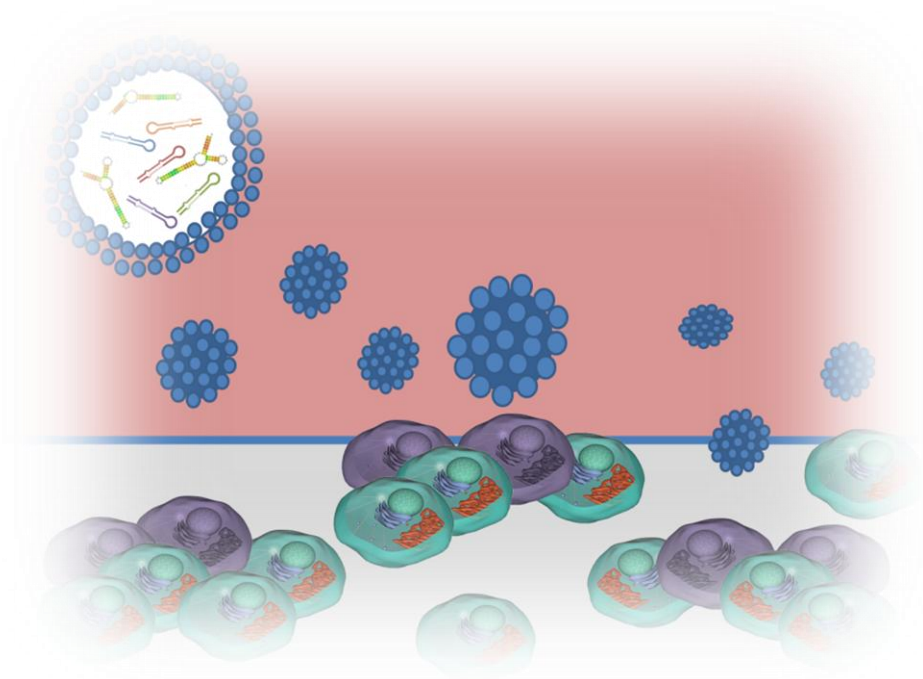


PEDECIBA - BIOLOGÍA

TESIS DE DOCTORADO

en Ciencias Biológicas

**Pequeños RNAs reguladores en el medio extracelular:
*secreción, estabilidad, transporte y potencial en la
comunicación intercelular***



Juan Pablo Tosar

Dirección: Dr. Alfonso Cayota

Tribunal:

Dr. Otto Pritsch

Dr. Ricardo Ehrlich

Dr. José Tort

XII-IV-MMXVI

*A Juan Francisco, que creció en la
panza de mamá mientras esta tesis juntaba
kilobytes en la computadora, pero que es
infinitamente más importante.*

*Y a la Bela, porque no debe haber
nadie en este mundo a quien esta tesis signifique
tanto.*

Esta tesis doctoral (PEDECIBA-Biología) se desarrolló principalmente entre mediados de 2012 y finales de 2015 en el Institut Pasteur de Montevideo (Laboratorio de Genómica Funcional), bajo la dirección del Dr. Alfonso Cayota. El autor es Magíster en Ciencias Biológicas y docente de la Universidad de la República (Asistente de la Unidad de Bioquímica Analítica, CIN, Facultad de Ciencias). Contó con apoyo financiero de la ANII. El autor disfrutó mucho escribiéndola y espera que pueda ser disfrutable también para el lector. Tal vez útil, en el mejor de los casos. AMDG

ÍNDICE

INTRODUCCIÓN GENERAL	1
A) Regulación de la expresión génica mediada por pequeños RNAs	1
B) Procesamiento de ncRNAs y nuevas familias de RNAs reguladores	3
C) RNAs pequeños en el medio extracelular	7
i. Asociados a vesículas extracelulares	7
ii. No asociados a vesículas extracelulares	10
OBJETIVOS GENERALES Y ESPECÍFICOS	11
CAPÍTULO I:	
<i>La dieta como fuente posible de miRNAs extracelulares, ¿Comemos información, o nos comimos un artefacto?</i>	
1.1 INTRODUCCIÓN	12
1.2 MATERIALES Y MÉTODOS	16
1.3 RESULTADOS	18
1.4 DISCUSIÓN	26
CAPÍTULO II:	
<i>Caracterización del conjunto de pequeños RNAs secretados por células mamarias en distintas fracciones extracelulares</i>	
2.1 INTRODUCCIÓN	31
2.1.1 Modelo de estudio	34
2.2 MATERIALES Y MÉTODOS	35
2.3 RESULTADOS	43
2.4 DISCUSIÓN	71
CAPÍTULO III:	
<i>Estabilidad de las mitades de tRNA en la fracción soluble (S100) del medio extracelular y sus proteínas asociadas</i>	
3.1 INTRODUCCIÓN	78
3.2 MATERIALES Y MÉTODOS	79
3.3 RESULTADOS	83
3.4 DISCUSIÓN	107
3.5 PERSPECTIVAS A CORO Y MEDIANO PLAZO	117
DIFUSIÓN Y PRODUCCIÓN DE ESTA TESIS	119
AGRADECIMIENTOS	121
REFERENCIAS BIBLIOGRÁFICAS	125
ANEXO METODOLÓGICO	134

Los llamados pequeños RNAs reguladores (20-30 nucleótidos) son moléculas capaces de reconocer mensajeros específicos y mediar su silenciamiento, dirigiendo a ellos una maquinaria proteica especializada. Los microRNAs (miRNAs), por ejemplo, regulan a nivel post-transcripcional la expresión de aproximadamente un 60% de los genes que en humanos codifican para proteínas. Nuestro laboratorio ha estado históricamente interesado en el estudio de otras familias de RNAs pequeños, tales como las mitades de tRNA, de función menos conocida. Además de ser fundamentales en el control de la homeostasis celular, los pequeños RNAs reguladores pueden ser secretados hacia el medio extracelular, ya sea protegidos en vesículas extracelulares (exosomas y microvesículas) o formando complejos con proteínas. Existe evidencia de que las células son capaces de captar estos RNAs reguladores del medio extracelular y, una vez internalizados, estos mantienen su capacidad de silenciamiento génico en la célula receptora. Esto quiere decir que existen posibles mecanismos de comunicación intercelular mediados por RNAs pequeños, y fue a lo que quisimos contribuir con este estudio.

En primer lugar, estudiamos la viabilidad de la comunicación mediada por miRNAs entre especies de reinos biológicos diferentes (plantas – mamíferos) a través de la dieta, según sugería un muy difundido y polémico reporte publicado en 2012. Nuestros resultados apuntan a que las conclusiones de dicho trabajo son inválidas, pues el laboratorio en cuestión estaría contaminado con miRNAs vegetales. Aprovechamos la ocasión para estudiar los efectos de la contaminación en los estudios realizados por secuenciación masiva, un aspecto no debidamente tomado en cuenta por la mayoría de los investigadores. En una segunda etapa secuenciamos todo el conjunto de pequeños RNAs reguladores secretados por una línea celular de cáncer de mama (MCF-7) y su contraparte no maligna (MCF-10A), en distintas fracciones extracelulares: exosomas, microvesículas, complejos ribonucleoproteicos. Encontramos sesgos marcados entre los perfiles intracelulares y extracelulares, e intentamos elucidar los procesos de secreción involucrados (selectividad vs. no selectividad). Entre los resultados que más llamaron nuestra atención, se encuentra el remarcado enriquecimiento de la fracción ribonucleoproteica en mitades 5' de tRNAs aceptores de Glu y Gly. El final del trabajo se centró en estudiar cómo estas mitades de tRNAs son transportadas en el medio extracelular, en qué forma molecular se las encuentra, y cuáles son sus – posibles – proteínas asociadas. Podría, dentro de algún tiempo, hablarse de la “biología de mitades de tRNA extracelulares”... se precisa para eso mucho más que una Tesis, pero puede que se encuentren aquí algunas bases.

INTRODUCCIÓN GENERAL

A) Regulación de la expresión génica mediada por pequeños RNAs

Aunque los primeros descubrimientos pueden rastrearse años antes, la “fiebre del oro” de los RNAs pequeños reguladores de la expresión génica puede decirse que comenzó en 1998, con la publicación del artículo que años más tarde (2006) le permitiría a Andrew Z. Fire y Craig Mello immortalizar su nombre entre la lista de laureados con el Premio Nobel (Fire et al., 1998). En cuatro carillas y tres figuras los autores introdujeron las bases del fenómeno de interferencia mediada por RNA de doble hebra (RNAi).

La RNAi existe principalmente como mecanismo de inmunidad celular innata contra los virus de RNA, o la de aquellos que generan intermediarios de RNA de doble hebra durante su ciclo replicativo. Una ribonucleasa del tipo III llamada Dicer actúa como una regla molecular que reconoce los híbridos de RNA a nivel citoplasmático y los corta en unidades de 21 nucleótidos. Dichas unidades, denominadas siRNAs (*small interfering RNAs*), son moléculas de RNA de doble hebra perfectamente complementarias, con protrusiones simple hebra de dos bases en los extremos 3'. Los siRNAs se unen posteriormente a proteínas de la familia Argonauta, para formar el corazón de lo que se conoce como complejo RISC (*RNA-induced silencing complex*). Únicamente una de las dos hebras del siRNA queda finalmente incorporada en el complejo RISC, la cual actúa como guía para llevar al complejo hacia RNAs complementarios. En el contexto de la respuesta antiviral, dichos RNAs complementarios corresponden al genoma o a los intermediarios del propio virus. Algunas proteínas Argonauta (p.ej.: AGO2 en humanos) poseen actividad endonucleolítica “*slicer*”, que les permite cortar el RNA viral una vez formado el complejo ternario AGO/siRNA/RNA.

Estos mecanismos no funcionan únicamente en el contexto de la respuesta contra virus. También funcionan en algunos organismos como métodos de control de la propia expresión génica de una célula. En algunas circunstancias las células eucariotas pueden producir moléculas de RNA de doble hebra (por transcripción bidireccional convergente, transcritos anti-sentido, formación de horquillas debido a secuencias repetidas invertidas, etc.), las cuales son clivadas por Dicer para generar los llamados endo-siRNAs, que silencian luego otros transcritos celulares con secuencias de ARN complementarias al mismo. Los endo-siRNA fueron primero descritos en plantas¹ y en nemátodos, y más tarde en moscas, ratones y humanos (Nilsen, 2008).

Una versión refinada, más ubicua en el universo eucariota y mucho mejor estudiada, la constituyen los llamados micro-RNAs (miRNAs). A diferencia de los endo-siRNAs, los miRNAs son moléculas de RNA de doble hebra con una media de 22 nucleótidos, pero

¹ Es de destacar la importante contribución del uruguayo Omar Borsani en este campo: Borsani, O., Zhu, J., Verslues, P.E., Sunkar, R., and Zhu, J.K. (2005). Endogenous siRNAs derived from a pair of natural cis-antisense transcripts regulate salt tolerance in *Arabidopsis*. *Cell* 123, 1279-1291.

poseen la característica de que ambas hebras no son perfectamente complementarias entre sí. Sus precursores (pri-miRNAs) están codificados en el genoma en secuencias capaces de adquirir una estructura característica en forma de horquilla una vez transcritos por la RNA polimerasa II. Estos pri-miRNAs son procesados primero a nivel nuclear por un complejo denominado micro-procesador y compuesto por la ribonucleasa del tipo III Drosha y la proteína DGCR8. El RNA resultante (pre-miRNA) es exportado del núcleo, y es blanco de la acción citoplasmática de Dicer, la cual genera el miRNA maduro. Al igual que los siRNAs, los miRNAs se unen a proteínas de la familia Argonauta para conformar el complejo RISC (miR-RISC en este caso). De nuevo, una sola de las dos hebras permanecerá unida a la Argonauta, y dicha hebra actuará como guía para el reconocimiento de RNAs celulares endógenos parcialmente complementarios, típicamente mRNAs.

En los animales, el apareamiento entre los miRNAs y sus mensajeros blancos suele darse en la región 3' UTR de estos últimos, y las secuencias reconocidas son solo parcialmente complementarias entre sí. De hecho, la especificidad suele estar conferida por las últimas siete u ocho primeras bases del miRNA, con aportes menos importantes de las bases del extremo 3' (Brennecke et al., 2005). Dado que los requerimientos para el reconocimiento no son muy estrictos, cada miRNA es capaz de reconocer una gran cantidad de blancos diferentes, generando complejos entramados de regulación génica. La última versión de la base de datos miRBase (v.21) registra en humanos 1881 *loci* codificantes para miRNAs (542 de ellos altamente confiables), y se calcula que al menos un 60% de los genes codificantes para proteínas se encuentran sometidos a su control (Friedman et al., 2009). Dado que modulan la expresión de la mayoría de las proteínas de la célula, no debería sorprender que estén implicados en una gran cantidad de procesos patológicos. Algunos miRNAs silencian oncogenes y otros silencian supresores tumorales, y es frecuente observar que varios tipos de cáncer presentan una sub-expresión de los primeros o una sobre-expresión de los segundos.

Otra particularidad que distingue a los miRNAs de los siRNAs – además de su complementariedad imperfecta y del reconocimiento parcial que efectúan de sus blancos – es que el silenciamiento no ocurre necesariamente por degradación directa mediada por la proteína Argonauta. La inhibición del comienzo o de la etapa de elongación de la traducción, el “secuestro” de los mensajeros a gránulos de estrés, la pérdida del 5' cap y/o el acortamiento de la cola poli-A del mRNA y su posterior degradación exonucleolítica, son mecanismos por los cuales los miRNAs también ejercen su silenciamiento post-transcripcional (Fabian and Sonenberg, 2012). De hecho, la mayoría de las proteínas Argonauta humanas carecen de la actividad endonucleolítica “*slicer*” necesaria para escindir mRNAs.

Los primeros miRNAs fueron descritos antes del mencionado artículo de Fire y Mello, pero la universalidad del fenómeno no estaba por entonces de manifiesto. El sistema se fue primero descubriendo y disecando en el gusano *Caenorhabditis elegans* por los grupos de Victor Ambros (Lee et al., 1993) y de Gary Ruvkun (Reinhart et al., 2000; Slack et al., 2000; Wightman et al., 1993). Pero el “nacimiento” de los miRNAs como una familia ubicua y diversa de pequeños RNA reguladores se suele fechar en octubre de 2001, cuando la revista *Science* publica los artículos de Thomas Tuschl (Lagos-Quintana et al., 2001), David Bartel

(Lau et al., 2001) y del propio Ambros (Lee and Ambros, 2001), y el término “micro-RNA” queda acuñado. En 2016 se cumplen 15 años de estas publicaciones, así que vaya por aquí un pequeño homenaje. Una interesante cronología con varios de los descubrimientos que marcaron hitos en el estudio de los miRNAs puede encontrarse en el link provisto al pie de página². El último de los hitos presentados en dicha *web* corresponde al artículo de Chen-Yu Zhang y colaboradores sobre la transferencia horizontal de miRNAs a través de la dieta (Zhang et al., 2012a), el cual será objeto expreso de análisis en el primer capítulo de esta tesis.

En 2016 se cumplen además 10 años del descubrimiento de los piRNAs, una tercera clase de RNAs pequeños reguladores. Cuatro artículos publicados casi en simultáneo en las revistas *Nature* y *Genes & Development* describieron la asociación de pequeños RNAs de un largo mayor del esperado para miRNAs (26-31 nucleótidos) con las proteínas MIWI y MILI – ortólogos de la proteína PIWI de la mosca – en células germinales de ratón (Aravin et al., 2006; Girard et al., 2006; Grivna et al., 2006; Watanabe et al., 2006)³. Desde un punto de vista filogenético, las proteínas Argonauta se dividen en dos subfamilias: Ago-like y Piwi-like, y mientras que las primeras participan en los mecanismos relacionados al silenciamiento por miRNAs/siRNAs, las segundas no tenían por entonces una función conocida. Estos pequeños RNAs recibieron el nombre de piRNAs, de *piwi-interacting small RNAs*, y su función en este caso estaba relacionada al silenciamiento de elementos transponibles del genoma. La activación de esta maquinaria en células germinales tenía sentido, pues es entonces cuando el cuidado de la integridad genómica se vuelve crítico.

El descubrimiento de los piRNAs fue muy importante porque hasta entonces se consideraba que la única clase de pequeños RNAs reguladores que actuaba en mamíferos eran los miRNAs. El fenómeno de RNAi se creía únicamente como un sistema de respuesta ante dsRNA exógeno, a diferencia de lo que ocurría en plantas y nemátodos. Fue así que los siRNA, miRNA y piRNA se consolidaron como las tres grandes familias de pequeños RNAs reguladores, y un esquema básico de sus mecanismos de biogénesis se volvió figura obligada de tesis y presentaciones. La que aquí se muestra fue extraída de (Siomi et al., 2011) (Figura 1).

B) Procesamiento de ncRNAs y nuevas familias de RNAs reguladores

A fines de 2008 comencé estudios de maestría en el Laboratorio de Genómica Funcional del IPMon. Me vinculé a la línea de investigación entonces activa en el protozooario patógeno *Trypanosoma cruzi*. Llamaba la atención que dicho organismo careciera de la maquinaria necesaria para las vías de RNAi o de miRNAs, mientras que su pariente cercano, *Trypanosoma brucei*, contaba con dichos mecanismos para su supervivencia. Además, *T. cruzi* mantenía en su genoma una proteína de la familia Argonauta, que aunque distinta a las de eucariotas superiores, debería tener algún sentido biológico. Esto provocó que este laboratorio que venía enfocándose en el estudio de los miRNAs en leucemia (Marton et al.,

²[http://www.rnaiweb.com/RNAi/microRNA/microRNA_Timeline/\(al 27/01/2015\)](http://www.rnaiweb.com/RNAi/microRNA/microRNA_Timeline/(al%2027/01/2015))

³ Tomas Tuschl puede jactarse de haber participado en el descubrimiento tanto de los miRNAs (Lagos-Quintana et al. 2001) como de los piRNAs (Aravin et al. 2006).

2008) comenzó a clonar y secuenciar pequeños RNAs en *T.cruzi*, buscando algo que pudiera parecerse a un miRNA.

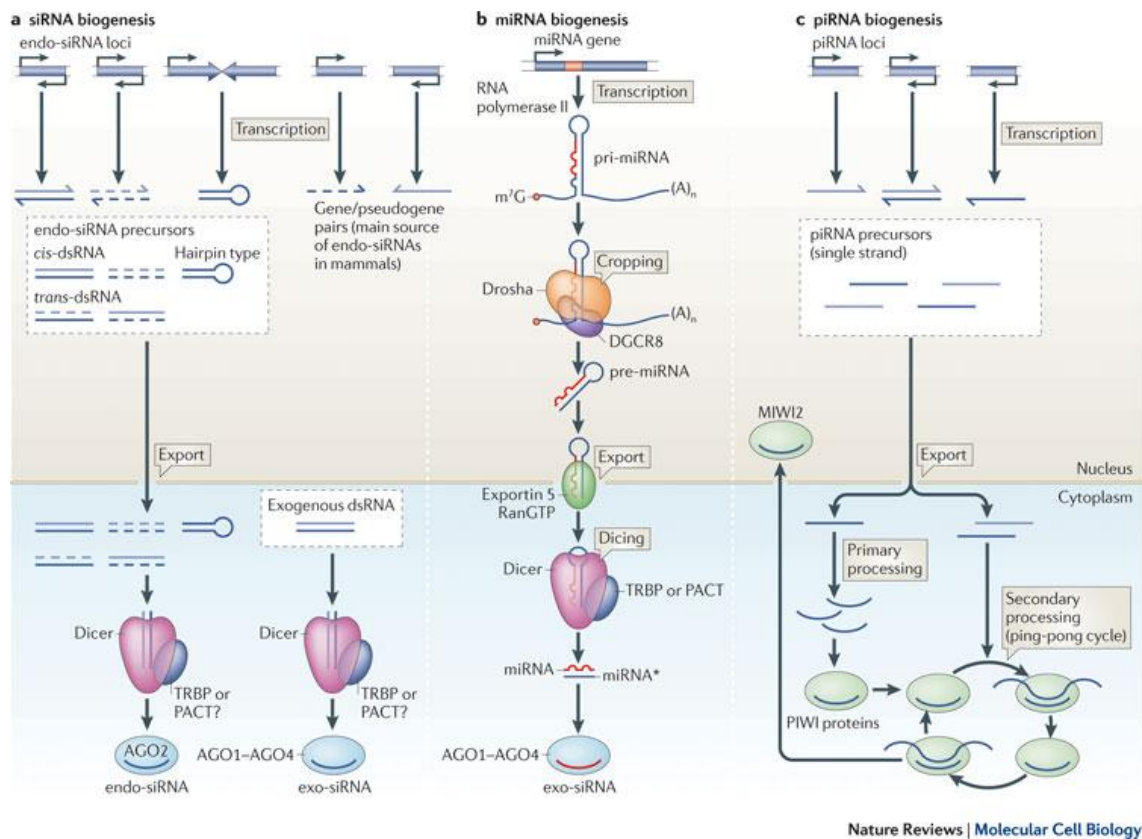


Figura 1: regulación endógena de la expresión génica en mamíferos mediada por pequeños RNAs reguladores: siRNAs (A), miRNAs (B) y piRNAs (C). (A) Los siRNAs se originan a partir de dsRNAs generados por transcripción bidireccional (en cis o en trans) o a partir de transcritos auto-complementarios que forman estructuras de horquilla. En el citoplasma son clivados por Dicer para obtener los siRNA maduros, de 21 nucleótidos. Una de las dos hebras del siRNA quedará retenida a la proteína Argonauto 2 (AGO2), conformando el complejo RISC. Dicha hebra dirigirá al RISC hacia RNAs complementarios, que serán cortados por AGO2. Esta vía también se alimenta de dsRNAs exógenos (de origen viral, o introducidos artificialmente en el laboratorio) en lo que es conocido como el fenómeno de RNAi. (B) Los miRNAs están codificados en el genoma y la RNA polimerasa II produce transcritos capeados y poliadenilados que presentan una estructura de bucle y horquilla parcial (pri-miRNAs). Estos transcritos sufren la acción del complejo micro-procesador a nivel nuclear (Drosha + DGCR8) que libera los pre-miRNAs, los cuales son exportados al citoplasma vía Exportina 5. El mecanismo a nivel citoplasmático es compartido con la vía de los siRNAs, pero además de AGO2 también pueden participar otras proteínas Argonauto sin actividad endonucleolítica (AGO1, AGO3 y AGO4). (C) Los piRNA son transcritos a partir de loci dispersos entre las regiones intergénicas del genoma y típicamente agrupados en clusters, y a diferencia de los anteriores, los precursores no forman estructuras de doble hebra. Se asocian en el citoplasma a distintas proteínas Argonauto de la subfamilia “Piwi-like”. Existe un mecanismo de amplificación interno conocido como “ciclo de ping-pong” que implica corte de precursores mediado por piRNAs, y asociación de los nuevos piRNAs generados a otras proteínas Piwi para a su vez procesar nuevos precursores. Finalmente, los piRNAs secundarios generados por esta vía se re-importan al núcleo donde participan en el silenciamiento de elementos transponibles. Figura extraída sin modificaciones de (Siomi et al., 2011)

Con los miRNA no hubo suerte, pero pronto empezó a quedar claro que un porcentaje muy alto de las secuencias clonadas correspondían a mitades 5’ de tRNAs. Más aún, existía un sesgo marcado hacia las mitades de algunos tRNA particulares, principalmente los aceptores de Asp y Glu. Dichos fragmentos, al analizarlos por inmunofluorescencia, parecían incrementarse en número y en intensidad en condiciones de estrés nutricional, y aparecían en

el citoplasma del parásito en forma de gránulos. Las secuencias obtenidas y los patrones de corte no eran consistentes con degradación al azar, sino que parecía haber algún tipo de proceso regulado. La generación de estas mitades de tRNA tampoco afectaba significativamente al *pool* de los tRNA maduros correspondientes, por lo cual no cabía pensar en una función de freno abrupto de la traducción; como ocurre con algunas toxinas bacterianas (Winther and Gerdes, 2011).

Para cuando nuestro reporte fue finalmente aceptado y publicado (Garcia-Silva et al., 2010), la generación de mitades de tRNAs frente a variados estímulos ya había sido descrita en otros organismos eucariotas (Cole et al., 2009; Fu et al., 2009; Jochl et al., 2008; Kawaji et al., 2008; Lee and Collins, 2005; Li et al., 2008; Thompson et al., 2008; Yamasaki et al., 2009).

Hoy día los fragmentos de tRNA son reconocidos como una nueva familia de RNAs reguladores, aunque los procesos efectores involucrados siguen sin estar del todo claros. Además de la diversidad de moléculas de tRNA de partida (en humanos existen 606 genes de tRNA, e incluso algunos de los tRNA que decodifican los mismos codones presentan secuencias diferentes... y todos ellos con un patrón variable de modificaciones post-transcripcionales) los fragmentos pueden ser producidos de distintas regiones y presentan distintos tamaños. A los fragmentos más cortos se les ha asignado el nombre de tRFs (por *tRNA-regulatory fragments*) (Lee et al., 2009), pero esta categoría incluye tanto a las secuencias producidas a partir de las regiones 5' y 3' de los tRNA maduros (tRF-5 y tRF-3, respectivamente) como a las secuencias correspondientes al *trailer* 3' que es eliminado durante el proceso de maduración del tRNA (tRF-1). Algunos de los tRF-5 y tRF-3 se originarían por la acción directa de la enzima Dicer (Cole et al., 2009), y pueden asociarse con proteínas Argonauta (Haussecker et al., 2010). Al menos uno de estos fragmentos, correspondiente al extremo 3' del tRNA^{Gly}_{GCC}, actúa como modulador de la proliferación celular y la respuesta al daño del DNA mediante el silenciamiento post-transcripcional específico del gen RPA1 (*replication protein A 1*) (Maute et al., 2013). Estos resultados conllevan a racionalizar a los tRNAs como precursores alternativos de miRNAs.

Además de estos fragmentos cortos (~ 22 nt) con características de miRNAs, los tRNAs suelen ser cortados a nivel del *loop* anticodón para dar lugar a mitades 5' y 3' del tipo de las encontradas en *T.cruzi*. En mamíferos se describió que la enzima responsable de realizar estos cortes frente a condiciones de estrés era la Angiogenina (Fu et al., 2009; Yamasaki et al., 2009), una ribonucleasa de la familia A vinculada a la angiogénesis tumoral, y cuya especificidad por los tRNAs era de largo tiempo conocida (Lee and Vallee, 1989). No existe aún una función clara y común para todas estas mitades, sino que parecerían participar en variados mecanismos cuya relevancia biológica recién comienza a ponerse de manifiesto. Algunos autores han señalado a la tRNaseZ^L, una versión larga y citoplasmática de la enzima responsable de la maduración de los precursores de tRNA en el núcleo, como una proteína capaz de unir mitades 5' de tRNAs y utilizarlas como guía para el clivaje de mRNAs parcialmente complementarios (Elbarbary et al., 2009). Las mitades de tRNA también podrían unirse al Citocromo c y proteger a las células del estrés hiperosmótico (Saikia et al., 2014). Resultados muy interesantes son los del grupo de Paul Anderson en Boston, que mostraron

cómo las mitades de algunos tRNAs específicos eran capaces de ensamblar gránulos de estrés y desplazar al complejo eIF4G/A del extremo 5' de mRNAs capeados y no capeados para de este modo inhibir la traducción. Únicamente aquellas mitades con trectos ricos en guanina en su extremo 5' (derivadas de los tRNA^{Ala} y tRNA^{Cys}) eran capaces de ejercer dicha función. Los autores fueron más allá e identificaron a la proteína YB-1 como la responsable de unir estas mitades y desensamblar los complejos iniciadores de la traducción (Ivanov et al., 2011). En un artículo posterior publicado en PNAS, encontraron que el tracto rico en guaninas era necesario para ensamblar una estructura conocida como “G-quadruplex”, la cual permitiría la unión a YB-1. Estos mecanismos estarían alterados en la esclerosis lateral amiotrófica debido a la sobre-expresión de un transcrito rico en guaninas que compite con las mitades de tRNA e impide la formación de gránulos de estrés (Ivanov et al., 2014).

La asociación entre los fragmentos de tRNA y la proteína YB-1 parece ser robusta. Un artículo de 2015 publicado en *Cell* muestra cómo YB-1 colabora con las células tumorales al estabilizar los mRNAs de una serie de genes oncogénicos. Algunos fragmentos de tRNA inducibles por hipoxia (sin observarse preferencia hacia un tracto poli-G en este caso) serían capaces de unir YB-1, desestabilizando así los mensajeros en cuestión. Las células tumorales más metastáticas responderían por su parte inhibiendo la inducción de los fragmentos de tRNA (Goodarzi et al., 2015). Pero no solo el estrés induce a la Angiogenina a generar las mitades de tRNA. Dichos mecanismos también se inducen en células mamarias y prostáticas por acción de estrógenos y andrógenos, respectivamente. Además, estos fragmentos inducidos por hormonas sexuales estimulan la proliferación celular y son abundantes en las pacientes con tumores mamarios positivos para el receptor de estrógeno (Honda et al., 2015).

De considerarse meros sub-productos de degradación sin una función intrínseca, los fragmentos de tRNA se han convertido ahora hasta en una marca epigenética heredable. Dos artículos publicados en *Nature* en enero de 2016 muestran cómo las mitades de tRNA presentes en los espermatozoides se transfieren al cigoto luego de la fecundación, donde silencian la expresión de algunos elementos transponibles (Sharma et al., 2016) pero pueden también transmitir a la prole un síndrome metabólico adquirido (Chen et al., 2016).

Se puede decir que de los fragmentos de tRNA se sabe lo suficiente como para comprender lo poco que se sabe. La fragmentación del tRNA puede estudiarse pero no puede entenderse como un mecanismo aislado de una compleja red de regulación y señalización que involucra a los tRNAs: su re-importación al núcleo y re-exportación al citoplasma, su de-aminoacilación, aminoacilación aberrante, corte de la secuencia CCA, y desde luego, todo el complejo y dinámico sistema de modificación de bases. Mecanismos y redes que exceden al alcance de este trabajo, pero adecuadamente resumidos en (Kirchner and Ignatova, 2015). Un ejemplo ilustrativo es el de los mutantes de NSun2, una enzima capaz de metilar citosinas en la posición 5 (m⁵C), y responsable de la metilación de C34 (anticodón) y de C48 y/o C49 en varios tRNAs. El estado de hipo-metilación provocado por el silenciamiento de NSun2 aumenta la generación de mitades de tRNA mediada por la Angiogenina, lo que provocaría las anomalías neurológicas observadas en los mutantes de Nsun2 en humanos y en ratones (Blanco et al., 2014).

Los tRNAs son en definitiva los grandes transductores de la célula, pues son la propia interfaz entre el mundo de los codones y el mundo de los aminoácidos. No debería sorprender que estén en el centro de una compleja red de señalización y respuesta. Lo extraño sería lo contrario.

El auge de las técnicas de secuenciación profunda ha puesto de manifiesto que, además de los tRNAs, otros RNAs no codificantes para proteínas (ncRNAs: ribosomales, nucleares y nucleolares pequeños, etc.) también presentan enriquecimiento en fragmentos de secuencia y largo determinados no compatibles con degradación al azar (Li et al., 2012). Si bien es siempre un desafío demostrar una función biológica intrínseca para cada fragmento⁴, comienza a quedar claro que el esquema de la Figura 1 no agota la complejidad de las vías mediadas por RNAs pequeños en mamíferos. Además, tiene su lógica pensar que la evolución se haya valido de ncRNAs ya presentes como sustrato de nuevas moléculas, antes que generar las mismas *de novo*. Desde ese punto de vista nada importa si los fragmentos son generados por la maquinaria de degradación natural de los precursores: basta con que algún fragmento así generado se estabilice (por unión a una proteína, por ejemplo) para que sea posible de adquirir nueva función.

C) RNAs pequeños en el medio extracelular:

i. asociados a vesículas extracelulares

La observación de que algunas células eran capaces de expulsar vesículas rodeadas de membrana al medio extracelular data de los años 60, pero durante muchos años estas vesículas se racionalizaron como un mecanismo de excreción de desechos celulares. Sin perjuicio de lo anterior, el tiempo fue ampliando los horizontes funcionales de estas partículas, que hoy se reconocen como actores claves en las vías de comunicación intercelular.

En función de su tamaño y sus mecanismos de biogénesis se distinguen dos grandes grupos de vesículas extracelulares: los exosomas y las microvesículas (denominadas también ectosomas), aunque los métodos de purificación más habituales no permiten una separación a homogeneidad de ambas poblaciones (J. Simpson and Mathivanan, 2012). Los exosomas son pequeñas vesículas de $\approx 40\text{-}150$ nm de diámetro que derivan de los cuerpos multivesiculares (MVB). Estos se forman por invaginaciones “hacia adentro” de la membrana de los endosomas tardíos, formándose una gran vesícula (el MVB) con vesículas más pequeñas en su interior. Cuando los MVB se fusionan con la membrana plasmática de la célula se liberan

⁴los ncRNAs suelen poseer muchas secuencias auto-complementarias, de modo que la introducción exógena de un determinado fragmento puede actuar a altas concentraciones como un competidor que altera la estructura de sus precursores... aunque esa podría ser también una función fisiológica

las vesículas intraluminales, que en el medio extracelular adquieren el nombre de exosomas. Si se repasa su biogénesis se comprenderá que estas partículas transportan por el medio extracelular una muestra del citoplasma de la célula. Aún así, es importante señalar que la formación y el transporte de proteínas hacia el interior de las vesículas intraluminales es un proceso regulado, que involucra al complejo ESCRT (*endosomal sorting complex required for transport*). Tal es así que muchos de los marcadores moleculares que se utilizan para distinguir exosomas son proteínas de dicho complejo (TSG101, Alix). También son marcadores exosomales clásicos algunas proteínas transmembrana de la familia de las tetraspaninas: CD9, CD63 y CD81. En microscopía electrónica de transmisión (tinción negativa) adquieren una morfología característica, que algunos autores describen como de “copa”, aunque el paralelismo es un poco arbitrario. Las microvesículas, en cambio, se producen por protrusión hacia el exterior de la membrana plasmática de las células. Poseen un tamaño mayor y más heterogéneo (100 nm – 1 µm), y una morfología mucho más irregular.

La transformación de los exosomas, de meros surtideros de desechos celulares a intermediarios de comunicación intercelular, comenzó con la elucidación de su rol en la respuesta inmunológica. En 1998 se reportó que los exosomas secretados por células dendríticas poseían en su superficie moléculas de MHC I y MHC II funcionales, así como moléculas co-estimuladoras de linfocitos T. Más aún, los exosomas secretados por células dendríticas pre-estimuladas con péptidos correspondientes a antígenos tumorales eran capaces de activar células T *in vivo* y dirigir una respuesta T-citotóxica efectiva hacia tumores murinos (Zitvogel et al., 1998). Estos y otros resultados, sumados a la detección de exosomas en el plasma de individuos sanos (Caby et al., 2005), fueron moldando la visión de que las vesículas extracelulares, más que desechos, transportan acción e información.

Descubiertos nuevos roles fisiológicos para las vesículas extracelulares, la manipulación de estos mecanismos por las células cancerígenas era algo predecible. Con el paso de los años se han descrito una gran cantidad de ejemplos ilustrativos. Muchos exosomas tumorales expresan en su superficie los ligandos FAS y TRAIL, que provocan la apoptosis de linfocitos T activados, y de este modo contribuyen a burlar la vigilancia inmunológica (Huber et al., 2005). También es conocida la capacidad de los exosomas tumorales de modular la respuesta inmune, favoreciendo la diferenciación de linfocitos T reguladores (Yang and Robbins, 2011). Pueden regular la angiogénesis, modular las células del estroma, o remodelar la matriz extracelular. Permiten además una acción "cooperativa" entre células de un mismo tumor. Por ejemplo, se ha demostrado que, entre los tumores cerebrales agresivos que expresan la forma truncada y oncogénica del receptor EGFR (EGFRvIII), son pocas las células tumorales que poseen la correspondiente mutación en su genoma. No obstante, estas pocas células logran transportar el receptor mutante al resto de las células del tumor, secretándolo en microvesículas que se fusionan con la membrana plasmática de las células receptoras (Al-Nedawi et al., 2008). Imposible omitir los aportes del grupo de David Lyden en lo que respecta al rol de las vesículas extracelulares en el acondicionamiento del nicho pre-metastático (Costa-Silva et al., 2015; Hoshino et al., 2015; Peinado et al., 2012).

