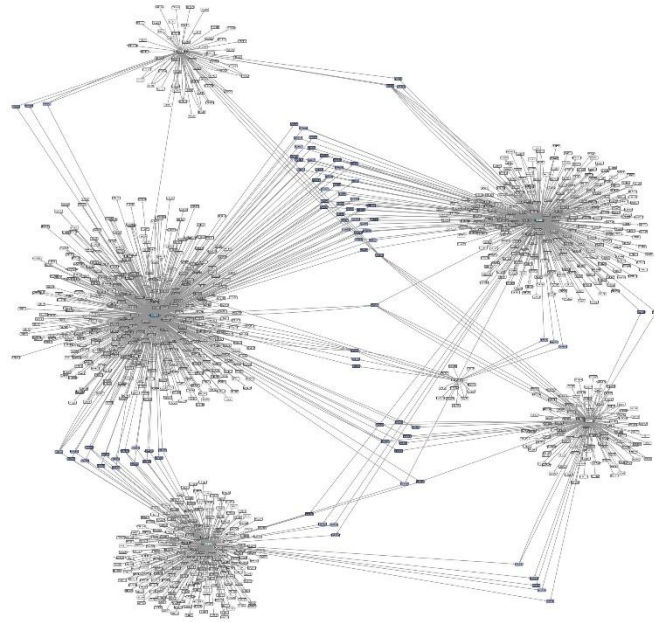
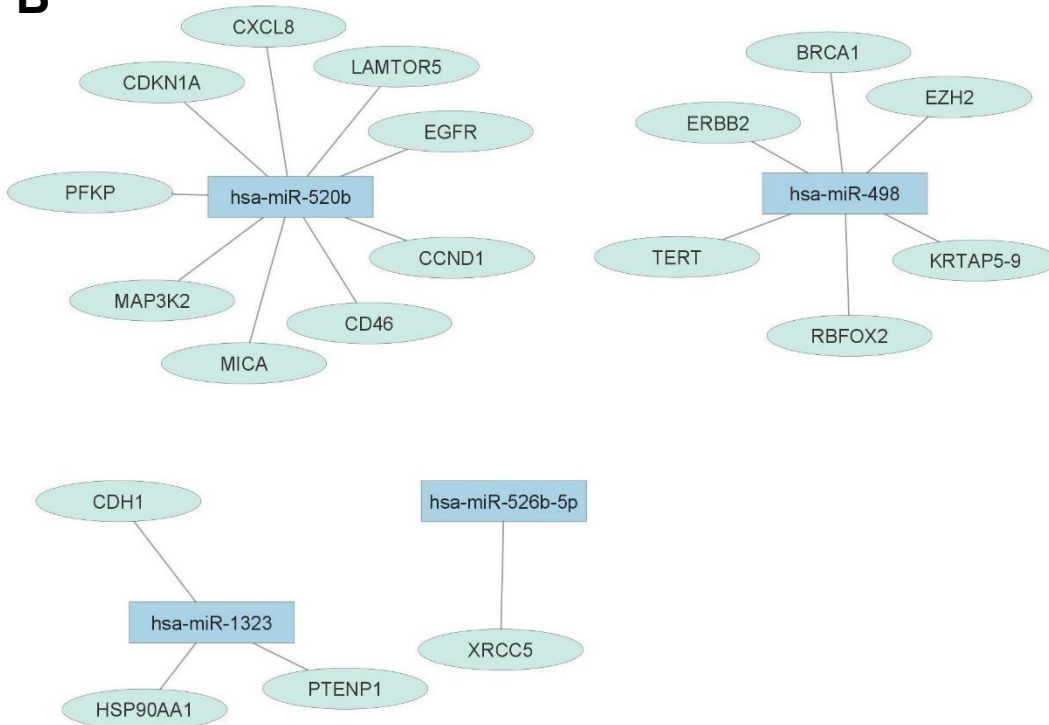


**Figura 2. Los miRNAs del C19MC se asocia significativamente con el crecimiento fetal.** A) Superposición obtenida (flechas) y superposición esperada (histogramas grises) entre C19CM y C14 CM y los seis miRNAs mayormente correlacionados con el crecimiento fetal. Correlación media de C19 (B) y C14 (C)

(flechas) contra el crecimiento fetal, en comparación con la distribución de correlación esperada (histograma gris).

Una de las más grandes perspectivas actuales de los c-miRNAs es poder inferir a través de ellos los posibles procesos biológicos afectados en diferentes condiciones fisiopatológicas. Dentro de los métodos propuestos para la determinación de blancos génicos para miRNAs, se han postulado diversos algoritmos que determinan la probabilidad de asociación de un miRNA con un blanco génico por complementariedad de bases dentro de estas herramientas la más utilizada es TargetScan (McGeary et. al., 2019). No obstante, observamos que por este método la detección de blancos presenta un gran sesgo ya que la probabilidad de encontrar complementariedad de base en alguna región del mRNA es muy alta. Por lo tanto, decidimos apoyarnos en el segundo método de detección de blancos génicos que se apoya en la base de datos de miRTarBase (Huang et. al., 2022). Esta base consiste en la recopilación y registro manual de toda la evidencia experimental disponible. Encontrando que en conjunto los seis miRNAs pueden estar regulando hasta 1,129 blancos génicos diferentes (**Figura 3.A**).

Debido a lo anterior, restringimos la búsqueda a únicamente aquellas asociaciones que presentaban fuerte evidencia experimental (Western-Blot, qRT-PCR y ensayos de reporte de luciferasa o GFP) y así eliminar aquellas asociaciones que puedan ser falsas por aproximaciones metodológicas poco precisas. De este modo, determinamos un total de 19 blancos génicos asociados a la firma de CF. Además, observamos que ninguno de los miRNAs que conforman la red presenta blancos génicos en común, lo cual sugiere que pueden no estar involucrados en la regulación de los mismos procesos biológicos.

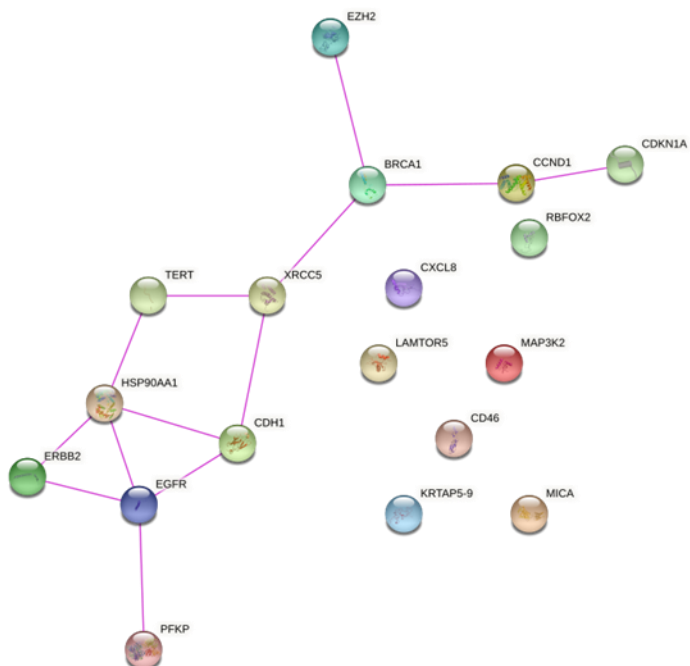
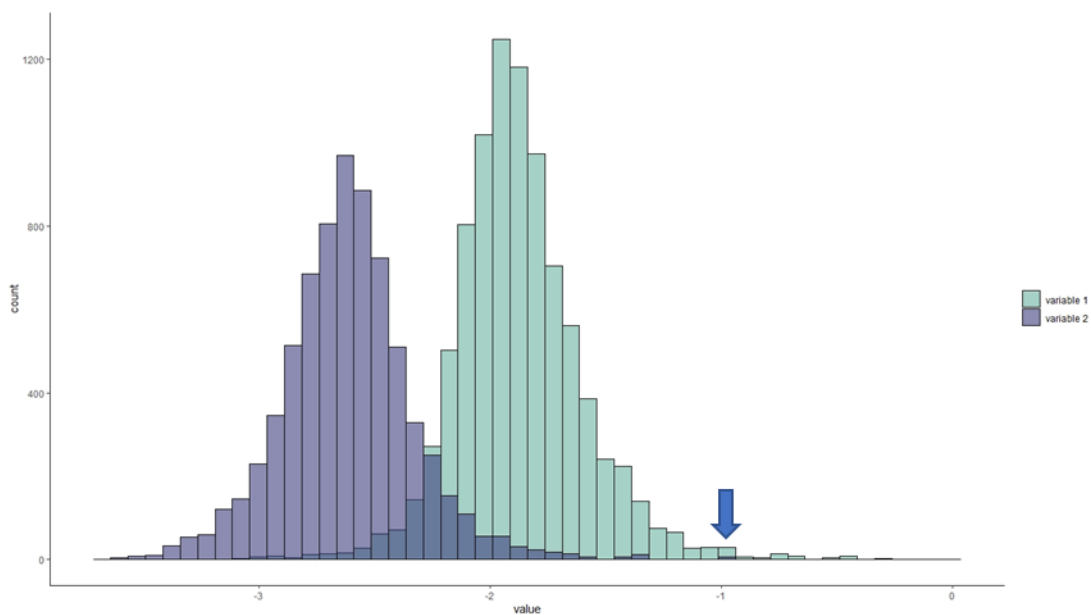
**A****B**

**Figura 3. Blancos génicos obtenidos para la firma de crecimiento fetal.** En cada agrupación se muestran a cada uno de los seis miRNAs asociados a crecimiento fetal, estos se encuentran unidos a sus respectivos blancos génicos obtenidos de la base de miRTarBase. **(A).** Red miRNA-Targets totales anotados en miRTarBase.

**(B).** Red miRNA-Targets con fuerte evidencia experimental (Western-Blot, qRT-PCR y ensayos de reporte de luciferasa o GFP) anotados en miRTarBase

No obstante, para determinar si existe relación funcional entre los blancos génicos conservados construimos una red de interacción proteína-proteína. Observamos que al menos 11 de estas proteínas presentaban interacciones físicas validadas experimentalmente (**Figura 4.A**). Inicialmente determinamos si el número de interacciones que conformaban la red presentaba significancia estadística a través del algoritmo de análisis presente en la interfaz gráfica de STRING donde obtuvimos un valor de significancia igual a  $5.15 \times 10^{-06}$ .

Sin embargo, al determinar la veracidad de este método para calcular la significancia estadística, a través de simulaciones con conjuntos de blancos génicos regulados por grupos de seis miRNAs aleatorios, observamos que la red siempre presenta significancia. Lo cual nos indica que independientemente de la relación de la firma con el CF, los blancos génicos regulados por algún miRNA en común presentan más IPP de lo esperado al azar. Por lo tanto, se realizó un ajusto independiente a través de simulaciones randomizadas y se recalculo el valor de p como se indica en los métodos en la **Figura 4.B** se muestra en color violeta la distribución de proporción de interacciones a partir de blancos aleatorios en color morado, donde, se observa un sesgo en la distribución hacia el lado izquierdo, mientras que la distribución en color verde muestra la proporción de interacciones de blancos con al menos uno de seis miRNAs en común. Observamos una clara tendencia significativa en la que la distribución verde tiende a presentar más IPP que conjuntos de blancos aleatorios. Además, observamos marcada con una flecha en color azul que la proporción de IPP que presenta la firma de CF es significativamente mayor a la presentada por blancos regulador por miRNAs aleatorios.

