

Tesis de Maestría, Área Biología Molecular y Celular PEDECIBA-
Biología

“Análisis de pequeños ARNs como biomarcadores en Cáncer de Pulmón”

Lic. Julia Sanguinetti

Orientador: Dr. Alfonso Cayota

Co. Orientador: Dr. Guillermo Dighiero

Tribunal

Dra. Adriana Geisinger

Dra. María Ana Duhagon

Dr. Mauricio Cuello

Montevideo, Octubre de 2016, IPMON

Índice

Agradecimientos.....	1
Abreviaturas.....	3
Resumen	5
1 Introducción	6
1.1 Los ARN no codificantes.....	6
1.1.1 Pequeños ARNs no codificantes	7
1.1.2 Pequeños ARNs emergentes.....	12
1.2 ARNs circulantes	16
1.3 Cáncer.....	19
1.3.1 Cáncer de pulmón.....	21
2 Hipótesis y Objetivos	26
3 Materiales y métodos	28
3.1 Voluntarios sanos y pacientes.....	28
3.2 Obtención de plasma y suero de sangre periférica.....	28
3.3 Aislamiento de pequeños ARN a partir de tejido.....	29
3.4 Aislamiento de pequeños ARN de plasma y suero.	29
3.4.1 Tratamiento con heparinasa I.....	29
3.5 Secuenciado de pequeños ARNs de tejido tumoral y peri tumoral.....	30
Construcción de librerías y secuenciado.....	30
3.6 Análisis bioinformático del secuenciado.....	31
3.7 Validación de candidatos por <i>stem-loop</i> RT-qPCR.....	32
3.8 Validación de candidatos con <i>primers</i> LNA.	35
3.9 Obtención de exosomas.	36
3.9.1 Extracción de exosomas de plasma.....	36
3.9.2 Extracción de exosomas de sobrenadante de cultivo celular.	36
3.9.3 Microscopía Electrónica de Transmisión (MET).....	37
3.9.4 Cuantificación proteica de exosomas.....	37
3.9.5 Detección de CD9 y HSP70 por Western Blot.....	37
4 Resultados.....	40
4.1 Construcción y caracterización de librerías de pequeños ARNs tisulares.....	40

4.2	Análisis bioinformático del secuenciado.....	43
4.3	Selección de pequeños ARNs diferencialmente expresados entre tejido tumoral y peri tumoral.	48
4.4	Validación de candidatos seleccionados por PCR cuantitativa basada en <i>stem-loop</i> RT-qPCR.	50
4.5	Análisis de expresión diferencial con <i>primer</i> LNA.	53
4.6	Análisis de plasma y suero como fuente de ARNs circulantes de interés como biomarcadores.....	54
4.7	Búsqueda de pequeños ARNs en exosomas.....	56
4.7.1	Obtención de exosomas	57
5	Discusión.....	60
6	Conclusiones	69
7	Perspectivas.....	69
	Bibliografía	71

Agradecimientos

Dr. Alfonso Cayota por las oportunidades en estos años y la orientación en esta etapa formativa.

Dr. Guillermo Dighiero por apoyar y formar parte de este proyecto.

Lic. Tamara Fernandez por el análisis de las secuencias.

Dr. Siul Salysbury, Dr. Mauricio Cuello y la Dra. Clarita Rodríguez por colaborar en la obtención de muestras de los pacientes.

Ratas del Laboratorio de Genómica Funcional del IPMON: Lic. Fabiana Gámbaro, Dr. Juan Pablo Tosar, Dra. María Rosa García, Msc. Braulio Bonilla, Msc. Florencia Cabrera, siempre a disposición, aportando y haciendo el día a día agradable y divertido.

A la salita: Sofía Horjales, Lorena Tomé, Federico Carrión, Natalia Olivero, Sergio Bianchi, compañeros de muchas tertulias.

Personal, técnicos e investigadores del Instituto Pasteur de Montevideo.

Al tribunal de tesis: Dra. Adriana Geisinger, Dra. María Ana Duhagon y Dr. Mauricio Cuello.

Al Programa de Desarrollo de las Ciencias Básicas (PEDECIBA), Agencia Nacional de Investigación e Innovación (ANII) e Instituto Pasteur de Montevideo (IPMON).

A mi familia y mis amigos.

Abreviaturas

ARN: ácido ribonucleico

ARNnc: ARN no codificantes

sARNs: pequeños ARNs no codificantes

miARN: micro ARNs

siARNs: pequeños ARNs interferentes

piARNs: ARN asociados a proteínas PIWI

nt: nucleótidos

ARNm: ARN mensajero

TE: elementos transponibles

ARNt: ARN de transferencia

ARNr: ARN ribosomal

sARNt: pequeños ARNs derivados de ARNs de transferencia

sARNr: pequeños ARNs derivados de ARNs ribosomales

sARNno: pequeños ARNs derivados de ARNs nucleolares

sARNn: pequeños ARNs derivados de ARNs nucleares

ARNv: ARN vault

sARNth: mitades de ARN de transferencia (tRNAs half)

tRFs: fragmentos derivados de ARN de transferencia

TKI: inhibidores de tirosin kinasas

Px: Paciente x (x: número de paciente)

TN: tejido peri tumoral sano

TT: tejido tumoral

Resumen

Los pequeños ARNs no codificantes (sARNnc) son los mayores reguladores de la expresión génica, regulando más del 50% de los genes que codifican proteínas. Además de su rol central en la fisiología celular, son liberados al medio extracelular (fluidos biológicos) en vesículas extracelulares o unidos a complejos ribonucleoproteicos, siendo captados e internalizados funcionalmente activos por células adyacentes o distantes.

A nivel mundial el cáncer de pulmón es la principal causa de muerte por cáncer. En Uruguay ocupa el primer lugar en hombres y tercer lugar en mujeres de las muertes por cáncer. A pesar de los grandes esfuerzos por combatir la alta mortalidad, no existen en la actualidad métodos que permitan una detección temprana ni un seguimiento poco invasivo de esta enfermedad. La diversidad y estabilidad de los pequeños ARN circulantes, la especificidad por sus blancos, la fácil accesibilidad de plasma o suero y el rol central en la iniciación y progresión del cáncer, los posiciona como eventuales candidatos a biomarcadores para diagnóstico, predicción de respuestas terapéuticas y como blancos terapéuticos en cáncer de pulmón.

Mediante la utilización de tecnologías de secuenciado masivo y PCR cuantitativa, este trabajo tiene como objetivo central el estudio de pequeños ARNs circulantes como potenciales biomarcadores diagnósticos en cáncer de pulmón. Nuestros resultados identificaron y validaron 7 sARNs como eventuales marcadores tisulares de cáncer de pulmonar (miR-21-5p, miR100a-5p, miR-148a-3p, Glu_TTC, Gly_GCC, Gln_CTG, His_GTG), de los cuales los fragmentos derivados de ARNt Gly_GCC y Glu_TTC se comportaron como potenciales biomarcadores en muestras de suero de pacientes. Estos resultados preliminares justifican la realización de estudios clínicos controlados para validar su capacidad diagnóstica mediante técnicas no invasivas.

1 Introducción

1.1 Los ARN no codificantes

Si se observa en una línea de tiempo cómo ha evolucionado el conocimiento en biología a nivel molecular, llama mucho la atención como se ha complejizado el dogma central de la biología molecular planteado en 1958 por Francis Crick (1). En los últimos 15 años, el impacto del conocimiento post-genómico asociado a los grandes avances tecnológicos han demostrado que la regulación de la expresión génica es mucho más compleja de lo que se pensaba en la era pre-genómica. En particular el análisis del transcriptoma humano muestra que el número de proteínas es menor del esperado, apenas el 1,8% del genoma codifica para ARNs que serán transcritos a proteínas (ARNs codificantes), y que más del 60% del genoma codifica para ARNs que no son traducidos a proteínas, ARNs no codificantes (ARNnc) (Figura 1) (2).

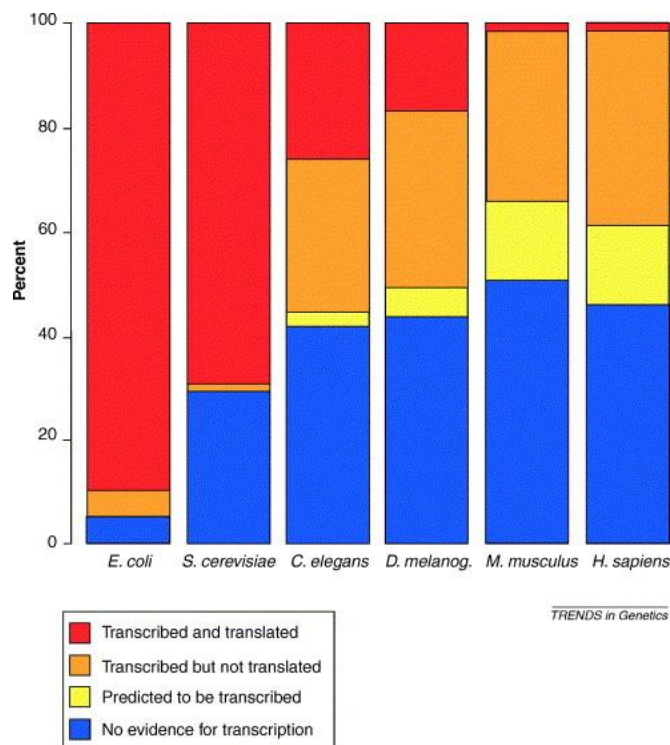


Figura 1. En esta figura se muestran los porcentajes de genes que son transcritos y traducidos junto con los que son transcritos y no traducidos en diferentes especies. Mostrando que a medida que se complejiza el organismo aumenta el porcentaje de transcritos no traducidos. Tomado de referencia (3).

A estos ARNnc se los reconoció rápidamente como ARN reguladores, con participación activa en la regulación de la expresión génica a nivel transcripcional y pos-transcripcional (2,3). Inicialmente se los clasificó según su longitud, identificándose varias categorías: pequeños ARNs de 18-40pb (sARNs) (4), ARNs largos mayores de 200pb (long non-coding ARNs) (5,6), ARNs de 500pb y hasta de 108kb (7).

El primer sARN descrito fue el microARN *lin-4*, identificado por primera vez en *Ceanorhabditis elegans*, en 1993 (8). En 1998 A. Fire y C. Mello describen el mecanismo por interferencia mediado por pequeños ARNs (9), lo que les valió el premio Nobel de Medicina en 2006. A partir de estos hallazgos se los ha encontrado en casi toda la escala evolutiva, lo cual muestra otra evidencia de la relevancia biológica y el rol central que cumplen estos sARNs en la regulación de la expresión génica y otros procesos celulares (10).

1.1.1 Pequeños ARNs no codificantes

Los pequeños ARNnc (sARNs, del inglés, *small RNAs*) son considerados reguladores mayores de la expresión génica a nivel transcripcional y postranscripcional, capaces de regular más del 50% de los genes (11). Se definen por su tamaño entre 18 y 40 nucleótidos, y por su asociación a proteínas de la familia Argonautas (AGO/PIWI) (12). El mecanismo de acción general de los sARNs, es a través de la unión a complejos proteicos e interactuando con sus blancos mayoritariamente por complementariedad total o parcial de bases. Los mecanismos efectores de estas interacciones dependerán del tipo de interacción y podrán según el caso, inhibir la traducción, degradar el ARN mensajero (ARNm) (13), promover la formación de heterocromatina, inducir modificaciones epigenéticas y re-arreglos cromosómicos (14,15).

Las alteraciones del perfil de expresión de sARNs conducen a una disfunción celular, provocando cambios que están en la base de múltiples procesos patológicos como el cáncer, enfermedades neurológicas, metabólicas o cardiovasculares, entre otras (16,17). El efecto de muchos sARNs en procesos como la transformación maligna ha llevado a catalogarlos como nuevos oncogenes o supresores tumorales (17,18).

Tomando en cuenta el origen de su biogénesis, el tipo de proteínas asociadas y su mecanismo de acción, los sARNs son clásicamente clasificados en tres grandes

familias: micro ARNs (miARNs), pequeños ARN interferentes (siARN) y ARNs asociados a proteínas piwi (piARNs) (Figura 2). Con la incorporación reciente de técnicas de secuenciación masiva han surgido nuevas clasificaciones en las cuales se han incorporado nuevas clases o categorías de sARNs. Estas nuevas clases se retomarán más adelante como sARNs emergentes.

1.1.1.1 siARNs, piARNs y miARNs

Los **siARN** son fragmentos de 21-25 nucleótidos (nt) aproximadamente. Los primeros reportes del silenciamiento génico por interferencia fueron en 1990 en plantas (19) y luego en *C. elegans* en 1998 (9). Derivan de largos fragmentos de ARN doble hebra endógenos (endo-siARNs) o exógenos (exo-siARNs) y son dependientes en su biogénesis de la endoribonucleasa tipo III "Dicer" (13).

Los fragmentos de ARN doble hebra procesados por la endoribonucleasa "Dicer" generan fragmentos doble hebra de 22nt con extremos característicos, un extremo 5' fosfato y 2 nucleótidos 3', sobresalientes en los extremos del ARN doble hebra (20). En cuanto a Dicer, en mamíferos y en *C. elegans* se ha identificado una única proteína, la cual también participa en el clivaje de los miARNs. Sin embargo, en *Drosophila* se conocen dos, una para los miARNs y otra para la producción de siARN (21). Una de las hebras de este fragmento de 22nt, generado por Dicer servirá como guía y será incorporado o "cargado" en el complejo multiproteico RISC (del inglés, *RNA-induced silencing complex*) que mediante la interacción con una proteína Argonata actuará como molécula efectora del complejo ribonucleoproteico (13,21). Si bien el RISC está compuesto por más de 15 proteínas, el ARN guía y las proteínas de la familia Argonata son los principales actores en este complejo (22). En humanos y *Drosophila* la proteína Argonata 2 (Ago 2), es la responsable de la degradación del ARNm blanco. En este caso el sARN permanece unido al complejo RISC permitiendo realizar múltiples rondas de corte de ARNm (23,24). En algunos eucariotas como plantas y nematodos la formación de ARN doble hebra requiere de proteínas ARN polimerasas dependientes de ARN (RdRP). Las RdRP son capaces de replicar el ARN teniendo como matriz una mono hebra de ARN complementario generando transcritos doble hebra que luego serán procesados por Dicer en la biosíntesis de siARNs (21).

Los siARN participan en la regulación de la expresión génica al igual que los miARN; sin embargo también están implicados en el silenciamiento de genes repetidos, de transposones y procesamiento de virus como mecanismo de defensa del genoma (13,21).

Los **miARNs** son fragmentos endógenos de 21-23nt y actualmente hay más de 2500 miARNs humanos catalogados en la versión 21 de la base de datos de referencia miRbase (<http://www.mirbase.org/>). Su principal función es la regulación de la expresión génica a nivel pos transcripcional uniéndose por complementariedad de bases a los ARNm blanco. Dependiendo del grado de complementariedad a sitios de unión en la región 3'UTR del ARNm los miARNs son capaces de inhibir la traducción (complementariedad imperfecta), o bien la degradación del ARNm a nivel co-traducciona (complementariedad perfecta) (25,26). Además también participan en procesos de control epigenético a través de la regulación de la metilación del ADN, la modificación de histonas y el remodelado de la cromatina (27–29).

Debido a que los miARNs son transcritos por la ARN polimerasa II, los transcritos tienen cola poliadenilada en el extremo 3' (cola poli-A) y una caperuza en el 5' (30,31). En ocasiones, los genes de los miARNs suelen ubicarse próximos unos de otros generándose transcritos policistrónicos (10). De acuerdo a la región del genoma donde se ubican estos genes se los agrupa en diferentes categorías: miARN intrónicos (suelen presentar el mismo patrón de expresión que el gen en el que reside), en exones de genes no codificantes y en intrones de genes codificantes (10). Aproximadamente el 40% de los miARN en animales están localizados en regiones intrónicas mientras que el 10% se localiza en regiones exónicas no codificantes (10).

La transcripción canónica o más "directa" es la generación de un transcrito primario denominado o bien pri-miARN que forma estructuras características en forma de horquilla o bien *stem-loop*. Esta estructura es reconocida por la RNasa III, denominada microprocesador o Drosha y un co-factor de Drosha que reconoce ARN doble hebra DGCR8 (del inglés, *Di George Critical Región 8*) (32). Los productos de corte son denominados pre-miARN y tienen aproximadamente 70nt. El pre-miARN

también adopta una forma de horquilla que se unirá a la exportina-V para ser exportados al citoplasma (33,34).

Otro mecanismo interesante es la generación de los mirtrons, los cuales son procesados juntos con el ARNm que los contiene, por el microprocesador o por el propio espliceosoma. Existen reportes donde se describe que otro grupo de sARNs son procesados por este mecanismo (35).

Una vez en el citoplasma el pre-miARN es reconocido y procesado por Dicer, generando un dúplex de 22nt ARN. Dicer puede actuar sola o unida a proteínas "*partners*". En la mayoría de los mamíferos se le conocen dos proteínas "*partner*": TRBP (del inglés, *transactivation-response RNA-binding protein*) y PACT (del inglés, *protein kinase R-activating protein*) (25). Continuando con la síntesis, una de las hebras del ARN dúplex se une a la proteína Argonauta (Ago 2 en humanos) mientras que la otra hebra es degradada ("hebra pasajera"). Ago 2 forma parte del complejo miRNP, que es el mínimo complejo activo, el cual puede estar formando parte del complejo RISC (el mismo que participa en el procesamiento de los siARNs) (25). El complejo ribonucleoproteico es guiado hacia el ARNm blanco, donde interactuará por complementariedad de bases con el miARN. En este caso la especificidad está dada por 6 bases del miARN ubicados en extremo 5', denominada también "región semilla". Frecuentemente estas regiones se unen a la región 3'UTR del ARNm blanco, disminuyendo la expresión del ARNm. Este mecanismo tiene el potencial de regular más del 50% de los genes (25,33). Otros mecanismos por los cuales actúan los miARNs son: deadenilación de la cola poli-A (induce degradación del ARNm), que sitúa el complejo en el 5'UTR, inhibe el inicio de la traducción, degrada el péptido en formación y recluta los mensajeros en cuerpos de inclusión o p-bodies para almacenamiento temporal o degradación (25).

En los últimos años, en diversos organismos como mamíferos, *Drosophila* y plantas se describieron diferentes estructuras citoplasmáticas que presentan miARN y proteínas de unión a ARNm (proteínas argonautas), que incluyen cuerpos P, GW y gránulos de estrés. Si bien se conocen como gránulos para la degradación, también se ha propuesto como un mecanismo de almacenamiento de ARNm para una eventual

respuesta rápida o como reserva en situaciones de estrés, por ejemplo, un contexto tumoral (23,25,26,36,37).

La otra clase de sARNs son los **piARNs**; fue la última clase de sARNs identificada dentro de las tres clásicas. Son fragmentos endógenos de aproximadamente 21-30nt, siendo los más largos de los sARNs. Se distinguen de los miARN y los siARN en que son procesados de manera independiente de Dicer y se unen específicamente a la subfamilia de Argonautas, PIWI. Los piARNs derivan de elementos genómicos repetitivos, *piARNs clusters* y su principal función es el silenciamiento de los elementos transponibles (TE) en la línea germinal, siendo imprescindibles para el correcto desarrollo de las gónadas y las células madre. Por otra parte, la presencia de los piARN está restringida a algunos organismos pluricelulares (28).

La vía que presentan los piARNs para discriminar a los elementos transponibles de los genes que codifican proteínas normales es perfectamente regulada y específica, y su regulación es principalmente a nivel de la cromatina (38). Durante su síntesis los piARN se asocian con proteínas PIWI en el citoplasma aunque también se ha reportado que algunas PIWI son translocadas al núcleo actuando a nivel transcripcional e induciendo la formación de heterocromatina. La vía de síntesis de los piARN se la puede dividir en dos partes. La primera está totalmente caracterizada, es un proceso en el cual se genera un pool de piARN primarios que marcan múltiples TE, de manera independiente de Dicer. Luego, en una segunda etapa, los piARN primarios (sin Uridina) se unen a la PIWI y a AUB (Aubergine) y los piARNs son amplificados en un ciclo dependiente de proteínas argonautas, denominado "piARN ping-pong" (28,38).

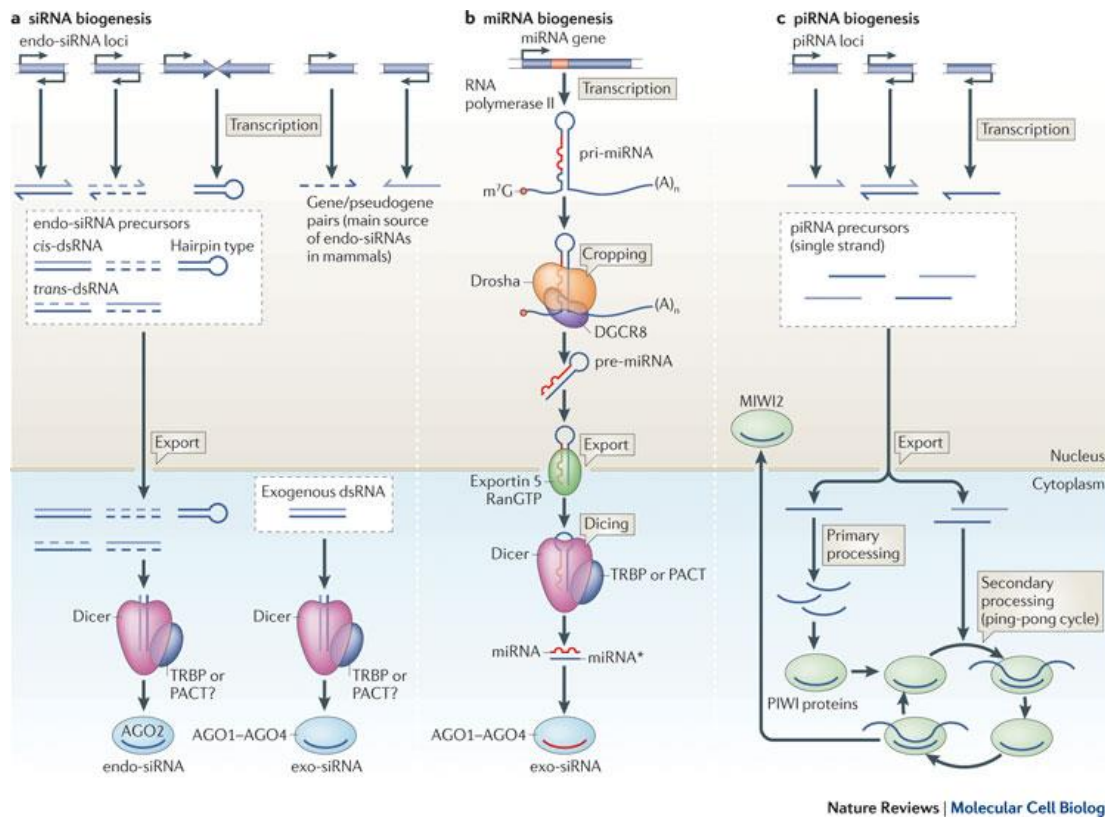


Figura 2. Esquema de las vías clásicas de sARNs. Tomado de referencia (38).