Un punto de inflexión ocurrió en el año 2007, cuando investigadores del grupo de Jan Lötvald en Suecia describieron la presencia de mRNAs y miRNAs en los exosomas. En vez de ser meros polizones en un gran transportador de lípidos y proteínas, el RNA reclamaba su cuota parte de protagonismo: los mensajeros eran traducidos y los miRNAs lograban silenciar sus blancos, una vez que las vesículas eran captadas por las células receptoras (Valadi et al., 2007). Poco después se reportó que una comparación de los niveles séricos de tres miRNAs (miR-155, miR-210 y miR-21) permitía distinguir entre pacientes con linfoma difuso de células B grandes (DLBCL) e individuos sanos (Lawrie et al., 2008). Si bien hoy día algunos autores relativizan la eficiencia de los fenómenos de transferencia funcional de miRNAs mediados por exosomas (Kanada et al., 2015), y la cantidad de miRNAs por vesícula parecería ser bastante baja (Chevillet et al., 2014), la “fiebre de los miRNA circulantes” ya se había echado a andar. Aún pese a los reparos anteriores, los años han acumulado evidencia suficiente para sostener el empaquetamiento de pequeños RNAs reguladores en vesículas extracelulares como un mecanismo de comunicación intercelular.

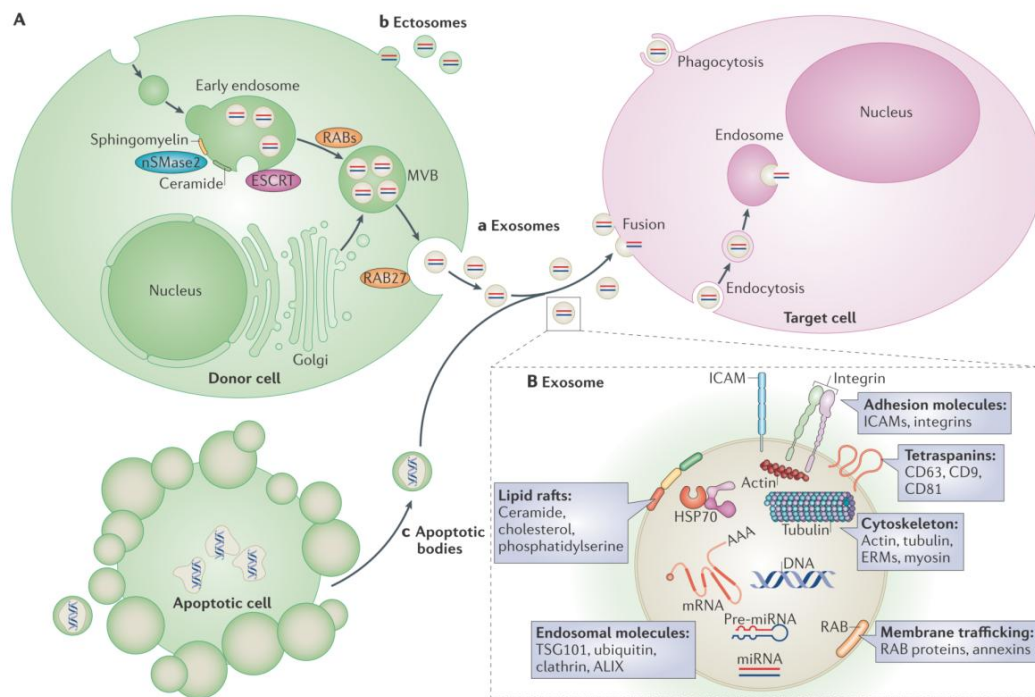


Figura 2: A) Esquema de la biogénesis de exosomas (a), ectosomas o microvesículas (b) y cuerpos apoptóticos (c), y algunas vías posibles de internalización por células receptoras (fusión, endocitosis, fagocitosis). B) Representación esquemática de un exosoma, con algunas de las macromoléculas potencialmente presentes en su interior o superficie. La figura fue extraída y sin modificaciones de (Mittelbrunn and Sanchez-Madrid, 2012).

Las células que ingresan en apoptosis también pueden liberar al medio extracelular vesículas grandes o “apoptosomas” (de diámetros en el entorno de la micra) que contienen DNA fragmentado y otros componentes celulares, entre los cuales se halla el RNA. Por su tamaño, no siempre es fácil separar estos cuerpos apoptóticos de las microvesículas (aunque sí de los exosomas), por lo que es menester monitorear el porcentaje de células apoptóticas cuando se trabaja con cultivos celulares. La Figura 2 muestra un esquema de la biogénesis de los exosomas (a), de las microvesículas (b; aquí llamados “ectosomas”), y de los cuerpos

apoptóticos (c). Faltaría mencionar que también se ha descrito el transporte de miRNAs en la sangre vía HDL (Vickers et al., 2011).

ii. no asociados a vesículas extracelulares

En pleno auge de las vesículas extracelulares como transportadoras de RNA, investigadores del grupo de Muneesh Tewari realizaron una separación cromatográfica del plasma humano, y llegaron a la conclusión de que solo el 10% de los miRNAs circulantes se encontraban realmente asociados a vesículas. El 90% restante viaja en forma de complejos ribonucleoproteicos (RNPs), y asociados a proteínas Argonauta (Arroyo et al., 2011). En simultáneo, Andrey Turchinovich llegaba a resultados similares en Alemania (Turchinovich et al., 2011). Ambos estudios mostraban como, en complejo con proteínas, los pequeños RNA circulantes adquieren una extraordinaria resistencia a la degradación, que les permite permanecer en la sangre mucho tiempo. También la Nucleofosmina 1 ha sido asociada al transporte de miRNAs en el medio extracelular (Wang et al., 2010).

OBJETIVO GENERAL

Contribuir al conocimiento sobre la presencia, rol, modo de secreción, transporte y estabilidad de los pequeños RNAs en el medio extracelular, con énfasis en las nuevas familias de pequeños RNAs reguladores.

OBJETIVOS ESPECÍFICOS

- 1) Estudiar la factibilidad de las vías de comunicación inter-especie mediadas por miRNAs captados a partir de la dieta (*Capítulo I*).
- 2) Obtener por secuenciación profunda una descripción del conjunto de pequeños RNAs reguladores secretados en distintas fracciones extracelulares (microvesículas, exosomas, complejos ribonucleoproteicos) por líneas celulares humanas, y evidenciar procesos de secreción selectiva o preferencial (*Capítulo II*).
- 3) Elucidar mecanismos de estabilización y transporte de pequeños RNAs en el medio extracelular, especialmente para las mitades de tRNA (*Capítulo III*).

CAPÍTULO I

La dieta como fuente posible de miRNAs extracelulares: *¿Comemos información, o nos comimos un artefacto?*

*“La ciencia humana consiste más en destruir errores que en descubrir verdades”. Sócrates
(470-399 A.C.)*

1.1 INTRODUCCIÓN

Cuando comenzamos (comienzos de 2012) a pautar y definir lo que sería más adelante el proyecto de esta tesis doctoral, un artículo publicado en los meses previos captó nuestra atención. Un grupo de investigadores de la Universidad de Nanjing (China) había publicado en la revista *Cell Research* (Zhang et al., 2012a) un reporte que llevaría el campo de los RNAs reguladores extracelulares a límites insospechados. Decidimos, en consecuencia, redefinir el alcance de nuestro proyecto de tesis a los efectos de abarcar también los nuevos horizontes que se abrirían con la publicación de este trabajo.

En el artículo en cuestión los investigadores realizaron un secuenciado masivo de RNAs pequeños en suero humano. Buscaban determinar el perfil de miRNAs circulantes en individuos sanos; seguramente para compararlo con el de algún tipo de pacientes. Lo sorprendente fue el hallazgo de cantidades biológicamente relevantes⁵ de miRNAs de origen vegetal ¿Cómo habrían llegado a la sangre humana? ¿Serían funcionales? El análisis de estas secuencias reveló que la especie más abundante era el osa-miR-168a, uno de los miRNAs más abundantes en el arroz⁶. Como los donantes de sangre eran individuos chinos, y como es sabido que el arroz es un componente fundamental en su dieta, los autores especularon que estos miRNAs exógenos podrían provenir de la ingesta de alimentos.

Si la presencia de RNAs reguladores exógenos en sangre humana era de por sí bastante sorprendente, mucho más lo era la fuente propuesta. Afirmar que los RNAs vegetales podrían ser incorporados a partir de la dieta era algo “difícil de asimilar”, principalmente porque implicaba una violación directa a lo que podríamos llamar el “dogma de la digestión gástrica”: las macromoléculas presentes en los alimentos son degradadas a sus estructuras monoméricas u oligoméricas básicas previo a su absorción intestinal. En este caso, los miRNAs deberían ser capaces de ser absorbidos sin que mediara ningún tipo de degradación; de lo contrario perderían su capacidad de regular la expresión génica.

Conscientes de que tal descubrimiento enfrentaría duras objeciones (o tal vez como respuesta a las objeciones presentadas durante la revisión del artículo), los autores abordaron la cuestión y realizaron algunos controles pertinentes. En primer lugar, los resultados fueron confirmados

⁵5% del total de miRNAs humanos

⁶la secuencia es en realidad específica de las gramíneas, no es posible catalogarla unívocamente como “de arroz”

por PCR cuantitativa y Northern Blot. La presencia de los miRNAs vegetales fue confirmada en el arroz, incluso después de un proceso de cocción. El RNA extraído del arroz fue posteriormente incubado por períodos variables de tiempo en HCl (pH = 2.0) a efectos de emular condiciones gástricas, sin pérdida observable de estabilidad. Según los autores, esta estabilidad a pH fuertemente ácido vendría conferida por el grupo 2'-o-metilo presente en el extremo 3' de los miRNAs vegetales, pues sondas sintéticas carentes de dicha modificación sí eran degradadas. En una serie de experimentos, los investigadores alimentaron ratones bajo dietas controladas libres de arroz, en base a arroz, o libres de arroz, pero con el agregado de miRNAs vegetales sintéticos. Únicamente en los dos últimos casos los ratones presentaron niveles elevados de los miRNAs vegetales miR-168a y miR-156a, tanto en el suero como a nivel hepático.

Los experimentos anteriores permitieron demostrar que al menos algunos mamíferos (seres humanos incluidos) eran capaces de absorber íntegramente miRNAs exógenos ingeridos en la dieta. Dichos miRNAs exógenos viajarían por el torrente sanguíneo y podrían ingresar a tejidos como el hígado. El siguiente paso era estudiar si dichos miRNAs exógenos eran capaces de efectuar una función biológica (es decir, silenciar algunos genes) en, por ejemplo, hepatocitos. Esto ya iría más allá de un simple problema digestivo: implicaría la existencia de mecanismos de regulación génica inter-especie (e inter-reino) mediados por la dieta.

Los autores analizaron la secuencia del miRNA vegetal más abundante en suero humano, osa-miR-168a, y descubrieron que era complementaria a una región del exón 4 del gen *LDLRAP1*, que en humanos codifica para la proteína adaptadora del receptor de lipoproteínas de baja densidad (LDL). Mediante estudios de gen reportero los autores demostraron la capacidad tanto del pre-miR-168a vegetal como del miRNA maduro de unirse a la secuencia predicha de *LDLRAP1* e inhibir la traducción de su mRNA, con la consecuente pérdida de actividad luciferasa en células HepG2 (línea celular de carcinoma hepático). Finalmente, demostraron un efecto *in vivo* ya que los ratones alimentados en base a arroz (o en dietas libres de arroz, pero conteniendo el miR168a sintético) presentaron menores niveles de la proteína LDLRAP1 en el hígado, y concentraciones elevadas de LDL en sangre.

No es la intención resumir aquí exhaustivamente el artículo en cuestión, y no se ha mencionado la inclusión de numerosos experimentos y controles complementarios, que le dan al artículo la solidez científica necesaria para sostener conclusiones tan provocativas y controversiales. Existen, sin embargo, algunos detalles técnicos y experimentales que es necesario mencionar por las implicancias que tendrían posteriormente en nuestro propio trabajo. La detección del osa-miR-168a en suero humano fue realizada mediante secuenciado masivo, y cabría la posibilidad de que se tratara en realidad de una secuencia humana modificada, que coincidiría aleatoriamente con un miRNA del arroz. Si bien hay varios argumentos en contra de esta objeción, comenzando por la alta cantidad de *mismatches*⁷ contra la secuencia humana más parecida, los autores recurrieron a una elegante prueba bioquímica.

⁷**Mismatch:** en inglés, término que se utiliza para indicar los sitios en los que, al alinear dos secuencias nucleotídicas, las bases correspondientes en cada una de ellas no coinciden.

A diferencia de lo que ocurre en mamíferos, los miRNAs vegetales poseen un grupo 2'-o-metilo en su extremo 3'. Aprovechando esta diferencia ya no solamente de secuencia sino también química, los investigadores trataron sus muestras de RNA con periodato de sodio (NaIO_4) previo a la secuenciación masiva. El anión periodato es capaz de oxidar dioles vecinos (tales como los encontrados en la posición 2 y 3 de las pentosas presentes en el extremo 3' del RNA) a aldehídos (Figura 3). Consecuentemente, un RNA tratado con periodato no es pasible de ser secuenciado por protocolos convencionales, basados en la ligación de un adaptador 3', para lo cual es necesario contar con un extremo 3'-OH en el RNA. Los RNAs modificados con un grupo 2'-o-metilo en su extremo 3' son resistentes al tratamiento con periodato y pueden, por tanto, secuenciarse.

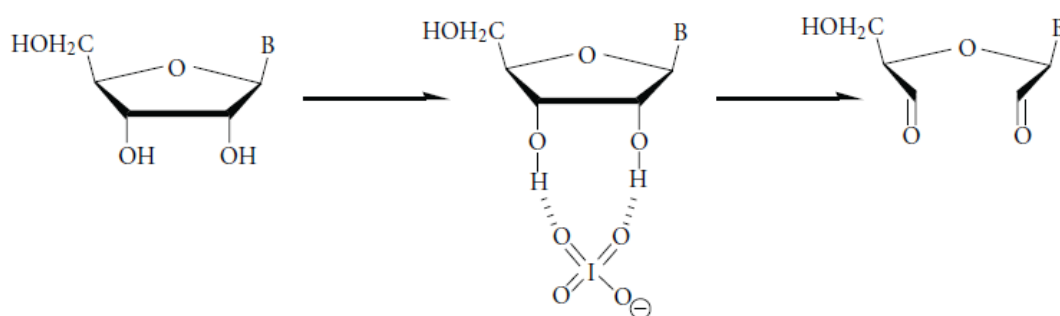


Figura3: diagrama de la reacción entre el anión periodato (IO_4^-) y los dioles vecinales presentes en ribonucleósidos, o en el extremo 3' del RNA. La pérdida del grupo 3'-OH evita que puedan ligarse adaptadores con extremo 5'-P al RNA tratado con periodato. Esto no ocurre en el caso de RNAs metilados en su extremo 3', pues la reacción requiere que los carbonos 2 y 3 estén en forma alcohólica.

En las muestras tratadas con periodato sódico, la abundancia relativa de miRNAs humanos decreció notoria y significativamente. En cambio, los niveles de los miRNAs vegetales previamente descritos permanecieron inalterados. Este elegante experimento echaba por tierra todo posible reparo sobre el origen no humano de estas secuencias exógenas: no solamente eran miRNAs vegetales por su secuencia nucleotídica, sino que también poseían marcas bioquímicas específicas.

Las repercusiones de tal controversial, pero bien fundado reporte no se hicieron esperar. En cuestión de meses ya habían sido publicados varios artículos de revisión con títulos provocadores como: “¿Comemos reguladores génicos?”, “Regulación de la expresión génica en mamíferos por miRNAs exógenos” o incluso: “más allá de los nutrientes: los miRNAs de los alimentos provocan regulación génica entre distintos reinos” (Jiang et al., 2012; Liang et al., 2012; Witzany, 2012). La publicación tuvo un impacto inmediato y global, y fue galardonada con el premio al mejor artículo de investigación de la revista del año 2012 (Li, 2013). Sin embargo, fue en el campo social, y particularmente en la llamada “blogósfera”, donde las repercusiones llegaron más lejos. No se necesitaba demasiada agudeza para predecir que la existencia de mecanismos fisiológicos para captar RNAs reguladores activos desde la dieta reforzaría el debate sobre la inocuidad o toxicidad de los alimentos genéticamente

modificados (máxime cuando muchos eventos transgénicos hoy día se basan en mecanismos de RNAi). *Scientific American* tituló: “*nuestra comida puede regular nuestros genes*”⁸. Esto a su vez dio pie a que la revista estadounidense “The Atlantic” publicase un artículo sobre “*El muy real riesgo de los alimentos genéticamente modificados*” (que fue luego enmendado para la versión web definitiva⁹), donde se habla de que además de proteínas y vitaminas, la dieta es fuente de información génica. La organización internacional *Greenpeace* publicó un informe técnico haciéndose eco de los hallazgos de Zhang¹⁰, así como numerosos portales de internet especializados en temas de ciencia, salud y/o alimentación.

Si nos detenemos en lo anterior no es para entrar en un debate que no hace al fondo de esta tesis, sino para rescatar un buen ejemplo de rápida apropiación social del conocimiento científico. Algo sumamente deseable en la medida que se logren poner de manifiesto las limitaciones intrínsecas de los estudios realizados, de modo tal que el debate pueda apoyarse en la evidencia y no termine haciendo de la ciencia un simple recurso legitimador. Es responsabilidad de la comunidad científica clarificar estos límites, y remarcar el concepto de que la “verdad científica” es una construcción que se asienta en la reproducibilidad.

Nuestro laboratorio ya venía trabajando en el campo de la comunicación entre organismos diferentes mediada por RNA, particularmente en el modelo hésped-parásito (Garcia-Silva et al., 2014a; Garcia-Silva et al., 2014b)¹¹. El trabajo de Zhang extendía estas vías a la comunicación entre nuestras células y la comida que ingerimos. De ser cierto, se abrirían interesantes posibilidades experimentales tales como la introducción sistémica de RNAs exógenos por vía oral en mamíferos, algo que ya era posible hacer en nemátodos (Conte and Mello, 2003).

Nuestra idea fue sumamente simple: si los miRNAs vegetales podían pasar a los tejidos humanos a través de la dieta, deberíamos poder encontrar miRNAs vegetales en las bases públicas de datos de secuenciación profunda de tejidos humanos. Cuando estas tecnologías alcanzaron un grado de desarrollo aceptable, numerosos investigadores comenzaron a secuenciar distintos tejidos, en trabajos meramente exploratorios. Si bien el interés de estos investigadores estaba centrado en la detección de miRNAs humanos, no habría razón para pensar que las secuencias vegetales – de estar presentes – no fuesen secuenciadas. Aquí entra una de las características del análisis de datos ómicos y de gran escala, y en particular, de la secuenciación profunda: es – casi – imposible analizar la totalidad de los datos. Si un estudio fue diseñado para el análisis de secuencias humanas, es probable que los investigadores hayan comenzado por hacer un mapeo de los datos al genoma humano, desechando todas las secuencias que no mapearan. A la hora de depositar los datos en servidores públicos, sin embargo, se suelen depositar los datos crudos, incluyendo las secuencias analizadas y las secuencias descartadas. Allí, pensamos, deberíamos encontrar miRNAs vegetales.

⁸<http://www.scientificamerican.com/article/vitamins-minerals-and-microrna> Visitada: 03/11/2015

⁹<http://www.theatlantic.com/health/archive/2012/01/the-very-real-danger-of-genetically-modified-foods/251051/> Visitada: 03/11/2015

¹⁰[http://www.greenpeace.org/australia/Global/australia/GM_Fact_Sheets/Read%20the%20full%20paper%20\(377kb\).pdf](http://www.greenpeace.org/australia/Global/australia/GM_Fact_Sheets/Read%20the%20full%20paper%20(377kb).pdf) Visitada: 03/11/2015

¹¹ En estos trabajos previos de nuestro grupo, se demostró que *T. cruzi* es capaz de secretar pequeños RNAs reguladores que luego son captados por células de mamífero susceptibles a la infección.

1.2 MATERIALES Y MÉTODOS

1.2.1. Reactivos

El reactivo TRI (análogo del Trizol), 2-propanol (>99%) y m-periodato de sodio fueron obtenidos de Sigma-Aldrich. Las librerías para secuenciación profunda fueron realizadas con el kit “*NEBNext Small RNA Library Prep Set for Illumina*”, de New England Biolabs. La transcriptasa reversa (Superscript II) fue de Invitrogen (Life Technologies). Las reacciones de secuenciación fueron realizadas con el *TruSeq SBS kit v5-GA* (Illumina) en un secuenciador Illumina Genome Analyzer IIx.

1.2.2. Meta-análisis de RNAs exógenos en tejidos humanos

Se descargaron datos de secuenciación profunda de RNAs pequeños obtenidos con tecnología Illumina como archivos *.fastq sin procesar. Los números de acceso de los datos analizadas refieren a la base SRA (*Sequence Read Archive*): SRP002347, SRP002272, SRP005294, SRP007825 y SRP002402. Una descripción detallada se brinda en el capítulo de resultados.

1.2.3. Secuenciación profunda de RNAs pequeños en espermatozoides

Muestras de espermatozoides humanos provenientes de tres donantes fértiles fueron obtenidas y purificadas por la clínica Fertilab de Montevideo. Los donantes otorgaron su consentimiento informado para participar en el estudio, y dichos consentimientos informados fueron aprobados por la Comisión de Ética del Hospital de Clínicas, acorde a los procedimientos establecidos para el trabajo con muestras humanas en nuestro medio. Las muestras de esperma fueron sometidas a un gradiente discontinuo de 90%, 70% y 45% (SpermGrad, Vitrolife) preparado en EBSS (*Earl's Balanced Salt Solution*). Las muestras fueron depositadas encima del gradiente y centrifugadas a 600g por 20 minutos. Los *pellet* fueron resuspendidos en 2mL de EBSS y centrifugados a 1200g por 5 minutos, luego de lo cual fueron resuspendidos en 0.5mL EBSS y separados en grupos según su motilidad y viabilidad.

Se partió de treinta millones de células del grupo de alta motilidad y viabilidad en cada caso, las cuales fueron lisadas en 0.5mL de reactivo TRI. Se extrajo el RNA en acuerdo a las especificaciones del fabricante, duplicando el volumen de isopropanol y lavando con etanol 80% a efectos de minimizar la pérdida de RNAs pequeños. Las muestras de RNA se resuspendieron en 80μL. Alícuotas de la mitad de este volumen fueron tratadas con NaIO₄ por 40 minutos a 4°C y en oscuridad (volumen final: 100 μL), manteniendo la relación: 10mM NaIO₄ cada 20μg RNA. El RNA tratado y sin tratar fue precipitado con etanol dos veces, resuspendido en agua, cuantificado por Qubit (Life Technologies) y sometido a secuenciación profunda.

El armado de las librerías se realizó acorde a las especificaciones del fabricante (New England Biolabs). Todas las reacciones se realizaron utilizando reactivos libres de nucleasas y tips con filtro. Brevemente, se ligaron adaptadores 3' y 5', el RNA se retro-transcribió, y se agregaron los índices para las reacciones multiplex (el “código de barras”) durante la amplificación por PCR (15 ciclos), utilizando *primers* complementarios a los adaptadores. Las librerías se corrieron en geles de poliacrilamida al 6% en *buffer* TBE (Tris-borato-EDTA), y se seleccionaron por tamaños en la región < 60 nucleótidos. Las librerías purificadas se analizaron luego utilizando un Bioanalyzer 2100 (Agilent), y se cuantificaron por espectrofluorimetría (Qubit, Life Technologies).

1.2.4. Análisis bioinformático

Se utilizó el mismo procedimiento independientemente del origen de los datos (muestras secuenciadas por nosotros o extraídas del SRA). En primera instancia se buscó y eliminó la secuencia correspondiente al adaptador 3', reteniendo únicamente aquellas lecturas que poseían dicha secuencia, y un largo remanente mayor a 15. Las secuencias obtenidas se alinearon al genoma humano (hg19) utilizando el programa Bowtie (semilla: 19, permitiendo un miss-match en la semilla, independientemente de la calidad de la base; todas las demás opciones por defecto) incluido en la interfaz Galaxy (<https://galaxyproject.org>). Se colapsaron las secuencias no alineadas, y todas aquellas con una cantidad de lecturas mayor o igual a 3 se sometieron a un segundo ciclo de mapeo al genoma humano utilizando un algoritmo distinto (Lastz).

Aquellas secuencias que no pudieron ser alineadas al genoma humano por ninguno de los dos métodos fueron clasificadas como “no alineadas”. De ellas, solamente se analizaron aquellas cuya representación fuese superior al 0.01% del total de secuencias “no alineadas”. Dichas secuencias fueron alineadas a la base nucleotídica no redundante (nr/nt) utilizando Blast (blast.ncbi.nlm.nih.gov), restringiendo la búsqueda a los mamíferos. Se eliminaron todas aquellas secuencias con menos de cuatro *mismatches* a cualquier secuencia de mamífero. Esto fue así dado que las diferencias inter-especie suelen ser mayores que los polimorfismos dentro de una misma especie. Por consiguiente, al eliminar toda secuencia con similitud a un mamífero se minimiza la posibilidad de que secuencias humanas polimórficas fuesen interpretadas erróneamente como exógenas. El costo de este criterio conservador es la pérdida total de la capacidad de evidenciar secuencias auténticamente exógenas provenientes de mamíferos.

Las secuencias no alineadas al genoma humano ni con similitud evidente al genoma de ningún mamífero fueron vueltas a alinear por Blast contra toda la base nr/nt. En esta oportunidad se retuvieron únicamente aquellas secuencias con un 100% de identidad en toda su extensión a algún genoma (salvo de mamífero). Este grupo constituyó el de las secuencias “exógenas putativas”. Cada una de estas secuencias fue anotada manualmente en función de su especie de origen predicha (o género, o familia, o clase; hasta encontrar una descripción unívoca) y del tipo de RNA del cual presumiblemente derivaría.

En la mayoría de los casos, la abundancia relativa de una determinada secuencia fue expresada como RPM (*reads per million*; lecturas en un millón de lecturas presentes en el secuenciado en cuestión). Para el análisis del porcentaje de secuencias exógenas respecto del total de miRNAs humanos, se identificaron estos últimos por alineamiento al miRBase (versión 20), permitiendo un *mismatch*.

Como método de verificación alternativo, las mismas series de datos fueron alineadas por Blast contra todas las especies contenidas en el miRBase (versión 20), y las secuencias correspondientes a miRNAs vegetales y humanos fueron extraídas. No hubo discrepancias entre ambas estrategias de análisis.

1.3 RESULTADOS

Se seleccionaron aleatoriamente 5 estudios que reuniesen las siguientes condiciones: i) secuenciación masiva de pequeños RNAs en distintos tejidos humanos, ii) plataforma Illumina, iii) datos crudos y métodos de armado de librerías disponibles, iv) al menos una de las muestras analizadas debería corresponder a tejidos en condiciones no patológicas, v) datos recientes (a 2012). Un primer estudio (SRA: SRP002347; (Faghihi et al., 2010) secuenció el contenido de RNAs pequeños en pulmón [LUN], corteza orbitofrontal [ORB], páncreas [PAN], bazo [SPL], riñón [KID], corazón y músculo de dos donantes cadavéricos. El segundo trabajo analizado (SRP002272:(Hou et al., 2011) correspondía a muestras de hígado en 5 donantes sanos [LIV], una muestra hepática de un paciente infectado con hepatitis B [HBV], y una muestra hepática de un paciente con hepatitis crónica severa [SCH]. De un tercer trabajo (SRP005294: (Schotte et al., 2011) se estudió médula ósea [B-M], linfocitos CD34+ [CD34], y timocitos [THYM]. Dos muestras de piel fueron incluidas también (SRP007825: (Joyce et al., 2011). Por último, y fortuitamente, se incluyeron 3 muestras de espermatozoides purificados provenientes de donantes sanos (SRP002402:(Krawetz et al., 2011).

Como se describe en el apartado metodológico, los *reads* de cada estudio se mapearon al genoma humano mediante dos métodos secuenciales, con parámetros lo suficientemente laxos como para que solo quedaran fuera las secuencias considerablemente distintas a toda secuencia humana a nivel genómico. Luego de sucesivos pasos de análisis descritos en la sección metodológica, se arribó a un catálogo de secuencias “exógenas putativas” que se presenta en la Figura 4 (clasificando según origen predicho) y Figura 5 (clasificando según qué tipo de RNA pequeño o fragmento de RNA representarían).

Como puede observarse, en casi todos los datos analizados aparece una predominancia del miR-168a vegetal. Esta secuencia es común a casi todas las gramíneas y otras plantas monocotiledóneas (que constituyen un grupo de plantas muy consumidas por los seres humanos: arroz, maíz, trigo, etc.) pero tiene dos bases de diferencia con la secuencia del miR-168a de plantas dicotiledóneas. Estos resultados nos entusiasmaron mucho, pues era lo esperado de acuerdo a nuestra hipótesis de trabajo: *si la transmisión de miRNAs vegetales fuese real, deberíamos encontrarlos en el vasto universo de datos de secuenciación masiva descartados.*

Además del miR-168a vegetal se detectaron otras clases o tipos de secuencias exógenas putativas ubicuamente representadas. No debería sorprender la presencia de secuencias bacterianas, pues estas pueden haber llegado a los datos por diversas razones: fenómenos de transferencia a través de la dieta o de la flora bacteriana, pero también mecanismos menos interesantes desde un punto de vista biológico – pero no por eso menos reales – tales como la contaminación bacteriana de las muestras o del material de laboratorio.

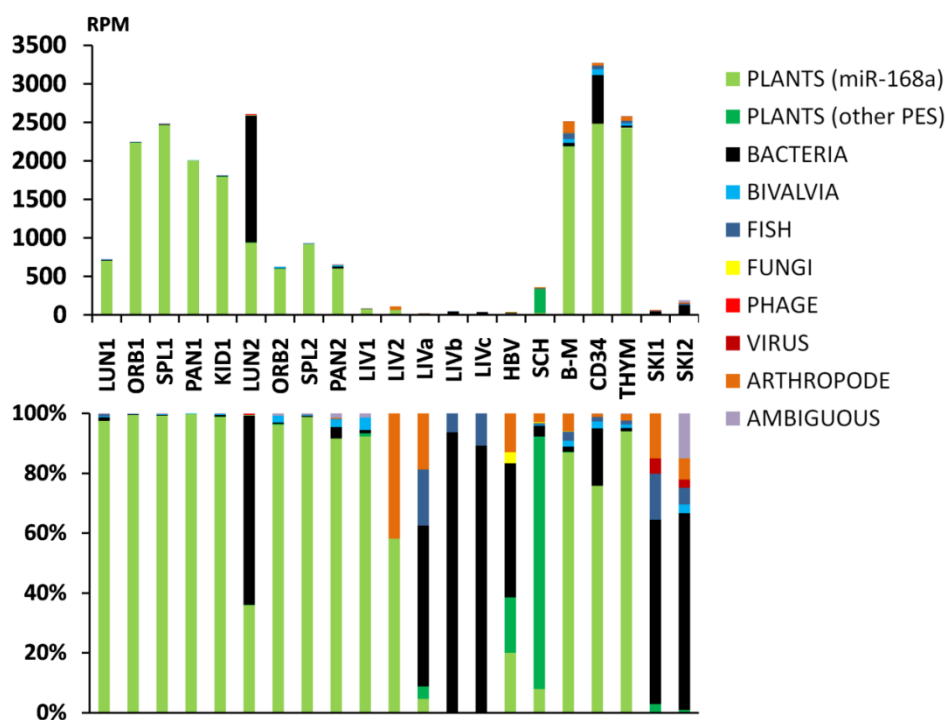


Figura 4: clasificación de las secuencias exógenas putativas detectadas en muestras humanas según su origen. El panel superior muestra los resultados expresados como RPM (lecturas en un millón de lecturas procesadas en la muestra correspondiente). El panel inferior muestra los porcentajes de cada origen en el total de secuencias exógenas putativas. Dado que la mayoría de las secuencias exógenas putativas de origen vegetal corresponden a osa-miR-168a, se habilitó un color especial (verde claro) para dichas secuencias, mientras que el verde oscuro corresponde a todo el resto de las secuencias vegetales. PES: putative exogenous sequences.

Más desconcertante eran las secuencias de bivalvos, encontradas en varios de los datos analizados, e incluso provenientes de distintos grupos de investigadores. Se detectaron diversos fragmentos de RNA ribosomal, algunos de un largo de hasta 30 nucleótidos, lo que permitió descartar que se tratase de alineamientos casuales. Algunas de las secuencias eran unívocamente derivadas de una determinada especie de mejillones: *Mytilus galloprovincialis* o *Mytilus edulis*. Estas secuencias aparecían en muestras donde era desafiante explicar su presencia, tales como el cerebro, los espermatozoides, y el páncreas (Figura 4). Es verdad que los mejillones forman parte de la dieta humana, pero de ahí a que su RNA pudiese atravesar tanto la barrera digestiva como la hemato-encefálica había un largo trecho. Por fin, años más tarde, se reveló el misterio al comprobar que casi todo el glucógeno comercial usado en procesos de precipitación y aislamiento de ácidos nucleicos proviene de los mejillones. Pero

para ese entonces ya quedaba claro que la gran sensibilidad de las técnicas de secuenciación profunda provocaba la secuenciación de toda clase de contaminantes.

En las Figuras 4 y 5 se omitió el análisis de los datos provenientes del estudio SRA: SRP002402, correspondientes a tres muestras de espermatozoides humanos provenientes de donantes fértiles. Esto fue así para evitar problemas de escala, ya que la proporción de secuencias exógenas en dichas muestras fue extremadamente alta. Por ejemplo, los niveles del miRNA miR-168a de gramíneas oscilaron típicamente por debajo del 1% del total de miRNAs humanos en casi todas las muestras (Figura 6, a); niveles similares a los reportados en el artículo original de Zhang (Zhang et al., 2012a). Lo mismo vale para el resto de las secuencias exógenas putativas. Por el contrario, los niveles del miR-168a vegetal llegaron a representar cerca del 1200% de los miRNAs humanos en una de las muestras de espermatozoides, y el resto de las secuencias exógenas varió en el intervalo 500-3000% (Figura 6, b).

Interpretar estos resultados a la luz de la hipótesis de que la presencia de miRNAs vegetales es consecuencia directa de la ingesta de los mismos, era una tarea difícil. Durante mucho tiempo se creyó que el DNA nuclear era el único tipo de material genético transferido por un espermatozoide a un óvulo durante la fecundación. Recientemente se ha descrito como algunos miRNAs provenientes del espermatozoide también pasan al cigoto, y son requeridos para las primeras divisiones celulares del nuevo individuo (Liu et al., 2012). ¿Cabría esperar un rol para los miRNAs vegetales – aparentemente mucho más abundantes - en la fecundación humana?

La pregunta anterior nos resultaba ciertamente inquietante, pero ese sentimiento convivía con la percepción – basada en el sentido común – de que tener a nivel de gametos doce veces más secuencias vegetales que la suma de miRNAs humanos era algo demasiado asombroso como para ser verdad. Difícil no pensar en un artefacto. En cualquier caso, evidencia empírica era lo que estaba faltando.

En colaboración con el Dr. Montes del laboratorio clínico Fertilab, obtuvimos – previa aprobación del comité de ética correspondiente – tres muestras de espermatozoides humanos altamente purificados, provenientes de tres donantes uruguayos. Nuestro objetivo fue realizar secuenciado profundo de pequeños RNAs en dichas muestras, y buscar posibles RNAs exógenos, particularmente vegetales. Adicionalmente, el RNA extraído de las muestras fue tratado con y sin NaIO₄, de modo tal de evidenciar si las secuencias vegetales detectadas poseían un grupo 2'-O-metilo en su extremo 3'.

Las tres muestras de espermatozoides secuenciadas en el presente trabajo no fueron obtenidas ni procesadas de manera simultánea. Nuestra intención inicial era estudiar una única muestra, que fue la aquí denominada como SPE4 (y su correspondiente contraparte tratada con periodato de sodio). En aquellos momentos, el único “secuenciador de nueva generación” disponible en el país era el Illumina Genome Analyzer IIx que el Institut Pasteur de Montevideo había adquirido. Las características de este equipo obligaban a que un gran número de muestras fuesen corridas en paralelo, de modo tal de amortizar los costos básicos de funcionamiento. En una de las primeras experiencias de secuenciación profunda realizadas

en el Instituto, construimos junto al Dr. Gonzalo Greif las librerías correspondientes a nuestras muestras, y a un gran número de muestras de diferentes usuarios centralizadas por la Unidad de Biología Molecular del IPMon. Enseguida se explicará por qué esta información anecdótica resulta relevante.