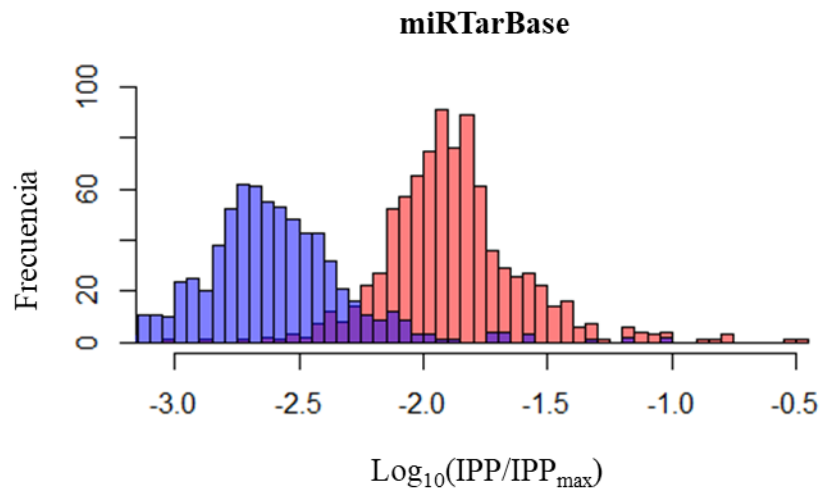
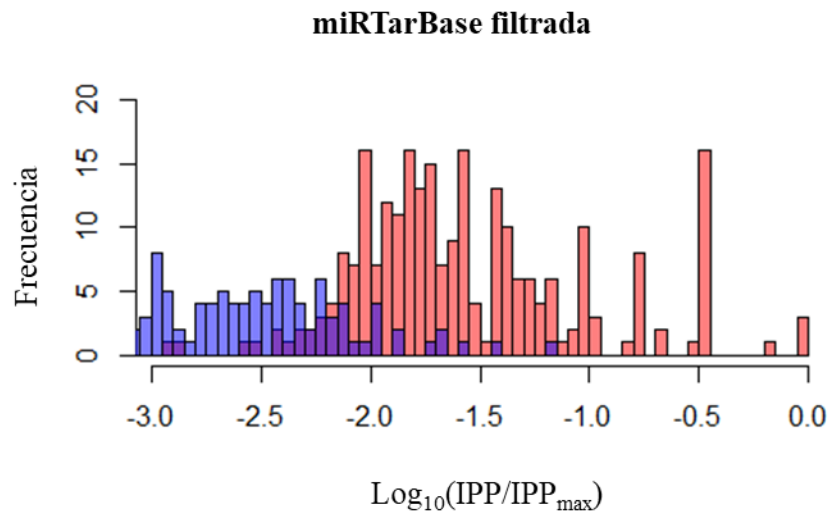
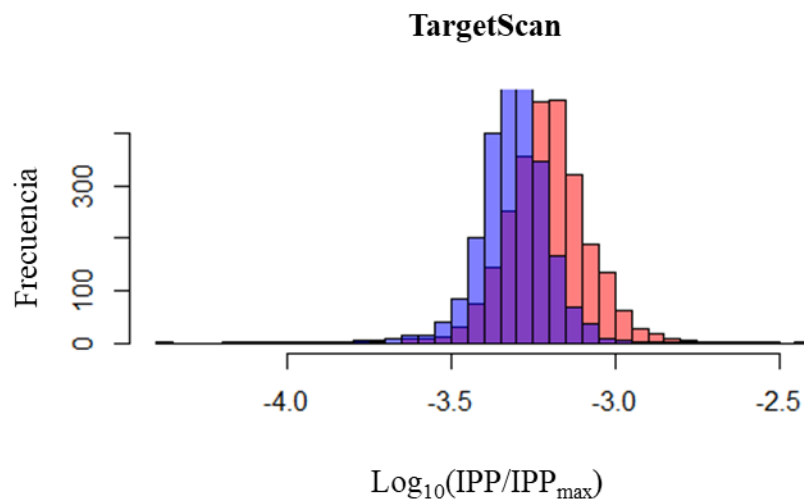
**A****B**

**Figura 4. Red de interacciones proteína-proteína de los blancos génicos obtenidos de la firma de crecimiento fetal. (A)** La red muestra las diferentes interacciones que se pueden presentar entre los diferentes blancos génicos identificados para la firma de crecimiento fetal  $p\text{-value} = 0.01$ . **(B)** Distribución de proporción de IPP de conjuntos aleatorios de blancos génicos (Morado) y proporción de IPP de blancos génicos regulados por conjuntos de seis miRNAs (verde).

Se determino la veracidad de este hecho, se realizan varias pruebas secundarias donde se evaluó en diferentes escenarios la capacidad de los miRNAs para regular blancos génicos cuyos productos proteicos tienden a estar relacionados entre sí por un mayor número de IPP de lo esperado al azar. Por un lado, evaluamos todos los blancos génicos anotados por miRNAs en la base de miRTarBase, observando nuevamente que la distribución tiene una mayor tendencia hacia las proporciones con más IPP, que blancos génicos aleatorios (**Figura 5.A**).

Además, corroboramos que el efecto observado no esté relacionado con el método de determinación de blancos empleado. Por lo cual, decidimos compararlo con un método de identificación de blancos diferente. Para ello, consultamos la base de datos de TargetScan que se basa en un algoritmo matemático que calcula un score de probabilidad basado en la estabilidad de la complementariedad de bases entre el miRNA y su mRNA blanco.

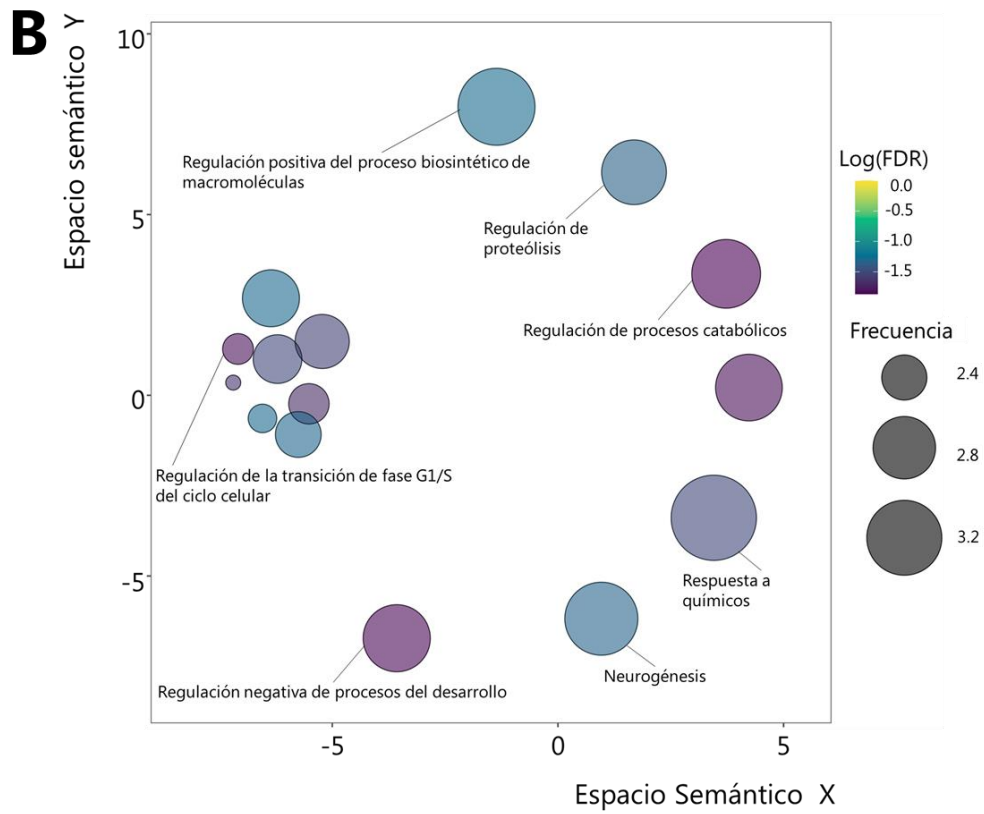
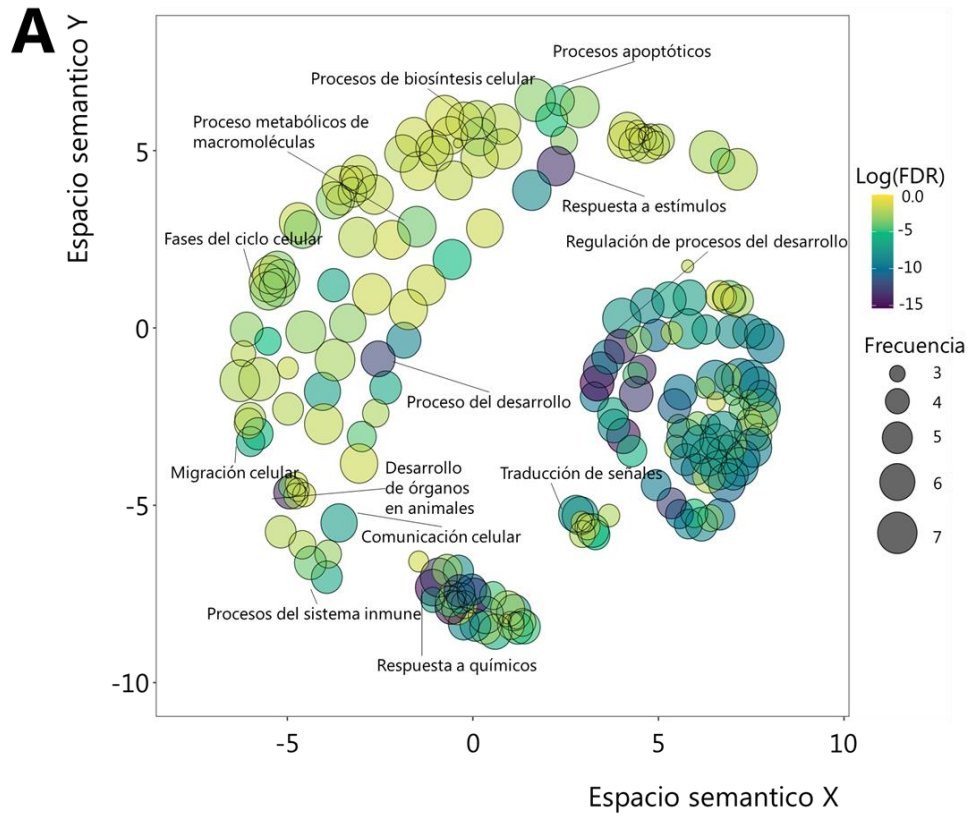
Al evaluar todos los posibles blancos detectados por este método, observamos que se presentaba gran imprecisión, debido a que el número de blancos detectados para un determinado miRNA era muy alto, además de compartir un gran número de blancos iguales entre diferentes miRNAs. Es decir, se presentaba una gran probabilidad de complementariedad de cada miRNA con cualquier gen. Por lo tanto, filtramos a aquellos blancos génicos con un score de contexto dentro del percentil 90. Finalmente, al construir y observar la distribución de número de interacciones para los blancos de cada miRNA aunado a una distribución aleatoria, observamos nuevamente que los blancos génicos regulados por miRNAs tienden a presentar un mayor número IPP (**Figura 5.B**) independientemente del método de predicción de blancos génicos.