1.1.2 Pequeños ARNs emergentes.

Además de los ARNnc conocidos (ARNt, ARNr, ARNno, ARNn, sARNs) el uso de tecnologías de secuenciado masivo ha revelado la presencia de otras clases de sARNs las cuales no son sintetizadas por las vías clásicas o canónicas. Dentro de estas nuevas clases se distinguen fragmentos derivados de ARN de transferencia (sARNt), de ARNs nucleolares (sARNno), ARNs ribosomales (sARNr), ARNs nucleares (sARNn) y vault ARNs (ARNv) (12,39).

En los últimos años, nuestro laboratorio ha estado trabajando en el estudio de estos nuevos sARNs emergentes, en particular, en los fragmentos derivados de ARNt (sARNt). Hemos reportado la presencia de sARNt en diferentes organismos a nivel celular y extracelular (36,40). Simultáneamente, comenzaron a aparecer reportes que describen sus potenciales funciones y algunas de las vías y componentes de sus rutas biosintéticas. Actualmente, casi no existen dudas de que los precursores de los sARNt son ARNt maduros, que comienzan a ser procesados en el núcleo por la RNasaZ y la

RNasaP (41). Se distinguen dos grupos principales de sARNt, las mitades de ARNt (del inglés, tRNAs halves) y fragmentos derivados de ARNt (de inglés, tRFs) (39).

1.1.2.1 tRFs

Los tRFs se clasificaron en tRF-1, tRF-3 y tRF-5 en función del segmento del transcripto de tRNA del cual derivan (Figura 3) (42). Los tRF-3 y tRF-5 son fragmentos de 13-22 nt, generados por clivaje del ARNt maduro a nivel del loop-D y loop-T respectivamente. El procesamiento en gran proporción es dependiente de Dicer (39,43), aunque se observan excepciones en la generación de los tRF-3 (44). En otros trabajos más recientes se postula a la Angiogenina como otra ribonucleasa, generadora de tRF-3 (43). Simultáneamente, se ha reportado en diferentes especies la unión de los tRFs a proteínas Argonautas (AGO 1-4) sugiriendo una alta conservación de estos ARNs a lo largo de la escala evolutiva con una relevancia biológica aún no elucidada ni valorada (43). Por último, los tRF-1, corresponden a los trailers 3' de los ARNt, los cuales tienen un repetido de Uridinas en el extremo 3', como consecuencia de la terminación de la polimerasa III en la transcripción. La proteína implicada en la generación de estos fragmentos es la endonucleasa tRNasaZ en el núcleo (42). El sitio de clivaje en el extremo 5' varía, por lo que se supone que alguna otra proteína, aún no identificada, también estaría actuando (39).

La asociación de los tRFs con proteínas que participan en las vías canónicas de sARNs (Dicer y Argonautas) y la semejanza en el tamaño sugieren la participación de los tRFs en vías de silenciamiento génico a nivel pos transcripcional semejantes a los planteados para los siARN/miARN. Otro ejemplo de la funcionalidad de los tRFs es la participación en vías de contención de virus, que se unen al ARN viral inhibiendo su amplificación (43). Por el contrario, se ha demostrado la participación del tRF-1, lo cual ha aumentado la proliferación celular, en líneas celulares de cáncer de próstata (42).

1.1.2.2 Mitades de ARNt ("tRNA halves")

Las mitades de ARNt son fragmentos que varían entre 30-35nt provenientes de ARNt maduros que son clivados a nivel o adyacente al anticodón (Figura 3). También son nombrados tiRNAs (del inglés, *stress-induced tRNA-derived fragment*) ya que son procesados por la ribonucleasa Angiogenina, la cual es activada y secretada bajo condiciones de estrés fisicoquímico o biológico (12). La activación de esta nucleasa en

esta vía bajo condiciones de estrés desencadena la inhibición de la traducción y la formación de gránulos de estrés a través de la interacción con la proteína de unión a ADN y ARN YB-1 (del inglés, *Y-box binding protein*) (45). En cuanto a la regulación del procesamiento de estos fragmentos se sabe que la angiogenina está regulada por la RNH1 (del inglés, *Ribonuclease/angiogenin inhibitor 1*) (46). A su vez existe una regulación independiente de angiogenina que depende de las proteínas Dnmt2 (del inglés, *DNA methyltransferase 2*) y NSun2 (del inglés, *NOP2/Sun RNA methyltransferase 2*) que modifican los ARNt mediante metilación para no ser clivados por la angiogenina durante situaciones de estrés (45).

Además de las vías de regulación mencionadas, recientemente se han descrito otras formas de regulación bajo ciertas condiciones patológicas. Un ejemplo es la descripción de los "SHOT-RNAs" (del inglés, *sex hormone- dependent tRNA derived RNAs*). Estos son mitades de ARNt que son inducidas por hormonas sexuales y tienen un rol central en las respuestas proliferativas de tumores de mama y próstata hormono-dependientes (45). Otro ejemplo es la participación de sARNt como moléculas de señalización inmunológicas propuesto por Dhabbi et al. (47).

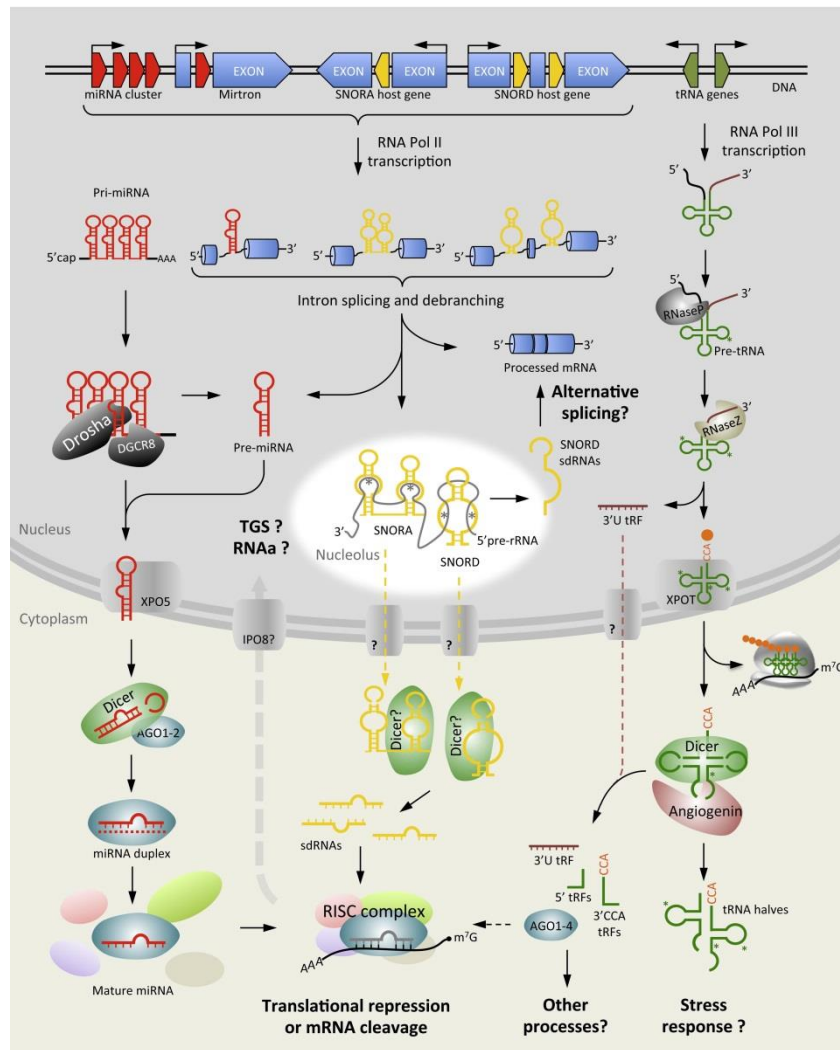


Figura 3. Esquema de las vías no canónicas que involucran a los sARNt (39).

Existe suficiente evidencia para pensar que la presencia de los fragmentos derivados de ARNt no es un proceso al azar o un producto de degradación (12,39):

- Los niveles de ARNt durante situaciones de estrés no se ven afectados mientras que sí se ven aumentados los sARNt.
- Hay secuencias específicas donde se da el clivaje en los ARNt maduros ya procesados.
- Las cantidades de cada sARNt (tRFs o mitades de ARNt) en una célula no se ven reflejados en la abundancia relativa de cada ARNt maduro.
- Depende del tipo celular la abundancia relativa de cada fragmento.
- Se asocian a proteínas Argonautas y se conocen funciones asociadas a estos fragmentos.

La heterogeneidad de sARNs entre los diferentes organismos dificulta el entendimiento de su rol, ya que muchos de estos carecen de alguna/s proteína/s implicada/s en estas vías. Sin embargo, es interesante resaltar que estos sARNs derivados de ARNt presentes en diferentes organismos, cumplen funciones comunes, como la regulación de la expresión génica. Esto ha llevado a diferentes autores a proponer que esta vía de sARNs representa un mecanismo ancestral de la regulación génica que se ha conservado y ha evolucionado hacia miARNs (39).

1.2 ARNs circulantes

La descripción de la presencia de ADN libre en la circulación sanguínea en 1948 por Mandel y Matteis (48), constituye el primer reporte formal de ácidos nucleicos circulantes en fluidos biológicos. Posteriormente, se demostró que un aumento de los mismos en circulación se asocia fuertemente con la carga tumoral en pacientes con cáncer. Actualmente, el ADN circulante es una fuente invaluable en análisis prenatales a partir de sangre materna (49,50). En cuanto a los ARNs circulantes, la presencia de ribonucleasas en fluidos biológicos y sobretodo en sangre hizo dudar mucho tiempo de su presencia. Actualmente, sabemos que moléculas intactas de ARNm, miARNs y otros sARNs reguladores son secretados activamente al espacio extracelular por diversos tipos celulares, siendo detectables en diversos fluidos corporales como suero, plasma, orina, saliva, fluido linfático, leche materna, entre otros. Estos ARNs son vehiculizados y protegidos de las ribonucleasas unidos a complejos ribonucleoproteicos (51,52), o en estructuras microvesiculares ("vesículas circulantes") como cuerpos apoptóticos, exosomas y ectosomas (53,54).

Simultáneamente, se ha demostrado la transferencia horizontal de ADN y ARN funcional entre células, lo cual es hoy concebido como un nuevo mecanismo de comunicación intercelular, capaz de inducir cambios significativos en las células receptoras (55–57). Esto es relevante para la homeostasis y fisiología tanto en condiciones normales como patológicas, entre ellas en cáncer (53,58,59).

Los **exosomas** son las vesículas circulantes mejor caracterizadas y su composición refleja el origen celular, el estado de la célula y el rol que puede llegar a cumplir en su célula blanco. Son vesículas con doble membrana lipídica, de entre 30-100nm de

diámetro y poseen una densidad entre 1,10-1,21 g/ml (60). Son sintetizados a través de la vía endosomal (endosomas tardíos) en los cuerpos multivesiculares, MVB (del inglés, *multivesicular bodies*), los cuales se fusionarán luego con la membrana plasmática, liberando los exosomas al medio extracelular. La formación de vesículas intraluminales es el comienzo de la formación de exosomas. Dos pasos claves parecen guiar el proceso. En el primero, se organiza la membrana endosomal, enriqueciéndose con tetraspaninas (principalmente con CD9 y CD63). A continuación, es reclutado el complejo ESCRT (del inglés, *endosomal sorting complex required for transport*), el cual parece ser clave en la vía de síntesis y en el contenido de los exosomas (60).

Si bien los exosomas tienen un perfil proteico similar al de la membrana plasmática, también poseen proteínas propias de exosomas que sirven como marcadores específicos. Por ejemplo: ALIX, CD63, CD81, CD9, LAMP1 y HSP70 (61,62). La interacción entre la célula blanco y los exosomas se da a través del reconocimiento de proteínas de la membrana del exosoma con receptores en la membrana plasmática, y activa así vías de señalización en la célula blanco. Otra forma de interacción es mediante la fusión del exosoma con la célula blanco, que libera el contenido en el citoplasma (61).

Las células tumorales tienen aumentada constitutivamente la generación y secreción de vesículas extracelulares que, por distintos mecanismos, promueven la progresión tumoral favoreciendo la proliferación celular, la angiogénesis, la migración celular y la generación de metástasis e inhibición de mecanismos de respuestas inmunológicas (54). Se han descrito ARNm, ARNr y sARNs como parte del contenido exosomal. Este material génico es transferido horizontalmente a otras células y son totalmente funcionales, induciendo cambios fenotípicos en la célula receptora (56,63). Un claro ejemplo lo aporta Peinado et al. (2012), demostrando como los exosomas de melanoma son capaces de preparar el nicho metastásico en ciertos órganos, como el pulmón, a través de la expresión de ciertas moléculas de adhesión y quimioquinas asociadas al reclutamiento de células progenitoras vasculares derivadas de médula ósea, que facilitan el proceso metastásico (63).

Los **ectosomas o microvesículas** son vesículas circulantes menos caracterizadas de 100-500nm liberadas por gemación directa de la membrana plasmática de las células. Se forman mediante el ensamble de complejos proteicos y lípidos reclutados específicamente en la membrana plasmática. Este complejo recluta lípidos y proteínas específicas de ectosomas tanto para la superficie como para el interior vesicular. Si bien aún no se conocen todos los mecanismos moleculares que guían este proceso, se sabe que proteínas como ESCRT juegan un rol fundamental en el reclutamiento de proteínas, formación y liberación de las microvesículas. A su vez aparecen moléculas comunes a los exosomas, como proteínas, y varios tipos de ácidos nucleicos, como ARNm, sARN y lncARN (64).

Sin embargo, la contribución de los tejidos periféricos a la población de vesículas circulantes (exosomas/microvesículas) se ve afectada por la alta proporción de vesículas aportadas por muchas de las células sanguíneas, como aquellas derivadas de plaquetas y cuerpos apoptóticos. Esto hace muy dificultoso poder identificar en plasma las pequeñas proporciones de este tipo de vesículas (exosomas/microvesículas).

Recientemente se demostró que los sARNs circulantes reportados previamente en vesículas extracelulares representan una pequeña proporción del total circulante. La mayor fracción de los sARNs está asociada a proteínas como la Argonauta 2 (AGO2) (51,65), a lipoproteínas de alta densidad (HDL) y a proteínas nucleares como la nucleofosmina-1 libres de vesículas (52). Estas proteínas circulantes se han propuesto como vehículos para internalizar los ARNs unidos en las células blanco (66). A esta fracción formada por complejos proteicos unidos a ARNs la denominamos **fracción ribonucleoproteica o no vesicular**. Inicialmente, se describieron miARNs asociados a estos complejos, pero actualmente se sabe que la mayor parte de los sARNs circulantes, tanto en sobrenadante de cultivo como en plasma, son mitades de ARNt y derivados de YARNs (36,40,67,68).

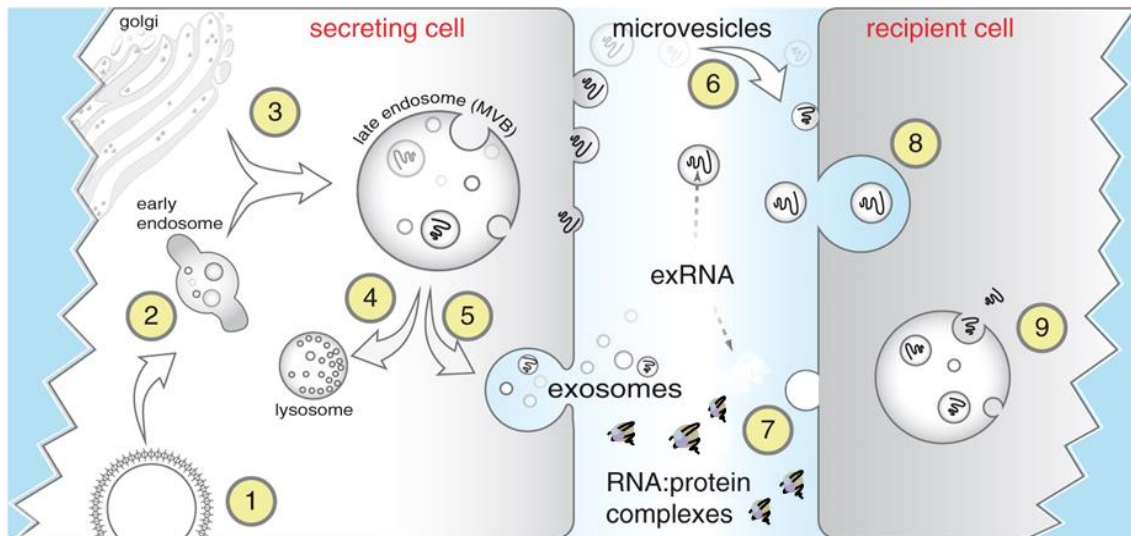


Figura 4. Esquema de secreción de ARNs. Fracción exosomal, microvesicular y complejos ribonucleoproteicos. Extraído y modificado de (69)

Actualmente, el campo de **sARNs circulantes** está en expansión lo cual es una gran promesa sobretodo en el área biomédica, ya que no solo denota un mecanismo más de comunicación entre las células, sino que también los posiciona como grandes candidatos a biomarcadores.

Existe una expresión tejido específica de los sARNs y cambios en los niveles de ARN circulantes se han asociado con la carga tumoral y la progresión del cáncer (70,71). Estos ácidos nucleicos circulantes actúan como “hormonas” o “mensajeros”, teniendo el potencial de influenciar la homeostasis normal e incidir en aspectos de la iniciación y progresión del cáncer, como la invasión y la metástasis (58,63). Desde el punto de vista médico y su aplicabilidad en la clínica, estos ARNs circulantes tienen como ventaja que pueden ser utilizados como biomarcadores para diagnóstico, seguimiento evolutivo y monitoreo terapéutico. Ofrecen la posibilidad de utilizar técnicas menos invasivas, fácilmente accesibles y de menor costo como fuente de biomarcadores diagnósticos y pronósticos.

1.3 Cáncer

El cáncer constituye un conjunto de enfermedades que se caracteriza por un crecimiento desregulado de las células que componen los diferentes tejidos del organismo, las cuales adquieren la capacidad de modificar el tejido de origen y de

propagarse invadiendo nuevos tejidos, tanto a nivel regional como a distancia. Las células afectadas presentan una acumulación de mutaciones en genes claves denominados genes supresores tumorales y oncogenes. En el proceso de transformación maligna las células adquieren una o más características conocidas como los "hallmark" del cáncer (figura 5). Estos son: el aumento en las señales proliferativas, no responden a las señales anti proliferativas, evaden la apoptosis, se replican indefinidamente, inducen la angiogénesis, adquieren la capacidad de invadir los tejidos y de hacer metástasis. Además, estas células presentan un aumento de la inestabilidad genómica, viéndose afectada la tasa mutacional, la activación de nuevas vías metabólicas, la evasión del sistema inmune y la generación de inflamación crónica local (72).

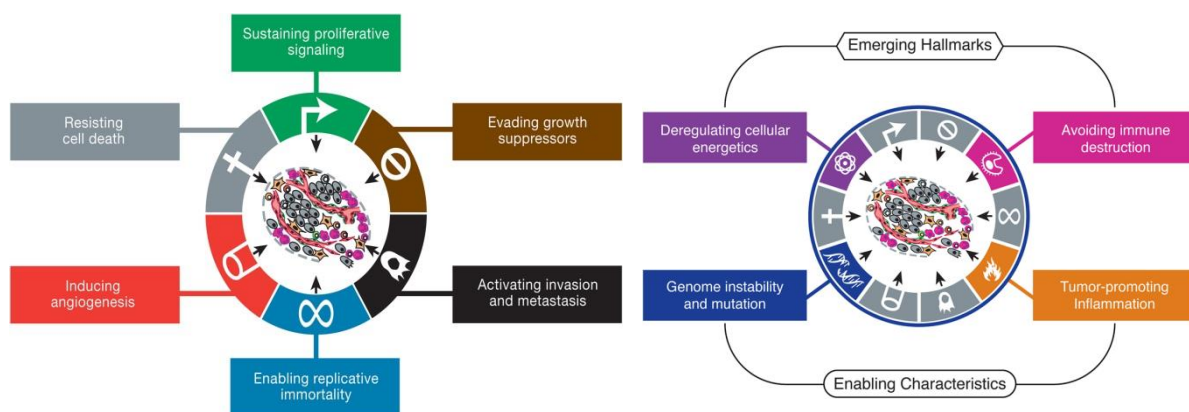


Figura 5. Hallmark del cáncer (Izquierdo). Nuevos hallmark propuestos por los mismos autores 10 años después (Derecho) (72).

Menos del 10% de los casos de cáncer presentan mutaciones heredadas en la línea germinal, y la mayoría de los casos son causados por mutaciones adquiridas (somáticas) a lo largo de la vida (cáncer esporádico), ya sean mutaciones espontáneas o causadas por agentes carcinogénicos ambientales (sustancias tóxicas, exposición UV entre otros) (73).

En cuanto a la biología molecular de los tumores, estos presentan un perfil mutacional propio, constituido inicialmente por una población homogénea de células tumorales. Debido a la alta inestabilidad genómica y la alta proliferación, durante el

desarrollo del tumor se generan nuevas subpoblaciones que generan múltiples y variados fenotipos. A su vez, dentro de un mismo tumor se ha identificado una población denominada "*cancer stem cells*" (CSC), la cual se distingue por presentar características de células pluripotentes y tener la capacidad de generar nuevos tumores cuando se inoculan en otros organismos (72). La inestabilidad genómica en estas poblaciones celulares modifica constantemente el perfil molecular de los tumores, lo cual constituye una de las razones en la generación de resistencia a los tratamientos y estimulando a que las terapias sean cada vez más dirigidas, es decir, con blancos moleculares específicos en cada tejido y para cada tumor. Actualmente existe un amplio número de mutaciones descritas, asociadas a distintos tipos de tumores, principalmente en genes oncogénicos y supresores tumorales (74).

Tabla 1. Vías asociadas a carcinogénesis por mutaciones en genes supresores tumorales y oncogenes. Adaptado de (74).