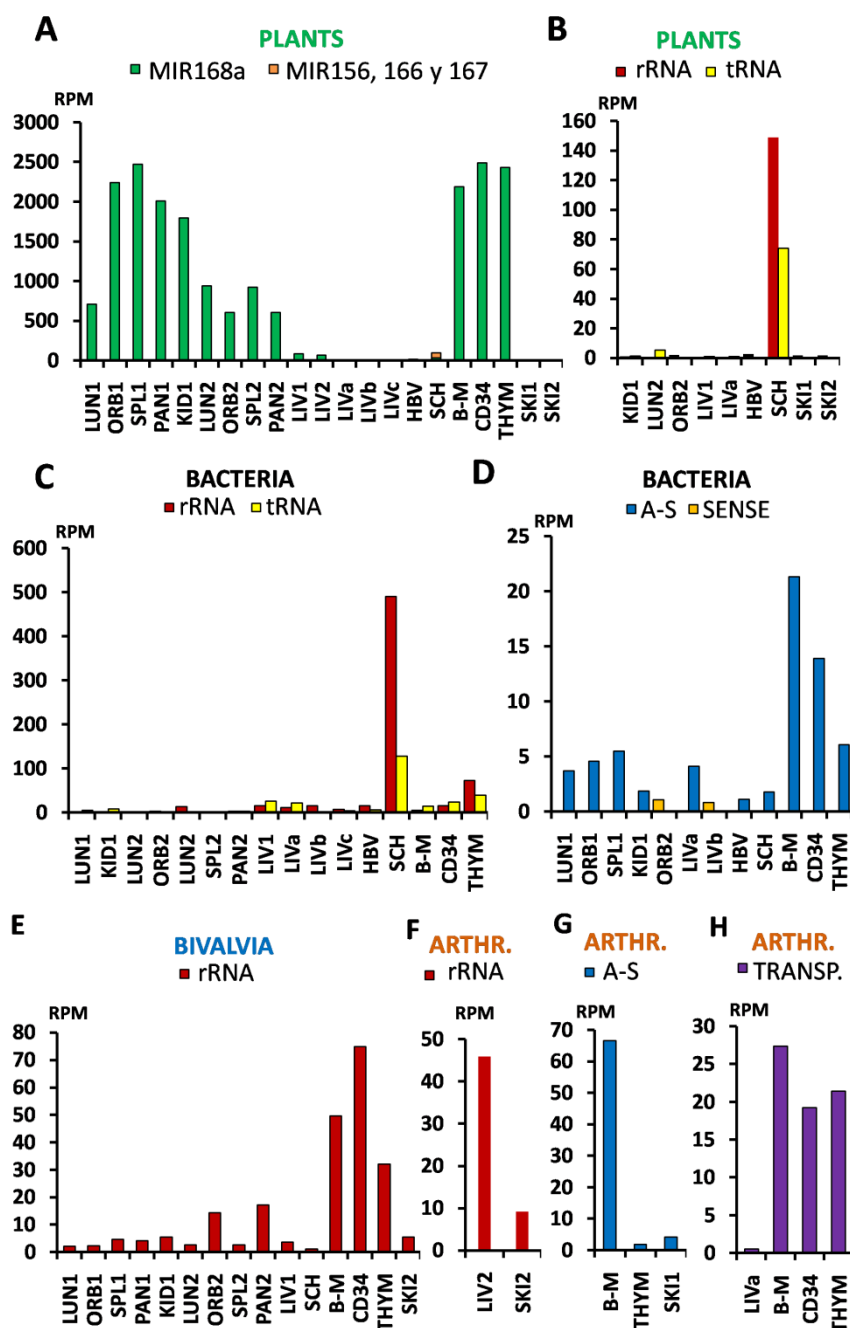


Figura 5: clasificación de las secuencias exógenas putativas detectadas en muestras humanas según el tipo de RNA del cual presumiblemente provienen. Solamente se representan aquellas muestras donde el número de lecturas correspondientes a las categorías graficadas fue mayor a cero. A.S: secuencias anti-sentido a mRNAs. SENSE: secuencias sentido a mRNAs. TRANSP: secuencias correspondientes a elementos transponibles anotados en el genoma.

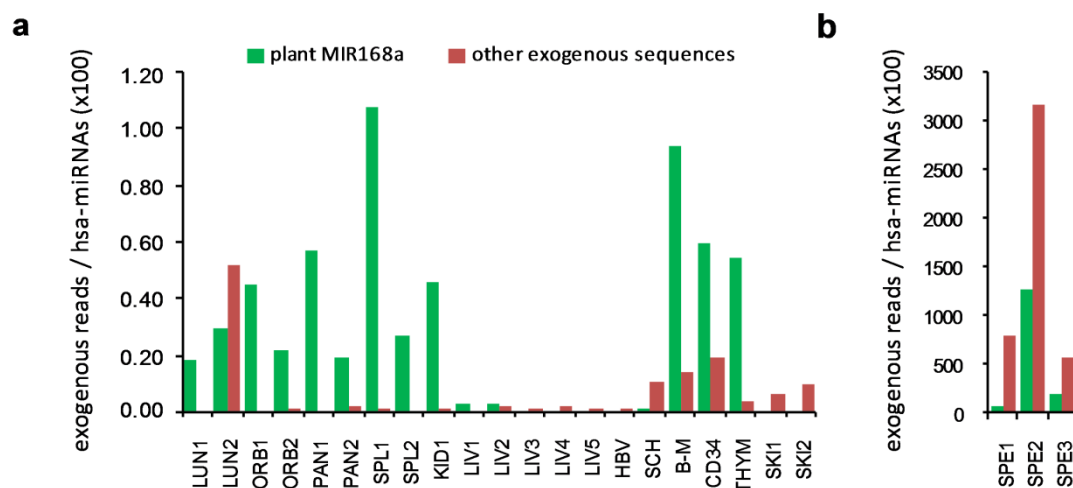


Figura 6:(A) Porcentaje del total de miRNAs humanos que representan los niveles de osa-miR-168a (verde) o el total de las demás secuencias exógenas putativas (rojo). (B) Las muestras SPE1-SPE3 (espermatozoides humanos) se representan en escala independiente.

Contrario a lo observado en los datos del equipo estadounidense (SRA: SRP002402), no se detectó ninguna secuencia correspondiente al miR-168a vegetal en los nuestros, ya sea con o sin tratamiento con periodato. Más aún, los niveles de secuencias vegetales en general fueron sumamente bajos (< 10 RPM), y ninguna de ellas correspondía a miRNAs. Es así que teníamos muestras de espermatozoides secuenciadas por distintos autores pero con los mismos métodos, y mientras que unas contenían cantidades abrumadoras de secuencias vegetales, las otras tenían niveles prácticamente nulos.

Algo intrigante ocurrió a la hora de analizar los miRNAs humanos. Como ya había sido descrito (Krawetz et al., 2011; Liu et al., 2012), se comprobó que el miR-34c-5p figuraba entre los más abundantes en espermatozoides, de algún modo apoyando la validez de los datos. Curiosamente, el perfil de miRNAs estaba completamente alterado en la muestra tratada con periodato. Los niveles de los tres miRNAs más abundantes en la muestra original eran muy bajos (miR-378a-3p, miR-769-5p y miR-34c-5p), mientras que otros como el miR-143-3p, y el miR-126-3p aparecían encabezando la lista. Esto era algo inesperado, pues la oxidación con periodato debería afectar por igual a todos los miRNAs humanos, salvo que alguno de ellos estuviese – y hubiese sido algo novedoso – modificado en su extremo 3'. Pero a esta altura ya todo parecía ser un artefacto.

Consideramos que una explicación posible podría ser que el periodato sódico hubiese oxidado a todos los miRNAs humanos volviéndolos no aptos para la secuenciación, y que la presencia del miR-143-3p y del miR-126-3p se debiese a un fenómeno de contaminación posterior al tratamiento¹². Pensando en fuentes posibles de contaminación, dirigimos la mirada a las otras

¹² En el primer secuenciado de exosomas que realizamos en paralelo a estas muestras, la casi totalidad de los miRNAs correspondían a miR-143-3p y miR-126-3p, lo cual era una coincidencia demasiado grande. Vale aclarar que dichos datos no fueron utilizados para el análisis realizado en el capítulo II de esta Tesis.

librerías¹³ construidas ese mismo día. La memoria indicaba que muchas de las muestras procesadas en paralelo con las nuestras correspondían a RNA de pulmón de rata. Solicitamos entonces a los investigadores pertinentes una lista de los 5 miRNAs más abundantes en dichas muestras. Como se puede apreciar en la Figura 7, esta lista fue un reflejo exacto de los miRNAs “humanos” más abundantes en SPE4 + NaIO₄. Cabe señalar que las secuencias humanas y de rata para estos miRNAs son exactamente iguales, dado el alto grado de conservación que por lo general presenta este tipo de genes.

Aunque parecía claro, decidimos desafiar la hipótesis de que lo que estábamos observando era una consecuencia directa de la contaminación cruzada con otras muestras procesadas en paralelo el mismo día. Además de muestras de rata, había también unas muestras de tejido de tortuga provenientes del Instituto de Investigaciones Biológicas Clemente Estable (IIBCE). Y si bien no encontramos miRNAs vegetales, pudimos detectar 9 secuencias distintas (29 reads; 11.5 RPM) correspondientes a fragmentos de RNA ribosomal de tortuga (género *Trachemys*) en la muestra SPE4. Más aún, 12 secuencias distintas de tortuga (77 reads, 108 RPM) fueron detectadas en SPE 4 + NaIO₄.

Llevamos la experiencia a n=3 sumando dos nuevas muestras de espermatozoides humanos (SPE 5 y SPE 6, y sus contrapartes tratadas con NaIO₄), pero extremando precauciones a efectos de minimizar la posibilidad de contaminación cruzada durante el armado de las librerías. En concreto, las muestras fueron procesadas aisladamente, en un laboratorio distinto, con material de laboratorio diferente, y adjudicando un “código de barra” distinto a cada muestra. Como cabía esperar, no se detectaron secuencias de tortuga en ninguna de las 4 muestras, ni aparecieron los miRNAs de rata entre los más abundantes luego de oxidar el RNA con periodato de sodio.

Volviendo al objetivo inicial, tampoco se encontraron en ninguna de estas muestras miRNAs vegetales. Ni siquiera el tratamiento con periodato de sodio – que debería amplificar la señal de los RNAs protegidos por metilación – permitió la detección de estas especies. Y no hay razón para dudar de la efectividad de dicho tratamiento ya que la proporción de piRNAs humanos aumentó significativamente ($P < 0.05$; prueba t a una cola en medidas no pareadas) en las muestras tratadas con periodato, consistente con el hecho de que los piRNAs están 2'-O-metilados en su extremo 3' (Ohara et al., 2007) (Figura 8).

Llegado a este punto, la evidencia acumulada permitía inferir o argumentar que los resultados del artículo de *Cell Research* (Zhang et al., 2012a) pudiesen ser de naturaleza artefactual. Esta afirmación podía sustentarse en las siguientes consideraciones:

- 1) La presencia de miRNA vegetales, y de osa-miR-168a en particular, fue ubicua en nuestro meta-análisis de muestras humanas. Ahora bien, dichas secuencias no estaban restringidas a la sangre o al hígado, sino que podían detectarse también en órganos como el cerebro y el páncreas.

¹³ De ADNc amplificado. Dado que el Illumina GAIIx posee 8 carriles en el secuenciador, los códigos de barra se pueden repetir siempre y cuando muestras con igual código vayan a parar a carriles diferentes.

- 2) Si tal presencia obedecía a mecanismos biológicos de captación a partir de la dieta, dichos mecanismos deberían ser eficientes para dirigir secuencias exógenas hacia los espermatozoides.
- 3) Y sin embargo, las muestras de espermatozoides secuenciadas por nuestro grupo (en un instituto donde no se trabaja con plantas) poseían cantidades nulas de miRNAs vegetales, en abierta contradicción con lo anterior.
- 4) Así mismo, mostramos cómo la contaminación cruzada durante el armado de librerías para secuenciación profunda era un mecanismo posible y válido para explicar la presencia de secuencias exógenas en muestras humanas.
- 5) Y todo esto sumado a los estudios sobre no reproducibilidad de los resultados de Zhang et al. (2012), que por esa época comenzaban a publicarse.

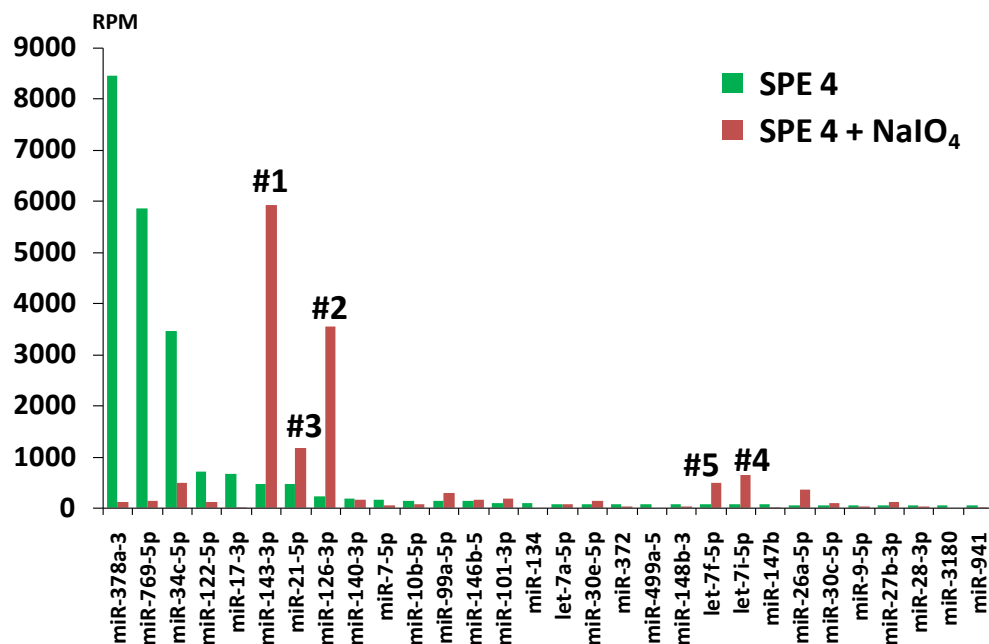


Figura 7: abundancia relativa (expresada como RPM) de los principales miRNAs en la muestra SPE4 (verde) y en la misma muestra tratada con periodato sódico (rojo). Por encima de las barras rojas se indica la posición de los respectivos miRNAs en un ranking decreciente (#1 correspondiente al miRNA más abundante) de miRNAs en muestras de tejido pulmonar de rata procesadas y corridas en paralelo.

Lo anterior podría ser suficiente para dudar de la realidad de los supuestos mecanismos de captación de miRNAs a partir de la dieta, dado que la contaminación aparecía como una interpretación alternativa, tal vez más razonable, de los mismos resultados. Pero ¿cómo demostrar que la presencia de miRNAs vegetales en los trabajos de Zhang et al. se debía, efectivamente, a contaminación cruzada? Por fin surgió una idea prometedora: recurrir una vez más a la minería de datos bioinformáticos ¿Qué más había publicado el grupo en cuestión en los últimos años? Tal vez allí encontraríamos la respuesta.

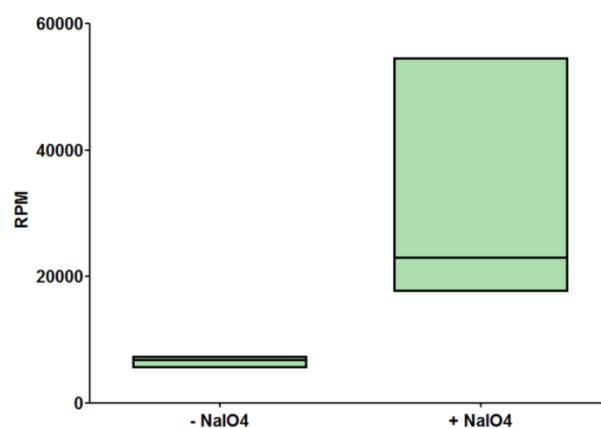


Figura 8: abundancia relativa (expresada como RPM) del total de piRNA humanos detectables en las muestras SPE4, SPE5 y SPE6 secuenciadas por nosotros, ya sea con (derecha) o sin (izquierda) tratamiento del RNA con periodato sódico.

Una búsqueda en PubMed nos puso en conocimiento de la publicación: “*identificación y caracterización de nuevos miRNAs en anfibios*”, que los mismos autores habían publicado en la revista *Genome Biology* en el año 2009¹⁴ (Chen et al., 2009). Los anfibios son pequeños animales del grupo de los Cefalocordados que viven enterrados en la arena de los mares templados. Desde luego, no se alimentan de gramíneas como el arroz o el trigo sino que, como se indica en el apartado metodológico de dicho artículo, fueron alimentados en base a algas. Por consiguiente, no cabría esperar la presencia del osa-miR-168a en estos organismos. Afortunadamente, y a diferencia del artículo de *Cell Research*, en esta oportunidad los datos crudos de la secuenciación estaban disponibles (SRA: SRP001014).

De un total de 1387630 lecturas totales correspondientes al anfibio *Branchiostoma belcheri*, encontramos 86 *reads* (61 RPM) correspondientes al osa-miR-168a. Verificamos que esta secuencia no tuviese ninguna similitud en el genoma de *Branchiostoma*, ni tampoco en los genomas de las algas utilizadas para alimentarlo.

Además del miR-168a, el trabajo de Zhang et al. (2012) había reportado la presencia de otros miRNAs vegetales. Ordenados según su abundancia relativa, como puede extraerse de la Tabla Suplementaria 1 de dicho artículo, estos eran: osa-miR168a, ath-miR156a, ath-miR167a, ath-miR166a, ath-miR169b y ath-miR172a, entre otros. En los datos de anfibio de los mismos autores, pudimos encontrar 86, 43, 23, 5, 1 y 2 *reads* correspondientes a cada uno de ellos de forma respectiva. La correlación de los niveles de miRNAs vegetales entre los datos de anfibio y de sangre humana fue sumamente alta (Pearson $r = 0.9874$; Spearman $\rho = 0.9439$), volviendo extremadamente improbable que esta asociación fuese aleatoria ($P < 0.0001$; Figura 9).

¹⁴ Si bien el artículo de *Cell Research* fue publicado on-line a fines de 2011, los autores contaron en una entrevista que estuvieron luchando por publicar el manuscrito por casi tres años, luego de la negativa de revistas como *Cell* y *Science* que lo consideraron un hallazgo “demasiado extraordinario”. Eso nos permite deducir que los datos fueron obtenidos en torno a los años 2008-2009.

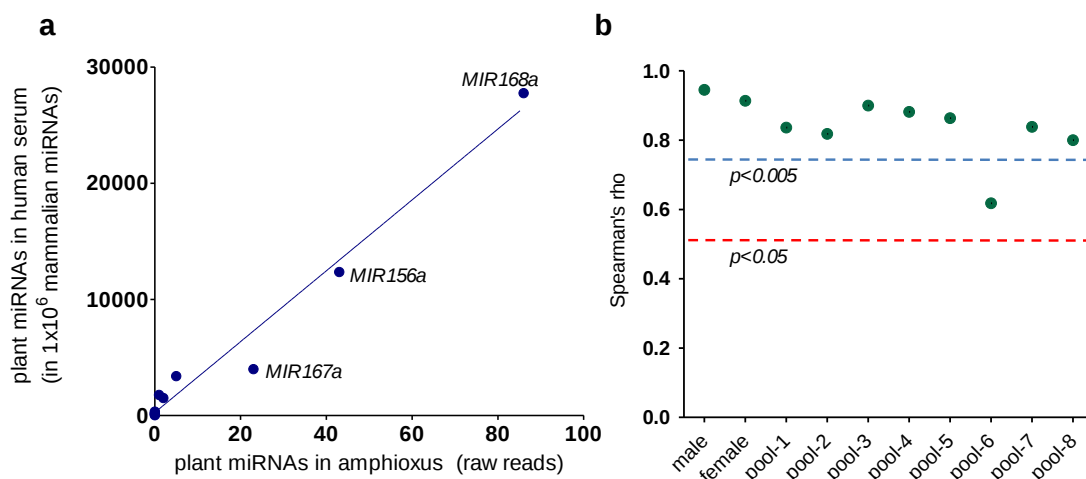


Figura 9: correlación entre los niveles de miRNAs vegetales detectados en plasma humano por Zhang et al. (2012) y los niveles de dichos miRNAs en un secuenciado de pequeños RNAs en anfioxos, realizado por los mismos autores. (A) Dado que no fue posible acceder a los resultados crudos del secuenciado en humanos, los niveles de miRNAs vegetales en dicho caso se representan de la manera como son suministrados en el artículo: reads en un millón de reads correspondientes a miRNAs humanos (eje vertical). Dado que los datos crudos del secuenciado en anfioxo sí eran de dominio público, en el eje horizontal se representa directamente el número de reads correspondiente a cada miRNA vegetal. (B) Coeficiente de correlación de Spearman para los miRNAs vegetales detectados en anfioxo, y cada una de las librerías analizadas en el artículo de Zhang et al (2012) en *Cell Research*. Las líneas punteadas muestran los niveles por encima de los cuales la probabilidad de que dicha correlación se haya alcanzado por azar tiene un $P < 0.05$ (rojo) y $P < 0.005$ (azul).

Ahora sí existía evidencia que permitía afirmar que la detección de miRNAs vegetales en sangre humana se debía con alta probabilidad a un fenómeno de contaminación a nivel de laboratorio. Insistir en una captación de los mismos a partir de la dieta implicaría aceptar el mismo tipo de fenómeno para el anfioxo, pero ¿De dónde podría éste obtener el mismo set de miRNAs vegetales, en exactamente el mismo orden de abundancias relativas, algunos de ellos siendo específicos incluso de gramíneas a las que el cefalocordado no debería jamás verse expuesto?

1.4 DISCUSIÓN

El mismo día que, buscando publicaciones de Chen-Yu Zhang (autor correspondiente del tan mencionado artículo de *Cell Research*), encontramos el artículo sobre los miRNAs en anfioxos, encontramos una publicación que dos días atrás había sido publicada en *Nature Biotechnology*: “Respuesta a: <<Falta de biodisponibilidad oral detectable para los microRNAs vegetales luego de realizar ensayos de alimentación en ratones >>” (Chen et al., 2013). Claramente, se tratan de los descargos de los autores a otro artículo publicado en la misma revista, cuyo título era el que se representa entre corchetes angulares.

Fue así que dimos con la publicación de Dickinson y co-autores (Dickinson et al., 2013), que fue el cuarto reporte publicado ese año sobre no reproducibilidad de lo reportado por Zhang et al., sumándose a las publicaciones de Zhang – otro Zhang distinto - (Zhang et al., 2012b),

Witwer (Witwer et al., 2013b) y Snow (Snow et al., 2013). El segundo de ellos, Kenneth Witwer, terminó siendo un estrecho colaborador en el trabajo que refiere al Capítulo II de esta tesis.

¿Qué tenía de particular el trabajo de Dickinson para justificar que un cuarto reporte sobre irreproducibilidad ocupase su lugar en una revista de la talla de *Nature Biotechnology*? Más aún, la publicación de este estudio motivó que dicha revista dedicase su editorial del mes de noviembre a abordar precisamente el asunto de la reproducibilidad, con el muy incitativo título de “[*estamos*] Receptivos a la Replicación” (Nat-Biotechnol, 2013).

La pregunta anterior puede contestarse desde el trabajo en sí, o desde los autores. El trabajo en sí fue bastante concreto y sólido: alimentaron ratones con raciones sintéticas conteniendo un elevado contenido de arroz, un moderado contenido de arroz, o completamente carentes de arroz, previa aclimatación a dichas raciones, y por períodos de tiempo variables (1, 3 y 7 días). Luego secuenciaron pequeños RNAs extraídos tanto del hígado como del plasma de dichos ratones con la tecnología HiSeq de Illumina, obteniendo una alta profundidad de datos (> 10 millones de *reads* por librería). Las cantidades de miRNAs vegetales detectados fueron despreciables (< 1 RPM). Más aún, estas pocas secuencias fueron detectadas en las distintas muestras, sin apreciarse ningún tipo de dependencia con el régimen de alimentación de los ratones. La mayoría de estas secuencias correspondían además al miR414 vegetal, el cual es no detectable en granos de arroz, sugiriendo que su aparición podría tratarse de un artefacto. Además de la secuenciación profunda, los autores evaluaron los niveles del osa-miR-168a por PCR cuantitativa (donde la presencia de la modificación 2'-o-metilo no debería generar sesgos), detectando altos niveles en las raciones ricas en arroz, y niveles no detectables en hígado o plasma. Por último, los autores suministraron evidencia para suponer que el aumento en los niveles de LDL plasmático reportados por Zhang et al. (2012) no se debían a la ingesta de miR168a *per se*, sino a una respuesta sistémica frente a una dieta basada exclusivamente en arroz, y consecuentemente más baja en lípidos y en colesterol que la de los ratones control.

Mirando la afiliación de los autores, el primero y último pertenecían a miRagen Therapeutics, una compañía con sede en Colorado (EE.UU.) interesada en utilizar miRNAs como agentes terapéuticos, y por consiguiente, interesada en la posibilidad de la incorporación de miRNAs por vía oral. Los restantes cuatro autores pertenecen a la compañía Monsanto, conocida por su dominio en el mercado de los alimentos transgénicos. Así, la objetividad del equipo de trabajo se lograba mediante la compensación entre quienes *a priori* preferirían ver la teoría de la ingesta de miRNAs funcionales confirmada, y quienes – puede suponerse – preferirían ver dicha teoría tan académicamente desprestigiada como fuese posible. En el medio estábamos nosotros, ajenos a toda esa polémica, tratando de publicar algo para rescatar el año y medio de esfuerzo que le habíamos dedicado a un artefacto. Y ahora que teníamos una prueba sólida de que todo se debía a fenómenos de contaminación, el tema había tomado bastante vuelo y parecía ya asunto cerrado.

En sus descargos (Chen et al., 2013), Zhang et al. se apoyaron fundamentalmente en sus experimentos basados en la oxidación con periodato sódico, argumentando que esa era la manera más eficiente para detectar los miRNAs vegetales; experimentos faltantes en el

artículo de Dickinson y colaboradores. Sin embargo, nuestros resultados mostraban como, paradójicamente, la oxidación con periodato era una forma de amplificar la representación relativa de los contaminantes en una muestra.

El hecho de haber podido detectar secuencias de *Trachemys* en muestras de espermatozoides humanos procesadas en paralelo con RNA de tortuga, o el hecho de que alguien más haya detectado secuencias vegetales en muestras de sangre humana y de hígado de ratón, podrían parecer fenómenos pertinentes únicamente a quienes investigan en el campo de los RNAs exógenos. Y sin embargo, los efectos más dañinos se darían en los casos donde muestras del mismo organismo se contaminasen de forma cruzada. Porque en estos casos, a diferencia de los anteriores, resultaría imposible advertir la contaminación *a posteriori*, es decir, al momento del análisis de los datos.

Uno podría pensar que el tema de la contaminación y la secuenciación profunda era algo propio para discutir entre usuarios del sur, que llegamos a la tecnología varios años después de su uso generalizado en institutos y universidades del llamado primer mundo. Y sin embargo, diversas señales parecían indicar que no era un tema que el mundo hubiese aún bien digerido. En primer lugar, la publicación de nuestro artículo mostraba como había lugar para que una revista como *RNA* admitiese un reporte cuyo título contenía la frase: “*efectos subestimados de la contaminación en secuenciación masiva*” (Tosar et al., 2014). Y eso salió a principios de 2014. En noviembre de ese mismo año, *Nature* publicó como noticia: “*Estudios sobre microbiota amenazados por la contaminación*”¹⁵, luego de que se mostrase que la diversidad de especies bacterianas detectables en diluciones seriadas de un cultivo de *Salmonella bongori* era proporcional al grado de dilución de las muestras (Salter et al., 2014).

En el estudio anterior, las fuentes de contaminación detectadas fueron los kit de extracción de DNA, que por lo general no se venden estériles. Otro ejemplo paradigmático es la historia asociada al virus XMRV, un retrovirus relacionado al virus de la leucemia murina descubierto en muestras de pacientes con cáncer de páncreas (Urisman et al., 2006). En 2009 se publicó en la revista *Science* la detección de dicho virus en células blancas de pacientes con síndrome de fatiga crónica, una enfermedad que afecta millones de personas en el mundo y cuya etiología era desconocida (Lombardi et al., 2009). Dos años después, dicho artículo sufrió una retracción total luego de que numerosos laboratorios fallaran en el intento de reproducir estos resultados, ni fueron capaces de detectar al virus XMRV en las mismas muestras que habían sido utilizadas en el reporte original. En 2012 también fue retractado el artículo de *Plos Pathogens* que había descrito el virus en cáncer de páncreas, al descubrirse que el mismo se habría formado por recombinación incidental de dos virus murinos en el laboratorio (Paprotka et al., 2011). Por consiguiente, la explicación más probable para la detección del XMRV en muestras humanas era contaminación del material de laboratorio con muestras infectadas.

En 2013, la revista *PNAS* publicó la detección por secuenciación masiva de un nuevo parvovirus denominado NIH-CQV, detectado en muestras de pacientes con hepatitis seronegativa (no A, no B, no C, no D, no E), otra enfermedad de pobre caracterización y etiología (Xu et al., 2013). Meses después, una carta al editor publicada en dicha revista

¹⁵<http://www.nature.com/news/microbiome-science-threatened-by-contamination-1.16327>

advertía que el NIH-CQV era probablemente un virus de algas, introducido en el laboratorio debido a las columnas de extracción en fase sólida para la purificación de ácidos nucleicos. La sílica presente en estos dispositivos – a la que se adsorbe el DNA – se extrae de la pared celular de diatomeas (algas unicelulares), explicando así el origen de la contaminación (Naccache et al., 2014).

No son pocos, pues, los casos en los cuales la secuenciación profunda ha llevado a conclusiones erróneas debido a una falta de concientización sobre los efectos de la contaminación y sus variadas fuentes. Desde luego, la culpa no es de la tecnología ni de su alto poder de análisis. Si estos son los casos dónde el origen de la contaminación pudo probarse y rastrearse ¿cuántas conclusiones basadas en artefactos continuarán aún auto-amplificándose en la literatura científica?

Cuando las fuentes de contaminación comienzan a ser los propios materiales de laboratorio, provenientes de los fabricantes más confiables y respetados, uno comienza a comprender que nadie está del todo libre de esta problemática. Y sin embargo, el sentido común indica que en algún momento las conclusiones erróneas asentadas sobre datos mal interpretados tendrían que caer por su propio peso. En algún momento los experimentos confirmatorios, si son realmente ortogonales a la técnica original, deberían dejar de cumplirse: controles positivos que no dan; controles negativos donde la contaminación se hace manifiesta. Es aquí donde aparecen las tensiones entre el método científico, la psicología humana, y los *curriculum vitae*: muchos de los artículos a los que se ha hecho referencia poseen correcciones y enmiendas debido a figuras cuya manipulación ha sido probada.

A modo de resumen de este capítulo, nuestros resultados y los de otros investigadores nos permiten afirmar que no existe evidencia suficiente para sostener la regulación inter-reino de la expresión génica mediada por miRNAs provenientes de la dieta, al menos para organismos con aparatos digestivos similares al nuestro. Si bien no puede descartarse la existencia de tales mecanismos, existen explicaciones alternativas para la interpretación de los resultados obtenidos en el artículo original de Zhang et al. (2012). Esto no significa que el debate no siga abierto: en junio de 2015 los mismos autores publicaron un artículo de título: “*Los RNAs pequeños pueden ser transferidos a través de la placenta de mamíferos y regulan directamente la expresión génica del feto*” (Li et al., 2015). Los autores detectaron los mismos miRNAs vegetales que habían descrito en humanos (y que nosotros detectamos en anfibios) en muestras de cordón umbilical y líquido amniótico. Uno pensaría que esta vez se tomaron toda suerte de recaudos para probar que la contaminación no era la fuente de los hallazgos, pero la realidad es que no: nuestro artículo no solo no es siquiera mencionado (ni los de ninguno de los otros detractores), sino que la palabra “contaminación” aparece únicamente una vez: “*se tomaron precauciones para prevenir la contaminación de los fetos con el plasma materno*”. No hay mención a la posibilidad de que cada vez que en dicho laboratorio se secuencian una muestra de lo que sea, aparecen estos miRNAs vegetales. Por el contrario, se propone que cuando una madre come arroz, los miRNAs vegetales pasan a su sangre y atraviesan la placenta, para luego regular la expresión génica de su hijo/a. Un caso de regulación de la expresión génica no solo entre reinos, sino también entre generaciones. Otros investigadores, incluso algunos muy respetados en su campo, también opinan que pese a las

críticas la evidencia disponible apoya la transferencia de miRNAs a través de la dieta en mamíferos (Fabris and Calin, 2016; Hirschi et al., 2015). El tiempo tendrá la última palabra.

Lo que sí llama la atención es la gran cantidad de fuentes posibles por las cuales secuencias contaminantes pueden aparecer en una muestra. Ya se mencionaron los kit de extracción y purificación de ácidos nucleicos. A ellos debería sumarse el glucógeno utilizado para la precipitación del RNA, que nosotros postulamos como una fuente de contaminación con secuencias de mejillones. Y aunque no lo hemos mencionado, también disponemos de evidencia sólida para afirmar que el Trizol – reactivo comúnmente utilizado para la purificación de RNA – es capaz de preservar ácidos nucleicos contaminantes por períodos de al menos dos años si es mantenido a 4°C. También está la contaminación cruzada durante el armado de las librerías. Y obviamente, la contaminación microbiológica de cualquier muestra no estéril. Una lista que sugiere un largo etcétera.

CAPÍTULO II

Caracterización del conjunto de pequeños RNAs secretados por células mamarias en distintas fracciones extracelulares

2.1 INTRODUCCIÓN

En el capítulo anterior se exploró la posibilidad de transferencia funcional de RNAs reguladores de la expresión génica entre organismos no emparentados, y en particular, la posibilidad de que la dieta fuese el proceso por el cual dicha transferencia pudiese llevarse a cabo. Nuestra conclusión es que no existe por el momento evidencia concreta que sostenga dicha hipótesis, pero es un campo de estudio abierto que podría arrojar nuevas sorpresas en el futuro.

En donde sí existen amplios consensos y numerosa cantidad de evidencia es en el campo de la transferencia funcional de RNAs reguladores entre:

- a) Organismos no emparentados, en el contexto de una relación parásito – hospedero.
- b) Células de un mismo organismo, ya sea espacialmente cercanas (es decir: formando parte de un mismo micro-ambiente) o ubicadas en distintos tejidos y regiones del cuerpo.

Nuestro laboratorio tuvo un interés temprano en el punto (a), y luego de haber descrito la presencia de mitades de tRNAs en el protozoo patógeno *Trypanosoma cruzi* (Garcia-Silva et al., 2010) y su posible asociación con proteínas Argonauta para conformar una nueva familia de RNAs reguladores (Garcia-Silva et al., 2014c; Garcia Silva et al., 2010), se advocó al estudio de la secreción de dichos RNAs en vesículas extracelulares y los efectos ocasionados luego de su captación por parte de células humanas susceptibles a la infección (Garcia-Silva et al., 2014a; Garcia-Silva et al., 2014b).

A pesar de los avances realizados en el modelo parasitario – realizamos una de las primeras descripciones sobre la presencia de mitades de tRNAs como posibles RNAs reguladores, – nuestro laboratorio posee un sesgo natural hacia la biología del cáncer, por lo que resultó de interés estudiar la secreción y transferencia de RNAs reguladores entre células humanas, particularmente en un contexto tumoral.

Como se explicó en la introducción general de esta tesis, las implicancias de las vesículas extracelulares en la comunicación intercelular y en la biología del cáncer son cada vez más claras. Materia más discutible es en qué medida dichos efectos son mediados por la presencia de RNAs reguladores contenidos en dichas vesículas.

La primera descripción de que los exosomas contienen miRNAs capaces de provocar respuestas funcionales en células receptoras data del año 2007, en un importante artículo del

grupo de Jan Lötvald que cuenta ya con más de 3000 citas académicas (Valadi et al., 2007). A partir de dicho reporte, numerosos laboratorios se lanzaron al estudio del rol de los miRNAs extracelulares contenidos en exosomas, y la evidencia al respecto comenzó a acumularse a un ritmo acelerado. Se llegó incluso a proponer el concepto de que los miRNAs pudiesen funcionar como una suerte de hormonas, viajando por el torrente sanguíneo y transfiriendo mensajes de un punto a otro del organismo (Cortez et al., 2011).