**A****B****C**

**Figura 5. Distribución de proporción de IPP por miRNA.** Las distribuciones muestran en color rojo la frecuencia de las proporciones de IPP blancos génicos por cada uno de los miRNAs reportados, en color azul se muestra la distribución correspondiente a la frecuencia de la proporción de IPP de conjuntos de genes aleatorios. **(A)** miRNAs anotados en miRTarBase. **(B)** miRNAs anotados en miRTarBase con fuerte evidencia experimental. **(C)** miRNAs anotados en TargetScan.

Finalmente evaluamos los procesos los enriquecimientos funcionales para los miRNAs de la firma de CF, del C14 y del C19. Sin embargo, para la firma de CF no se encontró ninguna categoría de procesos biológicos significativa por análisis de solapamientos. Por otra parte, se registraron un gran número de categorías enriquecidas para los miRNAs del C14 y del C19, por lo cual, se empleó un reductor de categorías por asociación de similitud entre los términos para determinar las categorías con mayor relevancia biológica. Dentro de las cuales, la regulación de procesos del desarrollo, el desarrollo de órganos, la respuesta a estímulos, la migración y comunicación celular son algunos de los procesos biológicos más destacados para los blancos génicos del C14. Mientras que para los blancos génicos de los miRNAs del C19, destacan los procesos de ciclo celular, procesos catabólicos y neurodesarrollo (**Figura 6**).





**Figura 6. Procesos biológicos enriquecidos en los miRNAs del C14 y del C19.** Las gráficas de burbujas muestran las agrupaciones de las categorías de procesos biológicos similares, el tamaño de la burbuja es proporcional al número de categorías agrupadas, el color más oscuro entre morado y azul son las agrupaciones con mayor significancia. **(A)** Principales procesos biológicos enriquecidos de los blancos génicos de los miRNAs del C14. **(B)** Principales procesos biológicos enriquecidos de los blancos génicos de los miRNAs del C19.

Además, buscamos identificar todos aquellos procesos biológicos modulados a lo largo del embarazo, dado el perfil de expresión de los c-miRNAs. Para ello determinamos los procesos biológicos enriquecidos en la circulación, empleando los blancos génicos de los 1449 miRNAs identificados en la circulación y manejamos como lista de referencia el transcriptoma humano, identificando un total de 3763 categorías enriquecidas. No obstante, al evaluar el FC de los c-miRNAs de las mujeres E respecto mujeres NE por categoría no encontramos ninguna categoría significativa por FDR.

## Discusión

Nuestro estudio determinó una asociación significativa entre un conjunto de c-miRNAs con el crecimiento fetal y placentario. Por un lado, verificamos por dos aproximaciones matemáticas diferentes la relación entre estos c-miRNAs con indicadores de crecimiento fetal medidos por USG comprobando que la firma de CF presenta una correlación significativa con el crecimiento del feto, además, de presentar un valor predictivo significativo para estimar el crecimiento del feto. Por otra parte, determinamos que en general el conjunto de miRNAs que integran al grupo del C19, grupo al que pertenecen los miRNAs de la firma de CF, se encuentran significativamente correlacionados con indicadores de crecimiento fetal como la LF y el PFE. Finalmente, determinamos que los miRNAs del C19 se encuentran enriquecidos en procesos biológicos tales como: la regulación del ciclo celular, regulación de procesos catabólicos y anabólicos, entre otros. Mientras que los miRNAs del C14 se encuentran principalmente en la regulación de procesos del desarrollo. No obstante, no logramos determinar con precisión los procesos

biológicos asociados a la firma de CF, sin embargo, a través de redes de IPP determinamos que los blancos génicos regulados por estos miRNAs presentan un alto número de interacciones lo que nos sugiere una relación funcional entre ellas que aún no se encuentra del todo esclarecida. Además, de forma intermedia identificamos que los blancos génicos predichos por miRNAs tienden a presentar significativamente un número mayor de IPP a diferencia de blancos génicos aleatorios, lo que nos sugiere una alta regulación por la cual los miRNAs selección a los mRNA que están regulando.

Estudios precedentes realizados por la UVC del INMEGEN determinaron que los c-miRNAs maternos reflejan eventos propios del embarazo. Dentro de esos hallazgos es que determinamos este conjunto de c-miRNAs significativamente correlacionados con indicadores del crecimiento fetal. Esta evidencia nos demuestra el potencial de los c-miRNAs para monitorear el crecimiento fetal a lo largo del embarazo (Chavira-Suárez et. al., 2023). No obstante, el estudio de c-miRNAs como biomarcadores es un campo en investigación que aún se encuentra en sus primeras etapas de desarrollo. Se han llegado a presentar discordancias entre los resultados presentados por diferentes grupos de investigación (Wu et. al., 2021), debido a que en la actualidad no existen protocolos consenso estandarizados tanto para las etapas preanalíticas como la recolección y almacenamiento de muestras, hasta el procesamiento de datos para las técnicas de cuantificación (Condrat et. al., 2020).

Por lo tanto, es importante destacar la relevancia de evaluar los diferentes factores clínicos, técnicos y bioinformáticos que se deben considerar dentro del estudio de miRNAs como biomarcadores. Cabe mencionar que dentro de los factores mejor caracterizados que afectan la expresión de c-miRNAs en un individuo son la actividad física y el curso de procesos inflamatorios (Condrat et. al., 2020; Cui et. al., 2017). Si bien, no evaluamos directamente la actividad física de las pacientes que participaron voluntariamente en este estudio se tuvo especial cuidado en que las condiciones socioeconómicas como: escolaridad, estado civil y actividad económica entre ellas fuera similar. Además, a través del seguimiento clínico a lo

largo de todo el embarazo descartamos a aquellas mujeres que desarrollaran complicaciones perinatales a lo largo de su embarazo o que presentaran niveles de pruebas bioquímicas básicos fuera de los límites de referencia de acuerdo con la edad gestacional. Esto con la finalidad de reducir la variabilidad de factores externos que afectarían el estudio en la medida posible.

Hasta nuestro conocimiento actual, la secuenciación sigue siendo la técnica mejor validada para la cuantificación de miRNAs, sin embargo, requiere de largos periodos de tiempo y equipos altamente especializados, además, de presentar costos elevados para su implementación. Por lo tanto, se encuentran desarrollando y perfeccionando métodos de alto rendimiento para cuantificar miRNAs por RT-qPCR con mayor eficiencia (Forero et al., 2019). No obstante, es importante definir claramente objetivos potenciales para evaluar la especificidad y sensibilidad de los c-miRNAs para el monitoreo del crecimiento fetal. Los ensayos *in silico* representan una excelente opción para buscar asociaciones confiables con mayor precisión.

Diversos estudios han demostrado la viabilidad de los algoritmos de Deep learnig para la detección de potenciales miRNAs biomarcadores de complicaciones perinatales (Yoffe et. al., 2019). Como se ha mencionado anteriormente la firma de CF se seleccionó a través de correlaciones significativas con indicadores de crecimiento fetal como la LF y PFE, por lo tanto, implementamos un análisis semi supervisado de random forest de regresión para evaluar a todo los miRNAs identificados en circulación materna. De esta manera buscamos verificar por una métrica matemáticamente diferente a las mediciones de correlación la asociación entre los miRNAs de la firma de CF con el crecimiento fetal. Encontramos un conjunto de 43 miRNAs de importancia para estimar la longitud de fémur por random forest. Es decir, al medir conjuntamente estos 43 miRNAs se puede predecir el crecimiento de la LF con un error absoluto medio de 1.86 grados. Es importante destacar que este es un estudio exploratorio y se tienen que realizar más pruebas experimentales y considerar un grupo mayor de personas para confirmar estos hallazgos. Encontramos a 4 miRNAs asociados con el crecimiento de la LF que se compartían por ambas aproximaciones, que se presentaban enriquecidos dentro de

los miRNAs identificados por random forest. Lo que nos indica que estos miRNAs presentan una fuerte asociación con el crecimiento fetal, además, de presentar un alto potencial para predecir el crecimiento del feto como una alternativa para complementar las limitaciones de la USG.

Es bien sabido que el desarrollo placentario y fetal se encuentran fuertemente vinculados, de modo que a medida que el feto crece y se desarrolla la placenta paralelamente también lo hace (Arenas et. al., 2005; Ashwal et. al., 2022). Además, numerosos estudios han asociado los resultados adversos en el embarazo para la madre como para el feto con disfunción placentaria. No obstante, en el tamizaje prenatal rutinario no se realizan evaluaciones placentarias más allá de la ubicación y apariencia placentaria (Salomon et al., 2019). Por lo tanto, nosotros decidimos evaluar la asociación de nuestra firma de CF con miRNAs del grupo del C14 y del C19, los cuales a través de estudios experimentales en animales (Mouillet et. al., 2020) e *in vitro* han demostrado que algunos de estos miRNAs presentan un papel vital en la regulación del desarrollo placentario, como el miR-515-5p que inhibe la diferenciación del citotrofoblasto en sincitiotrofoblasto (Zhang et. al., 2016). También se ha observado que otros miRNAs integrantes del C19 están implicados en la regulación de procesos proliferativos en cáncer lo cual sugiere su participación en torno a este proceso (Ren et. al., 2020). A través de la evaluación del enriquecimiento de nuestra firma de CF con el grupo de miRNAs del C14 y el C19 determinamos que dado el número de c-miRNAs 5 de los miRNAs postulados se presentaban significativamente enriquecidos en el C19 y que ninguno de ellos coincidía con los miRNAs del C14. Además, determinamos que los miRNAs del C19 en conjunto presentan una correlación media significativamente mayor a lo esperado al azar con indicadores de crecimiento fetal. Dadas las evidencias mencionadas nos sugiere que los niveles de expresión de estos miRNAs en circulación pueden estar reflejando eventos relacionados con el desarrollo placentario. Lo que nos indica que estos miRNAs en particular nos pueden brindar gran información sobre la evolución del embarazo, como lo es el crecimiento fetal y placentario normal.