Supresores tumorales
Inhibidores de las vías señalizadoras mitógenas (APC, PTEN)
Inhibidores de la progresión del ciclo celular (RB, CDKN2A)
Inhibidores de programas de metabolismo y angiogenia (VHL, STK11)
Inhibidores de la invasión y metástasis (CDH1)
Potenciadores de la estabilidad genómica (TP53)
Factores de reparación del ADN (BRCA1, BRCA2)
Oncogenes
Factores de crecimiento (TGF- α , HGF)
Receptores de factores de crecimiento (EGFR, HER, ALK, RET)
Proteínas que participan en transducción de señales (GTP, KRAS, BRAF, NOTCH1)
Proteínas reguladoras nucleares (MYC, NMYC)
Reguladores del ciclo celular (CCND1, CDK4)

1.3.1 Cáncer de pulmón

Según los últimos registros de la Organización Mundial de la Salud, el cáncer de pulmón es la principal causa de muerte por cáncer a nivel mundial y es uno de los cánceres con menor tasa de supervivencia con un 6% a los 5 años (75). En Uruguay, los datos reportados por la Comisión Honoraria de Lucha Contra el Cáncer hasta 2013,

indican que la tasa de incidencia de cáncer de pulmón cada 100.000 habitantes es de 47,9 en hombres y 10,7 en mujeres, ocupando el primer y tercer lugar respectivamente de las muertes por cáncer (76). Este tipo de cáncer se da frecuentemente entre los 40 y los 70 años.

La mayor parte de los casos de cáncer de pulmón se debe a la exposición a diversos carcinógenos. En este sentido, el 80% de los carcinomas se presenta en pacientes con una larga historia de tabaquismo. Otras sustancias reconocidas como carcinógenas se encuentran en el ambiente (radiación, smog) o en industrias donde sustancias como amianto, arsénico, cromo, uranio o níquel son los principales carcinógenos (73).

El 90-95% de los tumores pulmonares son carcinomas y se clasifican histológicamente según sus características clínicas y moleculares. Se dividen en cuatro categorías principales: adenocarcinoma (38%), carcinoma epidermoide (20%), carcinoma de células pequeñas (14%), carcinomas de células grandes (3%) y otros (25%) (73).

Los adenocarcinomas son los carcinomas más frecuentes, y la mayoría de los pacientes son fumadores. Presentan mutaciones con ganancia de función en genes que afectan las vías de señalización de receptores de factores de crecimiento. Entre estos están EGFR, ALK, ROS, MET y RET, todos receptores con dominios tirosin quinasa. También aparece mutado con alta frecuencia el gen KRAS (73).

Según su extensión anatómica en el momento del diagnóstico los tumores son estadificados utilizando el criterio TNM (T: de tamaño del tumor, N: de nódulos y M: de metástasis), detallado en la tabla 2. Este criterio es la base de inclusión de diferentes protocolos terapéuticos adaptados al grado de extensión loco-regional o a distancia y también permite establecer un pronóstico evolutivo.

Tabla 2. Sistema de estadificación internacional del cáncer de pulmón. Adaptado de (73)

Estadificación TNM (T: tumor, N: nódulos, M: metástasis).			
T1	Tumor ≤3cm, sin afectación de la plura o del bronquio principal (T1a, T1b).		
T2	Tumor 3-7cm, o afectación del bronquio principal		
T3	Tumor >7cm,		
T4	Cualquier tumor con invasión del mediastino, del corazón, de los grande vasos, de la tráquea, etc.		
N0	Sin metástasis en los ganglios linfáticos regionales.		
N1	Afectación ganglionar hiliar o peribronquial homolateral.		
N2	Metástasis en ganglios linfáticos mediastínicos o subcarinales homolaterales		
N3	Metastasis a ganglios linfáticos mediastínicos o hiliares contralaterales.		
M0	Sin metástasis a distancia.		
M1	Metástasis a distancia (M1a, M1b)		
Estadificación			
Estadio IA	T1a, T1b	N0	M0
Estadio IB	T2a	N0	M0
Estadio IIA	T2b	N0	M0
	T1a, T1b, T2a	N1	M0
Estadio IIB	T2b	N1	M0
	T3	N0	M0
Estadio IIIA	T1, T2	N2	M0
	T3	N1, N2	M0
	T4	N0, N1	M0
Estadio IIIB	Cualquier T	N3	M0
	T4	N2, N3	M0
Estadio IV	Cualquier T	Cualquier N	M1a, M1b

Actualmente la mejor forma de combatir el cáncer sigue siendo la prevención primaria o detección precoz. Se identifican dos factores claves que explican la alta mortalidad de esta neoplasia: la detección en estadios avanzados de la enfermedad y una amplia heterogeneidad histológica y molecular de los tumores. En los últimos años se han realizado grandes esfuerzos por revertir esta situación, enfocándose en la búsqueda de nuevos biomarcadores que permitan una detección más temprana y la identificación de nuevos blancos terapéuticos y de respuesta a los tratamientos.

Múltiples mutaciones somáticas se asocian a la transformación maligna en cáncer de pulmón, pérdida de heterocigosidad, translocaciones, deleciones y mutaciones puntuales. Tales mutaciones son utilizadas como marcadores para la clasificación histológica; para el pronóstico y seguimiento de la enfermedad y para la

elección de terapias específicas (77). Conocer el perfil mutacional de cada tumor se ha convertido en uno de los puntos centrales en el diagnóstico de cáncer de pulmón debido a que actualmente existen terapias diseñadas para tratar tumores con perfiles mutacionales específicos, al punto que pueden ser efectivas para determinada mutación y no responder con otra (78).

Un claro ejemplo son las terapias que utilizan como estrategia la inhibición de proteínas con dominios tirosin quinasa, las cuales participan en vías de proliferación, supervivencia y diferenciación celular. KRAS, EGFR y ALK, son proteínas claves para las terapias contra dominios tirosin quinasa (TKI) (78,79). Un 30% de los adenocarcinomas presentan mutaciones en KRAS, las cuales son resistentes a los tratamientos con TKI; mientras otro 25-30% presenta mutaciones en los exones 18 al 21 de EGFR y puede predecir resistencia o una buena respuesta, dependiendo de la mutación. De los pacientes que son tratados con TKI contra EGFR y que reincide el tumor, un 50% presenta mutaciones con resistencia en KRAS, mostrando una gran inestabilidad genómica en el tumor. Por otro lado, de aquellos pacientes con adenocarcinoma que no tienen mutaciones en KRAS ni EGFR, un 3-5% presentan la translocación EML4/ALK (78,79). Además de las terapias contra dominios tirosin quinasa, se ha reconocido una amplia gama de genes que participan en vías de proliferación, diferenciación y supervivencia de las células que podrían ser potenciales blancos terapéuticos o marcadores de respuesta a terapias específicas.

Por otra parte, los sARNs, mencionados en la sección 1.1.1, presentan una expresión tejido específica y cambios en los perfiles de expresión entre células normales y tumorales han sido reportados en diversos tumores hace más de 10 años (16). La disminución en la expresión del miARN let-7 en el tumor se observa en el 60% de los casos de cáncer de pulmón y se asocia a un mal pronóstico, debido a que normalmente actúa como supresor tumoral de varios oncogenes como KRAS, NRAS y HMGA2. A su vez miARNs como miR-34, miR-486 y miR-146a, cuya expresión también se ve disminuida, son predictores de un mal pronóstico. En sentido contrario, miR-21 cuyo mensajero blanco es PTEN, aparece sobre expresado en varios tumores sólidos, también asociado a un mal pronóstico (80).

Durante los últimos 5 años, una serie de reportes sugieren un gran potencial de los ARNs circulantes como nuevos marcadores en el diagnóstico y despistaje precoz del cáncer, como indicadores pronóstico-evolutivos y en el monitoreo de la eficacia terapéutica de las drogas antineoplásicas (63,71). Los transcritos de ARNm de genes vinculados con el proceso maligno han sido detectados en plasma o suero de pacientes con cáncer de mama (81). En estos casos, los niveles de ARNm de CCND1 (Ciclina D1) identifican a aquellos pacientes con diferente sobrevida global y con diferente respuesta terapéutica a la quimioterapia (82). Adicionalmente, los niveles de transcritos circulantes de TERT (del inglés, *telomerase reverse transcriptase*) y del receptor del factor de crecimiento epidérmico en pacientes con cáncer de pulmón se correlacionaron con el tamaño del tumor, la presencia de metástasis y mal pronóstico (83). Por otra parte, el miR-122 ha sido propuesto como un blanco terapéutico en el tratamiento de hepatitis C ya que es utilizado como cebador por el virus HCV para replicar su genoma. Actualmente el fármaco anti-miR-122 (Miravirsen) se encuentra en fase II de ensayos clínicos (84).

Existen varios marcadores proteicos circulantes en la rutina oncológica, pero su número y utilidad clínica es limitada (70,85). Los ácidos nucleicos circulantes permitirían una "biopsia líquida" de fácil acceso a lo largo de toda la evolución aun cuando ya se haya resecado el tumor o no se detecte la presencia de metástasis o recidivas del tumor. Debido a que su costo es bajo, el procedimiento es mínimamente invasivo y a que tiene gran potencial diagnóstico, es indudable que es ventajoso profundizar en el análisis de su capacidad diagnóstica.

2 Hipótesis y Objetivos

Hipótesis:

sARNs diferencialmente expresados por adenocarcinomas pulmonares, podrían ser detectables en sangre periférica y comportarse como potenciales biomarcadores de cáncer de pulmón.

Objetivos:

- Identificar sARNs diferencialmente expresados entre tejido tumoral y peri tumoral sano de pacientes con adenocarcinoma pulmonar mediante tecnologías de secuenciado profundo.
- Validación de los candidatos seleccionados por PCR cuantitativa en los tejidos secuenciados.
- Puesta a punto de RT-qPCR a partir de ARN de suero/plasma y detección de candidatos.
- Aproximación al estudio de la fracción exosomal plasmática como fuente de biomarcadores.

3 Materiales y métodos

3.1 Voluntarios sanos y pacientes.

En el presente estudio se incluyeron 3 pacientes con diagnóstico de adenocarcinoma pulmonar localmente avanzado (Estadio IIIa-IIIb) con indicación quirúrgica. En todos ellos se tomaron muestras de tejido tumoral y del margen quirúrgico de tejido sano peri tumoral en el momento de la cirugía (almacenadas a -80°C en RNAlater de Qiagen). Previo a la cirugía, se tomaron muestras de sangre venosa periférica.

También se obtuvo sangre de 3 voluntarios clínicamente sanos con previa firma de consentimiento informado. En todos los casos, se firmó previamente consentimiento informado de acuerdo a la normativa vigente y aprobado por el comité de ética del Hospital de Clínicas "Manuel Quintela".

De manera convencional se identificaron las muestras como Px (x, es el número de paciente), TN corresponde a los tejidos peri tumorales sanos y TT a los tumorales.

3.2 Obtención de plasma y suero de sangre periférica.

Para la obtención de plasma, se colectaron 15-20ml de sangre en tubos heparinizados a razón de 75 UI (Unidades Internacionales) por ml de sangre. La sangre fue centrifugada a 800g durante 20min a temperatura ambiente. Se extrajo el sobrenadante y se centrifugó nuevamente 15min a 2.000g a temperatura ambiente con el fin de eliminar restos celulares y/o remanentes de elementos formes de la sangre. A partir de los 20ml de sangre inicialmente utilizados se obtuvieron entre 9 y 10ml de plasma. El plasma se almacenó a -80°C en alícuotas de 0,5 ml. Cada alícuota descongelada previamente a ser utilizada se centrifugó a 2000g durante 5min a 4°C.

Para la obtención de suero se colectaron 10ml de sangre total en tubo seco y se dejó coagular por 30min a 37°C. Luego se centrifugaron las muestras a 1.500g durante 15min a temperatura ambiente. A continuación se extrajo el suero evitando el arrastre de la fase inferior, obteniéndose un promedio de 4-5ml de suero por cada 10 ml de sangre total.

3.3 Aislamiento de pequeños ARN a partir de tejido.

Se utilizaron 50-100mg de tejido triturado finamente con bisturí, los cuales se homogeneizaron en 500µl de Trizol con el homogeneizador "Pellet Pestle, Dispensable" (Sigma-Aldrich). A continuación, se agregó 500µl más de Trizol y se realizó la extracción de ARN de acuerdo a las instrucciones del fabricante. Se evaluó la pureza y se cuantificó con Nanodrop (NanoDrop® ND-1000, Thermo Fisher Scientific). El ARN se conservó en 100µl de agua ultrapura libre de RNAsas a -80°C. A partir del ARN total extraído se purificaron los pequeños ARNs utilizando el kit para purificación de pequeños ARNs, mirVana (mirVana™ miRNA Isolation kit, Ambion, Thermo Fisher Scientific). Luego se evaluó la calidad y el largo del ARN utilizando electroforesis capilar (Bioanalyzer2100, Agilent Technologies), con el kit para sARNs (RNA 6000 small kit, Agilent Technologies).

3.4 Aislamiento de pequeños ARN de plasma y suero.

Se realizó la extracción de ARN a partir de 250µl de plasma previamente tratado con heparinasa I (Sigma Aldrich) y 250µl de suero. Se utilizó Trizol LS (Sigma Aldrich) y seguidamente el kit para extracción de pequeños ARNs, mirVana (mirVana™ miRNA Isolation kit, Ambion, Thermo Fisher Scientific), según instrucciones del fabricante.

3.4.1 Tratamiento con heparinasa I.

La heparina es un potente inhibidor de la reacción de amplificación con polimerasas termoestables y debido a sus propiedades físico-químicas persiste como contaminante en los procedimientos de purificación de ARN. Por tal motivo, se utilizó el tratamiento del plasma con Heparinasa I de Flavobacterium (Sigma-Aldrich) a diferentes tiempos de incubación y a una concentración de 0.22U/µl. Alternativamente, el tratamiento con heparinasa I se realizó sobre 30µl de ARN previamente purificado a partir de 250µl de plasma, con 16µl de heparinasa I y 1µl de RNasin (Inhibidor de Ribonucleasas, Promega).

3.5 Secuenciado de pequeños ARNs de tejido tumoral y peritumoral.

Construcción de librerías y secuenciado.

El armado de librerías consiste en preparar el ARN para ser secuenciado. Se utilizó el kit NEBNext Multiplex Small RNA Library Prep Set for Illumina (Set 1) (NewEngland, Biolabs). En la figura 6 se esquematiza el flujo de trabajo utilizado. Inicialmente, se ligaron los adaptadores 3' y 5' al ARN para continuar con la retrotranscripción, de manera de obtener el ADN copia (ADNc). Finalmente, se realizó una PCR de 15 ciclos para amplificar el ADNc y ligar los fragmentos P5 y P7 que están incluidos en los *primers* de la PCR. P5 y P7 servirán de cebadores en las reacciones de "amplificación en puente" durante el secuenciado con tecnología Illumina.

Se partió de 500ng de ARN previamente purificado de tejido tumoral y peritumoral sano de 3 pacientes (P3, P4 y P5). Se continuó con el protocolo tal cual lo indica el fabricante y se chequeó la calidad de las librerías por electroforesis capilar (Bioanalyzer2100, kit DNA High Sensitivity).

Se prepararon las diluciones para el secuenciado y se cuantificaron las librerías por Qubit (Invitrogen). Luego se prepararon las diluciones según el protocolo de Illumina (MiSeq Illumina) y se secuenció en la plataforma de biología molecular del Instituto Pasteur de Montevideo con el secuenciador MiSeq. A continuación se muestra un esquema que muestra los pasos principales en la elaboración de las librerías (Figura 6).

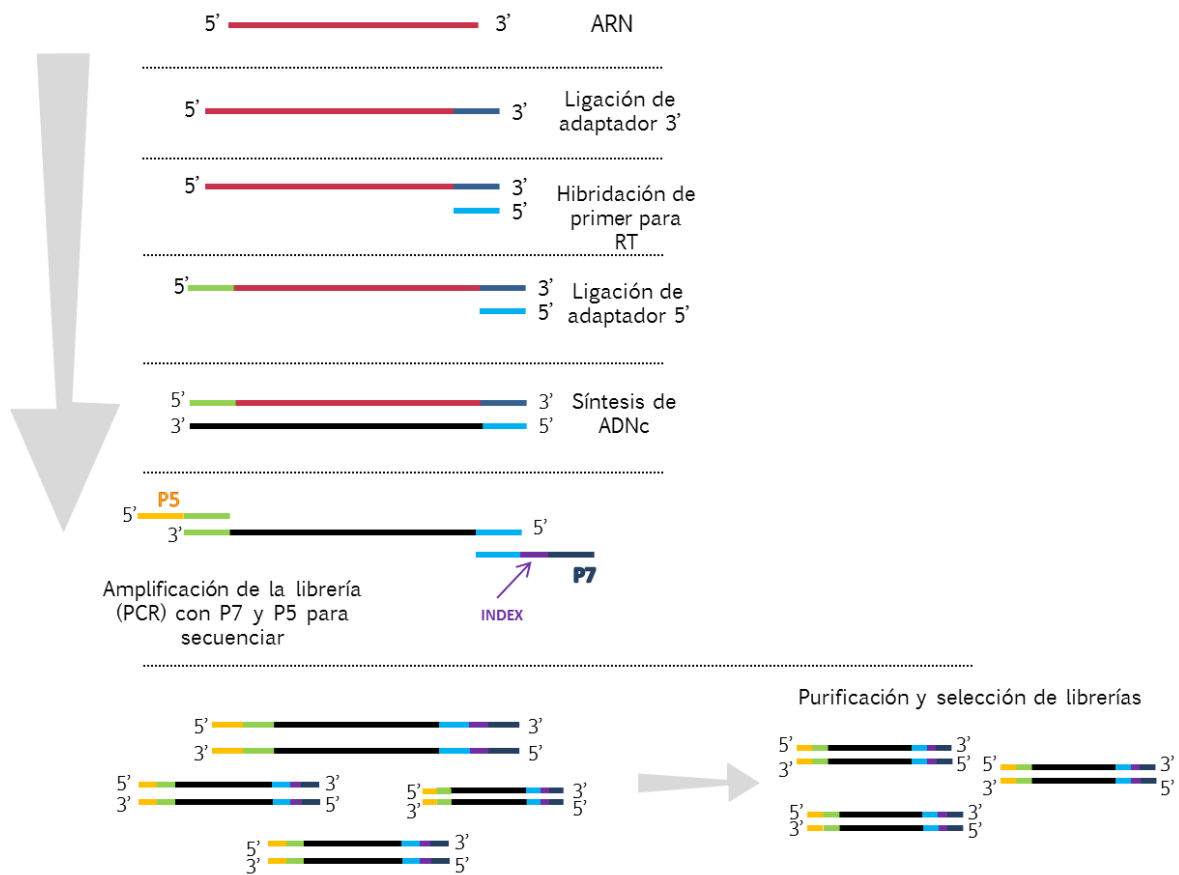


Figura 6. Esquema de construcción de librerías. Al ARN extraído de los tejidos se le liga el adaptador 3', luego se hibrida el *primer* 5' para la retro transcripción (RT) y se liga el adaptador 5'. Con el *primer* unido al ARN se hace la RT, sintetizando el ADN copia (ADNc). A continuación se amplifica el ADNc y en los *primers* se agrega P5 y P7. Como último paso, se purifica la librería seleccionando los fragmentos del tamaño de interés.

3.6 Análisis bioinformático del secuenciado.

El análisis del secuenciado se realizó en colaboración con la Unidad de Bioinformática del Instituto Pasteur de Montevideo y se tomó como referencia el artículo de Farazi et al. 2012 (86) para el análisis de todas las muestras. Se realizó el control de calidad de las librerías secuenciadas utilizando FastQC (87), se eliminaron las secuencias correspondientes a los adaptadores y se realizaron los controles de calidad correspondientes, filtrando las secuencias repetitivas y las secuencias menores a 16nt (87–90). Se colapsaron las secuencias únicas usando fastx_collapser de FASTX-Toolkit y se mapearon contra el genoma humano (Dec. 2013 (GRCh38/hg38) del UCSC Genome FTP

(<http://hgdownload.soe.ucsc.edu/goldenPath/hg38/bigZips/hg38.chromFa.tar.gz>)(91).

Se realizó el alineamiento utilizando el software bowtie permitiendo un único *mismatch* (92) y las secuencias que no mapearon con el genoma, fueron mapeadas contra bases de datos específicas de pequeños ARNs (detalladas más adelante). Este último mapeo se realizó utilizando el software Bowtie permitiendo hasta 3 *mismatches*.

Para el alineamiento y la clasificación de sARNs, en algunos casos surge la problemática de que algunas secuencias mapean en más de un tipo de ARN. Esto dificulta la designación de la categoría a la cual pertenece cada secuencia. Por tal motivo, primero se asignaron aquellos que mapearon contra un único gen. Luego, aquellos que mapearon en más de un sitio en el genoma se los clasificó utilizando una lista jerárquica de acuerdo a su abundancia dentro de la célula, y se les asignó una única categoría. El orden de abundancia dentro de la célula es el siguiente: rRNA, tRNA, snRNA, snoRNA, guideRNA, miRNA, miscRNA (vaultRNA, YRNA, srpRNA), telomerase RNA, piRNA, long non-coding RNA, repeat-masked: LINE, SINE, LTR, satélite y pseudogenes.

Para realizar el conteo por categoría se diseñaron *scripts* y se generaron tablas de conteo ya que existen ARNs que son codificados por genes que tienen varias copias (por ejemplo los ARN de transferencia); aquellas secuencias que mapearon con más de un gen dentro de una misma categoría —con la máxima calidad de mapeo— se contaron sumando $1/N$, siendo N el número de sitios al que mapeo.

Finalmente, se realizó el análisis de expresión diferencial, utilizando edgeR package y se seleccionaron candidatos a biomarcadores para ser validados por qPCR, primero en las muestras de los tejidos secuenciados y, luego, evaluarlos en plasma, suero y exosomas (93).

3.7 Validación de candidatos por *stem-loop* RT-qPCR.

Una de las técnicas utilizadas para la validación de los sARNs elegidos como candidatos de expresión diferencial en las muestras analizadas fue la RT-qPCR utilizando la técnica de "*stem-loop*". Esta técnica es específica para sARNs y consiste en utilizar un fragmento de ADN altamente específico (*stem-loop*) que hibridará con el ARN de interés. En la retro transcripción se utiliza como *primer* el *stem-loop* que tiene en su extremo 3', 6 bases complementarias al extremo 3' del sARN que queremos

amplificar. De esta manera se sintetiza un fragmento de ADNc con nuestro fragmento de interés más el *stem-loop* ligado (Figura 7). A continuación para la qPCR se utiliza un *primer forward* específico de nuestro ARN pequeño y un *primer* reverso "universal" que hibrida con el *stem-loop* (94).