Al margen de su rol fisiológico, la presencia de miRNAs en la sangre despertó un enorme interés por su posible uso como nuevos biomarcadores mínimamente invasivos. Ya era conocido que los distintos tipos de cáncer poseen una expresión alterada y particular de miRNAs (Lu et al., 2005). Por consiguiente, cabría esperar que la presencia de un tumor incipiente pudiese manifestarse por la presencia de miRNAs derivados del tumor en la sangre. Sería imposible citar aquí toda la bibliografía del tema, que hoy ronda los 400 artículos al año.

Dada la patología X, numerosos investigadores estudiaron los perfiles de miRNAs en sangre de individuos afectados por X e individuos control. Las típicas conclusiones son del tipo “*los miRs #a, #b y #c son más abundantes en sangre de pacientes con X, y, por consiguiente, se postulan como biomarcadores de X*”. Sin embargo, el bajo nivel de reproducibilidad entre estudios de este tipo ha ocasionado que no existan aún biomarcadores validados y altamente específicos basados en miRNAs extracelulares (Witwer, 2014). En muchos casos, los miRNAs afectados por la patología X parecerían deberse más bien a efectos inespecíficos indirectos y por tanto comunes a varias patologías (ej: derivados de un cambio en la composición de células sanguíneas a raíz de la irrupción de fenómenos de inflamación asociados a X).

Si se considera que la mayoría de los miRNAs presentes en el plasma provienen de células hematopoiéticas (Duttagupta et al., 2011) uno podría preguntarse qué tan sencillo pueda llegar a ser detectar el aporte de un pequeño tumor incipiente proveniente de un órgano interno, por ejemplo. El alto grado de dilución provocado por la liberación de miRNAs a la sangre podría contrarrestarse si se pudiera ya no analizar el plasma en su conjunto, sino pre-enriquecer la muestra en las vesículas extracelulares provenientes del tumor. Pero para esto sería necesario:

- a) Conocer si el perfil de RNAs que caracteriza a un determinado tipo de tumor se ve de algún modo reflejado en el perfil de RNAs secretados por las células tumorales.
- b) Si existen RNAs con potencial biomarcador: conocer el tipo de vesícula o vehículo extracelular en el cual estos son principalmente secretados.
- c) Disponer de marcadores de superficie para poder enriquecer la muestra en las vesículas extracelulares derivadas del tumor.

El punto (c) ha tenido un reciente empuje cuando la revista *Nature* publicó en julio de 2015 un reporte sobre la presencia de la molécula glipican-1 (GPC1) en la superficie de exosomas circulantes provenientes de cáncer de páncreas (Melo et al., 2015). Los autores mostraron cómo la cuantificación plasmática de exosomas GPC1+ permitía la detección del cáncer

pancreático con una mayor sensibilidad y especificidad que los métodos actualmente disponibles, en una etapa muy precoz, y con excelente discriminación entre tumores malignos y benignos.

Si bien el reporte anterior abre numerosas posibilidades experimentales, al momento de desarrollar esta tesis eran los puntos (a) y (b) los que requerían, a nuestro juicio, una respuesta más urgente. A pesar de años de investigación en la temática no quedaba claro, por ejemplo, si las células seleccionan un conjunto reducido de miRNAs que son posteriormente secretados de manera específica, o si por el contrario el secretoma de pequeños RNAs es una muestra estadísticamente representativa del contenido intracelular.

El segundo de los escenarios presentados en el párrafo anterior avalaría la búsqueda en sangre de RNAs expresados de manera aberrante en el tumor, muy por encima de los valores normales en el tejido correspondiente, y poco o nada expresados en células sanguíneas. Pero la mayoría de la bibliografía existente describe a la secreción de miRNAs en exosomas como un fenómeno altamente selectivo (van Balkom et al., 2015; Zhang et al., 2015), con la consecuente disrupción entre los perfiles de pequeños RNAs intra y extracelulares. Sin embargo, una exploración detallada permite evidenciar que casi todos los artículos citan a un conjunto reducido de reportes para sostener tal afirmación. Dichos reportes afirman que los perfiles de miRNAs en las vesículas extracelulares – estudiados por secuenciación profunda – difieren de los perfiles observados dentro de las células parentales, siendo esto evidencia indirecta de secreción selectiva (Guduric-Fuchs et al., 2012; Ji et al., 2014; Nolte-t Hoen et al., 2012). El asunto es que dichos reportes carecen de la inclusión de réplicas biológicas en la secuenciación profunda, por lo cual el soporte estadístico de las diferencias halladas en los perfiles intra y extracelulares es desconocido. Esta crítica no aplica a lo reportado por el grupo de Sánchez-Madrid (Mittelbrunn et al., 2011), quienes también describieron perfiles alterados entre células inmunes y sus respectivos exosomas, evaluados por microarreglos. Este grupo suministró luego una explicación mecanística, encontrado motivos de secuencia que distinguían a los miRNAs secretados de aquellos retenidos en la célula, y describiendo cómo la proteína hnRNPA2B1 era capaz de reconocer dichos motivos para dirigir a los miRNAs hacia los exosomas (Villarroya-Beltri et al., 2013).

Aunque con menos eco en la comunidad científica congregada en torno a las vesículas extracelulares, también existe evidencia sosteniendo un principio de secreción no selectiva para los miRNAs. Recientemente ha sido descrito cómo los niveles de miRNAs extracelulares están inversamente correlacionados a la abundancia intracelular de RNAs conteniendo sitios complementarios (Squadruto et al., 2014). Esto de algún modo sugiere que la secreción no es selectiva ya que un mayor “secuestro” de los miRNAs por RNAs complementarios implica una disminución en sus niveles intracelulares efectivos (“libres”). Y a menores niveles efectivos, menor sería la tasa de secreción de dichas especies. También existe un reporte que describe una fuerte correlación, en células MCF-7, entre los niveles de miRNAs dentro y fuera de la célula (Turchinovich et al., 2011). En resumen, si bien la “hipótesis de secreción selectiva” goza de una amplia aceptación, la evidencia disponible no es en absoluto definitiva.

Otro aspecto que se consideró relevante fue el conocer qué tipos de RNAs pequeños¹⁶ son secretados o están presentes en distintas fracciones extracelulares. Desde el punto de vista de su uso como biomarcadores, poseer una suerte de “atlas extracelular” de pequeños RNAs permitiría focalizarse y enriquecer en fracciones específicas en lugar de trabajar con el medio extracelular – o la sangre – en su conjunto. Lo que impactaría positivamente en el mejoramiento de la relación señal / ruido en este tipo de análisis. Pero más allá de sus posibles aplicaciones, la pregunta es importante desde el punto de vista biológico: existen diversas estructuras extracelulares capaces de transportar y proteger los RNAs de la acción de las nucleasas, y a su vez dichas estructuras son consecuencia de distintos mecanismos de tráfico intracelular y secreción. La presencia o ausencia de una determinada especie de RNA en alguna de dichas estructuras podría dar cuenta de los mecanismos y compartimientos en los que dicho RNA juega su rol dentro de la célula. Por otro lado, y pensando en la posibilidad de un rol del RNA como mediador de vías de comunicación intercelular, las estructuras extracelulares asociadas a cada especie pueden determinar el tipo de células capaces de captar dicho RNA y su destino intracelular una vez internalizado (Kanada et al., 2015).

2.1.1 Modelo de estudio

Para realizar una caracterización de la variedad de pequeños RNAs presentes en distintas fracciones extracelulares, decidimos en primer lugar trabajar con líneas celulares humanas. A diferencia de un estudio realizado en sangre, por ejemplo, el trabajo con líneas celulares permite conocer sin ambigüedades de qué tipo celular están proviniendo los RNAs estudiados¹⁷. Se optó por la línea celular derivada de cáncer de mama – adenocarcinoma metastático – MCF-7, por múltiples ventajas entre las que figura la existencia de estudios previos relacionados a la caracterización de su secretoma de miRNAs y su comparación con el contenido intracelular (Turchinovich et al., 2011). Pretendimos extender este tipo de abordaje experimental a todo el conjunto de RNAs pequeños – no solo miRNAs – y en distintos fraccionamientos del medio condicionado extracelular, aplicando para ello las tecnologías de secuenciación profunda. Además, se estudió en paralelo la línea celular MCF-10A, proveniente de tejido epitelial de la glándula mamaria y de características benignas (enfermedad fibroquística), por lo que puede racionalizarse como la contraparte no maligna de las células MCF-7. Ambas líneas celulares fueron obtenidas directamente de la ATCC, y se utilizaron en número de pasajes bajos (menor a 10).

Se consideraron tres poblaciones diferentes en las cuales, de acuerdo a la bibliografía, cabría esperar RNAs pequeños: microvesículas, exosomas y complejos ribonucleoproteicos. La biosíntesis y características de dichas estructuras ya fueron descritas en la introducción general. Se optó por un protocolo de purificación basado en centrifugación diferencial, por poseer buenos rendimientos, niveles de pureza adecuados – aunque no ciertamente los mejores – y por ser aún considerado como el “estándar de oro” para la purificación de

¹⁶ El estudio de RNAs de mayor tamaño se consideró fuera de los alcances de esta tesis

¹⁷ Se utilizaron líneas celulares homogéneas, y bien caracterizadas: http://iclac.org/wp-content/uploads/Cross-Contaminations-v7_2.pdf

vesículas extracelulares (Lobb et al., 2015; Van Deun et al., 2014). Si bien la viabilidad celular fue alta durante el curso de nuestros experimentos (>95% según evaluación por tinción con el colorante vital Azul de Tripán), la presencia de cuerpos apoptóticos no puede descartarse, aunque cabría esperar que la mayoría de ellos fuese eliminada en la centrifugación inicial de limpieza (30 minutos a 2,000g); el resto co-purificaría principalmente con la población de microvesículas.

En el campo de las vesículas extracelulares, lo más frecuente es que los investigadores realicen los cultivos celulares utilizando suero fetal bovino (FBS) depletado de vesículas; es decir: previamente centrifugado por varias horas a 100,000g. Esto permite mantener condiciones de cultivo similares a las condiciones estándar, pero minimizando el aporte de vesículas – y por tanto de RNA – provenientes del FBS. Sin embargo, este abordaje no nos resultó aplicable ya que nuestro estudio también se extendería a la fracción ribonucleoproteica. Por consiguiente, para eliminar toda posibilidad de secuenciar RNAs bovinos, decidimos realizar los cultivos celulares en medio definido y libre de suero. Esto no resultó un problema para la línea celular MCF-10A, ya que las condiciones de cultivo recomendadas por la ATCC implican de por sí el uso de un medio libre de FBS. Para el caso de las células MCF-7, se recurrió a la formulación presentada en un viejo artículo (Barnes and Sato, 1979) mediante la cual los autores obtuvieron células con características fenotípicas macroscópicas idénticas a las obtenidas con FBS 10%.

Se obtuvieron tres réplicas biológicas de cada fracción extracelular, y dos réplicas biológicas de cada fracción intracelular, para ambas líneas celulares (número de muestras secuenciadas: 22). Las réplicas biológicas correspondieron a células descongeladas de un mismo lote almacenado en nitrógeno líquido. Las células fueron descongeladas, puestas en cultivo, las distintas fracciones fueron purificadas, el RNA extraído, y las librerías y secuenciación masiva realizadas en fechas diferentes, con diferencias de meses entre una réplica biológica y la siguiente. Por consiguiente, nuestras réplicas capturan toda la varianza esperable en la manipulación de laboratorio del lote de células correspondiente, obtenido de la ATCC.

2.2 MATERIALES Y MÉTODOS

2.2.1 Reactivos

Las líneas celulares MCF-7, MCF-10A y NCI-H1299 fueron obtenidas de la American Type Culture Collection (ATCC) y usadas en pasajes bajos (<10). Las células HeLa fueron obtenidas de un colega del Instituto. El medio “*Eagle’s minimum essential medium*” (EMEM) fue obtenido de la ATCC. El medio “*Mammary epithelial growth medium*” (MEGM) fue obtenido de Lonza (Suiza). Los demás aditivos para cultivos celulares (Transferrina humana purificada, Factor de crecimiento epidérmico recombinante [EGF], sal de tris de la Prostaglandina F2 α [PGF2 α], Fibronectina humana purificada e Inhibidor de Tripsina

proveniente de la soja) fueron obtenidos de Sigma. El suero fetal bovino (FBS) y las disoluciones de tripsina-EDTA fueron de Gibco (Life Technologies, EEUU). Los frascos de cultivo de plástico fueron de marca Corning (EEUU). El Trizol y Trizol LS fueron obtenidos de Invitrogen (Life Technologies, EEUU). El glucógeno libre de RNasas y el isopropanol también fueron de Sigma, de calidad adecuada para la purificación de ácidos nucleicos. Los cebadores (“*primers*”) para RT-PCR fueron obtenidos en Integrated DNA technologies (IDT, EEUU). Las pruebas de hidrólisis utilizadas en PCR digital fueron adquiridas en Applied Biosystems (Life Technologies, EEUU).

2.2.2 Cultivos celulares

Todas las células fueron incubadas a 37°C, en ambiente húmedo conteniendo CO₂ al 5%, y sin el agregado de antibióticos ni antimicóticos. Las células MCF-7 se crecieron en EMEM + 0.01 mg/mL insulina humana recombinante + 10% FBS, hasta obtener el número de células adecuado. Posteriormente se eliminó el medio condicionado, se lavó con PBS (buffer fosfato salino) y se adicionó medio definido libre de suero (EMEM + 3.8µg/mL de insulina; 0.8 ng/mL EGF; 475pg/mL PGF2α; 5µg/mL de Transferrina y 20µg/mL de Fibronectina). Luego de un período de adaptación de 48hs, se removió el medio, se lavó con PBS, y se adicionó nuevo medio definido termostatzado a 37°C. Luego de un nuevo período de incubación de 48hs, se recolectó y almacenó el medio condicionado. Las células (típicamente al 70% de confluencia) fueron despegadas mediante tratamiento con tripsina-EDTA, contadas, y lisadas en Trizol. El mismo procedimiento se siguió para las células HeLa y NCI-H1299. Las células MCF-10A se crecieron en medio MEGM hasta obtener el número de células requerido. Posteriormente fueron despegadas con tripsina-EDTA (luego neutralizada con inhibidor de tripsina de soja) y un millón de células viables fue plaqueada en frascos de cultivo de 75cm². Luego de 24hs, se eliminó el medio, las células se lavaron con PBS, y se adicionó medio MEGM sin extracto de glándula pituitaria bovina (MEGM s/BPE). Luego de un período de adaptación de 48hs., se removió el medio, se lavó con PBS, y se adicionó nuevo medio MEGM s/BPE termostatzado a 37°C. Luego de un nuevo período de incubación de 48hs, se recolectó y almacenó el medio condicionado. Las células (típicamente al 70% de confluencia) fueron despegadas mediante tratamiento con tripsina-EDTA, contadas, y lisadas en Trizol.

El conteo mediante el colorante vital azul de tripán reveló que el número de células muertas al finalizar los experimentos fue inferior al 5% en todos los casos. De manera rutinaria y periódica, se analizó la presencia de *Mycoplasma spp.* en las líneas celulares (mediante un protocolo de PCR anidada), resultando siempre negativa.

2.2.3 Purificación de las fracciones extracelulares

El medio condicionado libre de suero (t = 48 hs) fue removido de los frascos de cultivo e inmediatamente centrifugado a 300g por 10 minutos (para eliminar las células desadheridas) y almacenado a -20°C. Una vez descongelados, los medios se centrifugaron a 2,000g y 4°C por

30 min para remover detritos celulares y cuerpos apoptóticos. El sobrenadante se centrifugó media hora a 16,000g y 4°C. El *pellet* (**fracción p16**) fue lavado dos veces en PBS y resuspendido en el volumen deseado de PBS o sometido directamente a extracción del RNA con Trizol. El sobrenadante se filtró con suavidad por un filtro estéril de acetato de celulosa de 0.22 µm y se centrifugó 2.5 hs a 100,000g y 4°C en una ultra-centrífuga Beckman Coulter Optima XPN-90, usando el rotor SW-40Ti (de ángulo móvil). El *pellet* (**fracción p100**) fue lavado con 10mL de PBS previo a su resuspensión en el volumen deseado de PBS (típicamente 100µL) o directamente sometido a extracción del RNA con Trizol. El sobrenadante de la primera centrifugación a 100,000g (**fracción S100**) fue concentrado hasta 250 µL mediante ultrafiltración (Vivaspin 20, Sartorius Stedim Biotech; MWCO 10 kDa) y sometido a extracción del RNA mediante Trizol LS.

2.2.4 Caracterización de las fracciones p16 y p100 por microscopía electrónica de transmisión (TEM): tinción negativa.

Las rejillas de cobre cubiertas de carbono y pre-tratadas con plasma (para aumentar su hidrofiliidad) fueron flotadas por 5 min (con la cara de carbono hacia abajo) sobre gotas de 10 µL correspondientes a las muestras p16 o p100, de modo de adsorber las vesículas. Luego, las rejillas fueron sucesivamente flotadas en 2 gotas de acetato de uranilo al 2% (30 seg en cada caso), con posterior remisión del líquido con papel absorbente. En casi todos los casos, las imágenes fueron obtenidas utilizando un microscopio electrónico Hitachi 7600 operando a 80 keV. En algunos casos – especificados – las imágenes fueron obtenidas con un microscopio electrónico JEOL JEM-1010.

2.2.5 Caracterización de las fracciones p16 y p100 por análisis de seguimiento de nanopartículas (NTE) y dispersión dinámica de luz (DLS).

Las fracciones p16 y p100 fueron analizadas por NTE utilizando un equipo NanoSight NS500 (Malvern Instruments) equipado con un láser de 405nm y una cámara sCMOS. Las muestras se diluyeron en PBS (Dulbecco's Phosphate-Buffered Saline; Life Technologies) hasta obtener la concentración óptima de partículas (típicamente diluciones 1/10 – 1/50). Las muestras se vortexearon y se inyectaron en el equipo realizando 5 videos de 20 seg cada uno. La concentración de las partículas por rango de tamaños se multiplicó por el inverso del factor de dilución correspondiente. Para DLS, se utilizó un equipo Brookhaven model Z Plus. Las funciones de correlación se analizaron con el software Particle Solutions v. 2.5. Las fracciones p16 y p100 se centrifugaron por 10 min a 2,000g y los sobrenadantes se diluyeron en PBS y se filtraron por 0.45 µm. Se realizaron 5 medidas consecutivas para cada muestra, cada una con una duración de 10 min.

2.2.6 Extracción y cuantificación del RNA.

El RNA se extrajo de las distintas muestras utilizando Trizol LS (Invitrogen, Life Technologies) de acuerdo a las especificaciones del fabricante, con mínimas modificaciones (precipitación con el doble de isopropanol y lavados con etanol 80% a efectos de minimizar la pérdida de RNAs pequeños). En p16 y p100, se agregaron 10 µg de glucógeno como agente colector. Las muestras de RNA fueron cuantificadas de acuerdo a su absorbancia a 260 nm (método poco confiable debido a la presencia de fenol, que absorbe a 270 nm) y utilizando un fluorímetro Qubit 2.0 (Life Technologies) y un kit “Qubit RNA high sensitivity kit”, de acuerdo a las especificaciones del fabricante.

2.2.7 Análisis por Western blot.

Las proteínas fueron purificadas de las fracciones intra- y extracelulares a partir de la fase orgánica del Trizol LS utilizado para las extracciones de RNA, de acuerdo a las especificaciones del fabricante. El pellet de proteínas fue resuspendido en tampón RIPA (20mM Tris-Cl pH 7.5; 150mM NaCl; 1mM Na₂EDTA pH 8.0; 1mM EGTA; 1% NP-40; 1% deoxicolato de sodio; e inhibidor de proteasas [cocktail, Sigma]) y cuantificado por el método del ácido bicinonínico. 20 µg de las muestras fueron corridos en un gel de poliacrilamida al 12% conteniendo SDS (SDS-PAGE) y electro-transferidos toda la noche a 9°C a membranas de PVDF Hybond-P (Amersham, GE Healthcare). Luego de un bloqueo con PBS-leche descremada al 5%, las membranas se incubaron toda la noche con alguno de los siguientes anticuerpos monoclonales murinos: anti-TSG101 (h) (4A10, Abcam; dilución 1/1000), anti-CD9 (h) (clon C-4, Santa Cruz Biotech.; dilución 1/300) y anti CD-63 (h) (MEM-259, Abcam; dilución 1/500). Para el caso de CD63, el gel de SDS-PAGE fue corrido en condiciones no reductoras, pues el epítipo reconocido por el anticuerpo utilizado requiere puentes disulfuro. Los blots fueron revelados utilizando un anticuerpo secundario policlonal contra las IgG de ratón (H&L, ab6728, Abcam) conjugado a peroxidasa de rábano. El sustrato quimioluminiscente fue obtenido de Viva Science (Uruguay).

2.2.8 Construcción de las librerías y secuenciación profunda

Se secuenciaron 3 réplicas biológicas de cada fracción extracelular (≈ 150ng RNA) y 2 réplicas biológicas de la fracción intracelular de cada línea celular (≈ 1.2µg). Se partió de mayor cantidad de RNA para las muestras intracelulares dado que la proporción relativa de RNAs pequeños es muy baja, debido a la presencia de especies de alto peso molecular (mayoritariamente RNAs ribosomales). Las librerías se construyeron mediante el kit *NEBnext Multiplexed Small RNA Library Prep Set for Illumina* (New England Biolabs), siguiendo las instrucciones del fabricante, y todos los recaudos y precauciones adoptados como consecuencia del capítulo anterior de esta tesis. Por mayor detalle, referir a la sección§1.2.3. La secuenciación se realizó en un Illumina Genome Analyzer IIx (primeras dos réplicas biológicas; 72 ciclos) y en un Illumina MiSeq para la tercera réplica de cada muestra (44

ciclos). En un principio se entendió necesario secuenciar al menos 72 nucleótidos de modo de poder leer desde el extremo 5' al extremo 3' de todos los RNAs de 65 nucleótidos para abajo (destinando las últimas bases a la identificación del adaptador 3'). Luego de corroborar la ausencia de poblaciones de RNA significativas por encima de los 40 nucleótidos, se optó por reducir el largo de la reacción de secuenciado (y por consiguiente los costos) de las sucesivas réplicas.

La Tabla 1 muestra un resumen de la profundidad de secuenciado obtenida para cada muestra y réplica correspondiente. Como puede observarse, se procesaron en esta oportunidad casi 90 millones de lecturas (*reads*). Los datos de secuenciación fueron alojados en la base de datos pública *Sequence Read Archive* (SRA; <http://www.ncbi.nlm.nih.gov/sra>) y se encuentran disponible bajo el código BioProject ID: PRJNA270876.

2.2.9 Análisis bioinformático de los datos de secuenciación

Luego de “*de-multiplexear*” los datos y eliminar las secuencias correspondientes a los adaptadores 3' (solamente se retuvieron aquellas secuencias de más de 15 bases que contuviesen un adaptador 3' identificable), los archivos FastQ se alinearon al genoma humano (versión hg19) utilizando el programa Bowtie (Langmead et al., 2009) incluido en la plataforma pública Galaxy (<https://galaxyproject.org/>). En este alineamiento se permitió un máximo de 1 *mismatch*. Las lecturas mapeadas fueron colapsadas a secuencias únicas con su correspondiente número de lecturas y alineadas con el programa Lastz (aceptando 1 *mismatch*) a librerías de referencia curadas a mano y sin ambigüedades, y correspondiendo a: miRNAs humanos descargados de la versión 20 del miRBase (<http://www.mirbase.org/>), tRNAs humanos maduros descargados de la “genomic tRNA database” (<http://gtrnadb.ucsc.edu/>), y RNAs humanos maduros ribosomales, nucleares pequeños, nucleolares pequeños, vault, y YRNAs descargados del GenBank (<http://www.ncbi.nlm.nih.gov/nuccore/>). Solamente se mantuvo una anotación por cada lectura, dándose prioridad a los alineamientos sentido sin *mismatches* y ocupando todo el largo del *read*. Dado que el tamaño mínimo de la región semilla utilizado por el programa Lastz es de 19 nucleótidos, el rango de tamaños de los RNAs analizados en este trabajo fue de 19-60 bases (en el formato de 72 ciclos). Las coordenadas genómicas correspondientes a exones codificantes para proteínas y “tráilers” 3' de los tRNAs fueron obtenidas del *UCSC Main Table Browser* y utilizadas para anotar las secuencias como CDS o tRF-1. Aquellas secuencias que fueron anotadas con más de una etiqueta (anotación ambigua) fueron manualmente resueltas considerando la presencia o ausencia de mismatches, el largo de la región alineada, y la probabilidad en función de la abundancia relativa.

La abundancia relativa de cada secuencia fue expresada como lecturas por millón de lecturas mapeadas (*reads per million mapped reads*; RPM). Los análisis más particulares se realizaron con algoritmos desarrollados ad-hoc. El análisis estadístico se realizó con el software GraphPad Prism 5 y se detallan los parámetros utilizados en cada caso.

Tabla 1: sets de datos de secuenciación obtenidos. Se indica la plataforma de secuenciación (“PLAT”), el tipo de células, tipo de fracción intra- o extracelular, el número de réplica (“REP”), el número de lecturas crudas obtenido (reads), el número de lecturas a las cuales se les pudo detectar y cortar la secuencia del adaptador 3’ (“+ADAPT”), el número de lecturas que pudieron alinearse al genoma humano (“MAP hg19”) y dicho número expresado como porcentaje del número de lecturas conteniendo adaptador (“MAP%”)

PLAT	CÉLULAS	FRACCIÓN	REP	LECTURAS	+ADAPT	MAP (hg19)	MAP(%)
IGA IIx	MCF-7	INTRA	1	1,359,296	1,146,123	1,089,987	95.1
IGA IIx	MCF-7	INTRA	2	3,340,856	2,337,941	2,215,459	94.8
IGA IIx	MCF-7	p16	1	3,194,756	2,225,449	1,854,902	83.3
IGA IIx	MCF-7	p16	2	3,597,278	2,718,592	2,043,866	75.2
miSEQ	MCF-7	p16	3	6,111,167	3,906,003	3,693,718	94.6
IGA IIx	MCF-7	p100	1	4,098,457	2,837,202	921,790	32.5
IGA IIx	MCF-7	p100	2	3,465,677	1,958,262	932,200	47.6
miSEQ	MCF-7	p100	3	3,883,674	2,375,492	2,238,953	94.3
IGA IIx	MCF-7	S100	1	3,777,176	2,625,730	2,300,006	87.6
IGA IIx	MCF-7	S100	2	14,303,439	7,988,465	6,849,645	85.7
miSEQ	MCF-7	S100	3	5,646,284	5,059,011	4,128,044	81.6
Σ				52,778,060	35,178,260	28,268,570	
IGA IIx	MCF10-A	INTRA	1	3,142,818	2,720,606	2,601,311	95.6
IGA IIx	MCF10-A	INTRA	2	4,080,019	2,261,473	2,166,827	95.8
IGA IIx	MCF10-A	p16	1	1,924,709	1,240,617	398,810	32.1
IGA IIx	MCF10-A	p16	2	3,857,137	540,462	191,753	35.5
miSEQ	MCF10-A	p16	3	1,282,539	602,715	332,905	55.2
IGA IIx	MCF10-A	p100	1	1,529,285	1,253,677	347,602	27.7
IGA IIx	MCF10-A	p100	2	4,147,778	3,201,490	1,193,638	37.3
miSEQ	MCF10-A	p100	3	2,150,827	106,938	45,224	42.3
IGA IIx	MCF10-A	S100	1	8,059,933	5,317,314	4,138,199	77.8
IGA IIx	MCF10-A	S100	2	1,988,548	1,630,760	1,485,025	91.1
miSEQ	MCF10-A	S100	3	4,119,105	3,428,244	2,822,142	82.3
Σ				36,282,698	22,304,296	15,723,436	

2.2.10 “Stem-loop” RT-PCR cuantitativa convencional y digital

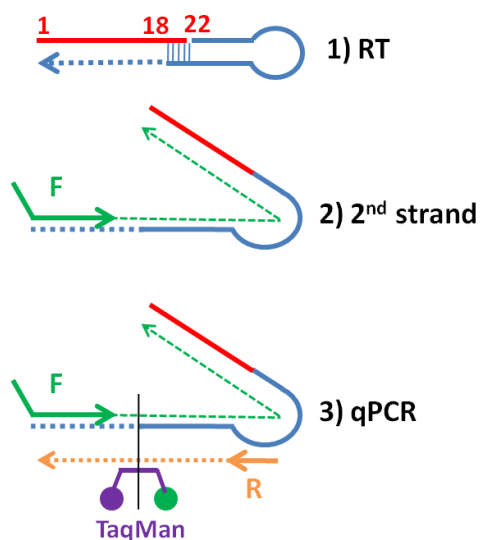
Utilizamos una modificación del popular método del “stem-loop” para la cuantificación de miRNAs (Chen et al., 2005), extendiéndolo a otras especies de RNAs pequeños. Para la detección por PCR digital (Hindson et al., 2013) utilizamos una prueba de hidrólisis (FAM-MGB-NFQ; Life Technologies) universal, correspondiente a una región presente en la horquilla del cebador de retrotranscripción “stem-loop”. Esta prueba de hidrólisis – universal, común a todas las especies a analizar, a diferencia del método original y comercial que propone sondas específicas para cada especie a analizar – permite obtener el mismo grado de sensibilidad y especificidad que el método original, ya que la especificidad en este método viene dado por el cebador de la RT y el primer *forward*, y no por la sonda de hidrólisis (Jung

et al., 2013). Una ilustración esquemática del funcionamiento de este método se suministra en la Figura 10.

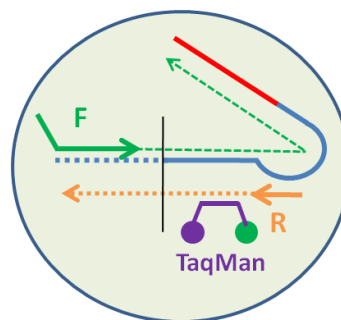
Las muestras de RNA se diluyeron a 10 ng/ μ L y 2 μ L fueron introducidos en el mix para la reacción de RT (Applied Biosystems; Life Technologies), junto con el cebador de stem-loop previamente plegado mediante un protocolo de temperaturas decrecientes (Kramer, 2011), a una concentración final de 10nM. La reacción de RT (12 μ L) se realizó según el siguiente perfil térmico: 16°C por 30 min, 42°C por 30 min y 85°C por 5 min. Se adicionaron 20 μ L de agua libre de nucleasas al cDNA, y 2 μ L de cDNA diluido fueron utilizados para obtener 20 μ L de reacción de PCR, conteniendo además 1X ddPCR master mix (Bio-Rad), 1.5 μ M del primer forward, 0.7 μ M del primer reverse, y 0.8 μ M de la sonda de hidrólisis. La emulsificación de la reacción de PCR y la lectura de la fluorescencia de las gotas fue realizada con un sistema QX100 Droplet Digital PCR System (Bio-Rad), de acuerdo a las especificaciones del fabricante.

Para las reacciones de qPCR convencional (“*real-time PCR*”) se utilizó un 2X qPCR Master Mix (Kapa Biosystems) y un sistema Rotor-Gene 6000 (Corbett Life Science). Los ciclos de cuantificación (Cq) fueron considerados relevantes si fueron al menos dos ciclos anteriores a los controles sin RNA (NTC) o sin retrotranscriptasa (cualquiera fuera el menor), y siempre y cuando las “*curvas de melting*” ofrecieran un pico único en la temperatura esperada. Las secuencias de los distintos cebadores utilizados se suministran en el Anexo Metodológico.

SL RT-qPCR (original method)



SL RT univ-ddPCR



SL RT-qPCR

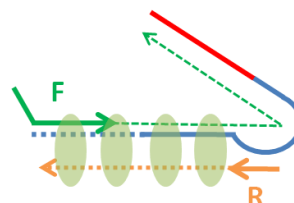


Figura 10: técnica original para la detección de miRNAs por “stem-loop RT-PCR” (Chen et al., 2005) y variantes utilizadas en este trabajo. El método original consiste en el uso de un cebador para la retrotranscripción (RT) que posee una estructura interna de “bucle y horquilla” (stem-loop). Al terminar la estructura de doble hebra, el cebador posee 6 bases desapareadas en su extremo 3'. Dichas bases deben ser exactamente complementarias a las últimas 6 bases del miRNA a analizar (posiciones 18-22, para un miRNA de 22 bases). Finalizada la retrotranscripción, el cebador stem-loop se abre por

acción de la temperatura, y la PCR se realiza mediante un primer forward (verde) cuya secuencia es igual al resto de la secuencia del miRNA a excepción de las últimas 6 bases (y una pequeña extensión 5' para llevar la T_m de dicho primer a 60°C), y un primer reverse (naranja) cuya secuencia es igual a una secuencia originalmente presente en el interior del cebador stem-loop. Para cuantificar la amplificación, se utiliza una sonda de hidrólisis del tipo TaqMan (violeta) cuya secuencia corresponde a la interfaz entre el cebador stem-loop y el extremo 3' del miRNA. Una de las modificaciones que se hicieron en este trabajo fue realizar la detección por PCR digital en gotas (ddPCR), corriendo la región de la sonda TaqMan a una región enteramente ubicada dentro del cebador stem-loop, y por consiguiente común a todos los sRNAs a analizar. En este caso, el uso de la sonda TaqMan no buscó ser una ganancia en especificidad, sino satisfacer un requerimiento técnico del equipo de ddPCR utilizado. En ensayos de qPCR convencional, se suprimió el uso de la sonda de hidrólisis, utilizando un reactivo fluorescente intercalante (SYBR Green) en su lugar.

2.2.11 Ensayos de protección a RNasaA y Proteinasa K

Para los ensayos de protección a RNasaA, se tomaron 3 réplicas de las fracciones p100 y S100 las cuales fueron a su vez divididas en dos alícuotas cada una. Una de dichas alícuotas recibió 4 kunitzU/mL de RNasaA (Sigma) mientras que la otra recibió un volumen equivalente de PBS. Luego de una incubación de 30 min a 37°C, se agregó un volumen adecuado de Trizol LS (que además inactiva a la RNasa) y se purificó el RNA.

Para los ensayos de protección a proteasa, se tomaron 3 réplicas de las fracciones p100 y S100 las cuales fueron a su vez divididas en dos alícuotas cada una. Ambas recibieron 64µg/mL de Proteinasa K (Sigma) y fueron incubadas por 30 min a 37°C, luego de lo cual se adicionó PMSF al 5mM para inhibir a la proteasa. Luego de la inhibición, se adicionó 4 kunitzU/mL de RNasaA a una de las alícuotas; la otra recibiendo igual volumen de PBS. Luego de una incubación de 30 min a 37°C, se agregó un volumen adecuado de Trizol LS y se purificó el RNA.

El RNA obtenido de ambos ensayos fue analizado por SL RT-qPCR específica para miR-21-5p [1-23], tRNA^{Gly}_{GCC}[1-30] y tRNA^{Glu}_{CUC} [1-31]. Se determinaron los ΔC_q para las muestras pareadas (+RNasa / -RNasa) y se compararon para evaluar el efecto del tratamiento con RNasa vs. Sin Tratamiento (ensayo de protección de nucleasa) y del tratamiento combinado con Proteinasa K y RNasaA vs. Proteinasa K únicamente (ensayo de protección de proteasa).

2.2.12 Sub-fraccionamiento de la fracción S100 por cromatografía de exclusión molecular.

La fracción S100 se concentró mediante ultrafiltración (Vivaspin 20, MWCO = 10 kDa; Sartorius Stedim Biotech) desde su volumen inicial (dependiente del número de frascos de cultivo utilizados para cada purificación; típicamente 40 mL) hasta $\approx$ 0.5mL, se diluyó hasta 10mL en PBS, se volvió a concentrar, se volvió a diluir, y se concentró una tercera vez. El objetivo de dicho procedimiento fue sustituir la matriz en la que se encontraban las proteínas y ácidos nucleicos por PBS 1X, sin pérdida apreciable de macromoléculas. Todas las centrifugaciones se realizaron a 4°C y a 5,000g. La muestra concentrada por ultrafiltración fue posteriormente centrifugada a 10,000g por 10 minutos y a 4°C, a efectos de eliminar el

material insoluble. Los sobrenadantes fueron inyectados en columnas de exclusión molecular (Superdex 200, 10/300 y Superdex 75 10/300; Amersham) previamente equilibradas en PBS, y la cromatografía se realizó utilizando un equipo de FPLC (Äkta Purifier; Amersham). La velocidad de flujo fue en todos los casos de 0.5mL/min y la detección se realizó utilizando el espectrofotómetro incluido en el equipo, y registrando los cromatogramas a 280 y 260nm. Se colectaron fracciones de 0.2mL, las cuales fueron posteriormente analizadas por SL-RT-qPCR.