No obstante, la interpretación de los niveles de c-miRNAs en el plasma sigue siendo ambiguo. Los estudios más actuales han presentado controversias con respecto a la expresión de estos grupos a lo largo del embarazo en la placenta. Por un lado, se postula que los miRNAs del C14 se encuentran mayormente expresados en el primer trimestre mientras que disminuyen su expresión para el tercer trimestre y viceversa para el grupo del C19 (Gonzalez et. al., 2021). Por el contrario, se postula un incremento de los miRNAs del C14 hacia mediados del embarazo y viceversa para el C19 (Smith et. al., 2021). Es importante considerar que el primer estudio evalúa únicamente placentas del primer y tercer trimestre, mientras que el segundo evalúa al primer trimestre y mediados del embarazo. Esto sugiere que pueden estar fluctuando a lo largo de la gestación, iniciando con una expresión base en el primer trimestre con un incremento para el segundo y finalmente decaer en el tercero. Por otro lado, también es importante considerar que para ambos estudios se realizaron protocolos diferentes para el procesamiento de datos por lo cual se sugiere hacer un análisis comparativo detallado para descartar que los resultados observados no se deban a discrepancias metodológicas sobre el manejo de datos. Sin mencionar, que igualmente puede estar infiriendo el propio diseño experimental, los criterios de selección de las muestras, así como, el dimorfismo sexual, tal y como comentan los autores, entre otras variables. No obstante, a diferencia de los estudios anteriores realizados en tejido placentario, nuestro estudio se centró en evaluar a estos miRNAs en la circulación materna y no propiamente en el tejido. Sin embargo, no se han realizado más estudios comparando los niveles de expresión de estos miRNAs en particular a lo largo del embarazo contra mujeres NE, por lo que, nos apoyamos en estos estudios para entender el contexto que rodea a la expresión de estos c-miRNAs. Estudios previos realizados en nuestro grupo de investigación muestran una tendencia creciente en los miRNAs del C19 a lo largo del embarazo, mientras que los miRNAs del C14 presentan una subexpresión sostenida significativa en los tres trimestres del embarazo y después del embarazo, observando que la mediana de expresión es más baja en el embarazo que en mujeres NE (Hernández Olvera et. al., 2021). Estudios comparativos entre placenta y circulación a lo largo del embarazo, sin embargo, se observado que la expresión

de estos miRNAs es significativamente más baja en circulación con respecto a la placenta, además de presentar un mayor índice de variación (Paquette et. al., 2018). Pese a ello, numerosos estudios de casos y controles han apoyado la viabilidad del uso de estos c-miRNAs concretamente los del C19 para el diagnóstico temprano de trastornos hipertensivos que asocian un incremento en este grupo de miRNAs en etapas tempranas de la gestación con la incidencia en el desarrollo de trastornos hipertensivos del embarazo (Kondracka et. al., 2022). Esta evidencia apoya que nuestra firma de CF puede convertirse una valiosa herramienta para monitorear el desarrollo placentario.

En nuestra búsqueda por determinar los procesos biológicos enriquecidos para esta firma de c-miRNAs, nos encontramos diversas limitantes metodológicas. Actualmente las asociaciones funcionales se realizan entre genes y procesos biológicos a través de la base de datos de gene ontology (GO) (Carbon et. al., 2021), sin embargo, estas asociaciones no consideran a los miRNAs dentro de sus anotaciones, una integración relativamente nueva denominada QuickGO en años recientes a integrado la asociación directa de procesos biológicos con miRNAs (Binns et. al., 2009). No obstante, al ser una integración nueva carece de solides por el rápido incremento de información por lo que decidimos no emplear esta metodología. Convencionalmente, la comunidad científica realiza enriquecimientos funcionales partiendo de la determinación de blancos génicos, donde un análisis previo determina la presencia de una relación entre un determinado miRNA y sus mRNA blanco con los que se ha demostrado fuerte evidencia experimental. Los miRNAs, al ser secuencias cortas de RNA que reconocen mRNAs por complementariedad de bases, presentan una alta probabilidad de regular simultáneamente una gran variedad de transcritos diferentes. Es por ello que la determinación de blancos se ha vuelto controversial encontrando un gran número de falsas predicciones (Fridrich et al., 2019). Si bien, dentro de nuestro método únicamente conservamos aquellas asociaciones miRNA-mRNA con fuerte evidencia experimental nos encontramos con un número muy reducido de posibles blancos génicos. De acuerdo con la red interacción miRNA-mRNA observamos que ninguno de ellos se compartía entre si lo que se interpretaría como poca

probabilidad de estar funcionalmente asociados. Sin embargo, la red IPP nos mostró un alto número de interacciones significativa, entre los productos proteicos de estos blancos, lo que sugiere una fuerte asociación funcional entre ellos. Esto nos indica que a pesar de que los miRNAs de la firma de CF fetal pueden estar regulando procesos biológicos comunes a través de blancos génicos distintos.

Como hemos mencionado los miRNAs de la firma de CF pertenecen al grupo de miRNAs del C19 y se encuentran altamente expresados en placenta, lo cual nos planteaba la posibilidad de que estos miRNAs en conjunto pueden estar regulando un mismo grupo de genes implicados en los mecanismos fisiopatológicos del desarrollo placentario. No obstante, esta evidencia sugiere que no necesariamente todos los miRNAs de este grupo estén regulando los mismos blancos génicos, si no, que la regulación de estos miRNAs conjunta de estos miRNAs en diferentes puntos de la cascada de señalización de manera simultánea son los responsables de los efectos biológicos observados.

Cabe mencionar que al validar la significancia de la red de IPP, observamos que independientemente del conjunto de miRNAs que evaluemos, ya sea, por blancos génicos predichos por homología de secuencias o por evidencia experimental, siempre tienden a regular genes que al traducirse en proteínas presentan un mayor número de interacciones entre si de lo esperado al azar. Lo cual sugiere, la probabilidad de que un factor aún no identificado intervenga en el mecanismo por el cual un determinado miRNA seleccione a los genes blanco que están regulando. Lo que nos plantea la hipótesis de que los miRNAs no regulen la expresión postranscripcional de forma independiente basados únicamente en la homología de las secuencias. Esta evidencia sugiere que los miRNAs de la firma de CF puede estar regulando vías de señalización de procesos biológicos involucrados en el desarrollo del feto y la placenta. No obstante, al determinar enriquecimientos funcionales para esta firma, no encontramos procesos biológicos significativos, lo cual puede deberse al bajo número de blancos génicos identificados que integramos al análisis de enriquecimiento funcional.

Posteriormente, decidimos evaluar por el mismo método no solo a los miRNAs de la firma de CF, si no, al grupo completo de miRNAs del C14 y C19. Curiosamente



a diferencia de lo observado para la firma de 6 en el caso de los miRNAs del C14 y C19 nos encontramos una gran variedad de procesos biológicos enriquecidos, razón por la cual implementamos un algoritmo de Reduce + Visualize Gene Ontology (REVIGO) para agrupar y reducir términos semejantes o redundante (Supek et. al., 2011). Dentro de los términos destacados para los miRNAs del grupo 19 destacaron los procesos relacionados con la regulación del ciclo celular lo cual coincide con las postulaciones de estar implicado en el proceso de proliferación y aunque propiamente no obtuvimos el proceso de diferenciación si encontramos categorías relacionadas con la biosíntesis, catabolismo y neurodesarrollo. En futuros estudios podemos analizar con mayor detenimiento aquellos genes que resultaron enriquecidos en estas categorías y su relación con el proceso de diferenciación del trofoblasto, y del mismo modo descartar asociaciones funcionales poco precisas. Por el contrario, los miRNAs del C14 presentaron un gran número de categorías enriquecidas y que además se encontraban asociadas entre si como: la regulación de procesos del desarrollo, el desarrollo de órganos, la respuesta a estímulos, la migración y comunicación celular dentro de los más destacados. Dentro de las aproximaciones que realizamos a los miRNAs de la firma de CF no encontramos ninguno perteneciente al C14, además, no identificamos estudios hasta el momento que caractericen mejor los procesos fisiopatológicos en los que encuentra implicado este grupo de miRNAs, por lo tanto, sugerimos estudiar con mayor hincapié este grupo de miRNAs que ha sido poco caracterizado y que aparentemente presenta un gran potencial en la regulación de procesos asociados al desarrollo fetal.

Es importante considerar que las anotaciones de los procesos biológicos presentan diversas limitaciones, si bien ciertas categorías están mejor anotadas y descritas, con más detalles y términos más precisos para procesos bien conocidos, muchas otras utilizan terminologías poco definidas o términos confusos (Wijesooriya et al., 2022), una de las áreas menos desarrolladas son los procesos biológicos relacionados con el embarazo, como el desarrollo embrionario, fetal y placentario (X. Li & Fang, 2019). Por lo que, identificar procesos funcionales relacionados con el embarazo es especialmente complicado. Por lo tanto

decidimos evaluar a través de los niveles de expresión de los c-miRNAs las posibles categorías activas durante el embarazo con respecto a mujeres no embarazadas. Sin embargo, al analizar detalladamente el proceso metodológico que estábamos realizando, descubrimos que no encontrábamos una señal enriquecida significativa debido a que no todos los miRNAs que regulan a un determinado blanco están cambian su expresión durante el embarazo. Es decir, si bien un mRNA puede tener un gran número de miRNAs que regulen su expresión, no todos ellos se están expresando al mismo tiempo, de acuerdo con la condición fisiológica evaluada este mRNA puede estar siendo regulado por una pequeña fracción de todos los miRNAs que pueden llegar a regularlo. Lo cual respalda la teoría de que existen otros mecanismos además de la homología de secuencias que pueden estar interviniendo en la regulación postranscripcional mediada por miRNAs.

A pesar de que nos encontramos con numerosas limitantes para la traslación de los miRNAs como biomarcadores del crecimiento fetal y placentario podemos observar igualmente numerosas ventajas en su investigación. El tamizaje prenatal se apoya fuertemente en las evaluaciones por USG para determinar el crecimiento y morfología adecuada tanto del feto como de la placenta (Salomon et. al., 2019). Si bien, los sofisticados equipos actuales permiten la detección temprana de malformaciones y estimar el crecimiento del feto a lo largo del embarazo, menos del 70% de las anomalías fetales se pueden diagnosticar en el primer trimestre de gestación, además, este se ve reducido únicamente a malformaciones graves o complicaciones perinatales severas (Kumar et. al., 2017). Además, numerosos estudios clínicos han demostrado que las biometrías fetales son insuficientes para diagnosticar los trastornos del crecimiento fetal (Gaccioli et. al., 2017; Kumar et. al., 2017). Asimismo, están sujetas a la disposición de un equipo de precisión y un personal clínico altamente capacitado para realizar la exploración. Por el contrario, la medición de miRNAs circulante se puede realizar en cualquier etapa del embarazo por cualquier personal capacitado para la toma de muestra sanguínea (Wu et. al., 2021). Además, de ofrecernos un panorama favorable para entender de manera concisa las etiologías de diferentes complicaciones perinatales y

resultados adversos al nacer y de esta manera desarrollar procedimientos médicos mejor focalizados para la intervención oportuna.