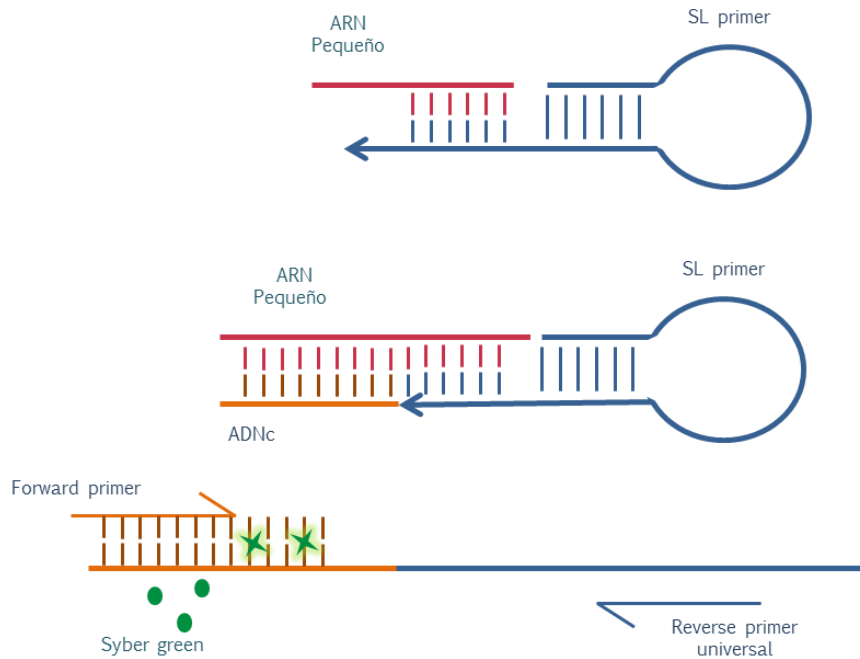


Figura 7. Esquema de estrategia experimental utilizando *stem-loop* RT-qPCR. Se utilizó syber-green como fluoróforo.

Se realizaron tres réplicas técnicas para cada réplica biológica y se utilizaron dos miARNs como normalizadores para evaluar la RT-qPCR en las muestras de tejidos. Las secuencias de los *primers* utilizados se muestran en la tabla 3.

Para la retro-transcripción de ARN de tejido se partió de 30-50ng de ARN. Para suero y el plasma, debido a la baja concentración, no se pudo cuantificar el ARN, de manera que se extrajo ARN siempre del mismo volumen de plasma/suero (ver sección 3.4) y se resuspendió el ARN también en el mismo volumen. Se utilizó 2µl de ARN molde para la retro transcripción para todas las reacciones analizadas, utilizando los mismos *stem-loop primers* que para las muestras de tejido, diseñados específicamente para cada sARN seleccionado.

Se utilizó la retrotranscriptasa SuperScript II, reverse transcriptase (TermoFisherScientific) y el termociclador Veriti 96 wells de AppliedBiosystems. Se realizó la qPCR en el equipo Eco de Illumina utilizando como fluoróforo Syber-green, que se une inespecíficamente a ADN doble hebra. Para la puesta a punto de la técnica y prueba de los *primers* diseñados se realizaron y chequearon las curvas de “melting” para descartar la formación de dímeros de *primers* o productos secundarios.

A continuación (tabla 4 y 5) se muestran las condiciones en que se realizó la retro transcripción y la qPCR respectivamente.

Tabla 3. Primers diseñados para *stem-loop* qPCR.

sARN	<i>Stem-loop primer</i>	Forward <i>primer</i> para qPCR
Glu_TTC	GTC GTA TCC AGT GCA GGG TCC GAG GTA TTC GCA CTG GAT ACG ACA GCG CC	TCC CTG GTG GTC TAG TGG TTA GG
His_GTG	GTC GTA TCC AGT GCA GGG TCC GAG GTA TTC GCA CTG GAT ACG ACA ACG CA	GCT GGC CGT GAT CGT ATA GTG GT
Gln_CTG	GTC GTA TCC AGT GCA GGG TCC GAG GTA TTC GCA CTG GAT ACG ACA GTC CA	CGA GGT TCC ATG GTG TAA TGG TTA GC
Gly_GCC	GTC GTA TCC AGT GCA GGG TCC GAG GTA TTC GCA CTG GAT ACG ACG GCG AG	GGG CAT TGG TGG TTC AGT GGT A
hsa-miR-148a-3p	GTC GTA TCC AGT GCA GGG TCC GAG GTA TTC GCA CTG GAT ACG ACA CAA AG	CAG CCA TCA GTG CAC TAC AGA AC
hsa-miR-21-5p	GTC GTA TCC AGT GCA GGG TCC GAG GTA TTC GCA CTG GAT ACG ACT CAA CA	CGC GCA TAG CTT ATC AGA CTG A
hsa-miR-100-5p	GTC GTA TCC AGT GCA GGG TCC GAG GTA TTC GCA CTG GAT ACG ACC ACA AG	CAC GCA AAC CCG TAG ATC CGA
hsa-miR-151a-3p	GTC GTA TCC AGT GCA GGG TCC GAG GTA TTC GCA CTG GAT ACG ACC CTC AA	CAC GCA CTA GAC TGA AGC TCC
hsa-miR-10a-5p	GTC GTA TCC AGT GCA GGG TCC GAG GTA TTC GCA CTG GAT ACG ACC ACA AA	CAC GCA TAC CCT GTA GAT CCG
Reverso universal		GTG CAG GGT CCG AGG T

Tabla 4. Condiciones de ciclado para la retro transcripción.

Temperatura (°C)	Tiempo
16	30''
42	30''
85	5''

Tabla 5. Condiciones de ciclado para qPCR.

Temperatura (°C)	Tiempo	Repeticiones
95	10''	1
95	15''	40
60	1''	

3.8 Validación de candidatos con *primers* LNA.

Los *primers* LNA (Locked Nucleic Acid) fueron diseñados originalmente para la amplificación de miARNs. Estos *primers* contienen un enlace entre los átomos 2'-O y 4'-C que bloquea la conformación característica del anillo de la ribosa. Este enlace hace más estables estos *primers*, aumentando significativamente las temperaturas de *annealing*.

Se diseñaron *primers* LNA para verificar el aumento de Gly_GCC y de Glu_TTC en suero y en los tejidos tumorales y peri tumorales (ver tabla 6). Se realizaron 3 réplicas técnicas para cada réplica biológica.

Para la retro-transcripción de ARN de tejido se utilizaron las mismas condiciones que para el *stem-loop* pero utilizando como cebadores "*random primers*" (SO142, Thermo Fisher Scientific). El hecho de que dichas reacciones nunca sobrepasan los 65°C, elimina la posibilidad de amplificar a partir de ARNt nativos. Se utilizó la retrotranscriptasa SuperScript II, reverse transcriptase (TermoFisherScientific) y el termociclador Veriti 96 wells de AppliedBiosystems.

A continuación, se realizó la qPCR en las mismas condiciones que para el *stem-loop*, variando la temperatura de *annealing* de acuerdo a los *primers* LNA diseñados.

Abajo en las tablas 7 y 8, se muestra las condiciones en que se realizó la retro transcripción y la qPCR respectivamente.

Tabla 6. *Primers* LNA diseñados para qPCR. (+) base anterior bloqueada.

sARNt	Forward <i>Primer</i>	Reverse <i>Primer</i>
Gly_GCC	G+CATT+GG+TGG	G+GC+GA+GAA+TT+C
Glu_TTC	TCC +CTG +GT+GG	A+GC+GCC+GAAT

Tabla 7. Condiciones de ciclado para la retro transcripción.

Temperatura (°C)	Tiempo
25	12''
42	50''
70	15''

Tabla 8. Condiciones de ciclado para qPCR.

Temperatura (°C)	Tiempo	Repeticiones
95	10''	1
95	15''	40
58	1''	

3.9 Obtención de exosomas.

3.9.1 Extracción de exosomas de plasma.

Para la extracción de exosomas se utilizaron 250µl de plasma y se diluyeron las muestras en 11ml de buffer fosfato salino (PBS). Luego las siguientes centrifugaciones se realizaron en la ultracentrífuga Optima™ XPN, Beckman-Coulter, utilizando el rotor SW 40 Ti (tubos de 14ml). Se centrifugó el plasma diluido en PBS a 16.000g, 30min a 4°C. En esta centrifugación se descartaron vesículas mayores a 150nm. El sobrenadante se centrifugó a 150.000g durante 12hs a 4°C. De esta centrifugación se guardó el sobrenadante (fracción ribonucleoproteica) a -80°C y se re suspendió el pellet (exosomas) en 11ml de PBS y se volvió a centrifugar a 150.000g durante 2hs a 4°C. El pellet de exosomas se re suspendió en 50µl de buffer de lisis (Urea 8M, SDS 2,5%, inhibidor de proteasa (Complete ULTRA, EDTA free de Roche), para hacer el análisis de proteínas por western blot. Para hacer el análisis por microscopia electrónica se re suspendió el pellet en 15µl de PBS filtrado.

3.9.2 Extracción de exosomas de sobrenadante de cultivo celular.

Como control positivo para el western blot, se extrajeron exosomas a partir de 20ml de sobrenadante de cultivo celular de la línea de cáncer de mama, MCF 7, ATCC HTB-22. Estas células fueron cultivadas 48hs en medio condicionado sin suero, para evitar el aporte de exosomas del suero. Al igual que para el plasma se utilizó la ultra centrifuga Beckman-Coulter y el mismo rotor. Por motivos de volumen se separó la muestra en dos tubos y se centrifugó a 16.000g durante 30min a 4°C. Se descartó el pellet y se centrifugó el sobrenadante a 100.000g a 4°C durante 2hs 30min. Se guardó el sobrenadante (fracción ribo proteica) a -80°C. Se realizó un lavado de los exosomas con PBS. Previamente se juntaron los pellet en un solo tubo y se llevó a 11ml con PBS nuevamente. Se centrifugó a 100.000g durante 2hs a 4°C, se re suspendió en 50µl de buffer de lisis y se guardó a -20°C. Los exosomas utilizados para la microscopía

electrónica no fueron congelados, sino que se guardaron 24 horas a 4°C y se observaron por microscopía electrónica de transmisión.

3.9.3 Microscopía Electrónica de Transmisión (MET).

La visualización de los exosomas se realizó por MET con tinción negativa. Se fijaron los exosomas con Glutaraldeído 2,5% en un soporte Grill Nickel con film de Carbon de 200 Mesh. Se colocó 10µl de exosomas fijados previamente, sobre la rejilla durante 30min, se secó la gota con papel absorbente y se le agregó 10µl de Acetato de Uranilo 2% durante 30seg. Se secó la gota con papel absorbente y se dejó secar unos minutos más.

Las láminas fueron preparadas y observadas en el microscopio electrónico de transmisión, JEOL JEM-1010 (Facultad de Ciencias, Universidad de la República). Las imágenes fueron analizadas y editadas con el programa de edición de imágenes Fiji, Image J.

3.9.4 Cuantificación proteica de exosomas.

Para la cuantificación proteica de exosomas se utilizó BCA (bicinchoninic acid, Sigma Aldrich). Se realizó la curva de calibración con albúmina bovina (BSA) y se siguió el protocolo como sugiere el fabricante. Se prepararon 3 diluciones para cada muestra y se realizaron los cálculos para sembrar la misma cantidad de proteína para los western blot de CD9 y HSP70.

3.9.5 Detección de CD9 y HSP70 por Western Blot.

Para la detección de CD9 y de HSP70 se evaluaron a partir de 2 volúmenes de plasma: una extracción de exosomas a partir de 250µl de plasma y otra a partir de 4 alícuotas de 250µl de plasma cada una, que se poolearon en el lavado con PBS durante la extracción de exosomas (1ml de plasma en total). Los pellet de exosomas se re suspendieron en el buffer de lisis, en el mismo volumen y se utilizaron 30µg de extracto proteico del pellet de la alícuota de 250µl de plasma y el mismo volumen de la alícuota de exosomas de 1ml de plasma.

Las muestras fueron preparadas en buffer de carga reductor (LaemmliSample Buffer: 50mM Tris-HCl pH 6,8, 2% SDS, 10% glicerol, 1% β-mercaptoethanol, 12,5mM

EDTA, 0,02% bromophenolblue). Se calentaron a 95°C durante 5 minutos y se sometieron a electroforesis en gel de poli-acrilamida 12%, en buffer Tris-glicina-SDS. Las proteínas se transfirieron a una membrana de PVDF (Amersham Hybond-P PVDF Membrane, GE Healthcare Life Sciences) mediante transferencia líquida a 40V durante 14 horas a 9°C, en buffer tris-glicina (buffer de transferencia: 25mM Tris-Cl pH 8,3, 192mM glicina, 0,03% SDS, 10% etanol). Para chequear la eficiencia de la transferencia se tiñeron las membranas con solución Rojo Ponceau (0,1% Ponceau, 5% ácido acético) y los geles con azul de Coomassie. Las membranas se bloquearon con BSA al 5% en PBS-Tween 0,05% durante 1 hora a 37°C. Luego se lavaron con PBS-Tween 0,05% durante 5min, 3 veces.

En el caso de CD9 se continuó con la incubación del anticuerpo primario anti-CD9 (sc-13118, Santa Cruz Biotechnology, INC.) diluído 1/300 en BSA 0,5%, PBS-Tween 0,05. Se incubó durante toda la noche a 4°C y luego se lavó 3 veces durante 5 minutos en PBS-Tween 0,1%.

El anticuerpo primario HSP70 (sc-59569, Santa Cruz) se utilizó diluido 1/100 en BSA 1%, PBS-Tween 0,1%. Se incubó durante 17 horas a 4°C y luego se lavó 3 veces durante 5 minutos en PBS-Tween 0,1%.

Para la detección de ambos anticuerpos primarios se utilizó como anticuerpo secundario un anti-IgG de ratón conjugado a peroxidasa; diluído 1/10.000 en PBS-Tween 0,05%, durante 1 hora a temperatura ambiente en agitación. Se lavó 3 veces con PBS-Tween 0,1% durante 5 minutos cada vez.

A continuación, se incubaron las membranas con sustrato quimio-luminiscente para revelado con peroxidasa (ECL-Signal, ViVoScience) durante 1 minuto para luego exponerlas a film fotográfico (Amersham® Hyperfilm® ECL™, GE Healthcare).

4 Resultados

4.1 Construcción y caracterización de librerías de pequeños ARNs tisulares.

Para las muestras P3 y P5 se prepararon las librerías genéticas con sus respectivas réplicas técnicas para el tejido tumoral y peri tumoral. La cantidad de material de biopsia obtenido para la muestra P4 no permitió la realización de réplicas técnicas.

Los ARNs purificados de todas las muestras se cuantificaron por Qubit y se evaluó la distribución de tamaños mediante electroforesis capilar. En la figura 8, se muestra un resultado representativo del perfil de tamaños para las muestras P3 de tejido tumoral y peri tumoral. Se observa una distribución de tamaños variada entre 20-150pb, con un patrón que no corresponde a ARN degradado sino que coincide con una distribución de tamaños característica de fragmentos menores a 200pb. Los picos de 60-70 nt se corresponden con la migración de ARNt maduros, los fragmentos de 90-120nt a pre-ARNt, ARNr5S y eventualmente a transcriptos cortos. Los fragmentos en la región entre 20-40nt corresponden a sARNs, los cuales son el objetivo principal del presente trabajo. Debido a que son la fracción minoritaria no se observa un pico tan marcado como en el caso de las demás poblaciones. Las variaciones en la intensidad de los picos y en el perfil del electroferograma para la muestra tumoral, se debe a que hay un aumento muy marcado en los fragmentos de 60-70nt que podrían corresponder a ARNt, lo cual saca de escala los picos anteriormente mencionados. Esto podría estar reflejando un aumento en el metabolismo celular típico de las células tumorales.

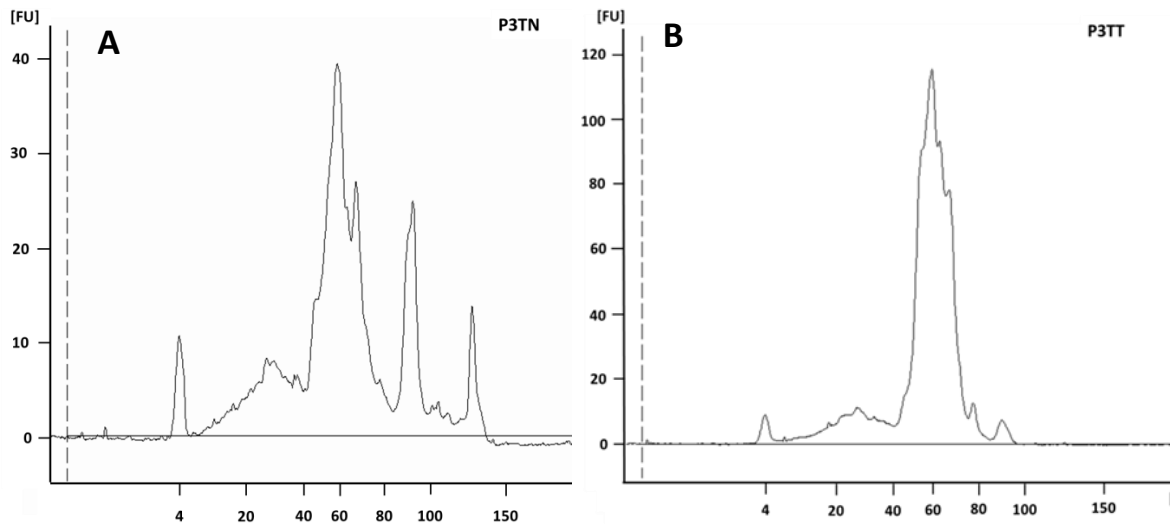


Figura 8. Electroferograma de sARNs de la muestra P3 obtenido por electroforesis capilar. (A) ARN de tejido peri tumoral de P3 (P3TN) y (B) ARN de tejido tumoral de P3 (P3TT). [FU] Unidades de Fluorescencia. [nt] nucleótidos.

Luego de evaluar la concentración y la calidad del ARN se construyeron las librerías para cada muestra de tejido. Como se muestra en la figura 9 se corrieron las librerías en un gel de poliacrilamida 6% para seleccionar los fragmentos de interés. Si bien nos interesan los fragmentos entre 18-40pb hay que considerar que las librerías tienen ligados adaptadores, lo que hace que nuestros fragmentos de interés sean de 130-190pb. Se marca entre dos líneas rojas la región donde realizamos los cortes en el gel que luego se chequearon por electroforesis capilar.

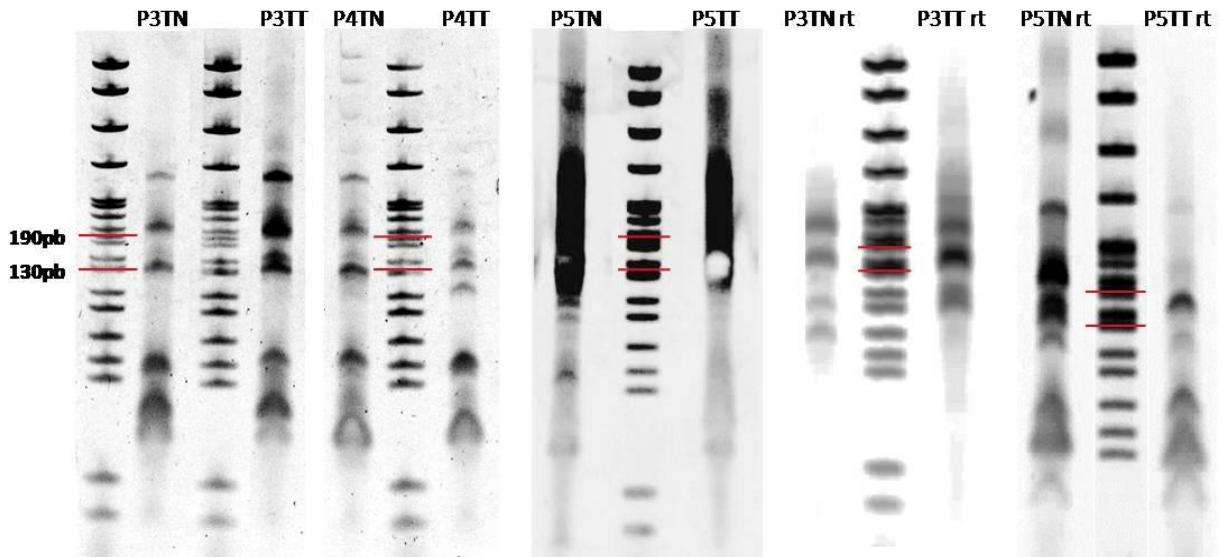


Figura 9. Librerías de ARN de tejido de P3, P4 y P5 para secuenciado masivo. Gel de poliacrilamida 6%. Región entre marcas rojas es la zona donde se cortaron los geles (130-190pb). TN: tejido peri tumoral TT: tejido tumoral. rt: réplica técnica.

A partir del gel se purificó el ARN y se cuantificó con Qubit (tabla 9). En la figura 10 se muestra a modo de ejemplo el perfil obtenido por electroforesis capilar para la muestra P5 (P5TN y P5TT), donde se observan bandas definidas correspondientes a fragmentos en el rango de 130 a 190pb.

Se prepararon las muestras con las diluciones correspondientes y se realizó el secuenciado en el secuenciador MiSeq de Illumina.

Tabla 9. Cuantificación de librerías de sARNs utilizando Qubit.

Muestra	Cuantificación Qubit (ng/μl)
P3TN	0,988
P3TT	5,23
P3TN rt	0,77
P3TT rt	1,66
P4TN	1,38
P4TT	0,742
P5TN	7,24
P5TT	20,8
P5TN rt	4,7
P5TT rt	0,738

P_TN: tejido peri tumoral de paciente x. P_TT: tejido tumoral de paciente x. rt: replica técnica.