2.3 RESULTADOS

2.3.1 Caracterización de las fracciones extracelulares purificadas

Si bien usamos un protocolo altamente extendido para la purificación de distintas poblaciones de vesículas extracelulares (§2.2.3), debemos reconocer que la centrifugación diferencial no es un método que permita una separación absoluta de distintas poblaciones vesiculares, ya que sus tamaños se encuentran parcialmente solapados (Witwer et al., 2013a). Por tal motivo, preferimos evitar nomenclaturas y definiciones mecanísticas (microvesículas, exosomas, complejos ribonucleoproteicos) y optamos – en un principio – por una nomenclatura meramente fenomenológica (*pellet*16,000g: “p16”; *pellet*100,000g: “p100” y sobrenadante 100,000g: “S100”).

Para evaluar el grado de pureza de dichas fracciones, y su correlato con los tipos de vesículas o estructuras presuntamente enriquecidos en las mismas, recurrimos en primera instancia a una visualización de las muestras por microscopía electrónica de transmisión (tinción negativa). Como puede observarse, la fracción p16 se caracterizó por poseer vesículas de tamaño considerablemente mayor (100nm-500nm; Figura 11, E) que las halladas en la fracción p100 (moda en 60nm; Figura 11, A-B). Si bien la distribución de tamaños pudo verse afectada por los distintos protocolos de fijación, almacenamiento y tinción utilizados, las vesículas de tamaño >200nm fueron observadas únicamente en la fracción p16, independientemente de los procedimientos experimentales. La morfología de las vesículas presentes en la fracción p100 fue reminiscente de la morfología característica de los exosomas (Figura 11, C-D). Las vesículas de mayor tamaño observadas en p16 fueron, por el contrario, de morfología irregular, consistente con una población de microvesículas (Figura 11, E-F). Cabe señalar que, si bien no se detectaron estructuras vesiculares grandes (>200nm) en la fracción p100, la fracción p16 fue una mezcla heterogénea de vesículas grandes y pequeñas (Figura 11, F).

Una técnica alternativa, ampliamente utilizada, es el análisis de seguimiento de nanopartículas (NTA), ya que permite obtener un perfil de tamaños representativo de toda la muestra. Esto también puede obtenerse mediante la técnica de dispersión dinámica de luz (DLS). Pero a diferencia del DLS, el NTA ofrece al mismo tiempo una medición individualizada de cada partícula, por lo que posee una mejor performance en muestras muy heterogéneas.

Básicamente, un equipo de NTA consiste en una cámara de video acoplada al ocular de un microscopio, y de un láser capaz de iluminar a la muestra presente en el mismo. Las partículas dispersan la luz del láser, ocasionando un cono de difracción, que es luego enfocado por el microscopio. Se obtienen así videos donde puede observarse una gran cantidad de puntos luminosos moviéndose a distintas velocidades. Desde luego, el diámetro de dichos puntos luminosos no guarda relación con el diámetro de las partículas que lo generaron; la microscopía óptica no posee resolución como para distinguir partículas de los tamaños aquí involucrados. Pero como dichas partículas se mueven por movimiento browniano y su trayectoria puede seguirse a partir de los videos, la ecuación de Stokes-Einstein ($D=kT/6\pi\eta R$)¹⁸ permite deducir el radio hidrodinámico de cada partícula cuya trayectoria permanezca en foco por algunos segundos.

La Figura 12 muestra los perfiles de NTA de las muestras p16 y p100. Como se desprende claramente, el perfil de tamaños en p16 fue mucho más heterogéneo, con una gran proporción de partículas en el rango 200-500 nm. Por el contrario, la inmensa mayoría de las partículas en p100 se encontraron en el rango 100-200 nm, lo cual es coherente con una población enriquecida en exosomas. Estos resultados fueron complementados con medidas del índice de polidispersión (PDI) evaluado por DLS. Valores de PDI bajos (típicamente en el rango 0.1-0.2) corresponden a muestras mono-dispersas (es decir: con un rango de tamaños relativamente acotado), mientras que valores mayores corresponden a muestras poli-dispersas. Consistente con los resultados de NTA y TEM, el PDI para la muestra p16 fue de 0.432 (muestra heterogénea), mientras que para la fracción p100 fue de 0.175 (muestra mono-dispersa).

Las fracciones vesiculares p16 y p100 también fueron analizadas por Western Blot (Wblot) para evaluar la presencia de los marcadores exosomales TSG101, CD9 y CD63. Como puede apreciarse en la Figura 13, los mismos solo fueron detectados en la fracción p100. Esto no excluye la presencia de exosomas en p16, pero se podría afirmar que a igual cantidad de proteínas la abundancia relativa de dichos marcadores sería mucho menor. Si bien estos marcadores tampoco fueron detectables en S100, dichos resultados no se muestran al entender que no son del todo concluyentes, pues fue muy difícil estimar cuál debería ser la cantidad de proteínas sembradas para un análisis comparable, dado que la mayor parte de las proteínas de dicha fracción corresponden a la Transferrina adicionada al medio de cultivo como aditivo.

Hasta aquí, los resultados de TEM, NTA, DLS y WBlot fueron coincidentes en mostrar que las fracciones p16 y p100 poseen composiciones diferentes, siendo p100 una fracción enriquecida en exosomas mientras que p16 estaría enriquecida en vesículas de mayor tamaño, posiblemente microvesículas.

¹⁸ *D*: coeficiente de difusión (observable); *k*: constante de Boltzmann; *T*: temperatura en escala absoluta; *η*: viscosidad de la disolución (se asume la del agua para muestras acuosas); *R*: radio hidrodinámico.

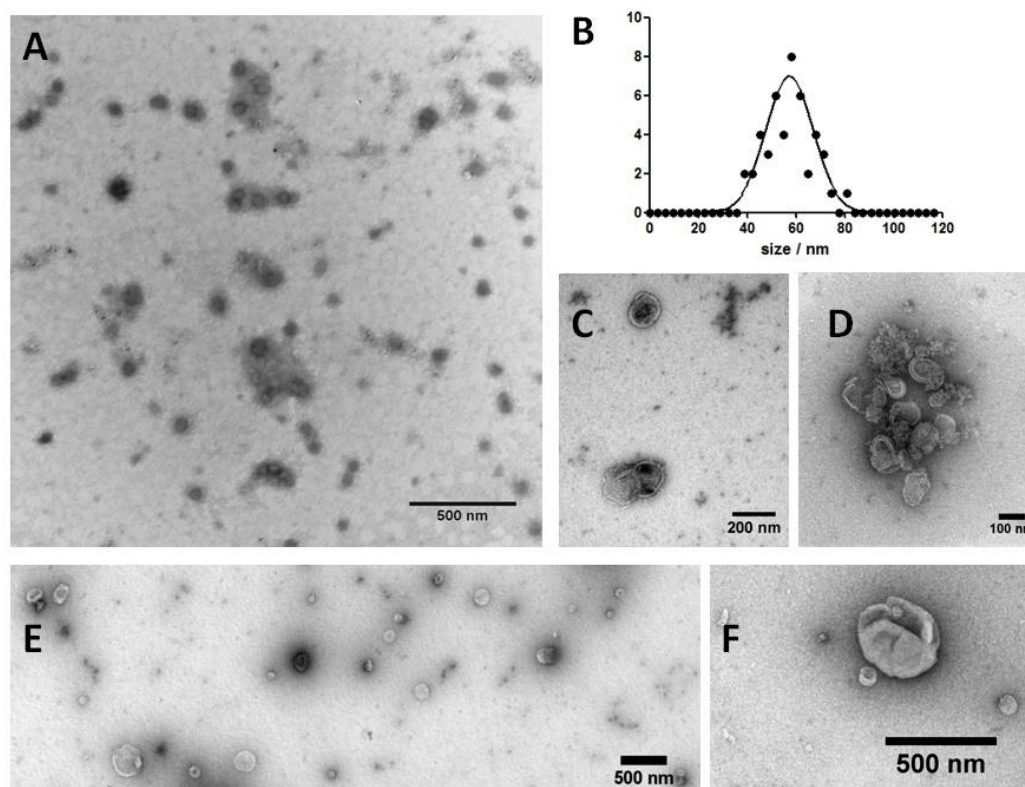


Figura 11:microscopía electrónica de transmisión (tinción negativa) de la fracción p100 (A, C, D) y de la fracción p16 (E y F) del medio condicionado de células MCF-7. Un histograma del tamaño de las partículas visualizadas en (A) se muestra en (B). En (D) puede observarse un “racimo” compuesto por varias estructuras que presentan la morfología característica de los exosomas, y otras tantas estructuras que podrían corresponder a exosomas rotos durante la preparación de los mismos. En la figura (A) se utilizó ácido fosfotúngstico al 2% (2 min de adsorción; 2 min de tinción) como agente contrastante, y un microscopio JEOL JEM-1010. Las muestras fueron, una vez obtenidas, fijadas con glutaraldehído al 1.75% y paraformaldehído al 3.6%, y almacenadas a 4°C por 72 hs previo a su tinción y observación al microscopio. Todas las demás imágenes fueron obtenidas con acetato de uranilo y un microscopio Hitachi 7600, según se indica en §2.2.4. En estos casos, las muestras no fueron fijadas sino congeladas a -80°C previo a su tinción y observación.

Para comparar las fracciones vesiculares con la fracción presuntamente ribonucleoproteica, se procedió a realizar ensayos de protección diferencial a nucleasas y proteasas. En un ensayo inicial (Figura 14, barras claras) tres réplicas independientes de las muestras p100 y S100 fueron analizadas por SL RT-qPCR para distintos pequeños RNAs contenidos tanto dentro como fuera de las vesículas (tRNA^{Gly}_{GCC} [1-30], tRNA^{Glu}_{CUC} [1-31], miR-21-5p [1-23]; ver sección §2.3.2), ya sea en muestras tratadas o no con RNasaA (endoribonucleasa). Al restar los valores de Cq en la situación control y en la situación tratada, se observó una variación casi nula en la muestra p100 (barra 1), lo cual es consistente con el hecho de que los RNAs allí presentes se encuentran en el lumen de los exosomas. En el caso de S100 (barra 3) el grado de protección fue más limitado, obteniendo una diferencia promedial de Cq equivalente a dos ciclos de PCR (no significativa respecto a p100, debido a la alta varianza). En su momento consideramos razonable pensar que los RNAs protegidos por proteínas deberían ser más accesibles a las RNasas que los RNAs protegidos mediante encapsulamiento en vesículas extracelulares. En el Capítulo III de esta tesis se volverá sobre este aspecto.

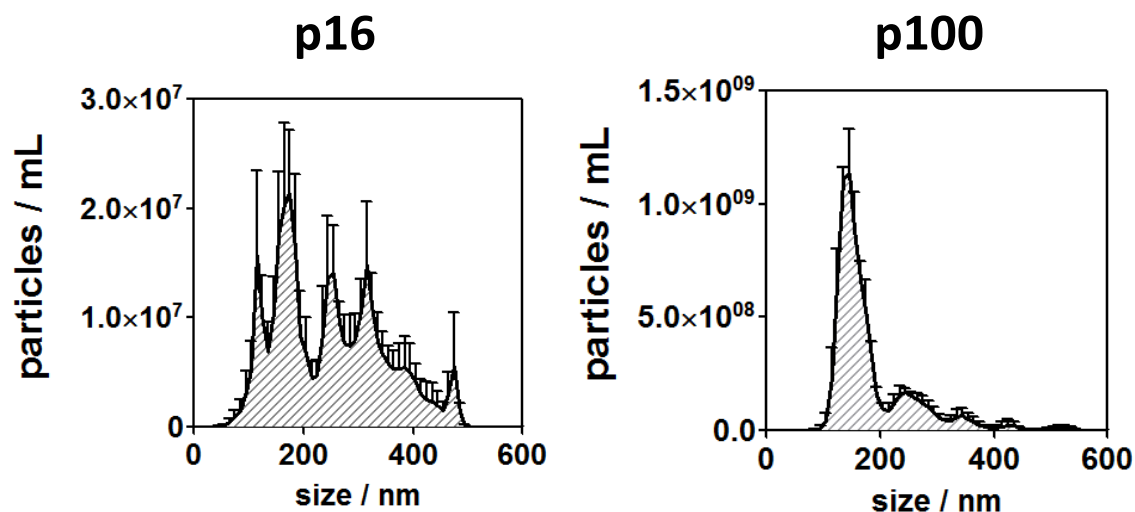


Figura 12: distribución de tamaños (medidas por la técnica de análisis de seguimiento de nanopartículas, NTA) de las fracciones p16 y p100. A la primera le corresponde un valor de índice de poli-dispersión medido por dispersión dinámica de luz (DLS) de 0.432, mientras que dicho valor descendió a 0.175 para la fracción p100.

Para evaluar la sensibilidad diferencial a proteasas (Figura 14, barras oscuras), se realizó un tratamiento en dos etapas. El mismo consistió en añadir en primer término Proteinasa K (PRK), la cual fue posteriormente inhibida. A continuación, se trató la muestra con RNasaA (o con buffer como control) y se purificó el RNA. En el caso de la fracción p100 (barra 2) se observó una cierta sensibilidad a proteasas (significativa respecto a la variación observada con RNasaA sin PRK). Dado que los exosomas son bicapas lipídicas cubiertas de proteínas transmembrana, era de algún modo esperable que la acción proteolítica pudiese tener un efecto sobre su integridad. En cualquier caso, el efecto de la PRK fue mucho más marcado para la fracción S100 (barra 4), obteniéndose una diferencia de 4 ciclos de PCR respecto al efecto provocado por la RNasa sola.

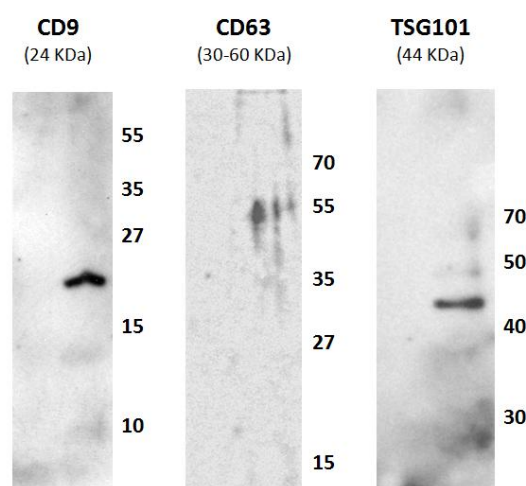


Figura 13: análisis por Western Blot de los marcadores exosomales CD9 (izquierda), CD63 (centro) y TSG101 (derecha). Para el caso de CD63, el anticuerpo utilizado requirió que la corrida electroforética se realizase en condiciones no reductoras. La proteína CD63 tiene 25kDa, pero se encuentra en la superficie de los exosomas presentando un grado de glicosilación variable, por lo que en WBlot suele detectarse entre los 30 y los 60kDa

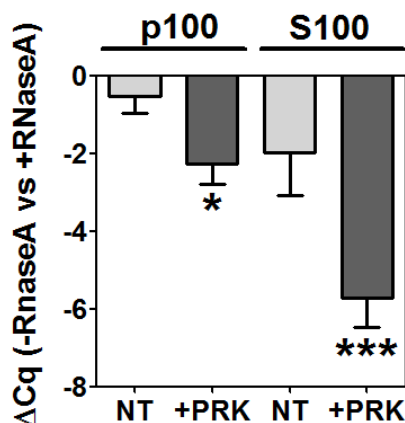


Figura 14: análisis de protección a RNaseA (barras claras) y de protección a Proteinasa K (PRK) + RNaseA (barras oscuras) en la fracción p100 y S100 de células MCF-7. Se determinaron por SL RT-qPCR los niveles de tres especies de RNAs presentes en el medio extracelular ($tRNA^{Gly}_{GCC}$ [1-30], $tRNA^{Glu}_{CUC}$ [1-31], miR-21-5p [1-23]) en tres réplicas independientes de cada fracción (n=9). Se realizaron posteriormente las diferencias entre los valores de Cq para los ensayos pareados con y sin tratamiento con RNaseA. Los asteriscos muestran niveles de significancia (*: $P < 0.05$; ***: $P < 0.001$) para la comparación entre las barras [2 vs. 1] y [4 vs. 3] (test t para muestras pareadas de una cola).

En suma, existe evidencia para suponer que el fraccionamiento del medio extracelular utilizado permitió una buena segregación de las microvesículas, exosomas y complejos ribonucleoproteicos en las fracciones p16, p100 y S100, respectivamente. Esto no implica que la separación haya ocurrido a homogeneidad o que no exista un cierto grado de solapamiento entre las fracciones, vale aclarar. Pero llegado a este punto las condiciones estaban dadas para proceder a realizar secuenciación profunda del contenido de pequeños RNAs en cada una de estas muestras.

2.3.2 Secuenciación profunda del contenido de pequeños RNAs en el medio intracelular y en las fracciones extracelulares (p16, p100 y S100) de las líneas MCF-7 y MCF-10A.

Antes de ingresar en lo que concierne a la secuenciación propiamente dicha, conviene introducir los rendimientos en la purificación de RNA obtenidos para las distintas fracciones. Al lisar la cantidad de células correspondientes a dos frascos de cultivo de 75cm² al 60-70% de confluencia ($\approx 3 \times 10^6$ células MCF-7; 1×10^6 células MCF-10A) se obtuvieron valores en torno a los 29 y 13μg de RNA total, respectivamente (Figura 15). Estos valores son consistentes con la cantidad de RNA presente en una típica célula de mamífero¹⁹. De dicha cantidad de células, y luego de 48 hs de incubación en medio definido, se obtuvo en promedio 155 ng (200 ng) de RNA en p16, 96 ng (57 ng) de RNA en p100, y 524 ng (68 ng) de RNA total en S100 (los valores fuera y dentro de paréntesis corresponden a MCF-7 y MCF-10A, respectivamente). Nótese que aunque la relación intracelular MCF-7 vs MCF-10A es aproximadamente 2.2:1 (en parte debida a la mayor cantidad de células a igual porcentaje de

¹⁹<https://www.qiagen.com/es/resources/faq?id=06a192c2-e72d-42e8-9b40-3171e1eb4cb8>. Al día 11/02/2016

confluencia), dicha relación crece a casi 8:1 para la fracción S100, y la diferencia es significativa ($P < 0.01$). Aun cuando se multipliquen los valores en MCF-10A por un factor de 2.2 para corregir por la diferencia de “carga”, la diferencia entre ambas líneas celulares sigue siendo igualmente significativa.

Obtenido el RNA, se procedió al armado de las librerías para secuenciación profunda. La Figura 16 muestra un esquema y algunos ejemplos representativos del proceso.

Como ya fue explicado, se secuenciaron 3 réplicas biológicas de las fracciones p16, p100 y S100 de las líneas celulares MCF-7 y MCF-10A, así como dos réplicas biológicas del contenido intracelular de pequeños RNAs de dichas células. Además, y a consecuencia de lo constatado y descrito en el Capítulo I de este trabajo, se realizó un secuenciado control “extrayendo RNA” a partir del medio no condicionado de las células MCF-7 (se colocó idéntico volumen en un frasco de cultivo sin células, y se incubó a 37°C por el mismo período, procediéndose luego al mismo tratamiento). De dicho secuenciado control o “blanco” vale decir que las cantidades de RNA obtenidas fueron despreciables, así como la visualización en gel de las librerías construidas. Una vez realizado el secuenciado, se corroboró que la mayor parte de las secuencias obtenidas fueron de origen bacteriano. La abundancia relativa de secuencias humanas fue despreciable. Por ejemplo, se detectaron 7 *reads* correspondientes a miR-21-5p (el miRNA humano más abundante en estas células y en las vesículas correspondientes) frente a los 4,646 *reads* detectados en la muestra p16 corrida y secuenciado en paralelo. La presencia de estas secuencias humanas en bajísima proporción es consecuencia probable del fenómeno de contaminación cruzada descrita en el capítulo I (Tosar et al., 2014), y no se le dio aquí mayor importancia.

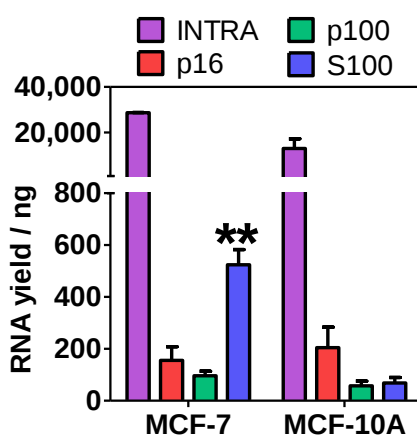


Figura 15: rendimientos a nivel de la purificación de RNA total en la fracción intracelular (violeta), p16 (roja), p100 (verde) y S100 (azul), de las células MCF-7 (izquierda) y MCF-10A (derecha). Los asteriscos señalan las fracciones con un rendimiento significativamente distinto entre ambas líneas celulares (test *t* de Student a dos colas para medidas no apareadas). La cuantificación del RNA se realizó mediante un fluorímetro Quibit (Life Technologies), de acuerdo a las instrucciones del fabricante.

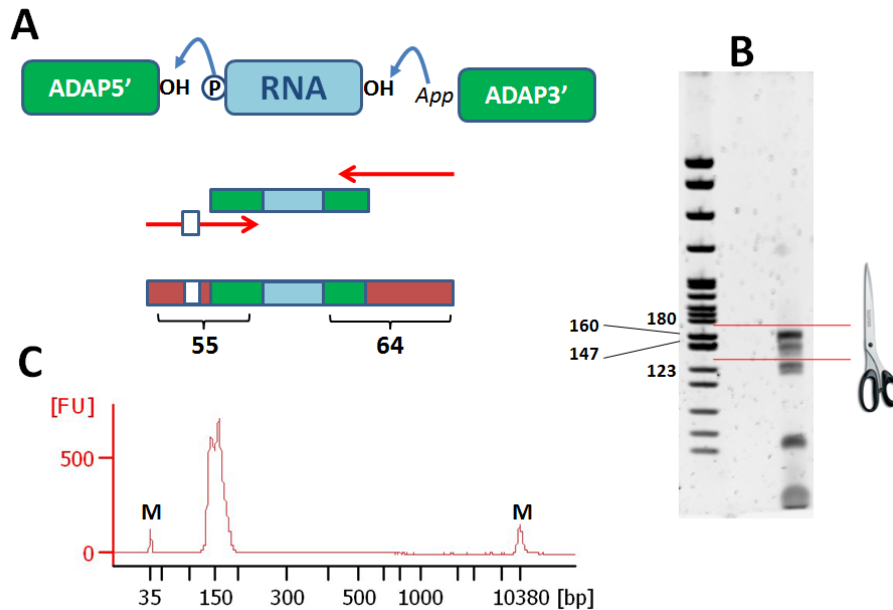


Figura 16: esquema del armado de librerías. A) El RNA purificado conteniendo un extremo 5' fosfato y un extremo 3' OH es ligado a adaptadores 5' y 3' (con un sitio App en 5', de modo de poder realizar la ligación en ausencia de ATP y evitar la formación de productos inespecíficos). Posteriormente se amplifica la construcción por PCR (pocos ciclos; en general menos de 15), utilizando primers (rojo) parcialmente complementarios a los adaptadores, pero teniendo largas colas 5' (donde se agregan los sitios que permitirán la captura en las “flow cell” de Illumina, y el código de barras específico para cada muestra que permite realizar el secuenciado de forma multiplex). Al finalizar, la construcción queda de 119 pb. (55 + 64) más el largo del RNA. B) la construcción se purifica en un gel de poliacrilamida al 6%. Se corta la librería, tomando la precaución de dejar afuera la banda de aproximadamente 120 pb (que corresponde a la auto-ligación de los adaptadores, o a la inclusión de muy pequeños fragmentos de RNA), y hasta un máximo de 180 pb (que corresponde a RNAs de 60 nucleótidos aproximadamente). C) la librería se purifica a partir del gel, se precipita el DNA, y se resuspende en un pequeño volumen. El último control de calidad consiste en una verificación mediante el Agilent Bioanalyzer (chip de DNA de alta sensibilidad) para verificar que la librería contiene únicamente poblaciones del tamaño escogido. M: marcadores internos.

Una vez concluida la etapa de obtención de los datos y finalizado su análisis, se corroboró que las muestras intracelulares estaban dominadas por la población de miRNAs, la cual era seguida – de lejos – por fragmentos de RNAs pequeños nucleolares (snoRNAs) y ribosomales (rRNAs) (Figura 17). En MCF-7, los miRNAs llegaron a representar entre un 15 y un 30% del total de pequeños RNAs mapeados al genoma humano (150,000-300,000 RPM).

Las muestras extracelulares, por el contrario, presentaron un perfil completamente diferente, caracterizado por la predominancia de fragmentos derivados de tRNAs. El caso extremo fueron las fracciones S100 de ambas líneas celulares, donde la suma de los fragmentos de 30 y 31 nt de tRNAs llegaron a constituir entre el 30% (en MCF-10A) y el 60% (en MCF-7) de las secuencias mapeadas. En claro contraste con el medio intracelular, la abundancia relativa de miRNAs en estas muestras estuvo siempre por debajo del 0.5%. También fue muy diferente la proporción de fragmentos derivados de snoRNAs, que en las muestras extracelulares se mantuvo por debajo del 0.1%, mientras que en el medio intracelular alcanzó valores entre el 1.4% (MCF-10A) y 5.4% (MCF-7) de los *reads* mapeados, en promedio.

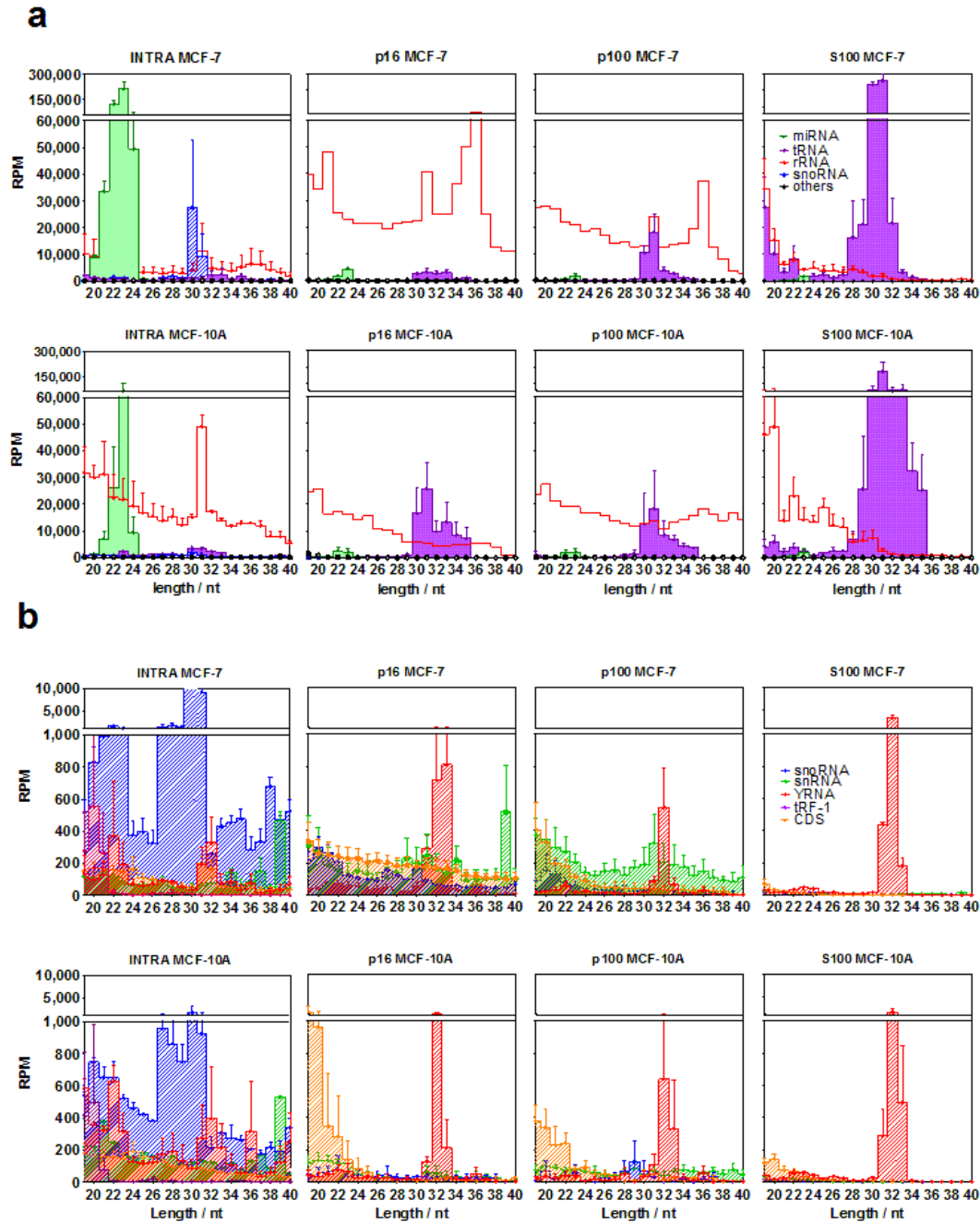


Figura 17: A) resumen de los resultados de secuenciación profunda obtenidos en la fracción intracelular (columna 1), p16 (columna 2), p100 (columna 3) y S100 (columna 4) de las líneas celulares MCF-7 (fila superior) y MCF-10A (fila inferior). Se muestra la abundancia relativa de las secuencias correspondientes (expresada como reads por millón de reads mapeados al genoma humano; RPM) en función del largo de dichas secuencias. Verde: miRNAs, rojo: fragmentos de rRNAs, azul: fragmentos de snoRNAs, violeta: fragmentos de tRNAs. B) La misma representación que (A), pero suprimiendo las especies mayoritarias para poder observar la distribución de tamaños de otros RNAs menos representados. Azul: fragmentos de snoRNAs, verde: fragmentos de snRNAs, naranja: fragmentos de exones, violeta: secuencias 3' tráiler de los tRNAs, rojo: fragmentos de YRNAs.

En secciones posteriores se profundizará en el análisis de las poblaciones de miRNAs (§2.3.3 y §2.3.4) y de los fragmentos de tRNA (§2.3.6y siguientes). De los fragmentos de rRNAs cabe decir que, si bien fueron detectados en altas cantidades en todas las muestras, presentaron la máxima variabilidad entre réplicas biológicas, lo que impidió un análisis más fino. Observaciones destacables son la preferencia marcada por fragmentos correspondientes al extremo 5', salvo para el rRNA 5S donde la mayoría de las secuencias comenzaron entre las posiciones +80 y +100. Tanto a nivel intracelular (en ambas líneas) como en las fracciones p16 y p100 de MCF-7, se observó un enriquecimiento en fragmentos de 31 y 36 nt derivados del rRNA 28S.

Además de las especies antes mencionadas (miRNAs, y fragmentos de tRNA, rRNA y snoRNA), nuestro análisis se extendió a los fragmentos derivados de RNAs pequeños nucleares (snRNAs), fragmentos de YRNAs, fragmentos de RNAs asociados a la partícula Vault, fragmentos de exones (CDS) y las secuencias 3' *trailer* de los genes de tRNAs (tRF-1). Si bien estas especies presentaron abundancias relativas mucho más bajas que las primeras, también hubo sesgos notorios, tales como un enriquecimiento de la fracción S100 en fragmentos derivados de YRNAs (Figura 17 B). Aunque dos órdenes de magnitud por debajo de los fragmentos de tRNAs, compartieron con éstos su estrecha distribución de tamaños, predominando las especies de 31-32 nucleótidos.

A este nivel global de análisis, no se observaron diferencias cualitativas evidentes entre los perfiles derivados de las células tumorales (MCF-7) y no tumorales (MCF-10A). Por el contrario, las mayores diferencias estuvieron presentes entre las muestras intracelulares y extracelulares de ambas líneas. Si bien los perfiles extracelulares fueron en todos los casos similares, las fracciones vesiculares (p16 y p100) fueron mucho más parecidas entre sí que cualquiera de ellas respecto de la fracción S100, donde la diversidad de secuencias se redujo a consecuencia del gran enriquecimiento en fragmentos derivados de tRNAs.

2.3.3 Una aproximación a la [no] selectividad en la secreción de miRNAs

Habiendo obtenido datos de este tipo, el análisis que podría resultar más inmediato sería un análisis de expresión diferencial de miRNAs entre muestras comparables (ej: exosomas) de las células tumorales y no tumorales. Pero ¿qué es lo que realmente arrojaría dicho estudio? Supongamos que se encontrara que los miRNAs A, B, y C fuesen 10 veces más abundantes en los exosomas de MCF-7, ¿podría de allí concluirse que parte de la fisiopatología del cáncer de mama implica la secreción de los miRNAs A, B y C al medio extracelular, tal vez procurando alterar la expresión génica de otros tipos celulares, tales como células del sistema inmunológico?

Dicha línea de razonamiento, no poco frecuente en la bibliografía científica en este campo de estudio, se asienta sobre la idea instalada de que las células secretan activamente aquellos RNAs que por algún motivo – y el motivo favorito es la comunicación intercelular – necesitan enviar al medio extracelular. Ahora bien: ¿será que las células están secretando la mayor parte de los miRNAs A, B y C sintetizados debido a la ganancia que le confiere la secreción de

dichos miRNAs para su propia proliferación? ¿O será que los miRNAs A, B y C están altamente expresados en la célula en cuestión, y aparecen en el medio extracelular en altas proporciones como consecuencia de su alta concentración intracelular? Ambos escenarios, bien distintos entre sí, pueden ser abordados de manera experimental: en el primer caso cabría esperar una alta proporción de A, B y C fuera de la célula, y una baja o moderada abundancia relativa en el medio intracelular. En el segundo escenario, en cambio, A, B y C serían los miRNAs principales tanto dentro como fuera de las células.

Las elucubraciones anteriores pretenden nada más que poner de manifiesto la importancia de, cuando se quieren estudiar los RNAs contenidos en fracciones extracelulares, contar también con la contraparte intracelular como referencia. Podemos adelantar que el perfil de miRNAs encontrados en los exosomas de las células MCF-7 fue distinto del encontrado en los exosomas de las células MCF-10A. Pero el punto argumental que intentamos desarrollar es que concluir sobre dichas diferencias no tiene sentido sin antes prestar atención al perfil intracelular de miRNAs en ambas células.

A nivel intracelular, los perfiles de miRNAs de las células MCF-7 y MCF-10A, si bien contaron con algunas notorias y destacables similitudes, fueron mayoritariamente diferentes. Esto no debería sorprender dado que en definitiva se trata de líneas celulares distintas, más allá del carácter tumoral o no tumoral de las mismas. Además, dado que las condiciones de cultivo ni siquiera fueron iguales²⁰, no tendría mucho sentido un análisis comparativo de este tipo.

Aclarado lo anterior, corresponde señalar al menos las características más sobresalientes. En ambas líneas celulares el miRNA mayoritario fue miR-21-5p, y sus niveles intracelulares – juzgados o estimados a partir de los valores de RPM – fueron en promedio 35 veces superiores (en MCF-7; 17 veces en MCF-10A) al segundo miRNA más abundante. Más aún, el miR-21-5p llegó a representar el 71% y el 58% de todos los miRNAs en MCF-7 y MCF-10A, respectivamente.

No debería sorprender el hecho de que hayamos detectado altos niveles de miR-21-5p en células MCF-7. Este miRNA es reconocido como un “onco-miR” dado que silencia una serie de blancos experimentalmente validados, entre los que se encuentran un número importante de supresores tumorales tales como PTEN, PDCD4 (Wickramasinghe et al., 2009), o TPM1 (Zhu et al., 2007). Además, desde hace años se conoce su sobre-expresión en varios tipos de cáncer humanos, incluyendo el cáncer de mama (Si et al., 2007). Más sorprendente es, tal vez, su detección en altos niveles en células MCF-10A, de origen no tumoral y no tumorigénicas. Aún así, debe considerarse que las células MCF-10A fueron obtenidas a partir de tejido mamario fibroquístico (tumor benigno), y fueron espontáneamente inmortalizadas en cultivo (Soule et al., 1990), por lo cual tampoco deben racionalizarse como tejido mamario normal.