## Conclusiones

En resumen, nuestro estudio verificó que la firma de CF presenta una asociación significativa con el crecimiento fetal y placentario. Además de que cuatro de estos miRNAs presentan un valor significativo para predecir el crecimiento fetal. Asimismo, identificamos que estos miRNAs pertenecían al grupo del C19 y que en general este grupo se encuentra correlacionado positivamente de manera significativa con indicadores de crecimiento como la LF y el PFE. No obstante, con las herramientas bioinformáticas actuales identificamos que la predicción de blancos génicos aún basados únicamente en evidencia experimental sigue siendo imprecisa, razón por la cual, no se observaron procesos biológicos enriquecidos para nuestro conjunto de blancos génicos. Sin embargo, a través de redes de IPP determinamos que los blancos génicos regulados por los miRNAs de la firma de CF presentan un alto número de interacciones significativas entre sí. Esto, nos sugiere una asociación funcional entre ellas que aún no se encuentra del todo esclarecida. Cabe mencionar, que a pesar de no identificar procesos biológicos significativos asociados a la firma de CF podemos inferir que los miRNAs de la firma de CF están participando en procesos relacionados con la regulación del ciclo celular dado que los miRNAs del C19 se encuentran enriquecidos en estos procesos biológicos principalmente.

## Perspectivas

- Determinar un proceso más eficaz para identificar blancos génicos de miRNAs.
- Identificar los procesos biológicos regulados por miRNAs tanto circulantes como placentarios.
- Identificar miRNAs que participan en la regulación del desarrollo placentario y fetal.
- Establecer posibles mecanismos por los cuales los miRNAs participan en la regulación del desarrollo placentario y fetal.

## Referencias

- Abbassi-Ghanavati, M., Greer, L., & FG, C. (2009). A Reference Table for Clinicians. *Obstetrics & Gynecology*, *114*(6), 1326–1331.
- Amadei, G., Handford, C. E., Qiu, C., de Jonghe, J., Greenfeld, H., Tran, M., Martin, B. K., Chen, D. Y., Aguilera-Castrejon, A., Hanna, J. H., Elowitz, M. B., Hollfelder, F., Shendure, J., Glover, D. M., & Zernicka-Goetz, M. (2022). Embryo model completes gastrulation to neurulation and organogenesis. *Nature*, *610*(7930), 143–153. <https://doi.org/10.1038/s41586-022-05246-3>
- Arenas, J., Otero, M. T., Duplá, B., Fernández-Iñarrea, J., Oviedo, E., & Rodríguez-Mon, C. (2005). The usefulness of placental biometry in predicting complications during pregnancy. *Clinica e Investigacion En Ginecologia y Obstetricia*, *32*(1), 10–13. [https://doi.org/10.1016/s0210-573x\(05\)73466-3](https://doi.org/10.1016/s0210-573x(05)73466-3)
- Ashwal, E., Ali-Gami, J., Aviram, A., Ronzoni, S., Mei-Dan, E., Kingdom, J., & Melamed, N. (2022). Contribution of Second Trimester Sonographic Placental Morphology to Uterine Artery Doppler in the Prediction of Placenta-Mediated Pregnancy Complications. *Journal of Clinical Medicine*, *11*(22). <https://doi.org/10.3390/jcm11226759>
- Babraham Bioinformatics - FastQC Una herramienta de control de calidad para datos de secuencia de alto rendimiento.* (n.d.). Retrieved May 25, 2020, from <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>
- Bartel, D. P. (2004). MicroRNAs: Genomics, Biogenesis, Mechanism, and Function. *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)
- Behm-Ansmant, I., Rehwinkel, J., Doerks, T., Stark, A., Bork, P., & Izaurralde, E. (2006). mRNA degradation by miRNAs and GW182 requires both CCR4:NOT deadenylase and DCP1:DCP2 decapping complexes. *Genes and Development*, *20*(14), 1885–1898. <https://doi.org/10.1101/gad.1424106>
- Binns, D., Dimmer, E., Huntley, R., Barrell, D., O'Donovan, C., & Apweiler, R. (2009). QuickGO: A web-based tool for Gene Ontology searching. *Bioinformatics*, *25*(22), 3045–3046. <https://doi.org/10.1093/bioinformatics/btp536>
- Carbon, S., Douglass, E., Good, B. M., Unni, D. R., Harris, N. L., Mungall, C. J., Basu, S., Chisholm, R. L., Dodson, R. J., Hartline, E., Fey, P., Thomas, P. D., Albou, L. P., Ebert, D., Kesling, M. J., Mi, H., Muruganujan, A., Huang, X., Mushayahama, T., ... Elser, J. (2021). The Gene Ontology resource: Enriching a Gold mine. *Nucleic Acids Research*, *49*(D1), D325–D334. <https://doi.org/10.1093/nar/gkaa1113>
- Chavira-Suárez, E., Hernández-Olvera, A. L., Flores-Torres, M., Celaya-Cruz, K. R., Gitler, S., De la Cerda-Ángeles, J. C., Espinosa-Maldonado, N. C., Flores-Jasso, C. F., Gutiérrez, H., & Vadillo-Ortega, F. (2023). Longitudinal large-scale changes in maternal circulating microRNAs associated with gestation-related compartments, fetal sex, and growth during and post-pregnancy. *Genomics*, *115*(3). <https://doi.org/10.1016/j.ygeno.2023.110628>
- Condrat, C. E., Thompson, D. C., Barbu, M. G., Bugnar, O. L., Boboc, A., Cretoiu, D., Suci, N., Cretoiu, S. M., & Voinea, S. C. (2020). miRNAs as Biomarkers in Disease: Latest Findings Regarding Their Role in Diagnosis and Prognosis. In *Cells* (Vol. 9, Issue 2). NLM (Medline). <https://doi.org/10.3390/cells9020276>
- Cui, S., Sun, B., Yin, X., Guo, X., Chao, D., Zhang, C., Zhang, C. Y., Chen, X., & Ma, J. (2017). Time-course responses of circulating microRNAs to three resistance training protocols in healthy young men. *Scientific Reports*, *7*(1). <https://doi.org/10.1038/s41598-017-02294-y>

- Daly, M., & Driscoll, L. O. (2017). *Chapter 5 MicroRNA Profiling of Exosomes*. 1509, 37–46. <https://doi.org/10.1007/978-1-4939-6524-3>
- Dennis Lo, Y. M., Corbetta, N., Chamberlain, P. F., Rai, V., Sargent, I. L., Redman, C. W. G., & Wainscoat, J. S. (1997). Presence of fetal DNA in maternal plasma and serum. *Lancet*, 350(9076), 485–487. [https://doi.org/10.1016/S0140-6736\(97\)02174-0](https://doi.org/10.1016/S0140-6736(97)02174-0)
- DeVilbiss, E. A., Mumford, S. L., Sjaarda, L. A., Connell, M. T., Plowden, T. C., Andriessen, V. C., Perkins, N. J., Hill, M. J., Silver, R. M., & Schisterman, E. F. (2020). Prediction of pregnancy loss by early first trimester ultrasound characteristics. *American Journal of Obstetrics and Gynecology*, 223(2), 242.e1-242.e22. <https://doi.org/10.1016/j.ajog.2020.02.025>
- Esteller, M. (2011). Non-coding RNAs in human disease. *Nature Reviews Genetics*, 12(12), 861–874. <https://doi.org/10.1038/nrg3074>
- Forero, D. A., González-Giraldo, Y., Castro-Vega, L. J., & Barreto, G. E. (2019). qPCR-based methods for expression analysis of miRNAs. In *BioTechniques* (Vol. 67, Issue 4, pp. 192–199). Future Science. <https://doi.org/10.2144/btn-2019-0065>
- Fridrich, A., Hazan, Y., & Moran, Y. (2019). Too Many False Targets for MicroRNAs: Challenges and Pitfalls in Prediction of miRNA Targets and Their Gene Ontology in Model and Non-model Organisms. In *BioEssays* (Vol. 41, Issue 4). John Wiley and Sons Inc. <https://doi.org/10.1002/bies.201800169>
- Friedländer, M. R., Mackowiak, S. D., Li, N., Chen, W., & Rajewsky, N. (2012). MiRDeep2 accurately identifies known and hundreds of novel microRNA genes in seven animal clades. *Nucleic Acids Research*, 40(1), 37–52. <https://doi.org/10.1093/nar/gkr688>
- Gaccioli, F., Lager, S., Sovio, U., Charnock-Jones, D. S., & Smith, G. C. S. (2017). The pregnancy outcome prediction (POP) study: Investigating the relationship between serial prenatal ultrasonography, biomarkers, placental phenotype and adverse pregnancy outcomes. *Placenta*, 59, S17–S25. <https://doi.org/10.1016/j.placenta.2016.10.011>
- Gebert, L. F. R., & MacRae, I. J. (2019). Regulation of microRNA function in animals. *Nature Reviews Molecular Cell Biology*, 20(1), 21–37. <https://doi.org/10.1038/s41580-018-0045-7>
- Georgadaki, K., Khoury, N., Spandidos, D. A., & Zoumpourlis, V. (2016). The molecular basis of fertilization (Review). In *International Journal of Molecular Medicine* (Vol. 38, Issue 4, pp. 979–986). Spandidos Publications. <https://doi.org/10.3892/ijmm.2016.2723>
- Gerri, C., Menchero, S., Mahadevaiah, S. K., Turner, J. M. A., & Niakan, K. K. (2020). Human Embryogenesis: A Comparative Perspective. *Annu. Rev. Cell Dev. Biol.* 2020, 36, 411–440. <https://doi.org/10.1146/annurev-cellbio-022020>
- Gluckman, P. D., Hanson, M. A., Cooper, C., & Thornburg, K. L. (2008). Effect of In Utero and Early-Life Conditions on Adult Health and Disease. *New England Journal of Medicine*, 359(1), 61–73. <https://doi.org/10.1056/nejmra0708473>
- Godoy, P. M., Bhakta, N. R., Barczak, A. J., Cakmak, H., Fisher, S., MacKenzie, T. C., Patel, T., Price, R. W., Smith, J. F., Woodruff, P. G., & Erle, D. J. (2018). Large Differences in Small RNA Composition Between Human Biofluids. *Cell Reports*, 25(5), 1346–1358. <https://doi.org/10.1016/j.celrep.2018.10.014>
- Gonzalez, T. L., Eisman, L. E., Joshi, N. v., Flowers, A. E., Wu, D., Wang, Y., Santiskulvong, C., Tang, J., Buttle, R. A., Sauro, E., Clark, E. L., Dipentino, R., Jefferies, C. A., Chan, J. L., Lin, Y., Zhu, Y., Afshar, Y., Tseng, H. R., Taylor, K., ... Pisarska, M. D. (2021). High-throughput miRNA sequencing of the human placenta: expression throughout gestation. *Epigenomics*, 13(13), 995–1012. <https://doi.org/10.2217/epi-2021-0055>