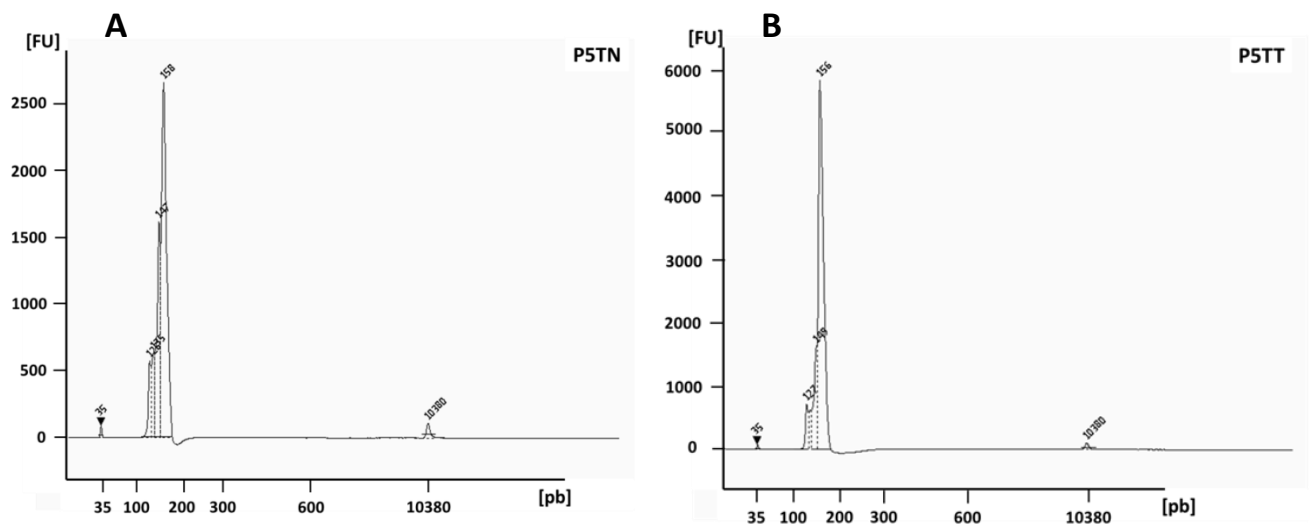


Figura 10. Electroferograma de librerías de sARNs de la muestra P5 obtenido por electroforesis capilar. (A) librerías de tejido peri tumoral (P5TN). (B) librerías de tejido tumoral (P5TT). Se muestran los picos del tamaño esperado, según los cortes realizados en la figura 9.

4.2 Análisis bioinformático del secuenciado.

A las secuencias obtenidas del secuenciado, se le realizaron los controles de calidad, se eliminaron los adaptadores y se clasificaron en categorías como se detalla en Materiales y Métodos. Se generaron tablas de manera de identificar los candidatos a

biomarcadores. Como se muestra en la tabla 10 la cantidad de secuencias crudas (secuencias sin filtrar generadas en el secuenciado) de cada muestra fue variable. Los porcentajes luego de eliminar los adaptadores y de aplicar los filtros de calidad fueron aceptables, exceptuando las muestras P3TT y P4TT que apenas pasaron el 50%, es decir que del total de las secuencias obtenidas menos del 50% fue aceptable para continuar con el análisis.

Tabla 10. Resumen del número de secuencias obtenidas por secuenciación masiva. Valores obtenidos durante el procesamiento de las mismas.

	Secuencias crudas	% de secuencias depuradas	Luego del filtro (% de sec. procesadas)	Secuencias Colapsadas
P3TN	4.119.488	85,2	74,4	257.853
P3TT	1.405.358	58,9	54,1	57.121
P3TNrt	1.384.885	98,6	87,6	84.512
P3TTrt	610.535	98,4	90,2	51.015
P4TN	4.727.611	92,5	82,0	279.651
P4TT	2.659.558	95,2	62,0	136.117
P5TN	1.611.144	95,6	72,2	179.789
P5TT	2.018.570	98,4	80,4	111.159
P5TNrt	5.234.073	98,5	77,3	313.827
P5TTrt	3.652.588	99,3	73,4	132.238

rt: replica técnica

La distribución por tamaño de las secuencias obtenidas para cada muestra se muestra gráficamente en la figura 11. En la misma se identifican claramente dos poblaciones principales, una entre 21-23nt y otra entre 31-34nt. En P3 y P4 se observa la población mayoritaria entre 21-23nt y para la muestra P5 se distinguen ambas poblaciones con proporciones similares, estando mayormente representada la población entre 31-35nt en las de tejido tumoral.

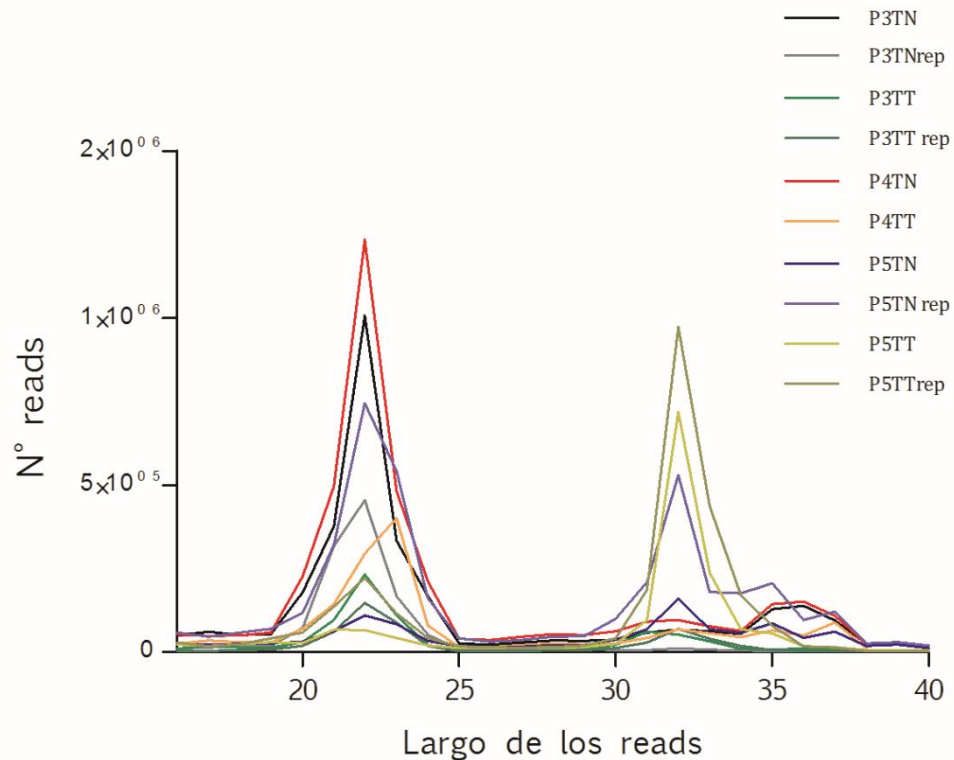


Figura 11. Representación gráfica de la distribución de tamaños de fragmentos secuenciados de sARN de tejidos tumorales y peri tumorales de los 3 pacientes.

En la figura 12 se muestra la distribución de las secuencias en las diferentes categorías funcionales. Como era de esperar en muestras tisulares, los más representados son los miARNs en P3 y P4 con un porcentaje que ronda el 60%, seguido de sARNr y sARNt con un 15-20% y 10-20% respectivamente. En las muestras P5 los sARNs que aparecen con mayor porcentaje son los sARNt con un 30% en el tejido peri tumoral y un 70% en tumoral, los sARNr con 20% en peri tumoral y 10% en tumoral, seguidos de los miARNs con un 30% en peri tumoral y un 10% en tumoral.

Las diferencias observadas entre las muestras P5 (P5TN y P5TT) y las correspondientes a P3 y P4 (P3TN, P3TT, P4TN y P4TT) podrían reflejar diferencias en sus perfiles moleculares no identificables por histopatología. En este sentido es útil recordar que es común observar una variabilidad propia de los tumores que reflejan la gran heterogeneidad de subtipos dentro de grupos clásicos convencionales según el estudio anatómo-patológico.

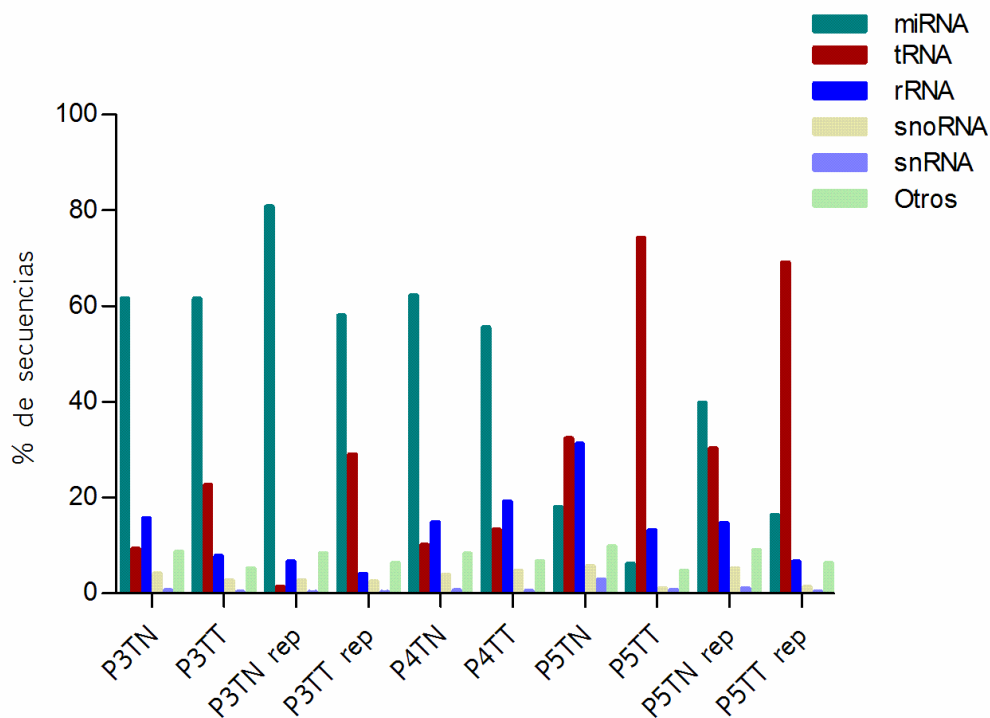


Figura 12. Clasificación de sARNs por categoría. Otros: piARN (piwiARNs), miscARN (ARNs misceláneos: ARNs citoplasmáticos, YARNs), lncARN (ARNs largos no codificantes, 200nt aprox.).

Para el análisis nos enfocamos específicamente en los miARNs y los sARNt como grupos mayoritarios, los cuales representan en conjunto más del 90% de sARNs secuenciados en todas las muestras. En el caso de los miARNs, es mejor conocida su función e implicancia en la progresión del cáncer en tejidos y también a nivel circulante en los fluidos corporales. En cuanto a los sARNt, han surgido recientemente como actores relevantes en la regulación de la expresión génica y respuestas a estrés celular independientemente de su rol conocido en la traducción (69).

En la figura 13 se muestra representada la distribución de tamaños de los miARNs, para cada muestra (P3, P4 y P5), presentando la misma distribución entre 20-23nt, con un pico en 22nt. Los sARNt presentan una población distribuida en un rango de 30-35nt, con una distribución diferencial entre el tejido peri tumoral y tumoral (figura 14). En los tejidos peri tumorales se observan dos picos, uno en 33nt y otro en 35nt, mientras que en los tumorales hay un corrimiento del pico de 35nt a un único

pico de 32nt. Esta diferencia se observa para las muestras P3 y P4 y en menor proporción en la P5. Sin embargo si analizamos la relación de las lecturas por millón entre los picos de 32 y 35nt en las tres muestras, en el tejido peri tumoral la relación 32/35 es menor que en el tumoral para todas las muestras. Siendo la población de 32nt mayor en relación a la de 35nt en el tejido tumoral para las tres muestras analizadas (tabla 11).

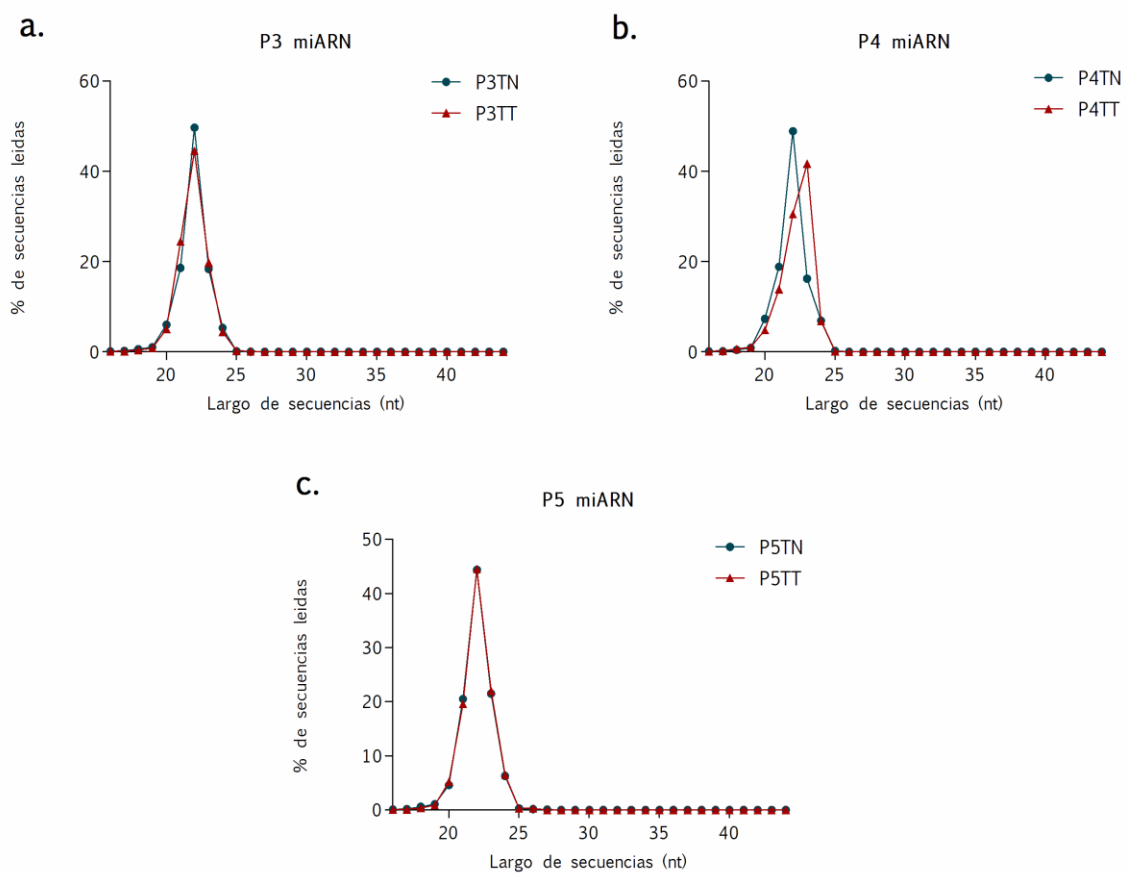


Figura 13. Representación gráfica de la distribución de tamaños de la categoría de miARN para las muestras P3 (a), P4 (b) y P5 (c).

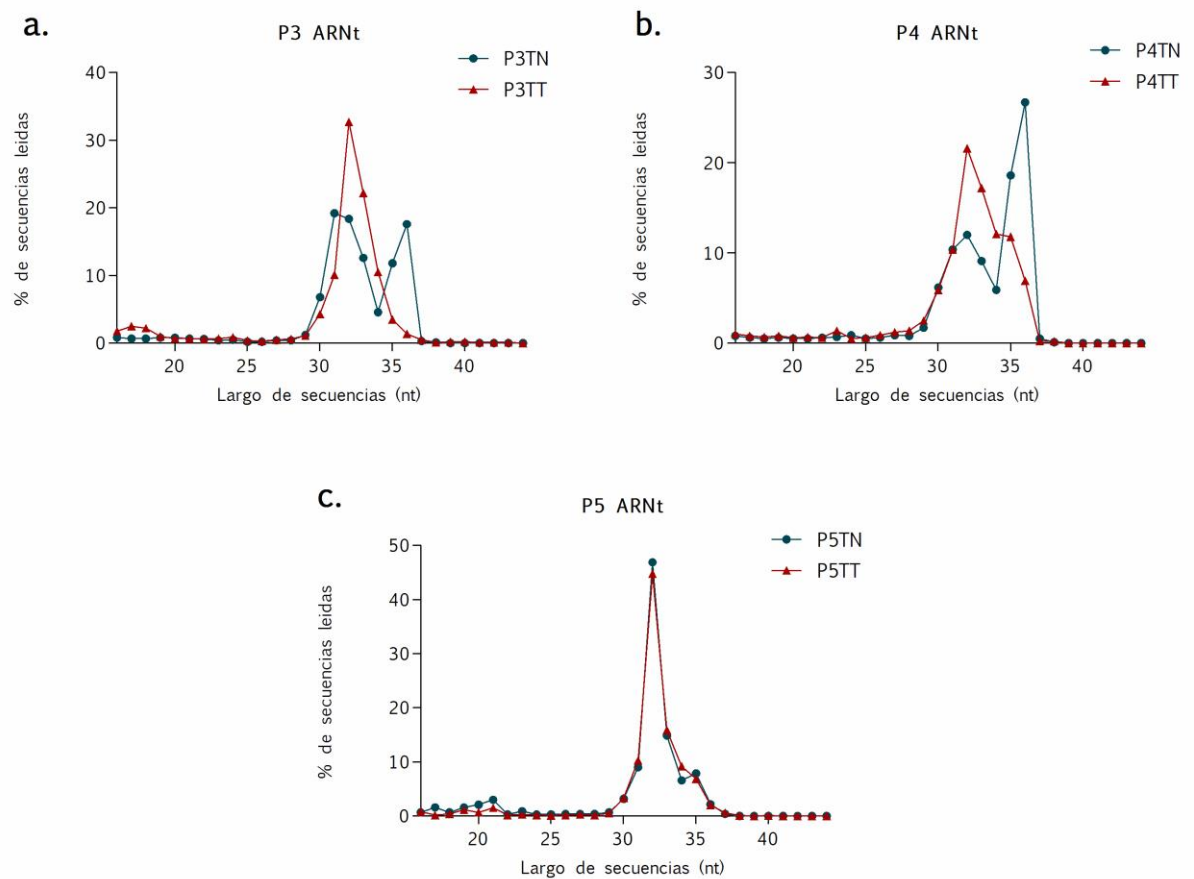


Figura 14. Representación gráfica de la distribución de tamaños para la categoría de sARNt para las muestras P3 (a), P4 (b) y P5 (c).

Tabla 11. Relación entre fragmentos de 32nt y 35nt de la categoría sARNt representados en la figura 14.

	Tejido peri tumoral	Tejido tumoral
P3	2,0	13,6
P4	0,6	1,8
P5	5,7	6,3

4.3 Selección de pequeños ARNs diferencialmente expresados entre tejido tumoral y peri tumoral.

Con la idea de identificar sARNs diferencialmente expresados y que eventualmente tengan valor como biomarcadores se analizaron las tablas de expresión

para los miARNs y los sARNt. Se seleccionaron aquellos candidatos que tuviesen una variación absoluta mayor a ± 1.5 veces entre ambos tipos de muestras y que tuviesen los mayores valores de lecturas por millón. La selección de los miARNs y sARNt se basó en que son la población mayormente representada a nivel tisular, lo que asociado a la elección de los que poseen mayor frecuencia, aumenta las probabilidades de detección en fluidos biológicos. En las tablas 12 y 13 se muestran resultados en azul, los sARNs seleccionados que cumplieron con estos criterios (4 sARNt y 3 miARNs). A fines de ser posteriormente validados por RT-qPCR, se seleccionaron 2 miARNs cuya expresión no variara significativamente entre tumor y tejido peri tumoral para ser utilizados como normalizadores en los 3 pacientes. Se detallan las secuencias de los candidatos seleccionados en la tabla 14.

Tabla 12. Lista de sARNt. Fragmentos entre 30-34nt. LPM: lecturas por millón.

	LPM									
	P3TN	P3TT	P3TN rt	P3TT rt	P4TN	P4TT	P5TN	P5TT	P5TN rt	P5TT rt
Glu_CTC_5half	43536	31134	2703	41694	44330	41029	25583	256862	22815	125090
Gly_GCC_5half	13842	152604	4643	131961	22815	36256	92226	691806	133754	521001
Glu_TTC_5half	2505	9921	478	11875	2741	6520	5530	26585	5393	11185
Glu_CTC_misc	2110	1616	290	1484	2735	4084	2383	9820	1310	2631
Gly_CCC_5half	2059	31360	699	25912	3884	6171	13452	154619	21434	131257
Lys_CTT_5half	1862	2554	490	14048	1971	7339	8934	47455	16982	67258
His_GTG_misc	1526	9194	334	6123	1149	2517	7554	61748	11767	48807
Pseudo_TTC_5half	1283	14919	265	14667	2486	4695	3755	48035	4740	67784
His_GTG_5half	1208	12406	258	19312	1286	1743	5263	79297	7363	181885
Gly_GCC_misc	725	2853	261	2245	1344	2646	2561	16496	2319	8242
Val_CAC_5half	690	761	46	2156	1294	614	3696	21158	5757	37537
Pseudo_CTC_misc	246	199	43	93	271	316	190	1154	128	262
Lys_CTT_misc	216	759	155	430	337	518	443	2948	627	1938
Gln_CTG_misc	214	873	253	156	228	410	57	110	160	154
Gln_CTG_5half	8	244	0	71	13	31	5	12	5	60

Resultado en azul: candidatos seleccionados.

Tabla 13. Lista de miARNs. Fragmentos entre 20-23nt. LPM: lecturas por millón.

	LPM									
	P3TN	P3TT	P3TN rt	P3TT rt	P4TN	P4TT	P5TN	P5TT	P5TN rt	P5TT rt
miR-143-3p	104392	30834	76314	36700	92722	36743	2876	3350	11706	10617
miR-148a-3p	74396	118655	66071	18957	57489	45551	3755	2754	7543	3462
miR-21-5p	46936	199763	55815	208636	55806	268202	36512	7869	122185	21536
miR-101-3p	20982	4920	9941	4717	13628	4090	257	286	1289	1129
miR-26a-5p	17405	13500	44835	26645	29738	11918	4384	7310	12178	34816
miR-100-5p	17348	194791	21182	82015	17173	5812	1527	3148	6208	9037
miR-27b-3p	12926	9421	11063	6501	16546	10445	1175	1034	4798	2169
miR-451a	3980	212	131364	808	4366	132	492	9640	1473	17532
miR-27a-3p	2212	1374	4663	1988	3865	1926	452	706	1244	927
miR-30b-5p	2097	523	2381	1455	2507	277	323	209	1283	783
miR-200b-3p	1542	4724	2378	2416	2126	5690	187	398	523	1025
miR-199b-5p	1402	724	1561	1627	1742	984	152	75	708	237

Resultado en azul: candidatos seleccionados.