Como puede apreciarse en la Figura 18 (A, C), las células MCF-7 presentaron un quiebre en las abundancias relativas de los 456 miRNAs en ellas detectados, donde el subconjunto de los

²⁰no se priorizó un diseño experimental que permitiese comparar estas líneas entre sí, sino buscar condiciones óptimas de cultivo de cada línea celular en medio libre de suero, de modo de mantener la liberación de cuerpos apoptóticos cerca de los mínimos posibles

16 miRNAs más abundantes (encabezados por miR-21-5p como ya se hizo referencia) llegó a representar el 91% del total. Este patrón de expresión bimodal no fue observado en MCF-10A (Figura 18, B). Además, el número de miRNAs detectados en dichas células (333) fue considerablemente inferior.

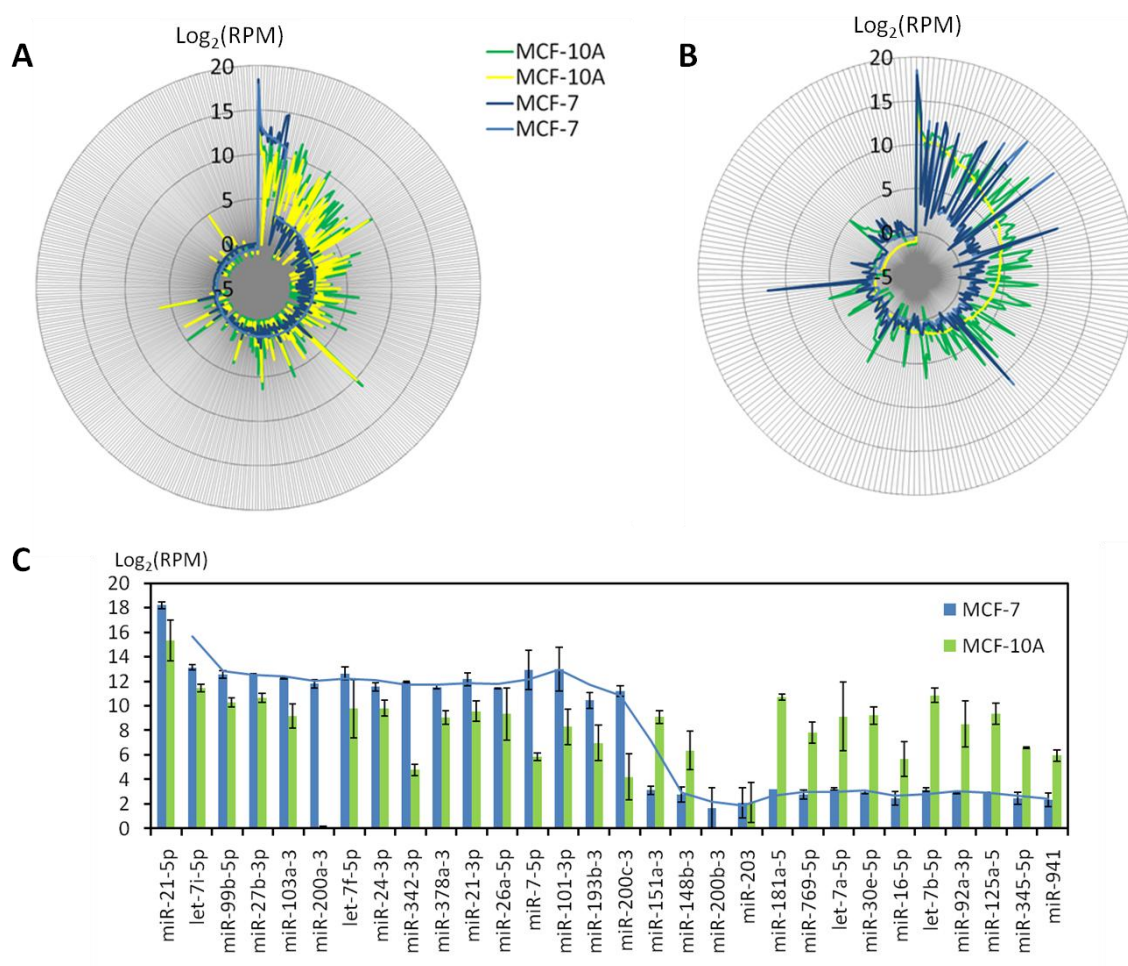


Figura 18: comparación de los perfiles de expresión de miRNAs (fracción intracelular) entre las células MCF-7 (celeste: réplica 1; azul: réplica 2) y MCF-10A (amarillo: réplica 1; verde: réplica 2). A) Gráfico de caracol, en el cual la expresión intracelular de cada miRNA se representa como una línea que se extiende desde el centro hacia el exterior, y cuyo largo corresponde al valor del logaritmo en base 2 de los reads por millón de reads mapeados [$\text{Log}_2(\text{RPM})$]. Partiendo del miR-21-5p (ubicado a las 12 en punto), los miRNAs están ordenados en sentido horario según su expresión en la réplica 1 de MCF-7. B) ídem, pero los miRNAs están ordenados según su expresión en la réplica 1 de MCF-10A. En ambos casos, se eliminaron aquellos miRNAs no expresados en las células utilizadas para el ordenamiento. C) Gráfico de barras representando los valores de $\text{Log}_2(\text{RPM})$ de los 30 miRNAs más expresados en MCF-7, y ordenados según su expresión en la réplica 1 de dichas células (celeste). También se muestra la expresión de dichos miRNAs en MCF-10A (verde). Las barras de error representan el error estándar de la media. La línea quebrada celeste corresponde a la media móvil (período: 2) de las barras correspondientes a MCF-7.

Como ya se ha explicado, nuestro objetivo no fue realizar una comparación detallada entre los perfiles de expresión de miRNAs de las células MCF-7 y MCF-10A, sino comparar los perfiles intra- y extracelulares en cada tipo celular. Para cualquiera de las tres fracciones

extracelulares estudiadas (p16, p100, S100), se observó una fuerte correlación con los perfiles intracelulares²¹. Las gráficas presentadas en la Figura 19 (A-C) muestran la abundancia relativa de cada miRNA (representado como un punto) en cada fracción extracelular, en función de sus abundancias relativas intracelulares. Como puede observarse, los miRNAs cuya expresión intracelular se encontró por debajo de los ≈ 5 RPM fueron indetectables (o detectados a niveles muy bajos) en las fracciones extracelulares. A partir de dicho punto, el incremento en la abundancia intracelular tuvo como correlato un incremento en la abundancia extracelular, pudiéndose ajustar los datos mediante una recta de pendiente 1 (considérese que las abundancias están expresadas en escala logarítmica). Ciertamente, este tipo de patrón era el esperado a partir de un modelo de secreción no selectiva a nivel global. A la luz de dicho modelo la secreción de cada miRNA sería independiente de su secuencia o de su función, y su nivel de expresión intracelular actuaría como la única variable capaz de predecir su abundancia relativa en el espacio extracelular. Dicho de otro modo, la probabilidad de secreción de cada molécula individual es la misma, y por tanto la tasa de secreción de cada especie (ej: miR-21-5p, miR-16-5p, let-7a-5p, etc.) es proporcional a la representación de dicha especie en el total de miRNAs intracelulares.

Para formalizar lo anterior se desarrolló un modelo teórico que asume que los miRNAs son secretados de forma pasiva. Tomando el valor observado – experimental – de representación intracelular de cada miRNA (p_i), se predijo su abundancia relativa extracelular. La misma se calculó de la siguiente manera:

$$p_{i,e} = p_i \times FS_e$$

Donde $p_{i,e}$ corresponde a la representación extracelular del miRNA “i” en la fracción extracelular “e”, p_i es su correspondiente representación intracelular (expresado en RPM), y FS_e es un “factor de secreción” que es común a todos los miRNAs en la fracción extracelular “e”. Precisamente, la hipótesis de secreción pasiva implica que FS_e debe de ser constante para una determinada fracción extracelular. Por el contrario, de existir selectividad en la secreción, FS_e debería modelizarse como una variable propia de cada miRNA.

El valor de FS_e se determinó realizando el siguiente cociente:

$$FS_e = \frac{\sum_i RPM_{i,e}}{\sum_i RPM_{i,INTRA}}$$

Es decir que el “factor de secreción” corresponde al cociente entre el total de los *reads* mapeados que son miRNAs en la fracción extracelular “e” y en el medio intracelular. Es un corrector, por consiguiente, de las grandes diferencias en la abundancia de miRNAs en las fracciones intra- y extracelulares (§2.3.2).

²¹ De aquí en adelante, el análisis se centra únicamente en la línea celular MCF-7 debido a que se obtuvo una fuerte correlación en el perfil de miRNAs intracelulares entre réplicas biológicas (Pearson $r = 0.9265$). Dado que este valor fue inferior a 0.9 en MCF-10A, las gráficas resultantes presentan demasiada dispersión, pero la tendencia es la misma.

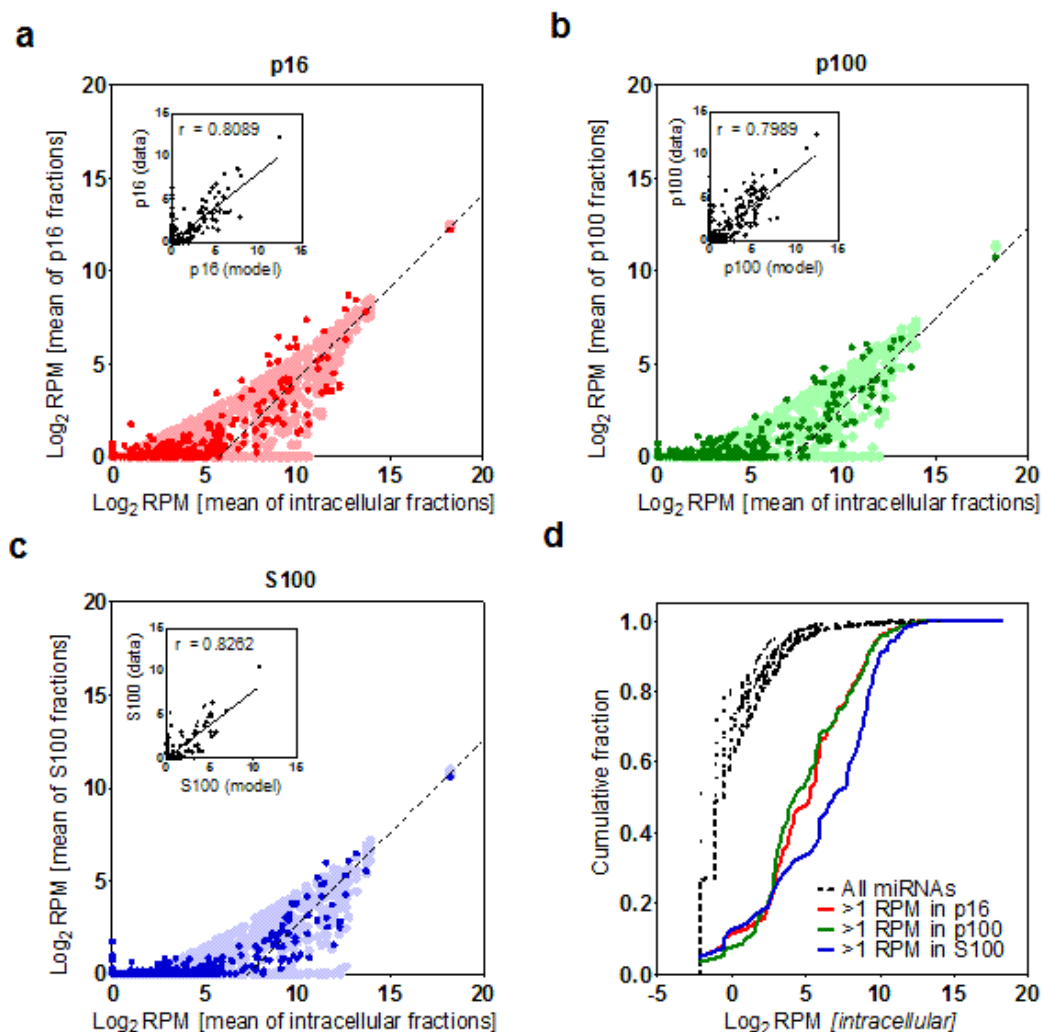


Figura 19: las abundancias relativas de los miRNAs en las fracciones extracelulares de las células MCF-7 son consistentes con un modelo de secreción no selectiva. La abundancia relativa promedio (en escala logarítmica de base 2) de cada miRNA en las fracciones extracelulares p16 (A), p100 (B) y S100 (C) se graficó en función de los niveles intracelulares correspondientes. Manualmente se agregaron líneas de pendiente igual a 1. Un modelo teórico que asume secreción pasiva fue superpuesto (sombras más claras). En los recuadros se muestra la correlación (Pearson) entre dicho modelo y los datos experimentales. D) Distribución acumulada de los valores de Log₂RPM para los miRNAs intracelulares. La línea negra punteada corresponde al total de miRNAs intracelulares en MCF-7 ($n = 456$). Se repitió el mismo análisis, pero incluyendo únicamente aquellos miRNAs con RPM > 1 en la fracción p16 (línea roja, $n = 119$), p100 (línea verde, $n = 114$) y S100 (línea azul, $n = 80$). Como control, 5 conjuntos de 80 miRNAs tomados de manera aleatoria fueron también analizados (puntos negros).

Llegado a este punto se realizó un contraste de hipótesis para evaluar si los datos experimentales se ajustaban o no al modelo propuesto. En primer lugar, se definió como hipótesis nula (H_0) la secreción no selectiva de los miRNAs. Posteriormente se evaluó la correlación entre los valores de $p_{i,e}$ experimentales y teóricos (es decir: la abundancia relativa de cada miRNA en el espacio extracelular obtenida a partir de los datos de secuenciado, o predicha a partir de los valores intracelulares y asumiendo la hipótesis nula). Como se muestra en los recuadros interiores de la Figura 19, A-C, los coeficientes de correlación de Pearson (r)

fueron: 0.8089, 0.7989 y 0.8262 para las fracciones p16, p100 y S100, respectivamente. Los P-valores (a dos colas) asociados a dichos coeficientes de correlación son todos < 0.0001 , por lo cual la similitud entre los valores teóricos y experimentales no puede suponerse obra del azar. Dado que la distribución predicha de los datos bajo la hipótesis nula coincide con los valores experimentales, concluimos que la misma no puede rechazarse. Es decir, no hay elementos para suponer que la secreción de los miRNAs no sea no selectiva²².

Para expresar lo mismo de manera gráfica, se tomaron los valores de $p_{i,e}$ predichos por el modelo y se les agregó ruido para reseñalar la varianza propia del muestreo. Se calculó una desviación estándar teórica para cada punto según una distribución binomial (σ_{binom}), y el valor graficado en el eje de las ordenadas se computó como un número aleatorio en el intervalo $p_{i,e} \pm 6\sigma_{\text{binom}}$. Superponiendo las distribuciones obtenidas para 7 repeticiones de este modelo al que se le agregó ruido, se definieron las áreas sombreadas que aparecen en la Figura 19, A-C. Como puede observarse, la mayoría de los puntos caen dentro de dichas áreas y siguen una tendencia similar.

Una última manera de analizar estos mismos datos, que pueda tal vez servir para esclarecer lo que intentamos comunicar: la Figura 19, D, muestra un análisis de frecuencia acumulada para la expresión intracelular de los miRNAs en MCF-7 (línea negra punteada). De la gráfica se desprende, por ejemplo, que un 60% de los miRNAs tienen una abundancia intracelular menor a 1 RPM (2^0), y cerca de un 90% están por debajo de los 32 RPM (2^5). Esto es así cuando se analiza el total de los miRNAs detectables en MCF-7 ($n = 456$). Si se toman, en cambio, únicamente aquellos que presentan una expresión extracelular elevada (>1 RPM), la distribución de frecuencias acumuladas cambia drásticamente y se mueve hacia mayores niveles de expresión intracelular. Por ejemplo, únicamente el 10% de los miRNAs con >1 RPM en S100 ($n = 80$) tienen niveles intracelulares menores a 1 RPM (en comparación con el 60% de la población global). Lo que esto quiere decir es que aquellos miRNAs detectados en niveles apreciables en las fracciones extracelulares se corresponden con aquellos de mayor expresión intracelular. Controles realizados con un número comparable de miRNAs tomados al azar permiten descartar que los corrimientos observados se debiesen únicamente a fenómenos de muestreo.

Otra consecuencia de la secreción no selectiva es que no deberían existir diferencias en el perfil de miRNAs entre las distintas fracciones extracelulares, ya que cada fracción actuaría como una pequeña muestra representativa del contenido intracelular. Efectivamente, los coeficientes de correlación de Pearson (r) para la comparación de las fracciones (promedios) p16 vs. p100, p16 vs. S100 y p100 vs. S100 fueron siempre > 0.9 . Más aún, el coeficiente de correlación p16 vs. p100 (0.9794) fue superior al obtenido entre las propias réplicas biológicas de p16 (0.9494) y de p100 (0.9348). Esto quiere decir que a pesar de las diferencias

²² Esta frase es un poco barroca, pues tiene incorporada tres negaciones. La opción simplista: “la secreción es no selectiva” puede resultar una alternativa más correcta desde el punto de vista lingüístico, pero es formalmente incorrecta: decidir no descartar la hipótesis nula no implica aceptarla. No estamos aquí descartando que haya mecanismos moleculares que provoquen que algunos miRNAs tiendan a ser secretados antes que otros. Pero de existir tales mecanismos, sus efectos no estarían logrando alterar significativamente el patrón global, y destruyendo la correlación entre las abundancias relativas intra y extracelulares.

estudiadas en §2.3.1, estas fracciones son indistinguibles entre sí en lo que respecta a su contenido de miRNAs.

2.3.4 Identificación de miRNAs candidatos a secreción selectiva

El subtítulo anterior puede parecer contradictorio con todo lo desarrollado en la sección §2.3.3, pero se trata en realidad de cosas distintas. Hasta ahora hemos venido analizando el perfil de miRNAs de una manera global, a efectos de elucidar la cuestión sobre si la secreción es o no mayoritariamente selectiva. Nuestros resultados no apoyan, en absoluto, lo que muchos autores sostienen en lo que respecta a la secreción de miRNAs. El proceso de secreción²³, según nuestros datos, sería mayoritariamente del tipo no selectivo, dado que los miRNAs más abundantes son los más representados en las fracciones extracelulares, y viceversa. Aún así, un pequeño número de miRNAs podrían ser secretados de manera selectiva (o selectivamente retenidos), y en la medida en que dicho número no fuese demasiado alto, esto no afectaría la distribución general. Dicho de otra manera: en un contexto donde la mayoría de los miRNAs trascienden los límites de la célula a una razón proporcional a su concentración citoplasmática, algunos de ellos podrían tener una tasa de secreción significativamente más alta o más baja. Esto podría ser consecuencia de mecanismos moleculares específicos para la discriminación, captura y secreción de algunos miRNAs particulares, o consecuencia indirecta de fenómenos no relacionados con la secreción (por ejemplo: un miRNA altamente expresado pero direccionado al núcleo presentaría una baja tasa de secreción²⁴).

Si se observa con detenimiento la Figura19, A-C, puede apreciarse que algunos puntos se ubican por fuera del área predicha por el modelo teórico. Esto de por sí solo no implica que los mismos sean excepciones del modelo. En dichas gráficas cada punto viene representado por el promedio de los valores intra- y extracelulares, y se pierde la información respecto de la dispersión entre réplicas asociada a dicho punto. Lo esperable para un miRNA candidato a ser selectivamente secretado no es que su extracelular sea mayor o menor a la predicha a raíz de su abundancia intracelular. Lo esperable es que dicho corrimiento sea además lo suficientemente grande para superar el ruido de fondo (dado por la variabilidad biológica, experimental, y la propia estadística de muestreo), y que sea además sistemático, reproducible.

²³ **Aclaración necesaria:** nuestro abordaje es meramente observacional y a posteriori, y no incursionamos en estudios mecanísticos. De aquí que cuando decimos “proceso de secreción” no estamos en realidad haciendo referencia a un proceso de secreción en el sentido biológico del término. Los miRNAs bien podrían acceder el espacio extracelular por mecanismos moleculares especializados, lo mismo que por otros fenómenos (ej: muerte celular necrótica y liberación de los complejos AGO2/miRNA al medio extracelular, fracción S100). En el caso de la fracción exosomal, uno tiende a pensar que el “proceso de secreción” coincide con el encapsulamiento dentro de los exosomas en formación (cuando se invaginan los endosomas para formar los cuerpos multivesiculares). Aún así, otros podrían ser los mecanismos involucrados y esto no afectaría nuestro mensaje. Lo importante es que la evidencia no respalda un modelo que asumiese que sólo unos pocos miRNAs son secretados en altas cantidades. En dicho caso, la correlación entre los perfiles intra- y extracelulares debería ser baja, cosa que no observamos.

²⁴ en este ejemplo se manifiesta una de las limitaciones de nuestro abordaje experimental: nuestro observable es la abundancia intracelular estimada a partir de los valores de RPM, cuando a los efectos de la secreción pasiva el parámetro más relevante sería la concentración citoplasmática.

Para abordar esta cuestión, se calcularon las abundancias relativas normalizadas al total de miRNAs en cada fracción²⁵ (RPM-miRNA), lo que permitió realizar comparaciones entre fracciones con alto (INTRA) y bajo (p16, p100, S100) porcentaje de miRNAs. Utilizando dicho parámetro, se calculó el “*fold-change*” (cociente) para cada miRNA entre las fracciones extracelulares y la fracción intracelular. A continuación, utilizando los valores de todas las réplicas biológicas, se computó el P-valor²⁶ correspondiente a la significancia estadística de que dicho “*fold-change*” sea distinto de 1. En la Figura 20 se muestran los “gráficos de volcán” obtenidos para la comparación de las distintas fracciones extracelulares con la fracción intracelular. Los mismos se obtienen al graficar el negativo del logaritmo decimal del P-valor en función del logaritmo en base 2 del “*fold-change*”. Cada punto corresponde a una determinada especie de miRNA. Las líneas verticales discontinuas indican la posición -2 y +2. Por consiguiente, cualquier punto hacia la derecha de la línea derecha o hacia la izquierda de la línea izquierda está – en promedio – más de 4 veces (2^2) más enriquecido o disminuido en la fracción extracelular respecto de la intracelular, respectivamente. La línea horizontal discontinua indica la posición +1.3, que es la que corresponde a un P-valor de 0.05 ($-\text{Log}_{10}0.05 = 1.301$) Por tanto, cualquier punto por encima de dicha recta se consideró estadísticamente significativo.

Las gráficas de volcán tienen la ventaja de que permiten rápida y visualmente identificar a los puntos relevantes como aquellos ubicados en los cuadrantes superiores izquierdo y derecho (definidos por las intersecciones de las líneas discontinuas antes mencionadas). De los 456 miRNAs detectados en MCF-7, solamente se encontraron 8 que estuvieron sobre-representados en las muestras extracelulares (3 en p16, 5 en p100, ninguno en S100) y 4 que estuvieron sub-representados (1 en p16, 1 en p100, 2 en S100). Es decir, solamente un 2.6% de los miRNAs presentaron niveles consistentes con la hipótesis de secreción selectiva. Más aún, si el nivel de confianza estadística se elevase un poco y solamente se considerasen significativos aquellos miRNAs con $P < 0.01$, únicamente el let-7d-5p y el miR-423-5p permanecerían como candidatos (0.4%).

¿Son estos los únicos miRNAs que cabría investigar si se pretendiese descubrir posibles mecanismos moleculares de secreción activa y específica (en MCF-7, al menos)? La respuesta es un contundente no. Evidentemente, la lista final de posibles candidatos se modifica en función de cómo uno defina los límites de los cuadrantes en las gráficas de volcán, lo que es de por sí una acción arbitraria. Que un miRNA esté presente en la lista de candidatos no implica que efectivamente haya sido secretado de manera selectiva (su enriquecimiento puede deberse a factores aleatorios, o a artefactos experimentales tales como la contaminación cruzada; Capítulo I). Análogamente, miRNAs selectivamente secretados pueden mostrar una abundancia extracelular mayor a la esperada en función de su grado de expresión intracelular, pero la variabilidad entre réplicas puede provocar que dicho resultado sea poco significativo.

²⁵ $\text{RPM} = \text{número de reads} / \text{sumatoria de reads mapeados al genoma humano} * 1 \times 10^6$

$\text{RPM-miRNA} = \text{número de reads} / \text{sumatoria de reads correspondientes a miRNAs} * 1 \times 10^6$

²⁶ Prueba t de Student-Welch, dado que las varianzas no fueron homogéneas entre las muestras intra- y extracelulares

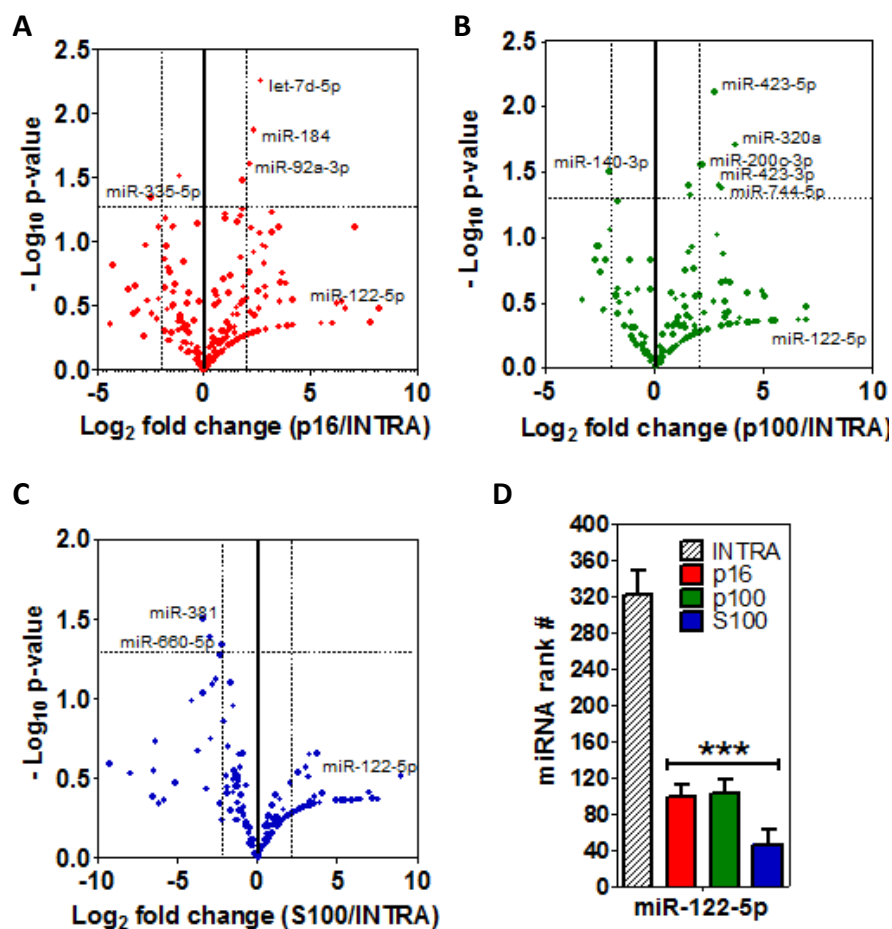


Figura 20: detección de miRNAs enriquecidos en las fracciones extracelulares. A-C) Gráficos de volcán mostrando aquellos miRNAs significativamente enriquecidos (cuadrante superior derecho) o menguados (cuadrante superior izquierdo) en las fracciones p16 (A), p100 (B) y S100 (C) respecto al contenido intracelular de las células MCF-7. D) Análisis por ranking (posición en una lista de miRNAs ordenados por abundancia) mostrando el enriquecimiento del miR-122-5p en las fracciones extracelulares. Las estrellas (***) denotan $P < 0.0001$ (ANOVA de una vía con test post hoc de Tukey).

Lo anterior aplica especialmente para aquellos casos que son tal vez los más interesantes: miRNAs de muy baja expresión intracelular y comparativamente alta – pero aún así baja en términos absolutos – abundancia extracelular. Cuando los valores son bajos, la variabilidad es mucho más grande (por numerosos motivos; siendo el más evidente el hecho de que los errores relativos aumentan a medida que disminuye la magnitud de la medida). Esto provoca que, por debajo de un cierto valor de RPM a nivel extracelular, sea muy difícil que un positivo verdadero llegue a alcanzar niveles de significancia (si no se aumenta el número de réplicas, o se disminuye la varianza de los datos de alguna manera).

Habiendo hecho estas consideraciones, nos llamó la atención que uno de los puntos que presentaban mayor *fold-change* en la Figura 20, A-C, era el miR-122-5p. Si bien los P-valores fueron demasiado altos (por lo que dicho *fold-change* bien podría ser obra del azar), el enriquecimiento promedio en la fracción S100 llegó a ser de 500 veces (2^9), lo cual no es un número despreciable. Pero lo que realmente nos llamó la atención es que miR-122-5p es famoso por ser uno de los miRNAs con mayor expresión tejido-específica y siendo la misma

restringida al hígado (Landgraf et al., 2007), donde alcanza altísimos niveles y actúa como regulador del metabolismo de los ácidos grasos.

Concordantemente, el miR-122-5p fue casi indetectable en células MCF-7 (0.45 RPM en réplica 1, 0 RPM en réplica 2). Sería esperable, por tanto, que tampoco fuese detectable en las muestras extracelulares, donde los valores de RPM asociados a cualquier miRNA fueron órdenes de magnitud más bajos. Sin embargo, alcanzó valores promedio de 1.4, 0.3 y 9.5 RPM en p16, p100 y S100, respectivamente. Por los reparos expuestos anteriormente, decidimos recurrir a un análisis que no fuese tan vulnerable a los valores de abundancia bajos. Es así que ordenamos jerárquicamente todos los miRNAs en cada muestra según su abundancia relativa, de mayor a menor. En el medio intracelular, a miR-122-5p le correspondió la posición #320. En cambio, ocupó las posiciones #100, #104 y #48 en p16, p100 y S100. Este aumento de *ranking* en las fracciones extracelulares sí fue altamente significativo ($P < 0.0001$; test de ANOVA con análisis *post hoc* de Tukey, Figura 20, D).

El resultado anterior fue posteriormente verificado por PCR cuantitativa, como se muestra en la sección §2.3.7. Si bien sus abundancias relativas, aún las extracelulares, estuvieron siempre en la “región de ruido” y no puede descartarse que su presencia extracelular fuese un artefacto - ¿puede alguna vez enteramente descartarse? -, no deja de ser un resultado inquietante. En 2012 se realizó un estudio de miRNAs enriquecidos en sangre de pacientes con cáncer de mama, y uno de los miRNAs más significativos fue precisamente el miR-122 (Wu et al., 2012). Los autores asumieron que podría tratarse de una respuesta hepática frente al tumor, pero ¿no permitirían nuestros resultados preguntarse si no sería la misma célula mamaria tumoral la responsable de liberar dicho miRNA al torrente sanguíneo? Después de todo, el miR-122-5p es un regulador importante del metabolismo energético y las células tumorales suelen ser ávidas de mayores fuentes de energía. En enero de 2015 (ya con nuestro manuscrito enviado y en revisión) la revista *Nature Cell Biology* publicó un artículo que lleva por título: “*Breast-cancer-secreted miR-122 reprograms glucose metabolism in premetastatic niche to promote metastasis*”²⁷ (Fong et al., 2015).

2.3.5 Enriquecimiento de las fracciones extracelulares en fragmentos derivados del RNA Y4

Los YRNAs son pequeños RNAs de 84-113 nt transcritos por la RNA polimerasa III que se asocian a las proteínas Ro y La para formar complejos ribonucleoproteicos cuya función no está clara todavía. Sí se sabe que, por su unión a la proteína Ro60, participan en el control del plegamiento de otras moléculas de RNA, y por mecanismos independientes también regulan la replicación del DNA (Kowalski and Krude, 2015). En determinadas circunstancias, estas moléculas pueden ser clivadas a fragmentos más cortos, de 22-27 nt y de 27-36nt, en un mecanismo independiente de la enzima Dicer (Nicolas et al., 2012). Recientemente, la presencia de dichos fragmentos ha sido descrita en vesículas extracelulares de células inmunes (Nolte-'t Hoen et al., 2012), en exosomas presentes en el semen (Vojtech et al.,

²⁷ El miR-122 secretado por las células de cáncer de mama reprograma el metabolismo de la glucosa en el nicho pre-metastático para de ese modo promover la metástasis.

2014), y en la sangre (Dhahbi et al., 2013a). De las cuatro especies de YRNAs que existen en humanos (Y1, Y3, Y4 e Y5), los autores encontraron que los fragmentos transportados por el torrente sanguíneo derivaban casi exclusivamente del RNA Y4.

En nuestro estudio, fragmentos de YRNAs fueron encontrados en todas las muestras, con abundancias relativas que variaron entre los 1,000 y los 10,000 RPM. En casi todos los casos, los fragmentos correspondieron al extremo 5' exacto (comenzaban en la posición +1). En un plano global, no se observaron diferencias cuantitativas entre las poblaciones intra- y extracelulares. Sin embargo, al clasificar los fragmentos según la especie de YRNA de origen, se observaron notables diferencias. Mientras que en el medio intracelular se detectaron cantidades similares de fragmentos derivados de los RNA Y3, Y4 e Y5, más del 85% de los fragmentos extracelulares correspondieron al RNA Y4 únicamente. Más aún, dichos fragmentos fueron casi exclusivamente de 31-32 nt, mientras que en la fracción intracelular la distribución de tamaños fue más heterogénea (Figura 21).

Resulta llamativo haber encontrado un enriquecimiento de los fragmentos del RNA Y4 en el medio extracelular, precisamente porque éstos han sido los únicos fragmentos hallados en grandes cantidades en la sangre (Dhahbi et al., 2013a). Las próximas secciones se centrarán en el estudio de otras especies muy enriquecidas en el medio extracelular: las mitades de tRNA. No son los signos de secreción diferencial lo único que tienen en común entre si estas especies. En bacterias, los YRNAs actúan como moléculas miméticas de los tRNAs, y son blanco de la misma maquinaria de procesamiento celular (Chen et al., 2014).

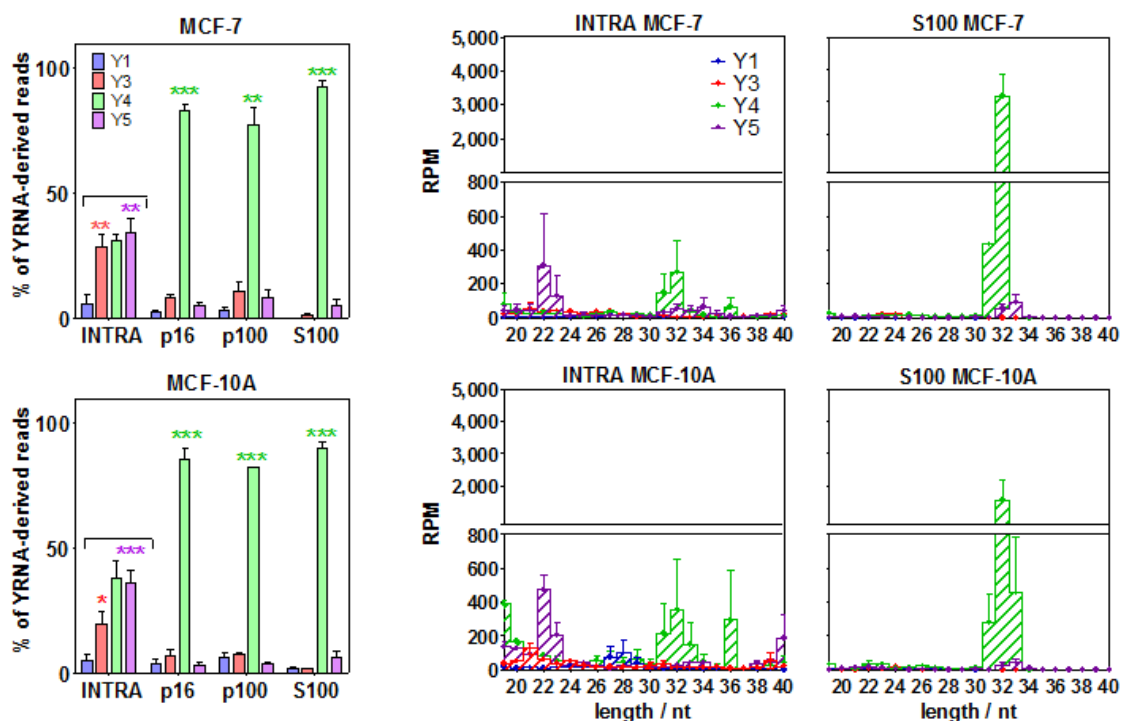


Figura 21: enriquecimiento de los fragmentos derivados de RNA Y4 en el medio extracelular. Izquierda: comparación del porcentaje de secuencias correspondientes a YRNAs derivadas del RNA Y1 (azul), Y3 (rojo), Y4 (verde) e Y5 (violeta). Las figuras superior e inferior corresponden a MCF-7 y MCF-10A, respectivamente. Cuando hubo muestras significativamente diferentes (ANOVA de una vía), se colocaron estrellas en las de porcentajes más altos. *: $P < 0.05$; **: $P < 0.01$; ***: $P < 0.001$

(test post hoc de Tukey). Gráficas del centro y derecha: distribución de tamaños y abundancias relativas de los fragmentos de YRNAs en el medio intracelular (centro) y en la fracción S100 (derecha). Las fracciones p16 y p100 no se muestran, pero cualitativamente fueron muy similares a la S100 en ambas líneas celulares.