- Goodall, G. J., & Wickramasinghe, V. O. (2021). RNA in cancer. *Nature Reviews Cancer*, 21(1), 22–36. <https://doi.org/10.1038/s41568-020-00306-0>
- Griffiths-Jones, S. (2006). miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic Acids Research*, 34(90001), D140–D144. <https://doi.org/10.1093/nar/gkj112>
- Hernández Olvera, A. L., Chavira Suárez, E., & Gutiérrez González, H. (n.d.). *Perfiles de miRNAs circulantes maternos que reflejan procesos asociados con el embarazo humano normal*. Universidad Nacional Autónoma de México.
- Hu, H., & Pasca, I. (2016). Management of Complex Cardiac Issues in the Pregnant Patient. In *Critical Care Clinics* (Vol. 32, Issue 1, pp. 97–107). W.B. Saunders. <https://doi.org/10.1016/j.ccc.2015.08.004>
- Huang, H. Y., Lin, Y. C. D., Cui, S., Huang, Y., Tang, Y., Xu, J., Bao, J., Li, Y., Wen, J., Zuo, H., Wang, W., Li, J., Ni, J., Ruan, Y., Li, L., Chen, Y., Xie, Y., Zhu, Z., Cai, X., ... Huang, H. da. (2022). MiRTarBase update 2022: An informative resource for experimentally validated miRNA-target interactions. *Nucleic Acids Research*, 50(D1), D222–D230. <https://doi.org/10.1093/nar/gkab1079>
- Huang, H. Y., Lin, Y. C. D., Li, J., Huang, K. Y., Shrestha, S., Hong, H. C., Tang, Y., Chen, Y. G., Jin, C. N., Yu, Y., Xu, J. T., Li, Y. M., Cai, X. X., Zhou, Z. Y., Chen, X. H., Pei, Y. Y., Hu, L., Su, J. J., Cui, S. D., ... Huang, H. da. (2020). MiRTarBase 2020: Updates to the experimentally validated microRNA-target interaction database. *Nucleic Acids Research*, 48(D1), D148–D154. <https://doi.org/10.1093/nar/gkz896>
- Jia-Yu Zhu, Zhan-Jun Pang, & Yan-hong Yu. (2012). *Regulation of Trophoblast Invasion: The Role of Matrix Metalloproteinases*. <https://doi.org/10.3909/riog0196>
- Jo, M. H., Shin, S., Jung, S., Kim, E., Jo, M. H., Shin, S., Jung, S., Kim, E., & Song, J. (2015). Human Argonaute 2 Has Diverse Reaction Pathways on Target RNAs Article Human Argonaute 2 Has Diverse Reaction Pathways on Target RNAs. *Molecular Cell*, 59(1), 117–124. <https://doi.org/10.1016/j.molcel.2015.04.027>
- Karin-Kujundzic, V., Sola, I. M., Predavec, N., Potkonjak, A., Somen, E., Mioc, P., Serman, A., Vranic, S., & Serman, L. (2019). Novel Epigenetic Biomarkers in Pregnancy-Related Disorders and Cancers. In *Cells* (Vol. 8, Issue 11). NLM (Medline). <https://doi.org/10.3390/cells8111459>
- Kim, S.-M., & Kim, J.-S. (2017). A Review of Mechanisms of Implantation. *Development & Reproduction*, 21(4), 351–359. <https://doi.org/10.12717/dr.2017.21.4.351>
- Kobayashi, H., & Tomari, Y. (2016). Biochimica et Biophysica Acta RISC assembly : Coordination between small RNAs and Argonaute proteins ☆. *BBA - Gene Regulatory Mechanisms*, 1859(1), 71–81. <https://doi.org/10.1016/j.bbagrm.2015.08.007>
- Kondracka, A., Jaszczuk, I., Koczkodaj, D., Kondracki, B., Frąszczak, K., Oniszczyk, A., Rybak-Krzyszowska, M., Staniczek, J., Filip, A., & Kwaśniewska, A. (2022). Analysis of Circulating C19MC MicroRNA as an Early Marker of Hypertension and Preeclampsia in Pregnant Patients: A Systematic Review. In *Journal of Clinical Medicine* (Vol. 11, Issue 23). MDPI. <https://doi.org/10.3390/jcm11237051>
- Kumar, M., Singh, S., Sharma, K., Singh, R., Ravi, V., & Bhattacharya, J. (2017). Adverse fetal outcome: is first trimester ultrasound and Doppler better predictor than biomarkers? *Journal of Maternal-Fetal and Neonatal Medicine*, 30(12), 1410–1416. <https://doi.org/10.1080/14767058.2016.1214709>
- Lee, R., Feinbaum, R., & Ambros, V. (1993). The *C. elegans* Heterochronic Gene *lin-4* Encodes Small RNAs with Antisense Complementarity to *lin-14*. *Cell*, 75(116), 843–854.

- Li, H., & Durbin, R. (2010). Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics*, 26(5), 589–595.  
<https://doi.org/10.1093/bioinformatics/btp698>
- Li, X., & Fang, Y. (2019). Bioinformatics identification of potential genes and pathways in preeclampsia based on functional gene set enrichment analyses. *Experimental and Therapeutic Medicine*. <https://doi.org/10.3892/etm.2019.7749>
- Marcel, M. (n.d.). *User guide Cutadapt 2.10*. Retrieved May 25, 2020, from <https://cutadapt.readthedocs.io/en/stable/guide.html>
- McGeary, S. E., Lin, K. S., Shi, C. Y., Pham, T. M., Bisaria, N., Kelley, G. M., & Bartel, D. P. (2019). The biochemical basis of microRNA targeting efficacy. *Science*, 366(6472). <https://doi.org/10.1126/science.aav1741>
- Milewski, R., & Ajduk, A. (2017). Time-lapse imaging of cleavage divisions in embryo quality assessment. In *Reproduction* (Vol. 154, Issue 2, pp. R37–R53). BioScientifica Ltd. <https://doi.org/10.1530/REP-17-0004>
- Mouillet, J. F., Goff, J., Sadovsky, E., Sun, H., Parks, T., Chu, T., & Sadovsky, Y. (2020). Transgenic expression of human C19MC miRNAs impacts placental morphogenesis. *Placenta*, 101, 208–214. <https://doi.org/10.1016/j.placenta.2020.09.069>
- Mugridge, J. S., Marcin, Z., Jacek, J., & Gross, J. D. (2016). Structural basis of mRNA cap recognition by Dcp1–Dcp2. *Physiology & Behavior*, 176(3), 987–994. <https://doi.org/10.1038/nsmb.3301>
- Muhr, J., & Ackerman, K. M. (2022). Embryology, gastrulation. In *StatPearls*.
- Neiger, R. (2017). Long-term effects of pregnancy complications on maternal health: A review. In *Journal of Clinical Medicine* (Vol. 6, Issue 8). MDPI. <https://doi.org/10.3390/jcm6080076>
- Niaz, S., & Hussain, M. U. (2018). Role of GW182 protein in the cell. *International Journal of Biochemistry and Cell Biology*, 101(April), 29–38. <https://doi.org/10.1016/j.biocel.2018.05.009>
- O'Brien, J., Hayder, H., Zayed, Y., & Peng, C. (2018). Overview of microRNA biogenesis, mechanisms of actions, and circulation. *Frontiers in Endocrinology*, 9(AUG), 1–12. <https://doi.org/10.3389/fendo.2018.00402>
- Ochoa-Bernal, M. A., & Fazleabas, A. T. (2020). Physiologic events of embryo implantation and decidualization in human and non-human primates. In *International Journal of Molecular Sciences* (Vol. 21, Issue 6). MDPI AG. <https://doi.org/10.3390/ijms21061973>
- Ojosnegros, S., Seriola, A., Godeau, A. L., & Veiga, A. (2021a). Embryo implantation in the laboratory: an update on current techniques. In *Human reproduction update* (Vol. 27, Issue 3, pp. 501–530). NLM (Medline). <https://doi.org/10.1093/humupd/dmaa054>
- Ojosnegros, S., Seriola, A., Godeau, A. L., & Veiga, A. (2021b). Embryo implantation in the laboratory: an update on current techniques. In *Human reproduction update* (Vol. 27, Issue 3, pp. 501–530). NLM (Medline). <https://doi.org/10.1093/humupd/dmaa054>
- OMS. (2019, September 9). *Mortalidad materna*.
- OMS. (2020a). *Cada 16 segundos se produce una muerte fetal, según las primeras estimaciones conjuntas de las Naciones Unidas*.
- OMS. (2020b). *Mejorar la supervivencia y el bienestar de los recién nacidos*.
- Panni, S., Lovering, R. C., Porras, P., & Orchard, S. (2020). Non-coding RNA regulatory networks. *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms*, 1863(6), 194417. <https://doi.org/10.1016/j.bbagr.2019.194417>
- Paquette, A. G., Chu, T., Wu, X., Wang, K., Price, N. D., & Sadovsky, Y. (2018). Distinct communication patterns of trophoblastic miRNA among the maternal-placental-fetal compartments. *Placenta*, 72–73, 28–35. <https://doi.org/10.1016/j.placenta.2018.10.004>