Tabla 14. Secuencia de candidatos seleccionados a validar a partir del análisis de expresión diferencial del secuenciado masivo de tejidos tumorales y peri tumorales.

sARNs candidatos	Secuencia más representada
Glu TTC	TCC CTG GTG GTC TAG TGG TTA GGA TTC GGC GCT
His GTG	GCT GGC CGT GAT CGT ATA GTG GTT AGT ACT CTG CGT T
Gln CTG	GGT TCC ATG GTG TAA TGG TTA GCA CTC TGG ACT
GlyGCC	GCA TTG GTG GTT CAG TGG TAG AAT TCT CGC C
hsa-miR-148a-3p	UCA GUG CAC UAC AGA ACU UUG U
hsa-miR-21-5p	UAG CUU AUC AGA CUG AUG UUG A
hsa-miR-100-5p	AAC CCG UAG AUC CGA ACU UGU G
hsa-miR-151a-3p	CTA GAC TGA AGC TCC TTG AGG
Hsa-miR-10a-5p	TAC CCT GTA GAT CCG AAT TTG TG

4.4 Validación de candidatos seleccionados por PCR cuantitativa basada en *stem-loop* RT-qPCR.

Como se detalla en Materiales y Métodos una de las técnicas utilizada fue el *stem-loop* RT-qPCR a modo de validar el secuenciado. Si bien es la técnica más recomendada y utilizada para la cuantificación de miARNs, una limitante de la misma es el diseño de cebadores. Esto es debido a que los extremos 3'-OH de muchos sARNs y

especialmente en los sARNt presentan variaciones en su punto de corte en ± 3 nt. Los *primers* de *stem-loop* se hibridan únicamente en las últimas 6 bases del extremo 3' del *stem-loop primer* y admite sólo una variante de extremo 3' OH. Por ello, para el diseño de cada *stem-loop primer* se tuvo en cuenta el pool de variantes de extremos 3'OH para cada categoría (miARN y sARNt) eligiéndose la secuencia más representada, para diseñar los mismos. Para el análisis por RT-qPCR de muestras intracelulares clásicamente se utilizan genes constitutivos como el pequeño ARN nuclear U6 (SNU6). En nuestras muestras se intentó realizar el análisis utilizando el mismo como normalizador pero se observó que presentaba diferencias cuantitativas entre el tumor y el tejido peri tumoral. Debido a este hecho decidimos utilizar dos sARNs sin variación significativa entre todas las muestras como normalizadores, el miR-151a-3p y miR-10a-3p. Como se detalla en la tabla 15 se chequearon mediante RT-qPCR, verificando a través de los $\Delta\Delta Ct$ que funcionan como normalizadores, ya que los valores son cercanos a cero. Posteriormente, se decidió continuar utilizando como normalizador el miR-151a-3p.

Tabla 15. Análisis de normalizadores por RT-qPCR.

	Ct miR-151a-3p	Ct miR-10a-3p	$\Delta\Delta Ct$
P3TN	30,51 \pm 0,26	29,27 \pm 0,28	
P3TT	28,46 \pm 0,36	27,54 \pm 0,20	-0,32
P4TN	30,14 \pm 0,16	29,81 \pm 0,24	
P4TT	27,65 \pm 0,32	27,77 \pm 0,39	-0,45
P5TN	30,42 \pm 0,21	30,38 \pm 0,15	
P5TT	27,98 \pm 0,17	27,57 \pm 0,39	0,38

De manera de poder comparar los datos obtenidos en el secuenciado y los de la PCR cuantitativa se elaboró un gráfico para cada categoría (Figuras 15 y 16). Del lado izquierdo de los gráficos está representada la variación cuantitativa ("foldchange") de cada sARN candidato en el secuenciado y en el panel derecho, se muestran los $-\Delta\Delta Ct$ de la RT-qPCR. Los datos son mostrados en este formato teniendo en cuenta que los obtenidos en el secuenciado son cuantitativos y los obtenidos de la RT-qPCR son

relativos a la expresión de un normalizador. De esta forma se podrá evaluar el aumento y/o disminución de los sARNs seleccionados en las muestras P3, P4 y P5.

Dentro de la categoría de los miARNs se evaluaron miR-21-5p, miR-100a y miR-148-3p. En el gráfico de la figura 15 se observa un aumento en la expresión de miR-100a y miR-21-5p en el tejido tumoral respecto al peri tumoral, verificándose tanto en el secuenciado como en la RT-qPCR. A su vez, miR-148-3p presenta una disminución en su expresión en el tejido tumoral respecto al peri tumoral.

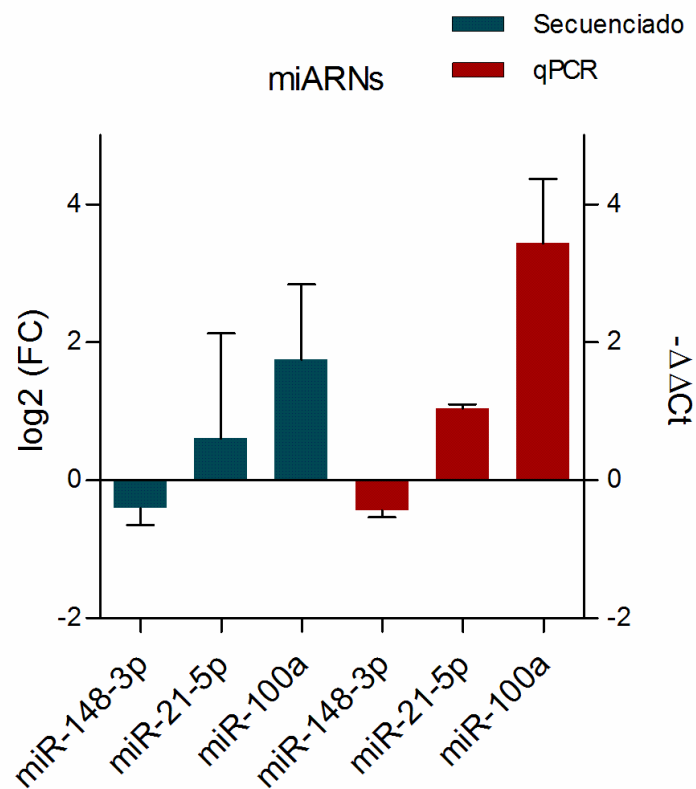


Figura 15. Representación gráfica de la variación de miARNs de tejido tumoral respecto al peri tumoral, medido por secuenciado masivo y por RT-qPCR. El cambio en la expresión para el secuenciado se muestra en azul como el $\log_2 (TT/TN)$ y en rojo el *stem-loop* RT-qPCR como $-\Delta\Delta Ct$.

Con respecto a los sARNt (figura 16A), se evaluaron 4 candidatos: Glu_TTC, Gly_GCC, His_GTG y Gln_CTG. Se observó que Gly_GCC y His_GTG aumentan en el tejido tumoral respecto al peri tumoral, según los datos del secuenciado y de la *stem-loop* RT-qPCR (Figura 16A). Por otro lado Gln_CTG y Glu_TTC muestran un aumento en la expresión en el secuenciado y una disminución en la expresión evaluada por *stem-loop* RT-qPCR. En estos últimos 2 casos el *stem-loop* RT-qPCR no permitió validar los mismos como candidatos. Las eventuales explicaciones a este hecho serán discutidas y analizadas más adelante.

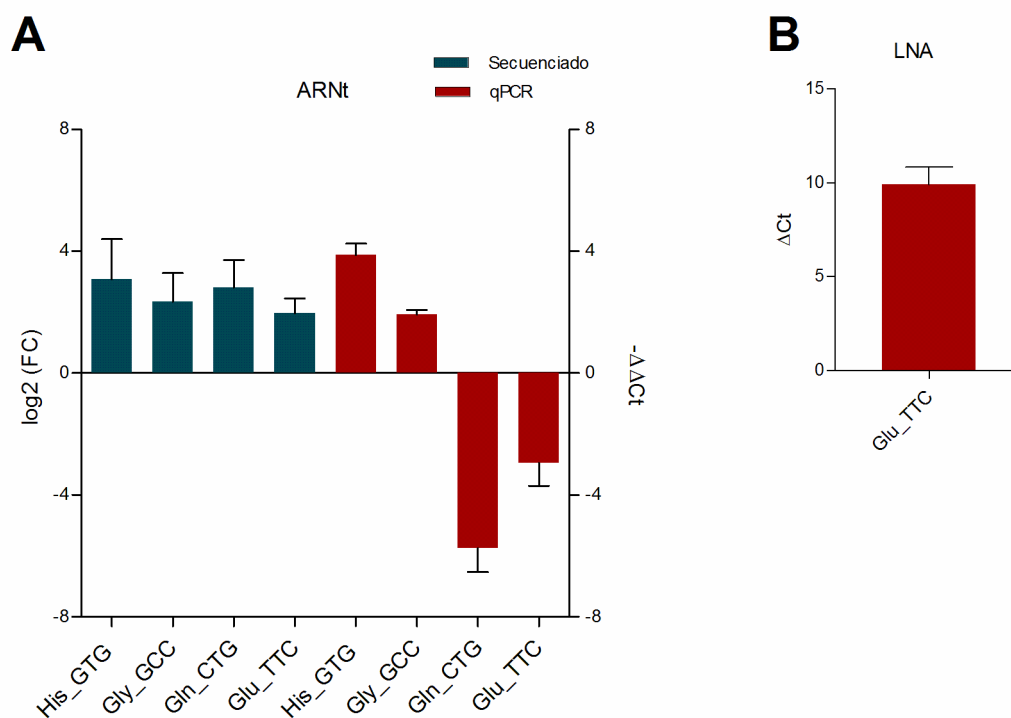


Figura 16. Representación gráfica de la variación de sARNt del tejido tumoral respecto al peri tumoral. (A) El cambio en la expresión en el secuenciado se muestra en azul como el \log_2 (TT/TN) y en rojo el *stem-loop* RT-qPCR como $-\Delta\Delta C_t$. (B) Representación gráfica de la variación de expresión de Glu_TTC por RT-qPCR utilizando LNA *primers* entre tejido peri tumoral y tumoral.

4.5 Análisis de expresión diferencial con *primer* LNA.

Una de las limitantes que presenta la metodología de "*stem-loop* RT-qPCR" es que no admite variaciones en los extremos 3' de los sARNs blanco. Según los datos del secuenciado, en la categoría de sARNt, existe una variación en el largo entre 31-35nt, sobretodo en el extremo 3'. Si bien se diseñaron cebadores teniendo en cuenta la

secuencia más representada no se logran amplificar aquellas para las cuales no se diseñó su cebador específico. Para verificar que efectivamente se trata de una limitante de la técnica se utilizaron *primers* LNA, los cuales permiten amplificar secuencias de fragmentos de ARNt independientemente de las variaciones en sus extremos.

A modo de validar la RT-qPCR utilizando *primers* LNA se evaluó Gly_GCC, verificando nuevamente su aumento en tejido tumoral respecto al peri tumoral. Paralelamente, se realizó la RT-qPCR con *primers* LNA para Glu_TTC (figura 16B), donde se observa una mayor expresión de Glu_TTC en el tejido tumoral respecto al peri tumoral. Esto sugiere que la discordancia observada entre el secuenciado y la validación por el método de *stem-loop* RT-qPCR se debe, en gran parte, a las diferentes variantes dadas por su terminación 3'. Esto podría explicar los casos contradictorios que existen en la literatura en cuanto a las variaciones cuantitativas de los sARNt según los autores, y marca la necesidad de sistematizar los métodos más adecuados para la determinación cuantitativa sobretodo de los sARNt.

4.6 Análisis de plasma y suero como fuente de ARNs circulantes de interés como biomarcadores.

Uno de los objetivos secundarios de este trabajo fue avanzar en la posibilidad de que los sARNs expresados pudiesen ser identificados en fluidos biológicos como el plasma, debido a que varias evidencias experimentales recientes, han sugerido su potencial como biopsias líquidas en diversas patologías humanas incluyendo el cáncer (69).

Es importante mencionar que tanto el plasma como el suero poseen un alto contenido proteico y de otras moléculas biológicas capaces de interferir con el aislamiento y/o amplificación del ARN/ADN asociado a una escasa concentración de los mismos. Por ello fue necesario analizar en profundidad muchos aspectos metodológicos para que los mismos puedan ser utilizados como una fuente viable y reproducible de determinación de ARNs circulantes con fines diagnósticos.

Respecto a la metodología utilizada no hay que perder de vista que para la extracción de ARN de muestras líquidas no es posible manipular grandes volúmenes, debido a la naturaleza de las muestras. Por tal motivo, se obtienen bajas

concentraciones de ARN circulante lo cual dificulta su cuantificación. Para remediar esta situación, tanto para plasma como para suero se tuvo especial cuidado en utilizar siempre los mismos volúmenes tanto en la cantidad inicial de plasma/suero, como en los volúmenes de elución del ARN y el material de partida en la RT-qPCR.

En primer lugar se optimizó la extracción de ARN de plasma utilizando Trizol LS y un kit de extracción de sARNs. A fines de evitar la fuerte inhibición de las polimerasas termoestables por anticoagulantes, especialmente la heparina utilizada en este trabajo, se realizó el tratamiento previo del plasma con heparinasa I en diferentes condiciones (diferentes concentraciones de heparinasa I y tiempos de incubación). Se extrajo el ARN como se describió en Materiales y Métodos de los tres pacientes y tres donantes sanos y se probaron los candidatos seleccionados por *stem-loop* RT-qPCR. Si bien se probaron varios métodos de purificación de ARN para eliminar posibles contaminantes y se realizaron diferentes tratamientos con heparinasa I, no se logró amplificar ninguno de los candidatos probados ni los controles a partir de plasma sanguíneo con heparina como anticoagulante.

Alternativamente, se evaluaron por *stem-loop* RT-qPCR tres sueros normales (sueros control) y los tres sueros de pacientes. Se probaron todos los candidatos mostrados en la tabla 14 y en la figura 17A se muestran los valores de "Ct" obtenidos de aquellos que se lograron amplificar (miR-21-5p, miR148-3p, Gly_GCC y Glu_TTC). En el caso de los 2 miARNs detectados, no se observan diferencias significativas entre los pacientes y los sueros control. Era de esperar que la presencia de miR148-3p en sangre no se viera aumentada, teniendo en cuenta la disminución en los tejidos secuenciados (figura 15). En cuanto a los sARNt, se observa un aumento significativo de Gly_GCC y Glu_TTC en el suero de los pacientes lo cual coincide con la variación observada a nivel del tejido tumoral. En el caso de Gly_GCC coincide con lo observado en los datos de la figura 16 del secuenciado y de la RT-qPCR con el mismo "stem loop". Sin embargo, Glu_TTC se corresponde únicamente con los datos del secuenciado y la RT-qPCR utilizando *primers* LNA, ya que ese candidato no se pudo validar por RT-qPCR en tejido mediante técnicas de *stem-loop*. Esto sugiere que, contrariamente a la heterogeneidad de los extremos 3' a nivel tisular, la especie mayoritaria de Glu_TTC en forma circulante es aquella con extremos 3' amplificables similarmente por *stem-loop* RT-qPCR.

A modo de validar lo observado en los tejidos con los *primers* LNA, se realizó la RT-qPCR con dichos *primers* (figura 17B) observando un aumento significativo de Glu_TTC en suero de los pacientes respecto a los sueros normales.

El valor potencial de estos dos sARNt (Glu_TTC y Gly_GCC) como marcadores deberá validarse en el futuro sobre una población de pacientes cuyo número permita evaluar adecuadamente el valor diagnóstico predictivo.

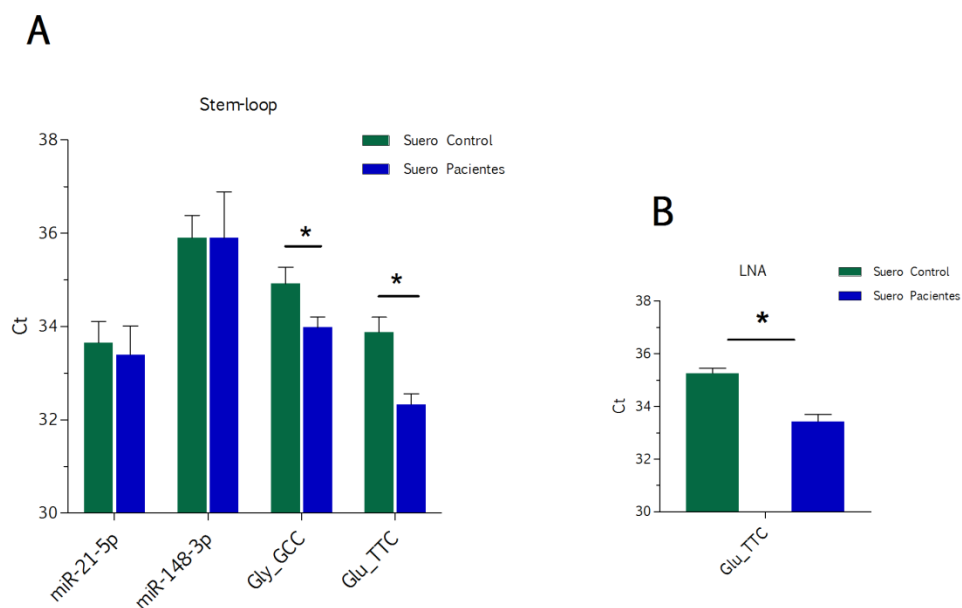


Figura 17. Resultados de RT-qPCR de sARNs seleccionados como candidatos en suero. (A) Gráfico de *stem-loop* RT-qPCR. (B) Gráfico de RT-qPCR utilizando *primers* LNA. Ct: cycle threshold.

4.7 Búsqueda de pequeños ARNs en exosomas.

Hoy en día existen diversos reportes de que los ARNs circulantes se encuentran asociados a complejos ribonucleoproteicos y a vesículas extracelulares. Se ha reportado que los pacientes oncológicos presentan un aumento y perfiles diferentes de sARNs circulantes, sin embargo no existe un consenso en cuanto a si son liberados selectivamente al medio, en que fracciones son vehiculizados e incluso donde son procesados. En trabajos realizados en nuestro laboratorio en sobrenadante de cultivos celulares hemos reportado que una fracción de miARNs (fracción minoritaria del total

de los ARN circulantes) está presente en la fracción exosomal y que la fracción mayoritaria son sARNt, unidos a complejos ribonucleoproteicos libre de vesículas (40). En este contexto nos propusimos ver si nuestros candidatos se encuentran en la fracción exosomal en exosomas aislados de plasma sanguíneo.

Se probaron diversas condiciones de RT-qPCR de los candidatos seleccionados a partir de ARN de exosomas de plasma y no se logró detectar ninguno de ellos. Con estos resultados nos preguntamos si a diferencia de sobrenadante de cultivo, la obtención de exosomas de plasma requería modificaciones en el protocolo o si efectivamente no estaban presentes nuestros candidatos en los exosomas sanguíneos.

4.7.1 Obtención de exosomas

Para la extracción de exosomas se tuvieron varias consideraciones: en principio se tomó como referencia el protocolo ya utilizado en nuestro laboratorio para la extracción de exosomas de sobrenadante de cultivo celular. Respecto al plasma hubo que considerar que contiene más exosomas que el sobrenadante de cultivo y que su composición es diferente, conteniendo mayor proporción de proteínas y lípidos. Teniendo en cuenta los aspectos mencionados se puso a punto la extracción de exosomas a partir de plasma sanguíneo.

Actualmente existen varias técnicas estándar que permiten la visualización y/o detección de exosomas. En este trabajo se utilizó microscopía electrónica de transmisión (MET) y western blot.

Para la MET se utilizó tinción negativa con acetato de Uranilo (2%). Tal como esperábamos, en la figura 18 se observan tamaños variables de exosomas y una tinción no homogénea entre las diferentes imágenes. Hay que tener en cuenta que el fijado y la tinción de las muestras pueden modificar los tamaños (disminuyendo y/o aumentando respectivamente) pudiéndose observar de un tamaño mayor o incluso menor al esperado. Se observan vesículas entre 30-150nm. En cuanto a la forma y estructura, en las figuras 18A-D se puede apreciar las membranas rodeadas de un halo negro que corresponde a precipitados que se forman sobre las vesículas durante la preparación de la muestra. Es por esta razón que no siempre se ven iguales, observándose en algunas imágenes el exosoma negro. En la figura 18B se puede ver

incluso una doble membrana, reflejando una de las características de los exosomas, la doble membrana lipídica.

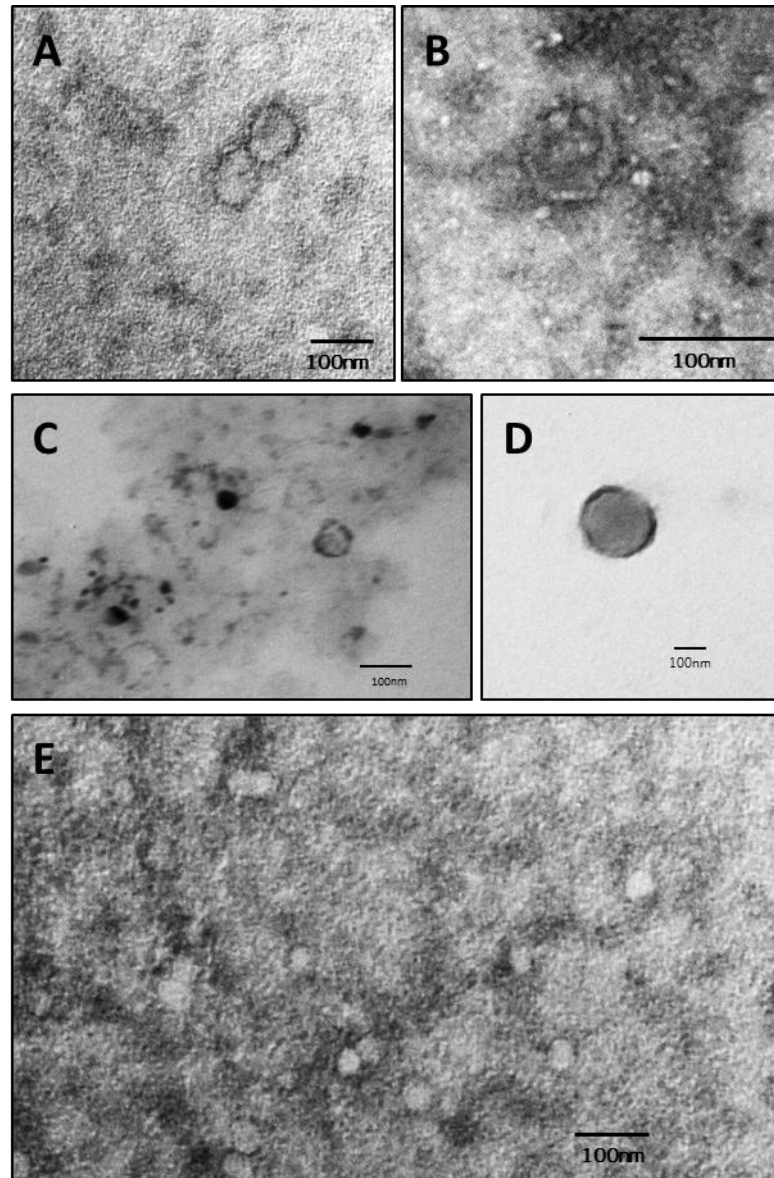


Figura 18. MET de exosomas de plasma obtenidos por ultra centrifugación.

Utilizando tinción negativa, se observó (A) con un aumento de 150.000x, (B) con un aumento de 120.000x, (C) y (E) a 100.000x y (D) a 80000x.