2.3.6 La fracción S100 está muy enriquecida en “complejos ribonucleoproteicos” conteniendo mitades 5' de tRNA-Glu y tRNA-Gly.

Como ya se adelantó en la sección §2.3.2., los fragmentos derivados de tRNAs fueron muy minoritarios en el medio intracelular, y sin embargo fueron una de las categorías principales en las fracciones extracelulares. Ya en reportes anteriores las mitades de tRNAs (en adelante tRNA-h, por *tRNA halves*) habían sido descritas en vesículas extracelulares de células inmunes (Nolte-t Hoen et al., 2012) y en exosomas presentes en el semen (Vojtech et al., 2014). Sin embargo, nuestros resultados apuntan a que la fracción S100, más que las fracciones vesiculares, es la que está fuertemente enriquecida en dichas moléculas ($P < 0.001$ en MCF-7; $P < 0.05$ en MCF-10A; Figura 22). Esto sugiere fuertemente que las tRNA-h pueden ser secretadas en forma de complejos ribonucleoproteicos (RNP), lo que está en consonancia con reportes recientes dónde se describió la presencia de complejos RNP conteniendo tRNA-h en sangre humana y de ratón (Dhahbi et al., 2014; Dhahbi et al., 2013b).

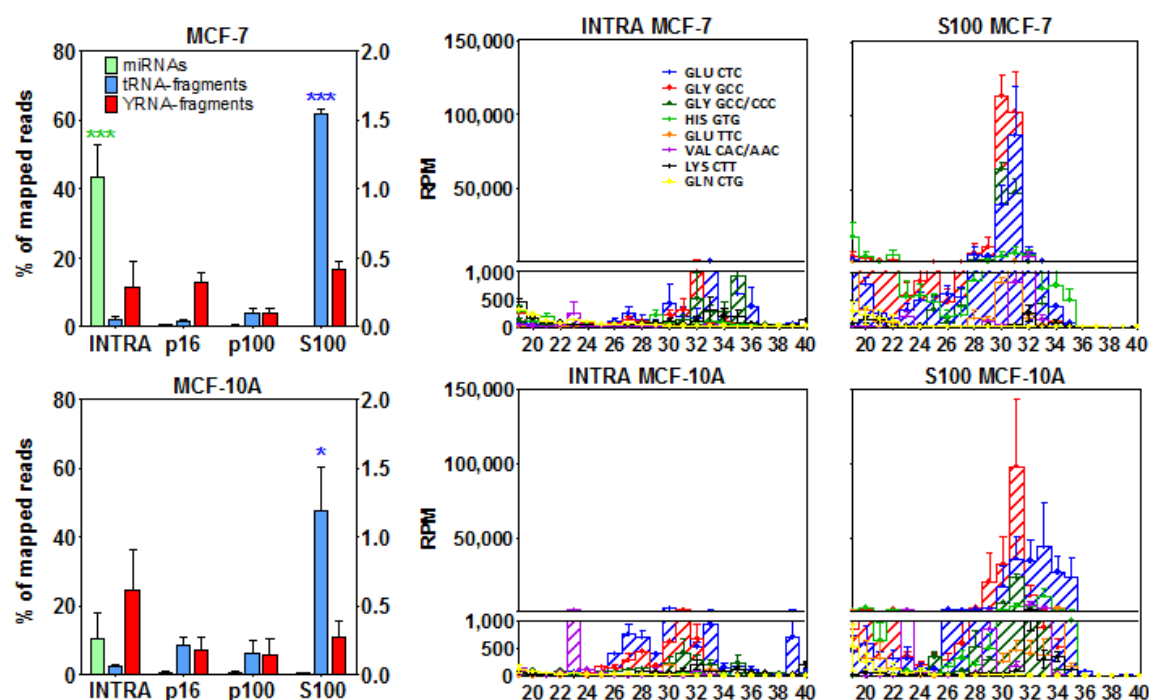


Figura 22: enriquecimiento de la fracción S100 en mitades de tRNAs. Izquierda: comparación del porcentaje de secuencias mapeadas correspondientes a miRNAs (verde; eje vertical izquierdo), tRNAs (azul; eje vertical izquierdo) e YRNAs (rojo; eje vertical derecho). Arriba: MCF-7; abajo: MCF-10A. Cuando hubo muestras significativamente diferentes (ANOVA de una vía), se posicionaron estrellas en las de porcentajes más altos. *: $P < 0.05$; **: $P < 0.01$; ***: $P < 0.001$ (test post hoc de Tukey). Gráficas del centro y derecha: distribución de tamaños y abundancias relativas de los fragmentos de tRNAs en el medio intracelular (centro) y en la fracción S100 (derecha). Los distintos colores corresponden a distintas especies de tRNAs (discriminados al nivel de isodecodificadores). Azul: tRNA^{Glu}_{CTC}; rojo: tRNA^{Gly}_{GCC}; verde oscuro: tRNA^{Gly}_{CCC}; por otros colores: ver leyenda. Las fracciones p16 y p100 no se muestran, pero fueron una situación intermedia entre el medio intracelular y la fracción S100.

Además de las grandes diferencias cuantitativas en la proporción de fragmentos derivados de tRNAs en las distintas fracciones, también se constataron diferencias cualitativas evidentes. En términos generales, los fragmentos más abundantes en el medio extracelular tendieron a ser también los más abundantes dentro de la célula, pero la complejidad y diversidad de secuencias fue mucho menor en el medio condicionado, y particularmente en S100. Por ejemplo, la suma de las secuencias derivadas de únicamente tres especies diferentes ($\text{tRNA}^{\text{Glu}}_{\text{CUC}}$, $\text{tRNA}^{\text{Gly}}_{\text{GCC}}$; y $\text{tRNA}^{\text{Gly}}_{\text{CCC}}$) llegó a corresponder al $89\pm 3\%$ y al $88\pm 3\%$ del total en MCF-7 y MCF-10A, respectivamente. Estas fueron también las especies más abundantes en el medio intracelular, pero representaron el $44\pm 9\%$ y el $63\pm 9\%$, respectivamente.

Además de la diversidad de especies, la complejidad también fue mucho más restringida en S100 cuando se observa la distribución de tamaños de los fragmentos presentes. En MCF-7, el $80\pm 5\%$ de los fragmentos tuvo entre 30 y 31 nt, vs. $9\pm 2\%$ en los lisados celulares ($P < 0.01$; test t de Student-Welch). El perfil de tamaños no fue tan estrecho en la fracción S100 de MCF-10A, pero también se observó una cierta condensación centrada en los 31-32 nucleótidos. Salvo unas pocas excepciones, todos estos fragmentos comenzaron en la posición +1 de los tRNAs maduros. Considerando su largo, corresponden a mitades 5'.

Además de las mitades, se ha descrito cómo el corte de los precursores de tRNAs por la RNasaZ libera los llamados *trailers* 3' (que no están presentes en la secuencia del tRNA maduro), los cuales pueden actuar también como RNAs pequeños reguladores (en ese contexto, se los llama tRF-1 o 3'U-tRFs) (Lee et al., 2009). Detectamos cantidades moderadas (1700 ± 600 RPM) en el medio intracelular de las células MCF-7. La mayoría de estas secuencias (74% en MCF-7; 63% en MCF-10A) correspondieron al tRF-1001, una secuencia de 20 nucleótidos derivada del trailer 3' del $\text{tRNA}^{\text{Ser}}_{\text{TGA}}$, y sobre-expresado en células altamente proliferativas (Lee et al., 2009). Sin embargo, a diferencia de las mitades de tRNAs, estos otros fragmentos estuvieron muy poco representados en las fracciones extracelulares. Por ejemplo, en la fracción S100 de MCF-7, el tRF-1001 estuvo por debajo de 20 RPM.

En resumen, de toda la diversidad de fragmentos generados a partir de transcritos de tRNA o de tRNA maduros, únicamente las mitades 5' de isoaceptores de Glu y Gly dominaron la fracción S100.

2.3.7 Validación por PCR digital y por PCR cuantitativa en tiempo real.

En cualquier estudio basado en secuenciación profunda es esencial la validación posterior por otras técnicas y de ser posible en réplicas diferentes. Lo más utilizado es la PCR cuantitativa en tiempo real (qPCR). Y como se desea típicamente evaluar varias especies de forma simultánea, lo más frecuente es utilizar una cuantificación relativa, es decir, determinando los valores de ΔCq respecto de un RNA referencia no afectado por la variable experimental en estudio.

¿Cuál sería, en nuestro caso, una adecuada referencia? Aquí no estamos comparando muestras iguales en distintas condiciones, donde se supone que es poco lo que se vería afectado, sino muestras bien distintas (un lisado celular vs. exosomas, o vs. complejos ribonucleoproteicos del medio condicionado extracelular). Absolutamente todos los RNAs deberían en teoría estar afectados por el proceso que está en estudio ya que éste no es ni más ni menos que la secreción del RNA. Si un RNA está realmente presente en el medio extracelular – descontando artefactos – es porque fue secretado, y por consiguiente, afectado por la variable en estudio. Sí se puede, al menos, distinguir dos categorías: aquellos RNAs secretados en forma selectiva (o que presentan una diferencia de estabilidad extracelular; lo que a fines prácticos es lo mismo) y aquellos secretados únicamente en función de su concentración intracelular. Se podría demostrar la selectividad utilizando como referencias RNAs cuya secreción no selectiva pueda darse por descontado. Pero dado que al momento de realizar este trabajo no había ninguna especie cuya secreción no selectiva fuese demostrada y reuniese amplios consensos, este camino no era aplicable.

Optamos entonces por recurrir a la técnica de PCR digital en micro-gotas (ddPCR), dado que es una técnica que permite una cuantificación absoluta de manera intrínseca, sin necesidad de utilizar estándares para cada especie a evaluar, y que presenta una gran sensibilidad y reproducibilidad²⁸. Su mayor desventaja es – a nuestro criterio – su no disponibilidad en el país. A diferencia de la qPCR o PCR en tiempo real, la PCR digital es una PCR a tiempo final, es decir, en condiciones de saturación de la señal. Para obtener información cuantitativa lo que se hace es realizar una dilución límite de la muestra, de manera que cada micro-reactor (micro-gotas de agua emulsionada en aceite – ddPCR – en los modelos de BioRad) únicamente reciba 1 o 0 moléculas de templado (cDNA). Luego de 40 ciclos de PCR, las gotas que recibieron templado serán fluorescentes, mientras que las que no lo hicieron presentarán intensidades de fluorescencias mucho más bajas y se considerarán negativas. Todas las gotas se hacen luego pasar de a una por un detector (en un sistema similar a lo que ocurre en citómetro de flujo) y se cuenta la proporción de gotas positivas y negativas. A partir de dichos cocientes, y conociendo las diluciones realizadas, se calcula la concentración original de moléculas de templado presentes en la muestra.

La Figura 23 muestra resultados crudos representativos para la cuantificación de un miRNA muy abundante en todas las muestras (miR-21-5p) y otro muy escaso (miR-122-5p). Cada punto representa una gota, la cual se clasifica como positiva o negativa en función de si su intensidad de fluorescencia supera o no un cierto umbral.

La Figura 24 representa la abundancia extracelular vs. intracelular (medidas como logaritmo en base 2 de las copias de cDNA por ng de RNA de partida) para distintos miRNAs (incluyendo aquellos de alta y baja expresión) y mitades de tRNAs. El tipo de sistema utilizado (*stem-loop* RT-PCR; ver materiales y métodos; §2.2.10) provoca que estas PCR sean específicas para las especies del largo estipulado (ej: si se procura amplificar un fragmento que va de la posición 1 a la 31 del tRNA^{Glu}_{CUC}, los fragmentos que llegan hasta la posición 30 o hasta la 32 no deberían en teoría interferir).

²⁸<http://www.sciencemag.org/custom-publishing/technology-features/digital-pcr-revolution>

Como puede apreciarse, los miRs: 21-5p, 16-5p, let-7i-5p y 17-5p – que por los datos de secuenciación deberían secretarse de manera no selectiva – tendieron a ajustarse a rectas de pendiente 1 en las distintas muestras, lo que implica que sus niveles extracelulares podrían predecirse a partir de su abundancia intracelular y el corte al eje de las ordenadas de dichas rectas. Esto es análogo a lo observado en la Figura 19 con los datos de secuenciación profunda, y valida por tanto dichos resultados. El miR-122-5p, en cambio, presentó una abundancia extracelular en S100 muy por encima de la “recta pasiva”, lo que de nuevo concuerda con los datos de secuenciado. La magnitud de la barra de error se debe a que esto ocurrió solamente en una de las dos réplicas biológicas analizadas, por lo que dicho resultado no sería del todo concluyente.

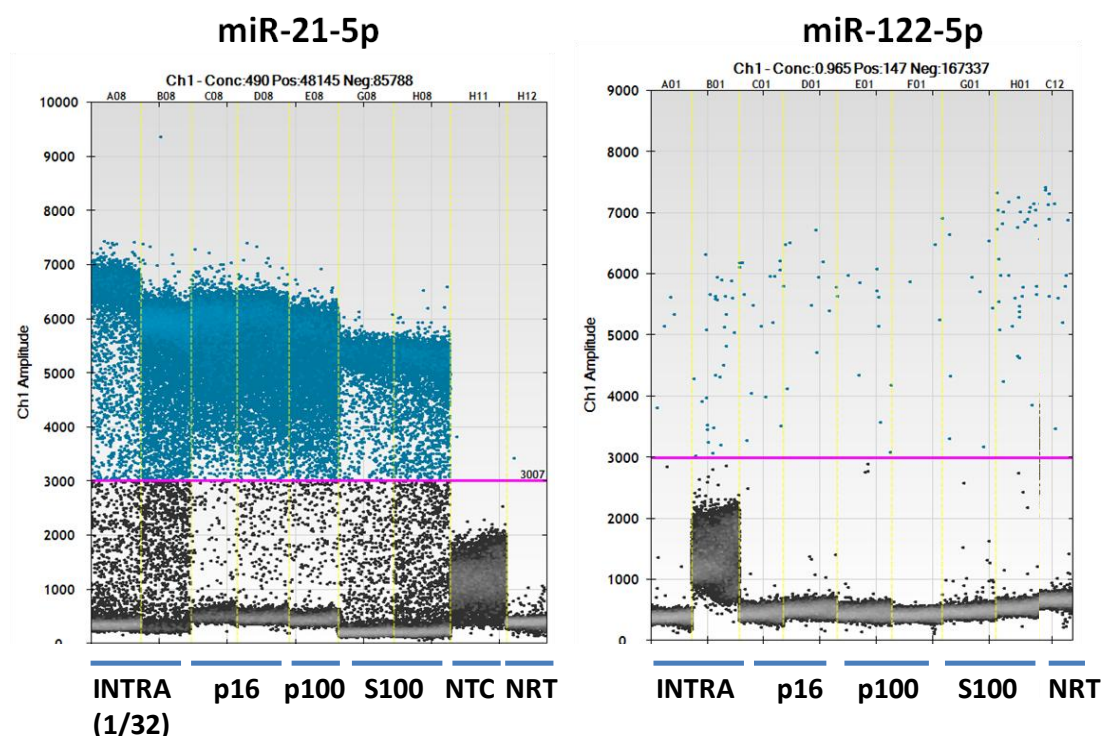


Figura 23: cuantificación absoluta por PCR digital en micro-gotas (ddPCR) del miR-21-5p (izquierda) y del miR-122-5p (derecha), en distintas réplicas biológicas de las fracciones intra- y extracelulares. En todos los casos se diluyó el RNA de partida a 10ng/μL, salvo en las fracciones intracelulares cuando se midió miR-21-5p, donde se realizó una dilución adicional indicada en la figura. Cada punto corresponde a una gota distinta, y en el eje vertical se representa la intensidad de fluorescencia. Se define automáticamente un umbral (rectas violeta horizontales) por encima del cual se considera que las gotas son positivas (señaladas en azul). NTC: control negativo; no se agrega RNA de partida (non-template control). NRT: control negativo; se agrega RNA pero no se adiciona la enzima transcriptasa reversa (permite descartar que la amplificación esté ocurriendo a partir de DNA genómico contaminante).

A diferencia de los miRNAs, las mitades de tRNAs aceptores de Glu (largo = 31 nt) y Gly (largo = 30 nt) se apartaron considerablemente de las rectas de pendiente 1 definidas por los primeros, y en particular en la fracción S100. Una forma rápida de ver esto es comparar los fragmentos del tRNA^{Gly}_{GCC} [1-31] con los del miR-21-5p [1-23]. Sus abundancias relativas a nivel intracelular fueron casi idénticas. Sin embargo, el fragmento de tRNA fue casi 32 (2⁵) veces superior en la fracción S100.

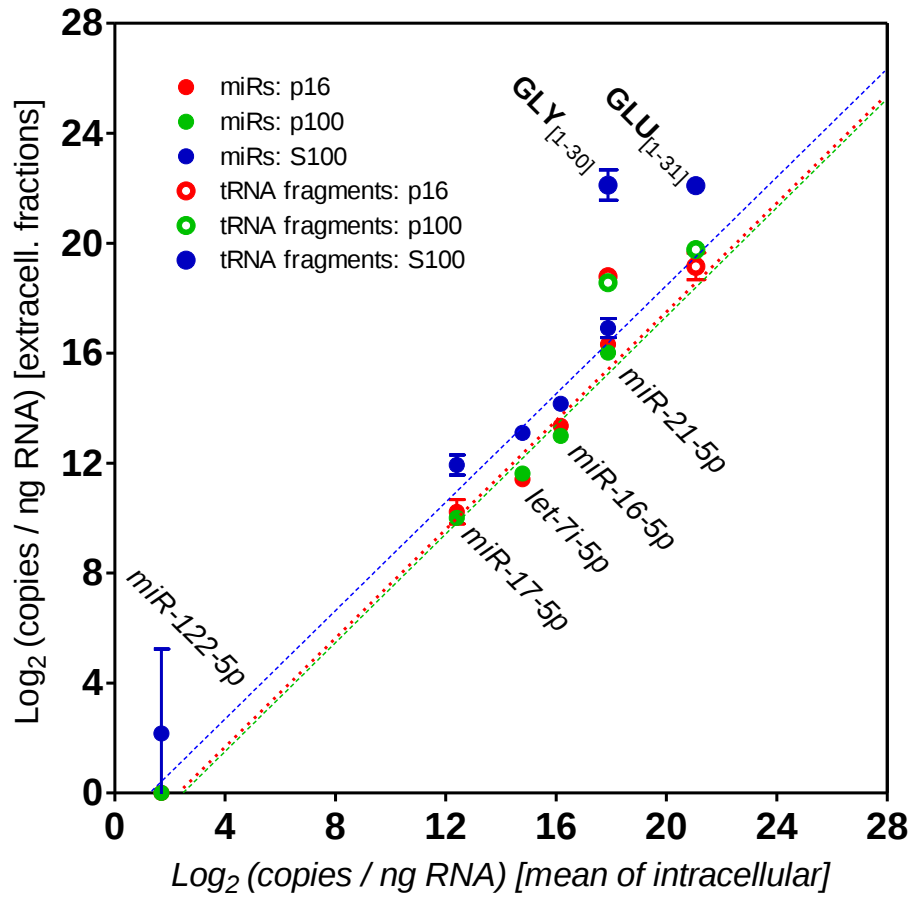


Figura 24: niveles extracelulares (medidos como Log_2 del número de copias de cDNA por ng de RNA de partida) de distintos miRNAs (círculos rellenos) y mitades de tRNAs (círculos abiertos) en función de sus respectivos niveles intracelulares promedio, según resultados de SL-RT-ddPCR. Las barras de error corresponden al error estándar de la media. Las líneas diagonales discontinuas de pendiente 1 fueron manualmente ajustadas a los miRs: 21-5p, 16-5p, let-7i y 17-5p. Los colores corresponden a las muestras p16 (rojo), p100 (verde) y S100 (azul).

Una formalización de estos resultados se muestra en la Figura 25, A, dónde se grafica la diferencia en el logaritmo del número de copias por ng de RNA entre los fragmentos de tRNA y algunos miRNAs utilizados como referencia (miR-21-5p y miR-16-5p), ya sea en el medio intracelular o en la fracción S100. Habiendo validado que la abundancia extracelular de dichos miRNAs es acorde a lo esperable para secreción no selectiva, los mismos pueden ser utilizados como referencia interna para evaluar el proceso de secreción de otras especies. Si un RNA cualquiera tuviese el doble de concentración que el miR-21-5p a nivel intracelular, también debería estar dos veces más concentrado a nivel extracelular, salvo que mediara algún proceso de secreción diferencial. Por tanto, el cociente de concentraciones - la resta cuando se emplean logaritmos - no debería modificarse significativamente. Como puede verse, este no es el caso.

Disponiendo de referencias internas ahora sí validadas por técnicas de cuantificación absoluta, consideramos que podíamos - ahora sí - recurrir a la cuantificación relativa propia de la PCR en tiempo real, y extender el número de réplicas biológicas a analizar. Como puede verse en

la Figura 25, B, los resultados fueron comparables a los obtenidos por ddPCR, con aún mayor confianza estadística. También se validó por esta técnica el enriquecimiento del miR-122-5p en las fracciones extracelulares (Figura 25, C).

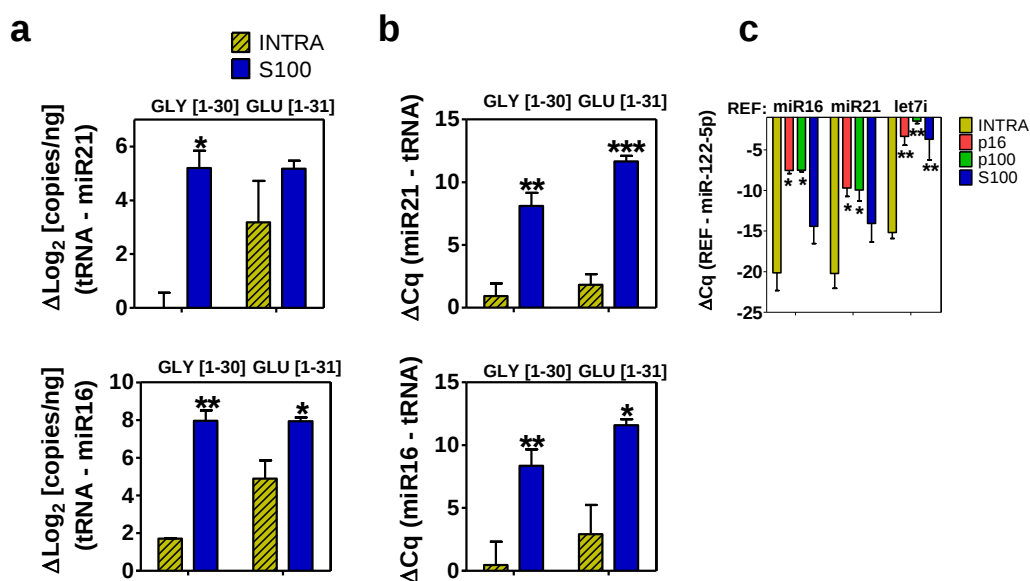


Figura 25: soporte experimental al enriquecimiento extracelular en mitades de tRNA (Glu-CUC y Gly-GCC) y del miR-122-5p, evaluado por SL-RT-ddPCR (A) y SL-RT-qPCR (B y C). A) se grafica la variación en los valores de Log_2 (copias/ng de RNA) de los fragmentos de tRNA menos miR-21-5p (arriba) o miR-16-5p (abajo). Se procura evidenciar un crecimiento en dichas diferencias en la fracción S100 (azul). $N = 2$. B) ídem, pero la variación viene ahora dada por el ΔCq entre el miRNA utilizado como referencia y el fragmento de tRNA en estudio. Nótese que se invierte el minuendo y sustraendo respecto de (A) de modo que barras más altas sigan queriendo decir lo mismo (dado que el Cq crece en forma inversa a la concentración). $N = 4$. C) validación por SL-RT-qPCR del enriquecimiento de miR-122-5p en las fracciones extracelulares. Se muestran los resultados obtenidos al utilizar como referencia miR-16-5p, miR-21-5p o let-7i-5p. $N = 4$. En (A) y (B) las estrellas indican significancia estadística para la diferencia entre las muestras INTRA y S100 evaluada por el test t de Student. En (C) corresponden a ANOVA de una vía. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

2.3.8 El enriquecimiento en mitades 5' de tRNAs de la fracción S100 es independiente de las condiciones de cultivo y común a varios tipos celulares.

Llegado a este punto, nos preguntamos sobre la generalidad de estos resultados ¿Serían una particularidad de las células MCF-7 y MCF-10A? ¿Serían una consecuencia del origen mamario de dichas células? ¿O será la secreción selectiva de mitades de tRNAs una propiedad general y común a varios tipos celulares? Se ha reportado que el corte de los tRNAs por parte de la ribonucleasa Angiogenina – que es la que generaría las mitades de tRNAs – es regulado y se incrementa en condiciones de estrés celular (Emara et al., 2010; Fu et al., 2009; Yamasaki et al., 2009). Por tanto, podría ser que las tRNA-h aparezcan en el medio extracelular como consecuencia del estrés provocado por las condiciones de cultivo en medio libre de suero, o debido a la formulación del medio de cultivo definido utilizado.

Para evaluar lo anterior, se cultivó a las células MCF-7 tanto en un medio definido de formulación distinta a la utilizada (MEGM; el mismo utilizado para las MCF-10A), o

siguiendo los estándares de la ATCC (EMEM + insulina humana + 10% suero fetal bovino). En ambas situaciones, los valores de ΔCq para las mitades de tRNAs (utilizando como referencias los miR-21-5p y miR-16-5p) fueron mayores en S100 respecto al medio intracelular, especialmente para el tRNA^{Glu}_{CUC} ($P < 0.05$; test t de Student; Figura 26, A-B). Esto muestra que la secreción de mitades de tRNA es independiente de las condiciones de cultivo, y no se asocia únicamente a estrés. Debe considerarse que tanto el miR-16-5p como las mitades 5' del tRNA^{Glu}_{CUC} fueron detectadas con bajos valores de Cq en los sobrenadantes a 100,000g del medio no condicionado, cuando este contenía 10% FBS. Esto no debería sorprender ya que se sabe que el miR-16-5p es el miRNA más abundante en sangre (en humanos, al menos), a la vez que subraya la importancia de haber trabajado en condiciones libres de FBS a lo largo de esta tesis. De todos modos, el incremento en los valores de ΔCq en la fracción S100 para el tRNA^{Glu}_{CUC} fue superior al explicable únicamente a causa del suero (ver líneas rojas discontinuas en dichas figuras).

Como una primera aproximación a la generalidad del fenómeno, extendimos el análisis de las mitades de tRNA-Glu y tRNA-Gly a otras líneas celulares provenientes de tejidos diferentes (HeLa: cáncer de cuello de útero; NCI-H1299: cáncer de pulmón). El enriquecimiento en la fracción S100 de las mitades de tRNA^{Gly}_{GCC} fue constatado en ambas líneas celulares. La línea NCI-H1299 mostró también enriquecimiento en tRNA^{Glu}_{CUC} (Figura 26, C-D).

En resumen, indicios de secreción selectiva (o de estabilidad extracelular diferencial) a la fracción S100 de las mitades de tRNA de aceptores de Glu y Gly fueron hallados en líneas celulares tumorales provenientes de distintos tejidos (MCF-7, HeLa, NCI-H1299) y en líneas celulares inmortalizadas no tumorales (MCF-10A), lo que da cuenta que se está ante un fenómeno fisiológico de carácter general.

2.3.9 Purificación cromatográfica de las mitades de tRNA presentes en la fracción S100

Habiendo visto y corroborado que las mitades de tRNAs (de Glu y Gly, al menos) se encontraban en el medio extracelular fundamentalmente en la fracción S100, restaba confirmar si dichos RNAs formaban parte de genuinos complejos ribonucleoproteicos. Existía la posibilidad de que muy pequeñas vesículas no precipitables a 100,000g pudieran ser las responsables de transportar – y proteger – estos fragmentos.

Una posible manera de zanjar esta cuestión era intentar purificar las mitades de tRNA presentes en la fracción S100. Para ello se concentró dicha fracción hasta un volumen pequeño (500 μ L) – después de cambiar la matriz de la muestra a PBS; ver materiales y métodos, §2.2.12 – y se inyectó en una columna de exclusión molecular pre-equilibrada en PBS. Se probaron columnas de igual geometría, pero conteniendo distintas matrices (Superdex 200 y Superdex 75; Amersham) de modo de poder separar vesículas extracelulares de complejos proteicos grandes y proteínas pequeñas. En ambos casos se obtuvieron cromatogramas similares, caracterizados por un alto pico en la absorbancia a 280nm cercano al volumen de elución esperado para la Transferrina (la cual está presente en la formulación

del medio definido utilizado, y se adiciona en grandes cantidades). Hasta dicho punto, la absorbancia a 260nm siguió la misma tendencia que la absorbancia a 280nm, de modo tal que el cociente 260/280 se mantuvo relativamente constante y cercano al valor esperado para proteínas (que, si bien suelen presentar un máximo de absortividad molar en 280nm, también absorben a 260nm). Sin embargo, esta tendencia comenzó a revertirse fracciones de mililitro después de eluída la Transferrina, definiéndose dos picos (a los que llamamos P1 y P2) en el cromatograma registrado a 260nm, y con una relación $260/280 > 1$ (Figura 27). Resultados similares fueron obtenidos para la fracción S100 de las células MCF-10A.

Nuestra interpretación fue que una alta cantidad de ácidos nucleicos eluídos en dichas fracciones podría ser la responsable de esta inversión en el cociente 260/280 nm. Para corroborar esta hipótesis analizamos por PCR cuantitativa las distintas fracciones colectadas en la búsqueda de las mitades de tRNA. Tal como era esperado, los picos de concentración (estimada por el valor de 2^{-Cq}) de las mitades de tRNA^{Glu}_{CUC} y tRNA^{Gly}_{GCC} se correspondieron perfectamente con los picos P1 y P2 definidos en el cromatograma. No se detectaron cantidades comparables en el volumen excluído (de ninguna de las dos columnas), lo que descarta que dichas mitades estuviesen presentes en vesículas extracelulares. Más aún, considerando que el límite de exclusión de las columnas Superdex 75 10/300 es de aproximadamente 100 kDa, estos resultados tampoco apoyarían la formación de complejos RNP con proteínas Argonauta. Por el contrario, los volúmenes de elución de los picos P1 y P2 corresponderían a proteínas globulares de 44 y 20 kDa, respectivamente. Considerando que las propias mitades de tRNA ya son de por sí bastante “pesadas” (rondan los 10 kDa), estos resultados implicarían la formación de complejos RNP con proteínas más bien pequeñas, salvo que las mitades de tRNA adoptasen una estructura muy compacta. Más aún, estos perfiles de elución tampoco se corresponderían con los tamaños de los complejos RNP conteniendo mitades de tRNA y reportados en sangre por Dhahbi y colaboradores (2013).

Por último, y para mostrar la presencia de proteínas asociadas a estos RNAs, procuramos pre-tratar las muestras con proteasas previo a su purificación por exclusión molecular. Luego de varios intentos fallidos con Tripsina, se logró obtener la desaparición completa del pico P1 mediante Proteinasa K. Estos ensayos, y consideraciones varias en torno a los mismos, serán objeto de análisis en el próximo capítulo de este trabajo.

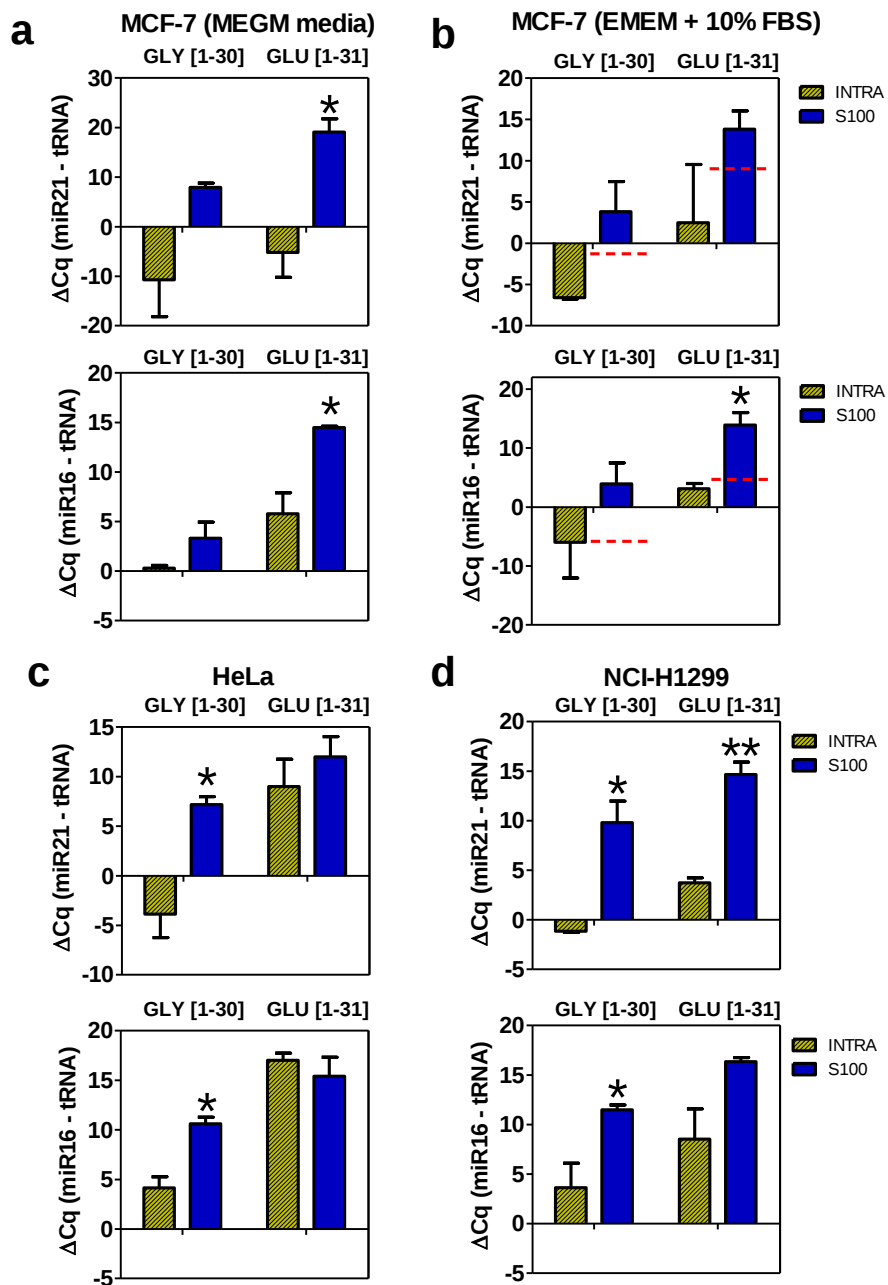


Figura 26: el enriquecimiento de la fracción S100 en mitades de tRNA es independiente de las condiciones de cultivo y no está restringido a células de origen mamario. Los valores de ΔCq (miRNAs – mitades de tRNA) fueron comparados entre las muestras intracelulares y S100, y los valores significativamente más altos (test t de Student; una cola) se muestran (*: $P < 0.05$; **: $P < 0.01$). A) Células MCF-7 crecidas en medio definido MEGM. B) Células MCF-7 crecidas en medio EMEM + insulina + 10% suero bovino fetal, FBS). Las líneas rojas discontinuas muestran los valores de ΔCq correspondientes en los sobrenadantes a 100,000g del medio no condicionado (conteniendo suero). C) Células Hela y (D) NCI-H1299 crecidas en medio definido sin suero (mismo utilizado para MCF-7 a lo largo de este trabajo). En todos los casos se utilizó medio condicionado a $t = 48$ hs. En todos los paneles se analizaron dos mitades de tRNAs, y se utilizaron dos miRNAs como referencia: miR-21-5p (gráficas superiores) y miR-16-5p (gráficas inferiores).