- Pijnenborg, R., Vercruyssen, L., & Hanssens, M. (2006). The Uterine Spiral Arteries In Human Pregnancy: Facts and Controversies. In *Placenta* (Vol. 27, Issues 9–10, pp. 939–958). <https://doi.org/10.1016/j.placenta.2005.12.006>
- Plotnikova, O., Baranova, A., & Skoblov, M. (2019). Comprehensive Analysis of Human microRNA–mRNA Interactome. *Frontiers in Genetics*, 10(October), 1–11. <https://doi.org/10.3389/fgene.2019.00933>
- Posfai, E., Rovic, I., & Jurisicova, A. (2019). The mammalian embryo's first agenda: Making trophoblast. *International Journal of Developmental Biology*, 63(3–5), 157–170. <https://doi.org/10.1387/ijdb.180404ep>
- Ren, L., Li, Y., Zhao, Q., Fan, L., Tan, B., Zang, A., & Yang, H. (2020). MiR-519 regulates the proliferation of breast cancer cells via targeting human antigen R. *Oncology Letters*, 19(2), 1567–1576. <https://doi.org/10.3892/ol.2019.11230>
- Rossi, G., Brogiere, N., Miyamoto, M., Boni, A., Guiet, R., Girgin, M., Kelly, R. G., Kwon, C., & Lutolf, M. P. (2021). Capturing Cardiogenesis in Gastruloids. *Cell Stem Cell*, 28(2), 230–240.e6. <https://doi.org/10.1016/j.stem.2020.10.013>
- Salomon, L. J., Alfirevic, Z., da Silva Costa, F., Deter, R. L., Figueras, F., Ghi, T., Glanc, P., Khalil, A., Lee, W., Napolitano, R., Papageorghiou, A., Sotiradis, A., Stirnemann, J., Toi, A., & Yeo, G. (2019). ISUOG Practice Guidelines: ultrasound assessment of fetal biometry and growth. *Ultrasound in Obstetrics and Gynecology*, 53(6), 715–723. <https://doi.org/10.1002/uog.20272>
- Sanghavi, M., & Rutherford, J. D. (2014). Cardiovascular physiology of pregnancy. *Circulation*, 130(12), 1003–1008. <https://doi.org/10.1161/CIRCULATIONAHA.114.009029>
- Say, L., Chou, D., Gemmill, A., Tunçalp, Ö., Moller, A., Daniels, J., Gülmezoglu, A. M., Temmerman, M., & Alkema, L. (2014). *Global causes of maternal death: a WHO systematic analysis*. 323–333. [https://doi.org/10.1016/S2214-109X\(14\)70227-X](https://doi.org/10.1016/S2214-109X(14)70227-X)
- Schäfer, I. B., Yamashita, M., Schuller, J. M., Schüssler, S., Reichelt, P., Strauss, M., & Conti, E. (2019). Molecular Basis for poly(A) RNP Architecture and Recognition by the Pan2–Pan3 Deadenylation. *Cell*, 177(6), 1619–1631.e21. <https://doi.org/10.1016/j.cell.2019.04.013>
- Semi, K., & Takashima, Y. (2020). Pluripotent stem cells for the study of early human embryology. *Development, Growth & Differentiation*, 63, 104–115. <https://doi.org/10.1111/dgd.12715>
- Slaby, O., & Calin Editors, G. A. (n.d.). *Non-coding RNAs in Colorectal Cancer*. <http://www.springer.com/series/5584>
- Smith, M. D., Pillman, K., Jankovic-Karasoulos, T., McAninch, D., Wan, Q., Bogias, K. J., McCullough, D., Bianco-Miotto, T., Breen, J., & Roberts, C. T. (2021). Large-scale transcriptome-wide profiling of microRNAs in human placenta and maternal plasma at early to mid gestation. *RNA Biology*, 18(S1), 507–520. <https://doi.org/10.1080/15476286.2021.1963105>
- Soma-pillay, P., Nelson-piercy, C., Tolppanen, H., & Mebazaa, A. (2016). *Physiological changes in pregnancy*. 27(2), 89–94. <https://doi.org/10.5830/CVJA-2016-021>
- Soma-Pillay, P., Nelson-Piercy, C., Tolppanen, H., & Mebazaa, A. (2016). Physiological changes in pregnancy. *Cardiovascular Journal of Africa*, 27(2), 89–94. <https://doi.org/10.5830/CVJA-2016-021>
- Sun, C., Groom, K. M., Oyston, C., Chamley, L. W., Clark, A. R., & James, J. L. (2020). The placenta in fetal growth restriction: What is going wrong? *Placenta*, 96, 10–18. <https://doi.org/10.1016/j.placenta.2020.05.003>
- Supek, F., Bošnjak, M., Škunca, N., & Šmuc, T. (2011). Revigo summarizes and visualizes long lists of gene ontology terms. *PLoS ONE*, 6(7). <https://doi.org/10.1371/journal.pone.0021800>



- Szkarczyk, D., Gable, A. L., Nastou, K. C., Lyon, D., Kirsch, R., Pyysalo, S., Doncheva, N. T., Legeay, M., Fang, T., Bork, P., Jensen, L. J., & von Mering, C. (2021). The STRING database in 2021: Customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Research*, 49(D1), D605–D612. <https://doi.org/10.1093/nar/gkaa1074>
- Taft, Ryan. (2010). *Non-coding RNAs: regulators of disease*. October 2009, 231–241. <https://doi.org/10.1002/path.2638>
- Vashukova, E. S., Kozyulina, P. Y., Illarionov, R. A., Yurkina, N. O., Pachuliia, O. V., Butenko, M. G., Postnikova, T. B., Ivanova, L. A., Ereemeeva, D. R., Zainulina, M. S., Bespalova, O. N., & Glotov, A. S. (2021). High-throughput sequencing of circulating microRNAs in plasma and serum during pregnancy progression. *Life*, 11(10). <https://doi.org/10.3390/life11101055>
- Vasudevan, S., & Steitz, J. A. (2007). AU-Rich-Element-Mediated Upregulation of Translation by FXR1 and Argonaute 2. *Cell*, 128(6), 1105–1118. <https://doi.org/10.1016/j.cell.2007.01.038>
- Wahid, F., Shehzad, A., Khan, T., & Kim, Y. Y. (2010). MicroRNAs: Synthesis, mechanism, function, and recent clinical trials. *Biochimica et Biophysica Acta - Molecular Cell Research*, 1803(11), 1231–1243. <https://doi.org/10.1016/j.bbamcr.2010.06.013>
- Wang, H., Peng, R., Wang, J., Qin, Z., & Xue, L. (2018). Circulating microRNAs as potential cancer biomarkers: The advantage and disadvantage. In *Clinical Epigenetics* (Vol. 10, Issue 1). BioMed Central Ltd. <https://doi.org/10.1186/s13148-018-0492-1>
- West, R. C., Ming, H., Logsdon, D. M., Sun, J., Rajput, S. K., Kile, R. A., Schoolcraft, W. B., Michael Roberts, R., Krisher, R. L., Jiang, Z., & Yuan, Y. (2019). Dynamics of trophoblast differentiation in peri-implantation–stage human embryos. *Proceedings of the National Academy of Sciences of the United States of America*, 116(45), 22635–22644. <https://doi.org/10.1073/pnas.1911362116>
- Wijesooriya, K., Jadaan, S. A., Perera, K. L., Kaur, T., & Ziemann, M. (2022). Urgent need for consistent standards in functional enrichment analysis. *PLoS Computational Biology*, 18(3). <https://doi.org/10.1371/journal.pcbi.1009935>
- Wu, Y., Li, Q., Zhang, R., Dai, X., Chen, W., & Xing, D. (2021). Circulating microRNAs: Biomarkers of disease. In *Clinica Chimica Acta* (Vol. 516, pp. 46–54). Elsevier B.V. <https://doi.org/10.1016/j.cca.2021.01.008>
- Xu, W., Lucas, A. S., Wang, Z., & Liu, Y. (2014). Identifying microRNA targets in different gene regions. *BMC Bioinformatics*, 15(Suppl 7), 6–8. <https://doi.org/10.1186/1471-2105-15-S7-S4>
- Yoffe, L., Polsky, A., Gilam, A., Raff, C., Mecacci, F., Ognibene, A., Crispi, F., Gratacós, E., Kanety, H., Mazaki-Tovi, S., Shomron, N., & Hod, M. (2019). Early diagnosis of gestational diabetes mellitus using circulating microRNAs. *European Journal of Endocrinology*, 181(5), 565–577. <https://doi.org/10.1530/EJE-19-0206>
- Zhai, J., Xiao, Z., Wang, Y., & Wang, H. (2022). Human embryonic development: from peri-implantation to gastrulation. In *Trends in Cell Biology* (Vol. 32, Issue 1, pp. 18–29). Elsevier Ltd. <https://doi.org/10.1016/j.tcb.2021.07.008>
- Zhang, M., Muralimanoharan, S., Wortman, A. C., & Mendelson, C. R. (2016). Primate-specific miR-515 family members inhibit key genes in human trophoblast differentiation and are upregulated in preeclampsia. *Proceedings of the National Academy of Sciences of the United States of America*, 113(45), E7069–E7076. <https://doi.org/10.1073/pnas.1607849113>

Zhao, Z., Moley, K. H., & Gronowski, A. M. (2013). Diagnostic potential for miRNAs as biomarkers for pregnancy-specific diseases. *Clinical Biochemistry*, 46(10–11), 953–960. <https://doi.org/10.1016/j.clinbiochem.2013.01.026>

## Apéndice A: Características clínicas de las participantes

**Tabla 1 características clínicas de mujeres embarazadas sanas**

SDG		1T 7.4 - 13.6	2T 19.5 - 25.1	3T 32.4 - 35.2	p
<b>Mediciones antropométricas</b>					
Edad (años)		24(21-34)			0.214
Peso (Kg)		62(46-90)	63(51-92)	67(58-93)	0.127
Talla (cm)		153(144-163)			0.065
IMC (kg/m <sup>2</sup> )		27(19-34)	27(21-35)	29(24-35)	<b>0.02</b>
PAS (mmHg)		110(100-121)	102(96-128)	104(99-118)	0.348
PAD (mmHg)		70(64-85)	60(55-70)	67(55-75)	<b>0.003</b>
PAM (mmHg)		68(35-84)	75(69-88)	79(70-89)	<b>0.004</b>
Historial obstétrico	Gravidez	3 (1-4)			
	Paridad	2 (1-4)			
Tipo de parto	Vaginal	7 (87.5)			
	Cesárea	1 (12.5)			
<b>Mediciones ultrasonográficas fetales</b>					
PFE (g)		98(58-99)	431(308-751)	2151(2004-2496)	<b>0.0001</b>
LF (cm)		1(0.7-1.4)	4(2.9-4.6)	6(5.8-6.8)	<b>0.0001</b>
Sexo fetal	Femenino	5 (62.5)			
	Masculino	3 (37.5)			
<b>Factores sociodemográficos</b>					
Nivel escolar	Básica	6 (77)			
	Media	2 (23)			
	Superior	0 (0)			
Estado civil	Soltera	1 (12.5)			
	Casada	1 (12.5)			
	Unión libre	6 (75)			
Económicamente activa		5 (55.6)			
Consumo de sustancias de abuso reportado		1 (12.5)			

Datos presentados como: mediana (rango intercuartil) o n (%). 1T: primer trimestre de embarazo, 2T: segundo trimestre de embarazo, 3T: tercer trimestre de embarazo, SDG: semanas de gestación, IMC: índice de masa corporal, PAS: presión arterial sistólica, PAD: presión arterial diastólica, PAM: presión arterial media, PFE: peso fetal estimado, LF: longitud de fémur.