Otra técnica ampliamente utilizada para la validación de extracción de exosomas es el western blot. Existe una variedad de proteínas que se utilizan como marcadores exosomales, CD9, CD63, *lamp-1*, HSP70, *caveolina-1*, entre otras.

Se probaron varios anticuerpos, en la figura 19 se muestran los Western blot de CD9 y HSP70. Para ambas proteínas se probaron dos volúmenes diferentes de plasma de donde se extrajeron los exosomas: 250µl y de 1ml de plasma. Para este ensayo se re suspendieron los dos pellet de exosomas en el mismo volumen de buffer de lisis. Se puede apreciar las bandas del peso molecular esperado, tanto para CD9 como para HSP70. Se observa que en el carril 2 de ambos western blot se ve aumentada la banda tal como se esperaba ya que el volumen inicial de plasma para extraer exosomas era 4 veces mayor que en el carril 1.

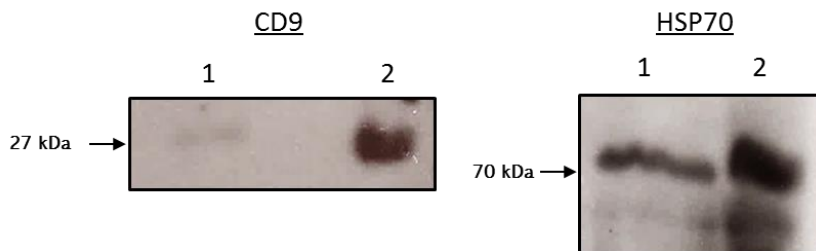


Figura 19. Western blot de CD9y HSP70 de exosomas de plasma. Se sembraron 30µg de extracto proteico de exosomas de 250µl de plasma (carril 1) y el mismo volumen de extracto proteico de exosomas de 1ml de plasma (carril 2).

Mediante estas dos técnicas se demostró la presencia de exosomas a partir de plasma sanguíneo, descartando la posibilidad de que las RT-qPCR no hayan amplificado debido a la ausencia de exosomas. Como técnica complementaria se podría agregar ensayos en nanosight, de manera de hacer una lectura más fina de la concentración y tamaños de las vesículas evaluadas.

5 Discusión

A pesar de los grandes esfuerzos por combatir las altas cifras de mortalidad por cáncer de pulmón, no se ha logrado identificar biomarcadores diagnósticos suficientemente sensibles y no invasivos que permitan una detección en estadios tempranos con el fin de mejorar los resultados terapéuticos. Es sabido que en la mayoría de los casos el cáncer de pulmón es diagnosticado cuando aparecen síntomas o signos clínicos sugestivos, los cuales aparecen muy tardíamente alejando las posibilidades de tratamientos curativos y comprometiendo severamente la sobrevida.

La descripción en los últimos años de nuevas familias de ARNnc, y su rol central en la regulación de la fisiología celular, los ha posicionado como grandes actores, no sólo en los procesos de transformación maligna, sino también como nuevos y potenciales blancos terapéuticos y biomarcadores diagnósticos en cáncer (95). En el caso particular del cáncer de pulmón, se ha sugerido el valor potencial de ciertos miARNs que, expresados en el tejido tumoral, podrían surgir como nuevos candidatos de utilidad como biomarcadores diagnósticos y pronósticos. En este sentido, se ha reportado que un aumento en la expresión de miR-146a parece tener una mayor sobrevida, sugiriendo un marcador de buen pronóstico ya que ha sido descrito como un inhibidor del crecimiento celular, de la migración e inducción de la apoptosis de células tumorales (96). También han sido reportados otros candidatos capaces de actuar como oncogenes o supresores tumorales, los cuales incluyen a miR-21, let-7, miR-29a, miR-34c, miR-205, miR-375, miR-125, entre otros (97).

Por otra parte, se ha demostrado la presencia de ARNs circulantes en diversos fluidos biológicos, los cuales son secretados por las células al medio extracelular ya sea en forma selectiva o no, y que son captados e internalizados funcionalmente intactos por células adyacentes o distantes (56,98). En el medio extracelular y fluidos biológicos, estos ARNs son transportados en vesículas o unidos a complejos de naturaleza proteica o lipoproteica (40,51,66,68). El aumento o la disminución de varios miARNs y ARNm circulantes se han asociado a diferentes neoplasias, proponiéndose diversos miARNs como candidatos a biomarcadores circulantes. Si bien para cáncer de pulmón no existe actualmente ningún ARN circulante validado para uso clínico, uno de los primeros y

más estudiados es el oncogén miR-21 que tiene como blanco PTEN. Altos niveles en suero de este miARN se asocia con mal pronóstico y estadios avanzados de la enfermedad (99,100). Sin embargo, los sARNt y los YARNs aparecen recientemente como una nueva clase de sARNs reguladores, cuyo poder predictivo como biomarcadores está siendo sugerido en base a recientes publicaciones. Reportes recientes de nuestro laboratorio y de otros grupos, muestran que tanto en sobrenadante de cultivo de líneas tumorales, así como en suero y plasma de sujetos sanos, la proporción de miARNs es minoritaria y que la mayor proporción de sARNs circulantes son sARNt e YARNs unidos a complejos ribonucleoproteicos. (40,67,68). En concordancia con estos hallazgos Chevillet y col., demuestran que hay menos de un miARN cada 100 exosomas, incluso del miARN más abundante (101), evidenciando que la mayor fracción de ARNs circulantes viaja libre de vesículas y que existe una población aún más representada que los miARNs que estaría participando en este eventual mecanismo de comunicación intercelular a distancia.

Debido a la alta diversidad y estabilidad que presentan los sARNs circulantes en sangre (51,52,102), la gran especificidad por sus blancos, la fácil accesibilidad de plasma o suero ("biopsias líquidas") y el rol central de los ARNnc en la iniciación y progresión del cáncer, es que actualmente se considera que los sARNnc circulantes podrían comportarse como eventuales candidatos a biomarcadores para diagnóstico, predicción de respuestas terapéuticas y como blancos terapéuticos en cáncer de pulmón (84,103).

Con el objetivo de poder aportar en este campo, nos propusimos identificar sARNs diferencialmente expresados en tejido tumoral y peri tumoral sano de tres pacientes con adenocarcinoma pulmonar y evaluar si dicha expresión diferencial podría reflejarse en suero/plasma. Como aproximación metodológica utilizamos secuenciado profundo sobre librerías de sARNs aislados y fraccionados a partir de tejido tumoral y peri tumoral.

De los datos obtenidos del secuenciado es interesante destacar que se identificaron dos poblaciones mayoritarias de sARNs, una entre 21-23nt y otra entre 31-35nt. Del análisis por categorías se identificaron en las muestras P3 y P4 aproximadamente un

60% de miARNs y un 10-20% de sARNt, mientras que en la muestra P5 un 30% de sARNt y de miARNs en el tejido peri tumoral y un 10% y 70% en el tejido tumoral. Las proporciones relativas de las distintas clases de sARNs obtenidas para las muestras P3 y P4 son las esperadas según la literatura, para muestras tisulares. Sin embargo en la muestra P5 se observan porcentajes de sARNt que se desvían de las proporciones esperadas. Si bien los tejidos fueron clasificados por histopatología como adenocarcinomas, existen diferencias moleculares no identificables. Esta variabilidad a nivel de la expresión de moléculas reguladoras refleja la heterogeneidad a nivel molecular de tumores que si bien, son homogéneos en el análisis histopatológico, los mismos reflejan una amplia diversidad de perfiles moleculares (79,97). Por otra parte no hay que perder de vista que se intentó trabajar con tejidos peri tumorales sanos. Sin embargo, existe la posibilidad que el tejido este afectado por el ambiente tumoral o mismo contaminado con células del propio tumor durante la propia biopsia. Sobre todo en la muestra P5, la cual muestra un perfil de expresión diferente tanto en el tumoral como en el peritumoral. A pesar de estas diferencias los sARNs que detectamos y seleccionamos como candidatos diferencialmente expresados entre tejido tumoral y peri tumoral variaron homogéneamente en las tres muestras tisulares.

De los datos obtenidos en el secuenciado se identificaron 7 sARNs como posibles candidatos a nivel tisular: miR-21-5p, miR-100a-5p, miR-148a-3p, Glu_TTC, His_GTG, Gln_CTG, Gly_GCC los cuales se validaron utilizando *stem-loop* RT-qPCR y *primers* LNA.

En concordancia con el secuenciado se verificó la disminución de miR-148a-3p y el aumento de miR-21-5p y miR-100-5p en el tejido tumoral respecto al tejido peri tumoral.

La desregulación de miR-21-5p no llama la atención ya que aparece desregulado frecuentemente en muchos tipos de tumores viéndose aumentada su expresión (80,97). Para cáncer de pulmón se ha descrito principalmente un aumento en la expresión de miR-21 que se asocia con un mal pronóstico (104). Trabajos recientes demuestran que la inhibición de miR-21 mediante la utilización de sARNs "esponjas" (son ARN circulares, complementarios a los miARNs, que actúan como "neutralizadores" del miARN sobreexpresado), reduce estos efectos (105). En cuanto a nuestros resultados, el

aumento de miR-21 en el tejido tumoral no nos llama la atención. Sin embargo, nos pareció interesante incluirlo ya que apoyo nuestros resultados en cierta forma. Por otra parte miR-148 aparece como un regulador a nivel epigenético, siendo uno de sus blancos el mensajero de la metiltransferasa Dnmt3b (DNA methyltransferase 3) (106). La sobreexpresión de esta proteína, cuya expresión se ve aumentada con la disminución de miR-148a, se ha asociado a cáncer hepático, de mama y de colon (107). En cáncer de pulmón se ha reportado su disminución tanto a nivel tisular como circulante proponiéndolo como un candidato a biomarcador y para diagnóstico (108,109). Esto iría en el mismo sentido de lo observado en nuestro trabajo. Por otra parte, miR-100-5p se ha encontrado en menor proporción que los candidatos anteriores, sobre expresado en ciertos tipos de neoplasias como, carcinomas hepáticos (110), en próstata (111) y en cáncer de ovario donde detectan una disminución de este miARN (112). Sin embargo, en cáncer de pulmón existen varios reportes que describen una baja en su expresión asociado con un mal pronóstico (113). A través del secuenciado de nuestras muestras observamos un aumento de miR-100-5p, lo cual es llamativo ya que no solo existen pocos reportes en cáncer de pulmón de este miARN sino que su expresión se ve disminuida. Es interesante mencionar al menos para los miARNs, que en todos los candidatos seleccionados, existen reportes que los describen tanto "up" como "down" regulados. Esto no llama la atención ya que los miARNs, en general, tienen varios blancos actuando tanto como oncogén y/o como supresores tumorales en diferentes tipos de tumores.

Respecto a los 4 sARNt seleccionados, se verificó por *stem-loop* RT-qPCR el aumento de His_GTG y Gly_GCC en el tejido tumoral respecto al peri tumoral. Sin embargo para Glu_TTC y Gln_CTG no se logró validar el aumento observado en el secuenciado utilizando esta técnica. Si bien el *stem-loop* es la técnica estándar de referencia ("gold-standard") para el análisis de miARNs por qPCR, existen ciertas limitaciones de la técnica, las cuales podrían estar explicando los resultados contradictorios en la literatura, incluyendo lo observado para Glu_TTC y Gln_CTG en nuestro caso. La principal limitante de esta técnica es que no admite ningún *mismatch* en las 6 bases del sARN que hibridan con el *stem-loop*, ni variaciones en el largo del extremo 3' del sARN. En contraparte, las mitades de ARNt analizadas, varían en su

tamaño entre 31-35nt, de manera que el estudio del pool de sARNt seleccionado no se adecuaría estrictamente a la técnica utilizada. Para verificar si efectivamente las variaciones observadas se restringían a las limitaciones del *stem-loop*, se realizó una RT-qPCR utilizando *primers* LNA. Esta técnica también es ampliamente utilizada para sARNs y si bien estos *primers* son "cortos", permiten amplificar fragmentos cuyos extremos posean diferencias en el largo de sus secuencias, de manera de amplificar casi por completo el "pool" de ARN de cada categoría (sARNt en nuestro caso). De la RT-qPCR realizada con *primers* LNA se logró verificar el aumento de Glu_TTC en tejido tumoral respecto al peri tumoral, demostrando que el *stem-loop* RT-qPCR no es la técnica que se adecua mejor para el estudio de sARNt ya que son muy variables en su largo.

Por otra parte no hay que perder de vista que los ARNt poseen muchas modificaciones postranscripcionales, lo cual dificulta enormemente el estudio de esta especie de ARNs. Desde la técnica utilizada en la preparación de las librerías hasta su análisis bioinformático son cuellos de botella a la hora de interpretar los resultados.

En cuanto a la búsqueda de los candidatos seleccionados en suero, plasma y exosomas, fue posible detectar ARN únicamente en suero. El plasma contiene grandes cantidades de inhibidores naturales como proteínas y lípidos (114). Y la heparina y otros anticoagulantes dejan rastros que dificultan su amplificación o bien la reproducibilidad de los resultados. Por otra parte los miARNs en exosomas, como se mencionó previamente, están presentes en muy bajas concentraciones (101), por lo que habría que procesar grandes cantidades de plasma o suero para realizar cuantificaciones tanto relativas como absolutas. En cuanto a lo informativo del contenido de este tipo de vesículas también es interesante mencionar que aún sigue en discusión que tan selectiva es la liberación de miARNs y otros ARNnc al medio extracelular (40,115).

La utilización de normalizadores en suero y plasma, es ampliamente discutida en los fluidos biológicos. En el caso del uso de los denominados "*spike-in*" (análogos sintéticos agregados exógenamente en cantidades conocidas) no existe consenso, ya que se discute su vida media en plasma o suero. Por otra parte los genes utilizados

generalmente como normalizadores a nivel intracelular (U6, globina, GAPDH, miR16 en suero) se ha mostrado que no están igualmente representados en sangre entre diferentes individuos teniendo cifras altamente variables (116). Hoy en día la técnica más adecuada para evaluar los ARNnc en fluidos biológicos es la PCR digital, la cual ha surgido recientemente como una técnica altamente sensible (117), permitiendo una cuantificación absoluta sin la necesidad de utilizar normalizadores. Dado que no contamos con el equipamiento para realizar dicha PCR, decidimos evaluar los candidatos en sangre sin utilizar normalizadores, optimizando los cuidados en la cantidad de tejido utilizado, la extracción del ARN, y los volúmenes utilizados para la RT-qPCR.

Se logró detectar miR-21-5p y miR-148a-3p sin observar diferencias significativas entre sueros normales y pacientes con adenocarcinoma pulmonar. Esta observación va en línea con lo observado en los tejidos secuenciados. Verificar un aumento de miR-148a-3p a nivel circulante estaría indicando que no sería aportado por el tumor, que como ya se comentó anteriormente gran parte de los ARN circulantes son aportados por plaquetas y células sanguíneas. En cuanto a los reportes de estos miARNs a nivel circulante por un lado se ha descrito un aumento de ambos miARNs en suero en pacientes con tumores en el timo (118). Por otra parte otros reportes indican en cáncer de pulmón, un aumento de miR-21 y una disminución de miR-148a-3p circulantes, postulándolos como predictores de respuesta a terapia y pronóstico (109,118).

De los sARNt analizados, se detectaron de manera aumentada en suero Gly_GCC y Glu_TTC. Actualmente existen pocos trabajos que asocien sARNt a cáncer en un contexto de ARNs circulantes. Se describieron inicialmente en varias especies como productos de respuesta a estrés, reclutados en gránulos de estrés (o p-bodies) y liberados al medio extracelular (36,44,119–122). Uno de los pioneros en esta temática fue Dhabbi y cols. quienes describen a las mitades de ARNt como potenciales moléculas señalizadoras en la respuesta inmune (47). Además este mismo grupo, propone a las mitades 5' de ARNt como potenciales biomarcadores en cáncer de mama, encontrando aumentadas las mitades cuyos isoaceptores son tRNA-Arg, -Asn, -

Cys, -Gln, -Gly, -Leu, -Ser, -Trp y -Val y disminuidas tRNA-Asp y -Lys en suero de pacientes (68).

Simultáneamente, en nuestro laboratorio se describieron como fragmentos mayoritarios *in vitro*, en sobrenadante de cultivo celular (MCF-7, HeLa y NCI-HI12299), las mitades 5' de ARNt-Gly y -Glu, no siendo representativo de lo observado a nivel intracelular (40). Estas mitades se encontraron mayoritariamente en la fracción ribonucleoproteica unidas a los complejos (40). Paralelamente *in vivo*, este trabajo con pacientes busco mostrar que tanto a nivel tisular como circulante, también se encuentran estos sARNt diferencialmente expresados. En particular en suero no logramos discriminar en que fracción se encuentran, pero si identificamos dos candidatos que aparecen aumentados en pacientes con adenocarcinoma pulmonar, ARNt-Gly y ARNt-Glu. Si bien no se trata de los mismos isoaceptores identificados en sobrenadante de cultivo, llama la atención que nuestros candidatos coincidan con los encontrados en sobrenadante de cultivo en líneas celulares derivadas de cáncer mamario. La acumulación de evidencias de la presencia de esta nueva clase de sARNs a nivel de ARN circulantes, estaría reflejando su importancia y cuyo rol no es aun del todo comprendido. No obstante es bastante claro su potencial como candidatos a biomarcadores diagnósticos y pronóstico.

Los biomarcadores identificados en este trabajo preliminar, justifican el diseño de un estudio a nivel clínico con un número de pacientes suficiente que permita validar estadísticamente el valor de los mismos así como los parámetros de sensibilidad y especificidad diagnóstica.

Al menos en los últimos 5 años se han destinado grandes esfuerzos para la identificación de biomarcadores circulantes, sin embargo se han planteado diversos problemas metodológicos, de los cuales no hemos sido ajenos. La falta de un consenso en la metodología utilizada tanto para las técnicas de extracción y cuantificación de ARN como de los controles utilizados en la búsqueda y validación de los sARNs, ha llevado a innumerables reportes contradictorios e irreproducibles (114,115,123). Witwer en su artículo de 2015 lo describe muy claramente, dejando en evidencia que es imprescindible llegar a un acuerdo en las metodologías utilizadas para la identificación

de biomarcadores circulantes. Por otra parte no hay que perder de vista que el cáncer es una enfermedad que genera a nivel sistémico, la activación de diversas vías metabólicas, de respuesta inmune, proliferación y angiogénesis entre otras (124), viéndose también afectada la vía de los ARNs circulantes. En su conjunto hacen aún más complejo el estudio de estas moléculas circulantes, lo cual nos lleva a pensar que estamos frente a una herramienta muy potente y prometedora de la cual actualmente se conoce muy poco.

6 Conclusiones

Se identificaron y validaron como posibles candidatos en tejido tumoral: miR-21-5p, miR-148-3p, miR-100a, His_GTG, Gly_GCC, Glu_TTC.

Se demostró que la fracción más adecuada para el estudio de ARNs circulantes es el suero, identificando y validando, un aumento de Glu_TTC y Gly_GCC en los pacientes analizados.

Se demostró que la técnica *stem-loop* RT-qPCR no es la técnica que se adecua mejor al estudio de sARNt, debido a la variación en el tamaño de los fragmentos, siendo más adecuada la utilización de *primers* LNA tanto en tejido como a nivel de ARNs circulantes.

Se demostró que los candidatos no se encuentran en la fracción exosomal de plasma.

7 Perspectivas

Dado el potencial diagnóstico, tanto en tejido como en suero es imprescindible aumentar el número de pacientes para darle un peso estadístico aún mayor a los candidatos identificados.

Realizar pruebas de validación utilizando PCR digital sería interesante ya que sería otra forma de validar nuestros hallazgos.

Para el caso de los miARNs no detectados en suero, habría que aumentar los volúmenes analizados teniendo en cuenta la baja concentración de miARNs circulantes.

Hacer un *screening* para ver si los candidatos están restringidos solo a sueros de pacientes con cáncer de pulmón o si aparecen en otras neoplasias, sobretodo en cáncer de mama, ya que fueron las líneas celulares utilizadas en el laboratorio.

Continuar trabajando en la elucidación del rol funcional de estas nuevas clases de sARNs a través de la identificación de proteínas que forman los complejos ribonucleoproteicos, mediante co-inmunoprecipitación.

Identificar blancos de estos sARNt aportaría en el entendimiento de la funcionalidad de esta nueva clase de ARNnc.

Bibliografía

1. Crick F. Central dogma of molecular biology. *Nature*. 1970 Aug 8;227(5258):561–3.
2. Carninci P, Hayashizaki Y. Noncoding RNA transcription beyond annotated genes. *Current opinion in genetics & development*. 2007 Apr;17(2):139–44.
3. Huttenhofer A, Schattner P, Polacek N. Non-coding RNAs: hope or hype? *Trends in genetics : TIG*. 2005 May;21(5):289–97.
4. Eddy SR. Non-coding RNA genes and the modern RNA world. *Nat Rev Genet*. 2001 Dec;2(12):919–29.
5. Ponting CP, Oliver PL, Reik W. Evolution and functions of long noncoding RNAs. *Cell*. 2009 Feb 20;136(4):629–41.
6. Massé E, Majdalani N, Gottesman S. Regulatory roles for small RNAs in bacteria. *Curr Opin Microbiol*. 2003 Apr;6(2):120–4.
7. Plath K, Mlynarczyk-Evans S, Nusinow DA, Panning B. Xist RNA and the mechanism of X chromosome inactivation. *Annu Rev Genet*. 2002;36:233–78.
8. Lee RC, Feinbaum RL, Ambros V. The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell*. 1993 Dec 3;75(5):843–54.
9. Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature*. 1998 Feb 19;391(6669):806–11.
10. Berezikov E. Evolution of microRNA diversity and regulation in animals. *Nat Rev Genet*. 2011 Dec;12(12):846–60.
11. Farazi TA, Juranek SA, Tuschl T. The growing catalog of small RNAs and their association with distinct Argonaute/Piwi family members. *Development*. 2008 Apr;135(7):1201–14.
12. Tuck AC, Tollervey D. RNA in pieces. *Trends in Genetics*. 2011 Oct 1;27(10):422–32.
13. Siomi H, Siomi MC. On the road to reading the RNA-interference code. *Nature*. 2009 Jan 22;457(7228):396–404.
14. Liu X, Fortin K, Mourelatos Z. MicroRNAs: biogenesis and molecular functions. *Brain pathology*. 2008 Jan;18(1):113–21.
15. Zhang R, Zhang L, Yu W. Genome-wide expression of non-coding RNA and global chromatin modification. *Acta biochimica et biophysica Sinica*. 2012 Jan;44(1):40–7.
16. Zhang B, Pan X, Cobb GP, Anderson TA. microRNAs as oncogenes and tumor suppressors. *Developmental Biology*. 2007 Feb 1;302(1):1–12.
17. Kent OA, Mendell JT. A small piece in the cancer puzzle: microRNAs as tumor suppressors and oncogenes. *Oncogene*. 2006;25(46):6188–96.