## Apéndice B: Código en Python modelado de random forest de regresión para predecir la longitud del fémur

```
# Importar paqueterías
import numpy as np
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.ensemble import RandomForestRegressor
from sklearn import metrics
import seaborn as sn
from matplotlib import pyplot as plt
import matplotlib.pyplot as plt
from sklearn.metrics import accuracy_score
from sklearn.tree import DecisionTreeRegressor
from dtreeviz.trees import *
import math

# Lectura de base de datos
features = pd.read_csv("/content/drive/MyDrive/Colab
Notebooks/RandomForest_CF/Base_randomforest_LFvsmiRNAs588.csv",
low_memory=False)

# Labels son los valores que deseamos predecir
labels = np.array(features['LF'])
# Remover columna de valores a predecir
# Eje 1 referencia para las columnas
features= features.drop('LF', axis = 1)
features= features.drop('ID', axis = 1)
# Guardar el nombre de las columnas o variables
feature_list = list(features.columns)
# Convertir base en arreglo de numpy
features = np.array(features)

# Usando Skicit-learn para dividir los datos en un set de
entrenamiento y un set de prueba
train_features, test_features, train_labels, test_labels =
train_test_split(features, labels, test_size = 0.25, random_state =
42)

# Instanciar modelo con 1000 árboles de decisión
rf = RandomForestRegressor(n_estimators = 1000, random_state = 42)

# Entrenar modelo sobre los datos de entrenamiento
rf.fit(train_features, train_labels);
```

```

# Usar el método de predicción de forest sobre el set de prueba
predictions = rf.predict(test_features)

# Definir función para calcular mse y rmse
def r_mse(pred,y): return round(math.sqrt(((pred-y)**2).mean()), 6)
def m_rmse(m, xs, y): return r_mse(m.predict(xs), y)

# Calcular mse y rmse
m_rmse(rf,train_features , train_labels)
m_rmse(rf, test_features, test_labels)

# Obtener variables numéricas de importancia
importances = list(rf.feature_importances)
# Lista de tuples con variables e importancia
feature_importances = [(feature, round(importance, 2)) for feature,
importance in zip(feature_list, importances)]
# Ordenar variables por importancia
feature_importances = sorted(feature_importances, key = lambda x:
x[1], reverse = True)
# Imprimir variables y su valor de importancia
[print('Variable: {:20} Importance: {}'.format(*pair)) for pair in
feature_importances];

#Guardar variables de importancia
df = pd.DataFrame(feature_importances)
df.to_csv('/content/drive/MyDrive/Colab
Notebooks/RandomForest_CF/Firma_RandomForest_CF.csv')

```

## Apéndice C: Código en R para determinar enriquecimiento

Se muestra el script representativo para los miRNAs del C19 con LF. La misma secuencia se siguió para C19 con PFE, C14 con LF y C14 con PFE, firma de CF con miRNAs del C19, miRNAs del CF con miRNAs del C14, miRNAs del CF en miRNAs de importancia por random forest

```
#####
####Lectura de datos####

#miRNAs C14 y C19 en circulación
C19.C14 <- read.csv(file = "C19_C14.csv", header = T, sep = ",",
stringsAsFactors = F)
C14 <- C19.C14$ID.miR.cir.14
C19 <- C19.C14$ID.miR.cir.19

#Id miRNAs secuenciados
ID_miRNAs <- read.csv(file = "miRNAs_1T-2T-3T-mediana-filtrado-2.csv",
header = T, sep = ",", stringsAsFactors = F)
ID_miRNAs <- ID_miRNAs$miRNA

#Firma de CF
Set_CF <- read.csv(file = "firma_miRNAs_LF-PF.csv", header = T, sep =
",",
stringsAsFactors = F)
CF <- Set_CF$miRNA

#Obtener index de miRNAs de la firma de CF
x <- match(CF, ID_miRNAs)

x14 <- na.omit(match(C14, ID_miRNAs))
x19 <- na.omit(match(C19, ID_miRNAs))

#Mediana de correlación miRNAs con LF
cor_miR <- read.table("Spearman_miRNAs588-LongFemur.csv", header = T, sep
= " ",
stringsAsFactors = F)

#Correlación media miRNAs del C19 con LF
u <- cor_miR$LF
p <- cor_miR[x19,]
p <- p$LF
s <- mean(p)

# Enriquecimiento miRNAs del C19 con LF
e <- 1000000
c <- c(1:e)*0

for (i in 1:e) {
  a <- sample(u, 18, replace = F)
  c[i] <- median(a)
}
if(s < mean(c)){
  d <- c<=s
  h <- c(1:e)*0
  h[d] <- 1
}
```

```
h <- sum(h)
pvalue <- h/e} else {
  d <- c>=s
  h <- c(1:e)*0
  h[d] <- 1
  h <- sum(h)
  pvalue <- h/e }
hist(c)

CLF.19 <- c
SLF.19 <- s
pvalue.LF.19 <- pvalue
```

## Apéndice D: Código en R predicción de blancos y construcción de redes de IPP

```
library(dplyr)
library(openxlsx)
library(tidyverse)

#####Prueba para determinar red IPP y anotaciones funcionales
#Lectura base de miRTarBase
url <-
"https://mirtarbase.cuhk.edu.cn/~miRTarBase/miRTarBase_2022/cache/
download/8.0/hsa_MTI.xlsx"

miRTarBase <- read.xlsx(xlsxFile = url, sheet = 1, skipEmptyRows =
F)

#####Prueba aleatoria
#Leer lista de miRNAs detectados en circulación

df <- read.csv("Lista_miRNAs_588_mod.csv")
u <- df$miRNA[!duplicated(df$miRNA)]

set.seed(1000)

#número de aleatorizaciones y miRNAs a evaluar
e <-1000
c <- c(1:6)*0
c <- as.data.frame(c)

### Permutar datos de miRNAs detectados
for (i in 1:e) {
  a <- sample(u, 6, replace = F) #extraccion de muestras aleatoria
  c[i] <- cbind(a) #Guardar muestras extraídas
}

#Filtrar base de blancos por fuerte evidencia experimental
(Western blot|reporter assay|qRT-PCR)
miRTarBase_filtrado <- miRTarBase %>%
  filter(grepl('Western blot|reporter assay|qRT-PCR',
Experiments))

#Extracción ID string
library(STRINGdb)

#Descargar datos de string para humano
string_db <- STRINGdb$new( version="11.5", species=9606,
input_directory="")

#Dada una lista de genes realizar el mapeo y la extracción de IDs
de string
Mapped_filter <- filter(Mapped, preferred_name %in% protein)
```



```

protein_stringID <- Mapped_filter$string_protein_id

string_filter <- filter(string, protein1 %in% protein_stringID,
                        protein2 %in% protein_stringID)

miRTarBase_Targets <- miRTarBase_filtrado$Target.Gene
miRTarBase_Targets <- data.frame(Target.Gene= miRTarBase_Targets)
Mapped <- string_db$map( miRTarBase_Targets,"Target.Gene",
                       removeUnmappedRows = TRUE)

miRTarBase_filtrado2 <- merge(Mapped, miRTarBase_filtrado)
miRTarBase_filtrado2 <-
miRTarBase_filtrado2[!duplicated(miRTarBase_filtrado2),]

#Crear función para identificar los blancos génicos dando una
lista de miRNAs
target_fun <- function(miR){
  Target <-filter(miRTarBase_filtrado2, miRNA %in% miR)
  Target <- Target$STRING_id
  Target <- Target[!duplicated(Target)]}

#Iterar la función para identificar blancos génicos
#a cada columna del data frame de muestras de miRNAs randomizada
df <- apply(c, MARGIN = 2, FUN = target_fun)

#Crear función para determinar el pValue de una red ppi
#determinar enriquecimiento de red y extracción de pvalue
String_Network_pValue <- function(IDs_string){
  ppi_enrichment <- string_db$get_ppi_enrichment(IDs_string)
  pValue <- ppi_enrichment[["enrichment"]]
  pValue <- as.numeric(pValue)}

####
#Determinar el pvalue de enriquecimiento de forma automática dada
una lista de genes
library(rlist)
q <- lapply(X=df, FUN = length)
q <- t(as.data.frame(q))
q <-data.frame(id=1:1000,
              length=q )
q <- dplyr::filter(.data = q, length<2)
q <- q$id

df1 <- list.remove(.data = df, q)

k <- lapply(X=df1, FUN = String_Network_pValue)
x <- t(as.data.frame(k))
x <- as.vector(x)
hist(x)

```

## Apéndice E: Código en R para validar significancia de redes de IPP por proporción de interacciones máximas

```
library(dplyr)
library(openxlsx)
library(tidyverse)

##### Lectura base de datos de miRTarBase
url <-
"https://mirtarbase.cuhk.edu.cn/~miRTarBase/miRTarBase_2022/cache/
download/8.0/hsa_MTI.xlsx"

miRTarBase <- read.xlsx(xlsxFile = url, sheet = 1, skipEmptyRows =
F)

miRTarBase_filtrado <- miRTarBase %>%
  filter(grepl('Western blot|reporter assay|qRT-PCR',
Experiments))

#####Predicción de blancos genéticos
df <- read.csv("Lista_miRNAs_588_mod.csv")
u <- df$miRNA

set.seed(1000)

e <-10
c <- c(1:6)*0
c <- as.data.frame(c)

for (i in 1:e) {
  a <- sample(u, 6, replace = F)
  c[i] <- cbind(a)
}

target_fun <- function(miR){
  Target <-filter(miRTarBase_filtrado, miRNA %in% miR)
  Target <- Target$Target.Gene
  Target <- Target[!duplicated(Target)]}

df <- apply(c, MARGIN = 2, FUN = target_fun)

library(rlist)
r <- lapply(X=df, FUN = length)
r <- t(as.data.frame(r))
r <-data.frame(id=1:e,
               length=r )
target.num <- r$length
r <- dplyr::filter(.data = r, length<2)
r <- r$id
```

```

df1 <- list.remove(.data = df, r)

####Lectura bases de datos de string

string <-
read.table("9606.protein.physical.links.detailed.v11.5.txt",
           header = T)
Mapped <- read.csv("9606_protein_info_v11_5.csv")

####Lectura blancos génicos de la firma de CF
protein <- read.csv("Target_CF.csv")
protein <- protein$x

####Mapeado blancos génicos de CF con string obtención de
interacciones
Mapped_filter <- filter(Mapped, preferred_name %in% protein)
protein_stringID <- Mapped_filter$string_protein_id

string_filter <- filter(string, protein1 %in% protein_stringID,
                       protein2 %in% protein_stringID)

interaction <- string_filter$experimental #Interacciones validadas
experimentalmente
interaction1 <- interaction>400 #indice de evidencia mayor a 400
interaction_num <- as.numeric(sum(interaction1, na.rm = TRUE))
interaction_num <- interaction_num/2

####Corrección de interacciones de acuerdo con el número máximo
de interacciones posibles dado el número de proteínas evaluadas
interaction_max <- as.numeric(length(protein_stringID))
interaction_max <- ((interaction_max^2)-interaction_max)/2

proporcion_interacciones <- interaction_num/interaction_max
s <- proporcion_interacciones

String.prop_interaction<- function(protein) {
  Mapped_filter_r <- filter(Mapped, preferred_name %in% protein)
  protein_stringID_r <- Mapped_filter_r$string_protein_id

  string_filter_r <- filter(string, protein1 %in%
protein_stringID_r,
                           protein2 %in% protein_stringID_r)

  interaction_r <- string_filter_r$experimental
  interaction1_r <- interaction_r>400
  interaction_num_r <- as.numeric(sum(interaction1_r, na.rm =
TRUE))
  interaction_num_r <- interaction_num_r/2

```

```

interaction_max_r <- as.numeric(length(protein_stringID_r))
interaction_max_r <- ((interaction_max_r^2)-interaction_max_r)/2

proporcion_interacciones_r <-
interaction_num_r/interaction_max_r

z <- proporcion_interacciones_r

}

#####Calcular proporción de interacciones IPP/IPPmax
q <- lapply(dfl, String.prop_interaction)
q <- as.data.frame(q)
q <- as.vector(t(q))
q <- na.omit(q)

hist(q, breaks = 50)

#####Calculo de pvalue corregido
if(s < mean(q)) {
  d <- q<=s
  h <- c(1:e)*0
  h[d] <- 1
  h <- sum(h)
  pvalue <- h/length(q)} else {
  d <- q>=s
  h <- c(1:e)*0
  h[d] <- 1
  h <- sum(h)
  pvalue <- h/length(q) }

```