18. Meltzer PS. Cancer genomics: Small RNAs with big impacts. *Nature*. 2005 Jun 9;435(7043):745–6.
19. Napoli C, Lemieux C, Jorgensen R. Introduction of a Chimeric Chalcone Synthase Gene into *Petunia* Results in Reversible Co-Suppression of Homologous Genes in trans. *The Plant cell*. 1990 Apr;2(4):279–89.
20. Zhang H, Kolb FA, Jaskiewicz L, Westhof E, Filipowicz W. Single processing center models for human Dicer and bacterial RNase III. *Cell*. 2004 Jul 9;118(1):57–68.
21. Ghildiyal M, Zamore PD. Small silencing RNAs: an expanding universe. *Nature reviews Genetics*. 2009 Feb;10(2):94–108.
22. Kawamata T, Tomari Y. Making RISC. *Trends Biochem Sci*. 2010 Jul;35(7):368–76.
23. Peters L, Meister G. Argonaute proteins: mediators of RNA silencing. *Mol Cell*. 2007 Jun 8;26(5):611–23.
24. Rana TM. Illuminating the silence: understanding the structure and function of small RNAs. *Nat Rev Mol Cell Biol*. 2007 Jan;8(1):23–36.
25. Filipowicz W, Bhattacharyya SN, Sonenberg N. Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? *Nat Rev Genet*. 2008 Feb;9(2):102–14.
26. Zamore PD, Haley B. Ribo-gnome: The Big World of Small RNAs. *Science*. 2005 Sep 2;309(5740):1519–24.
27. Gagnon KT, Corey DR. Argonaute and the nuclear RNAs: new pathways for RNA-mediated control of gene expression. *Nucleic Acid Ther*. 2012 Feb;22(1):3–16.
28. Luteijn MJ, Ketting RF. PIWI-interacting RNAs: from generation to transgenerational epigenetics. *Nat Rev Genet*. 2013 Aug;14(8):523–34.
29. Zardo G, Ciolfi A, Vian L, Billi M, Racanicchi S, Grignani F, et al. Transcriptional targeting by microRNA-Polycomb complexes. *Cell Cycle*. 2012 Oct 1;11(19):3543–9.
30. CAI X, HAGEDORN CH, CULLEN BR. Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs. *RNA*. 2004 Dec;10(12):1957–66.
31. Lee Y, Kim M, Han J, Yeom K-H, Lee S, Baek SH, et al. MicroRNA genes are transcribed by RNA polymerase II. *EMBO J*. 2004 Oct 13;23(20):4051–60.
32. Gregory RI, Yan K-P, Amuthan G, Chendrimada T, Doratotaj B, Cooch N, et al. The Microprocessor complex mediates the genesis of microRNAs. *Nature*. 2004 Nov 11;432(7014):235–40.
33. Ameres SL, Zamore PD. Diversifying microRNA sequence and function. *Nat Rev Mol Cell Biol*. 2013 Aug;14(8):475–88.
34. Bohnsack MT, Czaplinski K, Gorlich D. Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs. *RNA*. 2004 Feb;10(2):185–91.

35. Curtis HJ, Sibley CR, Wood MJA. Mirtrons, an emerging class of atypical miRNA. *WIREs RNA*. 2012 Sep 1;3(5):617–32.
36. Garcia-Silva MR, Frugier M, Tosar JP, Correa-Dominguez A, Ronalte-Alves L, Parodi-Talice A, et al. A population of tRNA-derived small RNAs is actively produced in *Trypanosoma cruzi* and recruited to specific cytoplasmic granules. *Mol Biochem Parasitol*. 2010 Jun;171(2):64–73.
37. Leung AKL, Calabrese JM, Sharp PA. Quantitative analysis of Argonaute protein reveals microRNA-dependent localization to stress granules. *Proc Natl Acad Sci U S A*. 2006 Nov 28;103(48):18125–30.
38. Siomi MC, Sato K, Pezic D, Aravin AA. PIWI-interacting small RNAs: the vanguard of genome defence. *Nat Rev Mol Cell Biol*. 2011 Apr;12(4):246–58.
39. Martens-Uzunova ES, Olvedy M, Jenster G. Beyond microRNA--novel RNAs derived from small non-coding RNA and their implication in cancer. *Cancer Lett*. 2013 Nov 1;340(2):201–11.
40. Tosar JP, Gámbaro F, Sanguinetti J, Bonilla B, Witwer KW, Cayota A. Assessment of small RNA sorting into different extracellular fractions revealed by high-throughput sequencing of breast cell lines. *Nucleic Acids Res*. 2015 Jun 23;43(11):5601–16.
41. Keam SP, Hutvagner G. tRNA-Derived Fragments (tRFs): Emerging New Roles for an Ancient RNA in the Regulation of Gene Expression. *Life*. 2015 Nov 27;5(4):1638–51.
42. Lee YS, Shibata Y, Malhotra A, Dutta A. A novel class of small RNAs: tRNA-derived RNA fragments (tRFs). *Genes Dev*. 2009 Nov 15;23(22):2639–49.
43. Shigematsu M, Kirino Y. tRNA-Derived Short Non-coding RNA as Interacting Partners of Argonaute Proteins. *Gene Regul Syst Bio*. 2015 Sep 10;9:27–33.
44. Emara MM, Ivanov P, Hickman T, Dawra N, Tisdale S, Kedersha N, et al. Angiogenin-induced tRNA-derived stress-induced RNAs promote stress-induced stress granule assembly. *J Biol Chem*. 2010 Apr 2;285(14):10959–68.
45. Honda S, Loher P, Shigematsu M, Palazzo JP, Suzuki R, Imoto I, et al. Sex hormone-dependent tRNA halves enhance cell proliferation in breast and prostate cancers. *PNAS*. 2015 Jul 21;112(29):E3816–25.
46. Yamasaki S, Ivanov P, Hu G-F, Anderson P. Angiogenin cleaves tRNA and promotes stress-induced translational repression. *J Cell Biol*. 2009 Apr 6;185(1):35–42.
47. Dhahbi JM. 5' tRNA Halves: The Next Generation of Immune Signaling Molecules. *Front Immunol* [Internet]. 2015 Feb 19 [cited 2016 Aug 8];6. Available from: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4333838/>
48. Mandel P, Metais P. [Not Available]. *C R Seances Soc Biol Fil*. 1948 Feb;142(3–4):241–3.
49. Lo YM. Circulating nucleic acids in plasma and serum: an overview. *Ann N Y Acad Sci*. 2001 Sep;945:1–7.

50. Lo YMD, Chiu RWK. Genomic Analysis of Fetal Nucleic Acids in Maternal Blood. *Annual Review of Genomics and Human Genetics*. 2012;13(1):285–306.
51. Arroyo JD, Chevillet JR, Kroh EM, Ruf IK, Pritchard CC, Gibson DF, et al. Argonaute2 complexes carry a population of circulating microRNAs independent of vesicles in human plasma. *Proc Natl Acad Sci USA*. 2011 Mar 22;108(12):5003–8.
52. Vickers KC, Palmisano BT, Shoucri BM, Shamburek RD, Remaley AT. MicroRNAs are transported in plasma and delivered to recipient cells by high-density lipoproteins. *Nat Cell Biol*. 2011 Apr;13(4):423–33.
53. György B, Szabó TG, Pásztói M, Pál Z, Misják P, Aradi B, et al. Membrane vesicles, current state-of-the-art: emerging role of extracellular vesicles. *Cell Mol Life Sci*. 2011 Aug;68(16):2667–88.
54. Raposo G, Stoorvogel W. Extracellular vesicles: Exosomes, microvesicles, and friends. *J Cell Biol*. 2013 Feb 18;200(4):373–83.
55. García-Olmo D, García-Olmo DC, Ontañón J, Martínez E. Horizontal transfer of DNA and the “genometastasis hypothesis.” *Blood*. 2000 Jan 15;95(2):724–5.
56. Valadi H, Ekström K, Bossios A, Sjöstrand M, Lee JJ, Lötvall JO. Exosome-mediated transfer of mRNAs and microRNAs is a novel mechanism of genetic exchange between cells. *Nat Cell Biol*. 2007 Jun;9(6):654–9.
57. Holmgren L, Szeles A, Rajnavölgyi E, Folkman J, Klein G, Ernberg I, et al. Horizontal transfer of DNA by the uptake of apoptotic bodies. *Blood*. 1999 Jun 1;93(11):3956–63.
58. Cortez MA, Bueso-Ramos C, Ferdin J, Lopez-Berestein G, Sood AK, Calin GA. MicroRNAs in body fluids--the mix of hormones and biomarkers. *Nat Rev Clin Oncol*. 2011 Aug;8(8):467–77.
59. Costa-Silva B, Aiello NM, Ocean AJ, Singh S, Zhang H, Thakur BK, et al. Pancreatic cancer exosomes initiate pre-metastatic niche formation in the liver. *Nat Cell Biol*. 2015 Jun;17(6):816–26.
60. Abels ER, Breakefield XO. Introduction to Extracellular Vesicles: Biogenesis, RNA Cargo Selection, Content, Release, and Uptake. *Cell Mol Neurobiol*. 2016 Apr;36(3):301–12.
61. Mathivanan S, Ji H, Simpson RJ. Exosomes: extracellular organelles important in intercellular communication. *J Proteomics*. 2010 Sep 10;73(10):1907–20.
62. Kourembanas S. Exosomes: vehicles of intercellular signaling, biomarkers, and vectors of cell therapy. *Annu Rev Physiol*. 2015;77:13–27.
63. Peinado H, Alečković M, Lavotshkin S, Matei I, Costa-Silva B, Moreno-Bueno G, et al. Melanoma exosomes educate bone marrow progenitor cells toward a pro-metastatic phenotype through MET. *Nat Med*. 2012 Jun;18(6):883–91.
64. Cocucci E, Meldolesi J. Ectosomes and exosomes: shedding the confusion between extracellular vesicles. *Trends Cell Biol*. 2015 Jun;25(6):364–72.

65. Turchinovich A, Weiz L, Langheinz A, Burwinkel B. Characterization of extracellular circulating microRNA. *Nucleic Acids Res.* 2011 Sep 1;39(16):7223–33.
66. Turchinovich A, Burwinkel B. Distinct AGO1 and AGO2 associated miRNA profiles in human cells and blood plasma. *RNA Biol.* 2012 Aug;9(8):1066–75.
67. Dhahbi JM, Spindler SR, Atamna H, Yamakawa A, Boffelli D, Mote P, et al. 5' tRNA halves are present as abundant complexes in serum, concentrated in blood cells, and modulated by aging and calorie restriction. *BMC Genomics.* 2013;14:298.
68. Dhahbi JM, Spindler SR, Atamna H, Boffelli D, Martin DI. Deep Sequencing of Serum Small RNAs Identifies Patterns of 5' tRNA Half and YRNA Fragment Expression Associated with Breast Cancer. *Biomark Cancer.* 2014 Dec 8;6:37–47.
69. Patton JG, Franklin JL, Weaver AM, Vickers K, Zhang B, Coffey RJ, et al. Biogenesis, delivery, and function of extracellular RNA. *J Extracell Vesicles.* 2015;4:27494.
70. Fleischhacker M, Schmidt B. Circulating nucleic acids (CNAs) and cancer--a survey. *Biochim Biophys Acta.* 2007 Jan;1775(1):181–232.
71. Schwarzenbach H, Hoon DSB, Pantel K. Cell-free nucleic acids as biomarkers in cancer patients. *Nat Rev Cancer.* 2011 Jun;11(6):426–37.
72. Hanahan D, Weinberg RA. Hallmarks of Cancer: The Next Generation. *Cell.* 2011 Mar 4;144(5):646–74.
73. Jon.C.Aster VK. Patología Estructural y Funcional. 9a ed. ELSEVIER; 2015. 712-719 p.
74. Jon.C.Aster VK. Patología Estructural y Funcional. 9a ed. ELSEVIER; 2015. 276-271 p.
75. Globocan 2012 - Home [Internet]. [cited 2016 Jul 25]. Available from: <http://globocan.iarc.fr/Default.aspx>
76. Comisión Honoraria de Lucha contra el cáncer [Internet]. [cited 2016 Jul 25]. Available from: http://www.comisioncancer.org.uy/categoria_53_1.html
77. Travis WD, Brambilla E, Riely GJ. New Pathologic Classification of Lung Cancer: Relevance for Clinical Practice and Clinical Trials. *JCO.* 2013 Mar 10;31(8):992–1001.
78. Nana-Sinkam SP, Powell CA. Molecular biology of lung cancer: Diagnosis and management of lung cancer, 3rd ed: American College of Chest Physicians evidence-based clinical practice guidelines. *Chest.* 2013 May;143(5 Suppl):e30S–9S.
79. Cooper WA, Lam DCL, O'Toole SA, Minna JD. Molecular biology of lung cancer. *J Thorac Dis.* 2013 Oct;5(Suppl 5):S479–90.
80. Zagryazhskaya A, Zhivotovsky B. miRNAs in lung cancer: A link to aging. *Ageing Research Reviews.* 2014 Sep;17:54–67.
81. O'Driscoll L, Kenny E, Mehta JP, Doolan P, Joyce H, Gammell P, et al. Feasibility and relevance of global expression profiling of gene transcripts in serum from breast cancer patients using whole genome microarrays and quantitative RT-PCR. *Cancer Genomics Proteomics.* 2008 Apr;5(2):94–104.

82. García V, García JM, Peña C, Silva J, Domínguez G, Lorenzo Y, et al. Free circulating mRNA in plasma from breast cancer patients and clinical outcome. *Cancer Lett.* 2008 May 18;263(2):312–20.
83. Miura N, Nakamura H, Sato R, Tsukamoto T, Harada T, Takahashi S, et al. Clinical usefulness of serum telomerase reverse transcriptase (hTERT) mRNA and epidermal growth factor receptor (EGFR) mRNA as a novel tumor marker for lung cancer. *Cancer Sci.* 2006 Dec;97(12):1366–73.
84. Janssen HLA, Reesink HW, Lawitz EJ, Zeuzem S, Rodriguez-Torres M, Patel K, et al. Treatment of HCV infection by targeting microRNA. *N Engl J Med.* 2013 May 2;368(18):1685–94.
85. Xiao D, Ohlendorf J, Chen Y, Taylor DD, Rai SN, Waigel S, et al. Identifying mRNA, MicroRNA and Protein Profiles of Melanoma Exosomes. *PLOS ONE.* 2012 Oct 9;7(10):e46874.
86. Farazi TA, Brown M, Morozov P, Ten Hoeve JJ, Ben-Dov IZ, Hovestadt V, et al. Bioinformatic analysis of barcoded cDNA libraries for small RNA profiling by next-generation sequencing. *Methods.* 2012 Oct;58(2):171–87.
87. Andrews S. Babraham Bioinformatics - FastQC A Quality Control tool for High Throughput Sequence Data [Internet]. 2010 [cited 2016 May 9]. Available from: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>
88. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal.* 2011 May 2;17(1):10–2.
89. Morgan M, Anders S, Lawrence M, Aboyoun P, Pages H, Gentleman R. ShortRead: a bioconductor package for input, quality assessment and exploration of high-throughput sequence data. *Bioinformatics.* 2009 Oct 1;25(19):2607–8.
90. R Core Team. R: The R Project for Statistical Computing [Internet]. 2015 [cited 2016 May 9]. Available from: <https://www.r-project.org/>
91. Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, et al. The human genome browser at UCSC. *Genome research.* 2002 Jun;12(6):996–1006.
92. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome biology.* 2009;10(3):R25.
93. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics.* 2010 Jan 1;26(1):139–40.
94. Chen C, Ridzon DA, Broomer AJ, Zhou Z, Lee DH, Nguyen JT, et al. Real-time quantification of microRNAs by *stem-loop* RT-PCR. *Nucleic acids research.* 2005;33(20):e179.
95. Fritz JV, Heintz-Buschart A, Ghosal A, Wampach L, Etheridge A, Galas D, et al. Sources and Functions of Extracellular Small RNAs in Human Circulation. *Annual Review of Nutrition.* 2016;36(1):301–36.

96. Chen G, Umelo IA, Lv S, Teugels E, Fostier K, Kronenberger P, et al. miR-146a inhibits cell growth, cell migration and induces apoptosis in non-small cell lung cancer cells. *PLoS ONE*. 2013;8(3):e60317.
97. Joshi P, Middleton J, Jeon Y-J, Garofalo M. MicroRNAs in lung cancer. *World J Methodol*. 2014 Jun 26;4(2):59–72.
98. Chim SSC, Shing TKF, Hung ECW, Leung T-Y, Lau T-K, Chiu RWK, et al. Detection and characterization of placental microRNAs in maternal plasma. *Clin Chem*. 2008 Mar;54(3):482–90.
99. Liu X-G, Zhu W-Y, Huang Y-Y, Ma L-N, Zhou S-Q, Wang Y-K, et al. High expression of serum miR-21 and tumor miR-200c associated with poor prognosis in patients with lung cancer. *Med Oncol*. 2012 Jun;29(2):618–26.
100. Schwarzenbach H, Nishida N, Calin GA, Pantel K. Clinical relevance of circulating cell-free microRNAs in cancer. *Nat Rev Clin Oncol*. 2014 Mar;11(3):145–56.
101. Chevillet JR, Kang Q, Ruf IK, Briggs HA, Vojtech LN, Hughes SM, et al. Quantitative and stoichiometric analysis of the microRNA content of exosomes. *PNAS*. 2014 Oct 14;111(41):14888–93.
102. Mitchell PS, Parkin RK, Kroh EM, Fritz BR, Wyman SK, Pogosova-Agadjanyan EL, et al. Circulating microRNAs as stable blood-based markers for cancer detection. *PNAS*. 2008 Jul 29;105(30):10513–8.
103. Garzon R, Marcucci G, Croce CM. Targeting microRNAs in cancer: rationale, strategies and challenges. *Nat Rev Drug Discov*. 2010 Oct;9(10):775–89.
104. Zhang J, Wang J, Zhao F, Liu Q, Jiang K, Yang G. MicroRNA-21 (miR-21) represses tumor suppressor PTEN and promotes growth and invasion in non-small cell lung cancer (NSCLC). *Clin Chim Acta*. 2010 Jun 3;411(11–12):846–52.
105. Yang Y, Meng H, Peng Q, Yang X, Gan R, Zhao L, et al. Downregulation of microRNA-21 expression restrains non-small cell lung cancer cell proliferation and migration through upregulation of programmed cell death 4. *Cancer Gene Ther*. 2015 Jan;22(1):23–9.
106. Duursma AM, Kedde M, Schrier M, Sage C le, Agami R. miR-148 targets human DNMT3b protein coding region. *RNA*. 2008 May 1;14(5):872–7.
107. Kanwal R, Gupta S. Epigenetic modifications in cancer. *Clinical Genetics*. 2012 Apr 1;81(4):303–11.
108. Li J, Yu T, Cao J, Liu L, Liu Y, Kong H-W, et al. MicroRNA-148a Suppresses Invasion and Metastasis of Human Non-Small-Cell Lung Cancer. *Cell Physiol Biochem*. 2015;37(5):1847–56.
109. Yang J, Li B, Lu H, Chen Y, Lu C, Zhu R, et al. Serum miR-152, miR-148a, miR-148b, and miR-21 as novel biomarkers in non-small cell lung cancer screening. *Tumor Biol*. 2014 Dec 14;36(4):3035–42.

110. Varnholt H, Drebber U, Schulze F, Wedemeyer I, Schirmacher P, Dienes H-P, et al. MicroRNA gene expression profile of hepatitis C virus-associated hepatocellular carcinoma. *Hepatology*. 2008 Apr;47(4):1223–32.
111. Leite KRM, Sousa-Canavez JM, Reis ST, Tomiyama AH, Camara-Lopes LH, Sañudo A, et al. Change in expression of miR-let7c, miR-100, and miR-218 from high grade localized prostate cancer to metastasis. *Urol Oncol*. 2011 Jun;29(3):265–9.
112. Nagaraja AK, Creighton CJ, Yu Z, Zhu H, Gunaratne PH, Reid JG, et al. A link between mir-100 and FRAP1/mTOR in clear cell ovarian cancer. *Mol Endocrinol*. 2010 Feb;24(2):447–63.
113. Liu J, Lu K-H, Liu Z-L, Sun M, De W, Wang Z-X. MicroRNA-100 is a potential molecular marker of non-small cell lung cancer and functions as a tumor suppressor by targeting polo-like kinase 1. *BMC Cancer*. 2012;12:519.
114. Jayachandran M, Miller VM, Heit JA, Owen WG. Methodology for isolation, identification and characterization of microvesicles in peripheral blood. *J Immunol Methods*. 2012 Jan 31;375(1–2):207–14.
115. Witwer KW. Circulating MicroRNA Biomarker Studies: Pitfalls and Potential Solutions. *Clinical Chemistry*. 2015 Jan 1;61(1):56–63.
116. Roberts TC, Coenen-Stass AML, Wood MJA. Assessment of RT-qPCR Normalization Strategies for Accurate Quantification of Extracellular microRNAs in Murine Serum. *PLOS ONE*. 2014 Feb 19;9(2):e89237.
117. Baker M. Digital PCR hits its stride. *Nat Meth*. 2012 Jun;9(6):541–4.
118. Bellissimo T, Russo E, Ganci F, Vico C, Sacconi A, Longo F, et al. Circulating miR-21-5p and miR-148a-3p as emerging non-invasive biomarkers in thymic epithelial tumors. *Cancer Biology & Therapy*. 2016 Jan 2;17(1):79–82.
119. Garcia-Silva MR, Cabrera-Cabrera F, Güida MC, Cayota A. Hints of tRNA-Derived Small RNAs Role in RNA Silencing Mechanisms. *Genes (Basel)*. 2012;3(4):603–14.
120. Li Y, Luo J, Zhou H, Liao J-Y, Ma L-M, Chen Y-Q, et al. Stress-induced tRNA-derived RNAs: a novel class of small RNAs in the primitive eukaryote *Giardia lamblia*. *Nucleic Acids Res*. 2008 Nov;36(19):6048–55.
121. Ogawa T, Tomita K, Ueda T, Watanabe K, Uozumi T, Masaki H. A cytotoxic ribonuclease targeting specific transfer RNA anticodons. *Science*. 1999 Mar 26;283(5410):2097–100.
122. Thompson DM, Parker R. Stressing Out over tRNA Cleavage. *Cell*. 2009 Jul 23;138(2):215–9.
123. Eldh M, Lötvall J, Malmhäll C, Ekström K. Importance of RNA isolation methods for analysis of exosomal RNA: evaluation of different methods. *Mol Immunol*. 2012 Apr;50(4):278–86.
124. Jon.C.Aster VK. *Patología Estructural y Funcional*. 9a ed. ELSEVIER; 2015. 265 p.