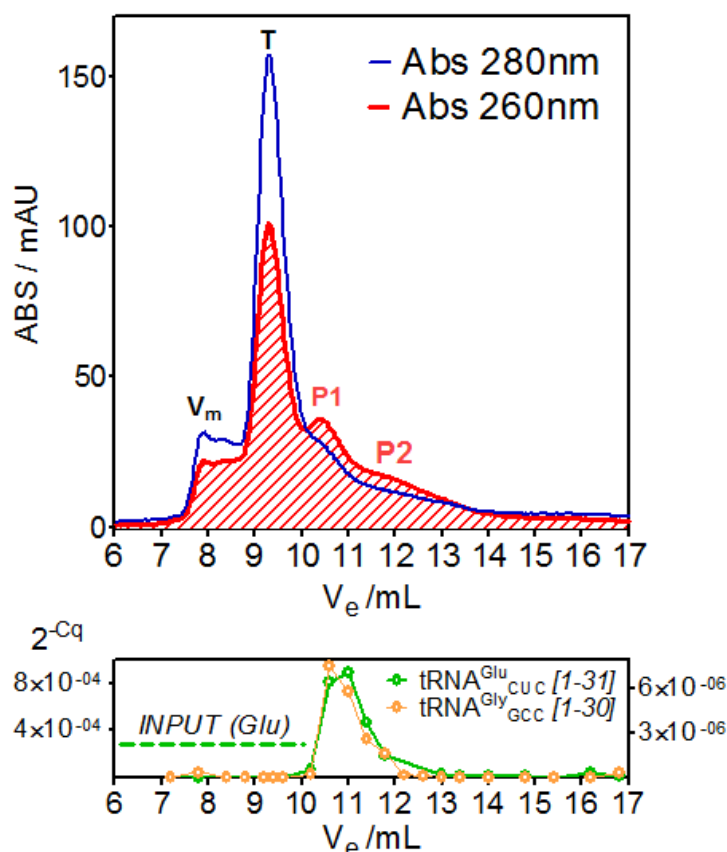


Figura 27: sub-fraccionamiento de la fracción S100 de las células MCF-7 por cromatografía de exclusión molecular (Superdex S75 10/300; Amersham) utilizando un equipo de FPLC (Äkta Purifier, Amersham). Arriba: cromatograma obtenido mediante monitoreo espectrofotométrico operando a 280nm (curva azul) y a 260nm (curva roja). V_m : volumen de exclusión; T: Transferrina; P1 y P2: picos donde la ABS 260nm > ABS 280nm. Abajo: se colectaron fracciones de 0.2mL, las cuales se analizaron por SL-RT-qPCR específicas para las mitades del $tRNA^{Glu}_{CUC}$ (31 bases; verde; eje vertical izquierdo) y del $tRNA^{Gly}_{GCC}$ (30 bases; anaranjado; eje vertical derecho). Se representa el valor de 2^{-Cq} en función del volumen de elución ($V=0$ corresponde a la inyección). “Input Glu” corresponde al valor obtenido para el $tRNA^{Glu}_{CUC}$ en una muestra del contenido inyectado, previo a su purificación. Los valores superiores implican una concentración de dicha especie provocada por su efectiva purificación.

2.4 DISCUSIÓN

Este trabajo no fue el primero en estudiar por secuenciación profunda los pequeños RNAs presentes en exosomas, microvesículas o en complejos ribonucleoproteicos en el medio extracelular de células tumorales o no tumorales. Tampoco fue el primero en trascender a los miRNAs y extender el estudio a otras especies menos estudiadas tales como las mitades de tRNAs o los fragmentos de YRNAs, rRNAs, snRNAs, snoRNAs, vault RNAs, etc. Ciertamente, no fue el primero en comparar los perfiles intracelulares y extracelulares para proponer modelos de secreción selectiva o no selectiva, ni en secuenciar varias réplicas biológicas para que los resultados tuviesen una base estadística más o menos confiable. Sí fue el primero en realizar todo lo anterior de manera conjunta y sistemática, y eso permitió apreciar la real dimensión de algunas observaciones cuya relevancia no había quedado aún debidamente calibrada.

A nuestro juicio, uno de los resultados más importantes fue el gran enriquecimiento de las mitades de tRNAs en la fracción S100. Al punto tal que el trabajo decantó en una profundización de dicho estudio que se presenta en el Capítulo III. La relevancia fisiológica intracelular de las tRNA-h está cada vez más clara: sus implicancias en la regulación de la expresión génica (Lee et al., 2009), su participación en respuestas de estrés (Emara et al., 2010; Li and Hu, 2012), su rol en el control del inicio de la traducción (Ivanov et al., 2011) y su papel en variadas enfermedades (Anderson and Ivanov, 2014; Blanco et al., 2014; Ivanov et al., 2014). En enero de 2016, la revista *Science* publicó dos artículos en los cuales se mostró como la dieta²⁹ de un ratón adulto puede modificar los niveles de las mitades de tRNA en sus espermatozoides, y esto afectar directamente a sus hijos, lo que implicaría un nuevo mecanismo de herencia epigenética mediada por mitades de tRNA (Chen et al., 2016; Sharma et al., 2016).

Los fragmentos de tRNA siguen – y seguirán – dando sorpresas, pero ¿qué hay de su presencia y rol en el medio extracelular? Previo a nuestro trabajo ya se había advertido la presencia de mitades de tRNA en la sangre (Dhahbi et al., 2013b), y en exosomas presentes en el semen y en el medio condicionado de células inmunológicas (Nolte-t Hoen et al., 2012; Vojtech et al., 2014). Nuestro estudio comparativo de distintas fracciones extracelulares permitió comprender que, si bien pueden detectarse fácilmente en preparaciones vesiculares, la mayoría de los fragmentos de tRNA no precipitan a 100,000g dado que formarían parte de complejos RNP de 20-40 kDa. Distintos, por tanto, de otros complejos de proteína y RNAs pequeños conocidos, tales como los complejos AGO2/miRNA (Arroyo et al., 2011).

Esta observación pone acento en el estudio de la fracción no vesicular del medio extracelular como rica en RNAs potencialmente reguladores, cuando casi toda la bibliografía científica desde el 2007 a esta parte se ha centrado casi exclusivamente en los exosomas. Esta visión “vesículo-centrista” del RNA extracelular ha sufrido ciertos “golpes” bajo su línea de flotación. Los primeros vienen de un artículo publicado en PNAS en 2011 en el que los autores reportaron que la mayoría de los miRNAs extracelulares se encuentran en la forma de complejos RNP libres, y no en vesículas (Arroyo et al., 2011). Tres años más tarde, el mismo grupo publicó en la misma revista una estimación de la estequiometría de los “miRNAs por exosoma” (Chevillet et al., 2014). Cuantificando exosomas por NTA y miRNAs por ddPCR, concluyeron que se deberían analizar al menos 100 exosomas para encontrar una copia del miRNA más abundante (0.01 copias por exosoma). Utilizando nuestros propios datos de NTA y ddPCR, llegamos al número de 98, 780 y 2058 vesículas para encontrar una copia del miR-21-5p, miR-16-5p y let-7i-5p, respectivamente. Nuestros datos concuerdan, pues, con lo reportado por estos autores. Conforme al enriquecimiento extracelular observado para las mitades de tRNA, la estequiometría fue más favorable en este caso, pero aún así lejos de la relación 1:1 (1 copia cada 7 y 16 exosomas para las mitades de tRNA^{Glu}_{CUC} y tRNA^{Gly}_{GCC}, respectivamente).

Si algo destruyó el reporte anterior fue la concepción del exosoma como una especie de caballo de Troya repleto de miRNAs capaces de regular la expresión génica de cualquier

²⁹La dieta es un tema recurrente, pero aquí no se habla de RNAs exógenos incorporados por la dieta, sino de una alteración de los niveles endógenos de los fragmentos de tRNA.

célula incauta que lo terminase captando. Sin embargo, esto de ningún modo implica que no haya motivos para considerar que la secreción de miRNAs en exosomas sea una forma biológicamente relevante de comunicación intercelular. Una copia del miR-21-5p cada 100 exosomas podrá parecer poco, hasta que se cae en la cuenta de que cada célula secreta vesículas en el orden de decenas de miles por cada 24 horas. En nuestro caso, la concentración medida de exosomas en el medio de cultivo condicionado a $t = 48$ hs. fue de 1.5×10^9 partículas / mL, lo que implica 1.5×10^{10} vesículas por frasco de cultivo, provenientes de aproximadamente 2.5×10^6 células (MCF-7). Esto nos da que en nuestras condiciones experimentales cada célula secretó en 48 horas unos 6000 exosomas, equivalentes a 60 copias de miR-21-5p. Esto puede parecer poco desde el punto de vista de la célula emisora, pero no necesariamente desde la perspectiva de una eventual célula receptora. Si suponemos el caso extremo de una única célula capaz de absorber todos los exosomas presentes en el frasco de cultivo, estaría incorporando 150 millones de copias del mencionado miRNA. A nivel comparativo, los hepatocitos presentan una altísima expresión del miR-122-5p y sus niveles rondan las 53,000 copias por célula (Bissels et al., 2009).

Los cálculos anteriores permiten comprender que los pequeños RNAs exosomales podrían mediar procesos de comunicación intercelular, pero que para ello es necesario que la célula receptora incorpore un número más bien alto de vesículas. Esto es fácil de imaginar en dos contextos: i) en el caso de células emisoras y receptoras que se encuentren próximas en el espacio, de modo tal que los efectos de dilución sean mínimos – ej: en el microambiente tumoral – ó ii) en el caso de células distantes en la medida que procesos de reconocimiento del tipo receptor/ligando permitan una concentración de las vesículas sobre la superficie de algunos tipos celulares específicos.

Prueba de la etapa incipiente – pero fermental – en que se encuentra todo este campo de estudio: no había hasta hace poco evidencia experimental que respaldase el punto (ii), por más lógico que fuese pensar que la captación de exosomas de la sangre tendría que ser un proceso específico y regulado. Al menos esa es la lección que nos dejan los virus, cuya conexión evolutiva con las vesículas extracelulares ha sido propuesta (Gould et al., 2003). Por fin, en octubre de 2015 la revista *Nature* publicó que “*las integrinas de los exosomas tumorales determinan la metástasis organotrópica*”³⁰ (Hoshino et al., 2015). Los autores mostraron cómo los exosomas secretados por células de distintos tumores presentan juegos particulares de integrinas en su superficie, lo que les permite concentrarse en tejidos específicos para efectuar la ya conocida – y descrita por el mismo grupo – preparación exosomal del nicho pre-metastático (Costa-Silva et al., 2015; Peinado et al., 2012). Con este descubrimiento, los autores se auto-adjudicaron el honor de haber brindado una explicación molecular a la “hipótesis del suelo y la semilla”³¹ propuesta por Stephen Paget en 1889. Podrá parecer un exceso, pero sin dudas que esta publicación dinamizará el estudio de los mecanismos de internalización de las vesículas extracelulares.

³⁰*Tumour exosome integrins determine organotropic metastasis*

³¹*The seed and soil hypothesis*, atribuida a Stephen Paget (1855 – 1926), médico inglés que, observando que los tumores de determinados órganos tienden a producir metástasis en órganos específicos, propuso que para que la metástasis tenga efecto no basta únicamente con las propiedades de las células tumorales (la semilla), sino que éstas deben a su vez ser capaces de interactuar positivamente con factores presentes en el órgano de destino (el suelo). Los trabajos del grupo de David Lyden apuestan a que son los exosomas los responsables de labrar la tierra y fertilizarla.

Nada hemos contribuido con este trabajo al aspecto anterior, pero sí hemos aportado a la comprensión de la otra cara de la misma moneda que es la secreción de las vesículas por parte de la célula que llamamos emisora. Aunque esto también suena un exceso dado que en los aspectos moleculares de la secreción tampoco nos hemos adentrado. Sin embargo, nuestro pequeño grano de arena en lo que hace al modelo de secreción de miRNAs no luce tan pequeño cuando se lo mira de cerca. Es cierto que proponer que la célula secreta miRNAs en función de su abundancia intracelular no resulta demasiado interesante, en parte porque suena a la opción por defecto, y en parte porque tendemos a valorar siempre más los fenómenos físicos (“*A es reconocido y movido por B*”) que los físico-químicos (“*A difunde a favor de gradiente*”). Además, la búsqueda de miRNAs derivados de tumor en la sangre como potenciales biomarcadores ¿no asume de entrada este modelo?³² Y sin embargo, la ortodoxia en el tema transita por el carril contrario.

Nuestros resultados no excluyen la posibilidad de que algunos miRNAs sean seleccionados para su secreción, y prueba de ello es que hemos procurado encontrar posibles candidatos, entre los que destacamos al miR-122-5p. Lo que sí indican es que la secreción selectiva es más la excepción que la regla – al menos en las células MCF-7 – o que sus efectos no son lo suficientemente grandes como para superar el “ruido” ocasionado por la secreción no selectiva. Es verdad que los casos más interesantes pueden ser los de aquellos miRNAs de baja expresión para los cuales la evolución ha tejido mecanismos específicos de secreción, pero dichos casos y mecanismos deben de ser estudiados de manera particular y debidamente validados. La secreción selectiva no debería suponerse la opción por defecto, y la presencia de un determinado miRNA en los exosomas no implica que dicho miRNA haya de tener un rol en la comunicación célula-célula. Tampoco es válido hacer generalizaciones por el lado contrario. Un miRNA bien puede “subirse de polizón” a los exosomas a causa de su alta concentración citoplasmática, y eso no impide que pueda actuar sobre otras células una vez captadas las vesículas. En definitiva, el modelo de secreción y la función biológica son cosas distintas. Tampoco puede perderse de vista que hay “función biológica” más allá de la comunicación intercelular: la célula sintetiza miRNAs para silenciar mensajeros, pero estos miRNAs tampoco pueden acumularse de manera descontrolada. Si bien bastante se conoce sobre el “turnover” de los miRNAs (Ruegger and Grosshans, 2012), los niveles intracelulares deberían surgir del balance entre síntesis, degradación, repartición intracelular y secreción (excreción).

Hemos hecho esfuerzo en hablar de secreción selectiva o no selectiva, y evitar hablar de secreción activa o pasiva. Lo primero puede entenderse como un parámetro meramente descriptivo y observacional, mientras que lo segundo hace a cuestiones de mecanismo que aquí no podemos abordar. Los miRNAs podrían ser secretados de manera no selectiva por medio de un proceso activo dependiente de ATP, en la medida en que se cumpla que la tasa de secreción de cada miRNA dependa únicamente de su concentración y no de su secuencia o

³²No necesariamente. También se realizan muchos abordajes “ciegos” que consisten en analizar sangre de pacientes y de controles, y buscar diferencias significativas que luego puedan validarse como marcadores o predictores. Y todo eso al margen de que provengan o no de la célula tumoral: si la respuesta al cáncer de páncreas implica una proliferación de linfocitos que vuelcan mayores cantidades del miR-XX-Xp a la sangre, y eso es específico de una determinada patología, entonces miR-XX-Xp puede servir como marcador. Pero este no era el punto que queríamos ilustrar.

identidad. Menos cuidadosos hemos sido con el empleo del término “secreción”. Estamos simplemente observando RNAs dentro y fuera de la célula, y el mecanismo de salida lo inferimos a partir del tipo de fraccionamiento extracelular que hemos realizado en cada caso. Es decir, asumimos que los miRNAs presentes en la fracción p16 fueron secretados durante la biogénesis de las microvesículas, por ejemplo. Pero también es posible que estuviesen presentes en cuerpos apoptóticos que co-purificasen con las microvesículas, y en dicho caso no sería correcto hablar de secreción. De todas formas, la alta viabilidad celular hacia el fin de los ensayos (> 95%) y la observación del mismo patrón de miRNAs en otras fracciones donde la presencia de cuerpos apoptóticos es menos probable (p100 y S100) hacen que nos inclinemos más por la primera opción.

Lo anterior vale también para las otras especies de RNA estudiadas. La caída en los niveles relativos globales de miRNAs se vio acompañada por un aumento en mitades de tRNA, especialmente en la fracción S100. Y si bien hemos hablado de secreción selectiva, tampoco tenemos evidencia directa de esto. Nuestro abordaje experimental tiene la gran limitación de que los efectos ocasionados por la secreción selectiva y la estabilidad extracelular diferencial no son distinguibles entre sí. Los fragmentos de 30 y 31 nucleótidos de los tRNAs de Glu y Gly, así como los fragmentos 5' del RNA Y4, bien podrían ser secretados con la misma tasa que los miRNAs y que los otros fragmentos de tRNA e YRNA. Con que se degradasen mucho más lento en el medio extracelular ya sería suficiente para que el efecto observado sea similar al de la secreción selectiva.

Aunque la cuestión anterior sigue abierta, algunos resultados preliminares nos incitan a pensar que las tRNA-h deben su alta abundancia extracelular a una tasa de secreción elevada, más que a cuestiones de estabilidad diferencial. Una cinética comparativa entre las tRNA-h de Gly y el miR-21-5p (evaluada por qPCR) permite apreciar que ya 1 hora después del cambio de medio, los Cq de los tRNA-h son muy inferiores a los del miR-21-5p, pese que a nivel intracelular se da exactamente lo contrario. Contrario a lo que cabría esperar en el caso de que las tRNA-h se terminasen concentrando debido a su menor tasa de degradación, los valores de ΔCq en el medio extracelular permanecieron relativamente inalterados en el tiempo (Figura 28).

Cualquiera sea el caso, el estudio de estos fragmentos nos resultó de particular interés. Si las mitades de tRNA se secretasen de manera selectiva a la fracción soluble del medio extracelular debería haber una proteína responsable de dicho mecanismo, y tal vez dicha proteína pudiese permanecer unida a los fragmentos una vez secretados. Si fuese el caso de la estabilidad diferencial ¿cómo podría un pedazo de RNA resultar tremendamente estable si no fuese por medio de su unión a proteínas? Así que todos los caminos nos llevaban a preguntarnos cuál sería la proteína responsable de unir y estabilizar a las mitades de tRNA en el medio extracelular. Parecía un objetivo accesible para la última etapa de este doctorado...

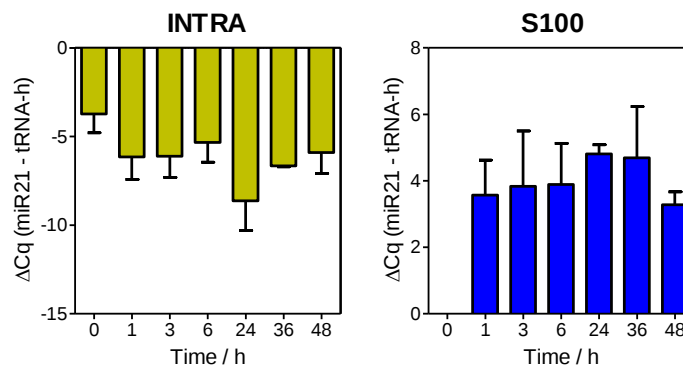


Figura 28: variación de la relación entre los tRNA-h (del tRNA^{Gly_{GCC}}) y el mir-21-5p, tanto en el medio intracelular (izquierda) como en la fracción S100 (derecha), a distintos tiempos post cambio de medio. A $t = 0$, los valores de Cq para ambas especies estuvieron cercanos a los de sus respectivos controles negativos (NTC) y por dicha razón no se grafica su diferencia.

Una última palabra debe decirse sobre la posibilidad que ofrecen las técnicas de secuenciación masiva para conocer todo el universo de moléculas de RNAs pequeños, contenidos dentro de una célula o de una vesícula. La secuenciación profunda es una técnica potente pero técnica al fin, y como cualquier técnica de laboratorio presenta una serie de limitaciones. La profundidad del secuenciado (el número de *reads*) no es el único problema, a pesar de ser cierto que en este caso trabajamos muy lejos de la saturación. Pero aún cuando se leyeran miles de millones de secuencias, habría limitaciones impuestas por la propia técnica utilizada para construir y secuenciar las librerías. En nuestro caso recurrimos a la ligación de adaptadores en los extremos 5' y 3' del RNA. Esto es muy provechoso pues el *primer* de la secuenciación se une a la región del adaptador 5', y uno luego lee la secuencia correspondiente a cada RNA hasta encontrarse con la secuencia complementaria a la del adaptador 3'. De ese modo, el comienzo y el final “real” de cada RNA ligado, retrotranscrito, amplificado y secuenciado queda debidamente establecido. Pero dicho proceso tiene sus sesgos. El más grande es el hecho de que solamente aquellos RNAs que presentan un extremo 5' fosfato y un extremo 3' OH son ligables con el protocolo utilizado. Luego están los problemas con las bases modificadas, algunas de las cuales provocan problemas de fidelidad durante la síntesis del cDNA, o directamente abortan el proceso.

Todos los escollos anteriores convergen cuando de secuenciar tRNAs se trata. En primer lugar, se ha visto que algunas mitades de tRNA no poseen extremos 5'fosfato o 3'OH, por lo que permanecen invisibles a las técnicas de secuenciado convencionales (Honda et al., 2015). Además, los tRNAs son reconocidos por poseer gran variedad y diversidad de bases modificadas, por lo que es necesario contar con protocolos específicos para poder obtener información cuantitativa comparativa a nivel de secuenciado (Cozen et al., 2015; Zheng et al., 2015). Llama la atención que los fragmentos mayoritarios que hemos secuenciado provienen de los tRNA^{Glu_{CUC}} y tRNA^{Gly_{GCC}}, que parecerían carecer de las bases modificadas más problemáticas para la transcriptasa reversa (Cozen et al., 2015).

Entonces ¿qué valor tienen nuestros resultados, en lo que a mitades de tRNA refiere? Opino que mucho, siempre y cuando uno esté dispuesto a renunciar a la pretensión de describirlo todo para focalizarse en cuestiones comparativas. Seguramente haya muchos otros fragmentos de tRNA que no estamos viendo, quizás más abundantes incluso que los que hemos estudiado. Pero las mitades 5' de 30 y 31 nucleótidos de Glu y Gly existen como tales, son entidades biológicas reales. Si fuesen producidos por un salto prematuro de la transcriptasa reversa no podríamos verlos, ya que el paso de ligación de ambos adaptadores es anterior a la RT. Además, hemos validado la presencia de dichas especies por SL-RT-ddPCR y SL-RT-qPCR, que son técnicas independientes de ligación, pero extremadamente dependientes de que el extremo 3' sea el correcto (una base más o una base menos y el juego de *primers* ya dejaría de ser efectivo). Redondeando la idea: no estamos secuenciando todos los fragmentos de tRNA que existen, pero aquellos fragmentos que sí podemos secuenciar existen. Y en el medio intracelular observamos mucha más diversidad de fragmentos, tanto en identidad como en tamaño, de lo que observamos fuera de la célula. Y fuera de la célula observamos que la abundancia relativa de los fragmentos de tRNA en general, y de la de las mitades de Glu y Gly en particular, son muy superiores a las intracelulares. Y dichos resultados no pierden validez por el hecho de que haya otros fragmentos que escapen a nuestro anzuelo.

CAPÍTULO III

Estabilidad de las mitades de tRNA en la fracción soluble (S100) del medio extracelular y sus proteínas asociadas

3.1 INTRODUCCIÓN

Habiendo estudiado los pequeños RNAs presentes en el medio extracelular de distintas líneas celulares, nos decidimos a proseguir con el resultado más sorprendente: la gran abundancia relativa de mitades 5' de tRNA^{Glu} y tRNA^{Gly} en el sobrenadante a 100,000g del medio condicionado de las células MCF-7. Fragmentos de RNA mucho más enriquecidos – en términos relativos al total de RNAs pequeños – fuera de la célula que dentro. Y aparentemente no tan asociados a vesículas extracelulares como a complejos ribonucleoproteicos solubles. Si bien dichos fragmentos fueron detectados en altas proporciones tanto en el medio condicionado de las células tumorales (MCF-7, HeLa, NCI-H1299) como de células no tumorales (MCF-10A), las cantidades – ahora sí absolutas – de RNA en la fracción S100 de las células tumorales MCF-7 fueron significativamente mayores a las de las células MCF-10A (Figura 15)³³, por lo que cabría pensar en diferencias ya no cualitativas sino cuantitativas entre células tumorales y no tumorales.

En el capítulo anterior se discutió que nuestros resultados podían explicarse por tres motivos diferentes: i) secreción selectiva de las mitades de tRNA; mejor dicho: de algunas mitades de tRNA; ii) secreción no selectiva pero estabilidad extracelular exacerbada de las mitades de tRNA frente a otros RNAs pequeños; iii) ambas. Pero que por muy distintos que los puntos (i) y (ii) pudieran parecer, ambos estimulaban la búsqueda de la “proteína de unión a las mitades de tRNA en el medio extracelular”. Por simplicidad, vamos a referirnos de aquí en adelante a dicha proteína(s) hipotética como ETBP (*extracellular tRNA-h binding protein*). La ETBP sería entonces la proteína responsable del transporte de las mitades de tRNA a través de la membrana plasmática de la célula, o la proteína responsable de captar dichas mitades en el espacio extracelular y protegerlas de la acción de las nucleasas ¿Qué podíamos esperar *a priori* de ETBP? Nada, a excepción de que su peso molecular debería ser tal vez inferior a los 40-50kDa, ya que de lo contrario los complejos ETBP/tRNA-h eluirían por exclusión molecular en fracciones anteriores a las detectadas.

Llegado a este punto tomamos una decisión importante; tal vez acertada, tal vez no. Recuérdese que la motivación de este trabajo era poder contribuir a la comprensión del rol de los pequeños RNAs en la comunicación intercelular. No quisimos comenzar con ensayos de transferencia entre células hasta no contar con una descripción detallada del repertorio de

³³ La Figura 15 muestra los rendimientos en la purificación de RNA, es decir, cantidad absoluta de RNA obtenido en cada caso. Si bien no refiere específicamente a las cantidades absolutas de tRNA-h, sabemos por nuestros datos de secuenciación que la mayoría del RNA en S100 son tRNA-h.

pequeños RNAs en distintas fracciones extracelulares. De lo contrario terminaríamos siguiendo el camino por el cual están muchísimos grupos de investigación de todo el mundo transitando: rol de los miRNAs en exosomas. Queríamos encontrar algo nuevo, algo distinto. Al cabo de cierto tiempo – como siempre, mucho más del previsto - teníamos el conocimiento de la abundancia de las tRNA-h en la fracción S100 del medio extracelular. Podíamos comenzar a estudiar entonces el rol de estos fragmentos en la comunicación célula – célula. Pero ¿cómo encarar experimentalmente esta tarea sin “haberle visto la cara” a la ETBP? Inyectar las tRNA-h puras, o agregarlas al medio extracelular: ¿garantizaba que los eventuales efectos observados fuesen los mismos que los obtenidos con la entidad biológicamente relevante, es decir, con el complejo ETBP/tRNA-h? Y si bien sabíamos como purificar dicho complejo, también sabíamos que las fracciones obtenidas no eran puras, sino que contenían otras proteínas ¿Cómo separar los efectos de unas y de otras? Conocer la identidad de la ETBP permitiría purificar el complejo por afinidad, reconstruir el complejo *in vitro*, silenciar ETBP y tener un control de las fracciones purificadas sin el complejo, bloquear su interacción con las células mediante anticuerpos, etc. Decidimos que necesitábamos conocer la identidad de ETBP antes de dar un siguiente paso. Y otra vez lo que parecía simple y rápido fue un mar de sorpresas que insumió tiempo, borrones y cuentas nuevas. Pero ahora sabemos algunas cosas que antes no sabíamos y, después de todo, de eso se trata.

3.2 MATERIALES Y MÉTODOS

La mayoría de los materiales y métodos utilizados en este capítulo son los mismos (con variantes especificadas caso a caso) que aquellos utilizados en el capítulo II. Se detallan a continuación los aspectos metodológicos de algunas técnicas utilizadas únicamente en esta parte del trabajo

3.2.1 Fases móviles y *buffers*

La formulación del “PBS” utilizado fue: Na_2HPO_4 8.059 mM; KH_2PO_4 1.4705 mM; KCl 2.7 mM; NaCl 137.93 mM; pH = 7.4. En ocasiones, se utilizó un *buffer* “PB” (Na_2HPO_4 10 mM; NaH_2PO_4 1.8 mM; pH= 7.4), en combinación con algunas sales, según se indica en cada caso (ej: PB - 150 mM NaCl)

3.2.2 Purificación de las tRNA-h por cromatografía de intercambio aniónico, y deconvolución de los cromatogramas de exclusión molecular.

Se purificaron los picos P1 y P2 de la SEC (a partir del medio condicionado de MCF-7), los cuales se trataron o no con RNasaA, según se indica en el apartado correspondiente. Las muestras (1mL) se inyectaron en una columna Resource Q (Amersham) de 1mL, utilizando

un *loop* de 1 mL y en un equipo FPLC (Äkta Pure, GE). La muestra fue inyectada y corrida en PBS, a una velocidad de flujo de 1 mL / min. Se colectó la fracción no unida (*flow-through*, FT) en los 1-3 mL post inyección. Posteriormente, se continuó el pasaje de fase móvil por al menos 9 volúmenes de columna adicionales. A partir de este punto, se eluyó en gradiente, pasando de 0% a 50% de *buffer* B (PB- 1M NaCl) en 10 min, manteniendo 50% de B por 3 min, y un nuevo gradiente de 50% a 100% de *buffer* B en 3 min. En todo momento se monitoreó la conductividad a la salida de la columna.

Para la deconvolución de los cromatogramas de exclusión molecular (SEC) se siguieron dos aproximaciones, según el caso. En los casos en que se analizó la fracción S100 proveniente del medio de cultivo, la muestra se trataba de una mezcla de RNA y proteínas. Ambas de las cuales absorben tanto a 280 nm como a 260 nm. Se realizaron las siguientes suposiciones/consideraciones:

- i) Para proteínas, se supuso un coeficiente de absortividad de $1.0\text{mg}^{-1}\text{mL cm}^{-1}$ a 280 nm (válido para proteínas con un contenido promedio de aminoácidos aromáticos), y de $0.6\text{ mg}^{-1}\text{mLcm}^{-1}$ a 260 nm (estimado de la relación 260/280 nm para el pico de la Transferrina).
- ii) Para el RNA, se supuso un coeficiente de absortividad de $30.1\text{ mg}^{-1}\text{mLcm}^{-1}$ a 260 nm (correspondiente al calculado para las mitades de tRNA-Glu mediante el programa IDT oligo Analyzer) y de $15.7\text{mg}^{-1}\text{mL cm}^{-1}$ a 280 nm (estimado de la relación 260/280 luego de inyectar oligonucleótidos puros).

Utilizando los parámetros anteriores, y los valores de absorbancia a 280 y 260 nm registrados, los cromatogramas se deconvolucionaron según el aporte estimado para las proteínas (generalmente representado en azul) y el aporte estimado para el RNA (suponiendo que todo el RNA fuese mitades 5' del tRNA-Glu, y generalmente representado en rojo). Si bien este procedimiento es solamente aproximado, permite una mejor visualización y comprensión de los cromatogramas.

En los casos en los que se inyectó RNA sintético, la deconvolución no corresponde ya a la segregación entre los efectos del RNA y las proteínas (que no las hay), sino el intento por predecir el área debajo de la curva para los picos P1 y P2, ya que ambos se encuentran considerablemente solapados. Para ello, se buscó el mejor *fitting* que, en base a la adición de dos o más campanas de Gauss, predijera la envolvente observada. Esto fue realizado mediante el Software Origin 8.

3.2.3 Preparación de las muestras para análisis por nanoLC-MS

Los picos P1 y P2 obtenidos a partir de la purificación cromatográfica de la fracción S100 del medio condicionado de células MCF-7 fueron concentrados hasta un volumen de 25 μL mediante ultrafiltración (Vivaspin 500; MWCO 5 kDa; Sartorius Stedim). En el caso de

muestras obtenidas a partir de los picos retenidos en las columnas de intercambio aniónico, se aprovechó este paso para volver la matriz a PBS, luego de dos ciclos de concentración, dilución, y re-concentración. Obtenida la muestra concentrada, se adicionó 1µL de Tripsina recombinante modificada para espectrometría de masas (0.5 µg/mL; Promega) y se digirió toda la noche a 37°C. Al día siguiente se purificaron los péptidos mediante el sistema de ZipTip (Millipore), utilizando una resina C18 y siguiendo las instrucciones del fabricante (activación de la resina en acetonitrilo [ACN] 50%, equilibrado en ácido trifluoroacético [TFA] al 0.1%, retención de la muestra mediante diez pasajes por la punta del ZipTip, lavado con la disolución de equilibrado, y elución en ACN 60%). El eluido se llevó posteriormente a sequedad mediante concentración en SpeedVac (30°C) y fue resuspendido en 10µL de ácido fórmico 0.1%, inyectándose en el equipo. El análisis inicial de los datos corrió por cuenta de la Unidad de Bioquímica y Proteómica Analíticas del IPMon, utilizándose el programa PatternLab para adjudicación de los espectros, y considerándose como presentes aquellas proteínas con al menos dos espectros.

Dado que el protocolo anterior implicaba exponer a la Tripsina a una disolución concentrada con presencia de inhibidores (ver: Resultados), se modificó el protocolo de modo tal de desnaturalizar todas las proteínas previo a la digestión triptica. Además, se eliminó el paso de purificación por ZipTip, por lo cual las muestras purificadas fueron invariablemente tratadas con RNasaA (320ng) por 30 min a 37°C. La nucleasa fue inhibida (DTT 1mM + 2µL de inhibidor de RNasa murino 40U/ µL [NEB]) y los nucleótidos o pequeños oligonucleótidos fueron separados de las proteínas mediante sucesivos pasos de dilución y concentración por ultrafiltración (Vivaspin 500; MWCO 5 kDa). Las muestras se concentraron hasta 50µL, luego de lo cual se adicionó DTT (10 mM) y se incubó por 20 min a 56°C. Posteriormente se adicionó iodoacetamida (50 mM), incubándose media hora a temperatura ambiente. De este modo los aminoácidos Cys quedaron alquilados (evitando la formación de futuros aductos con acrilamida), y la muestra fue mezclada con *buffer* de carga para SDS-PAGE (*buffer* Laemmli 4X, diluido ¼ con la muestra), previo a su corrida electroforética en geles al 12.5% (nota: preparados en flujo laminar horizontal y con todas las precauciones necesarias para evitar la inclusión de Queratina). Los geles se corrieron a 15mA y la corrida se cortó cuando el frente hubo ingresado aproximadamente 1 cm en el gel separador. El objetivo de este paso de purificación no fue, por tanto, separar las proteínas entre sí, sino eliminar posibles interferentes. Los geles se fijaron (50% etanol, 10% ácido acético) y se tiñeron brevemente (20 min) con *Coomasie* coloidal (*Coomasie* G250 preparado en etanol) y se destiñó parcialmente en agua ultrapura. El propósito de la tinción fue identificar el carril por donde migraron las proteínas, a efectos de poder cortar entre el comienzo del gel separador y el frente de corrida, llevando la menor cantidad de gel libre como fuese posible. Cortada la “banda”, se transfirieron a tubos *Eppendorf* de 1.5mL, y se cubrieron con una mezcla de *buffer* bicarbonato de amonio (AmBic) 0.2M pH = 8 (50%) y ACN puro (50%). Se incubó la mezcla a 30°C por 30 min con agitación, se descartó el líquido, y se repitió este paso. Llegado a este punto, la “banda” cortada del gel estaba completamente desteñida. Se procedió luego a deshidratarla mediante el agregado de ACN puro, incubando 5 min, quitando luego el ACN, y dejando los tubos abiertos en flujo laminar por 30 min para su completo secado. Las “bandas”

fueron entonces re-hidratadas agregando 100 μ L de una mezcla de buffer AmBic 1/3 y 1 μ L de Tripsina modificada (Promega), incubando toda la noche a 37°C. Al día siguiente se procedió a extraer los péptidos retirando el sobrenadante, y agregando 100 μ L de TFA 0.1% / ACN 60% y agitando 1 hora a 30°C. Este proceso fue repetido una vez, y los tres últimos sobrenadantes colectados fueron incorporados en un solo tubo, y llevados a sequedad mediante SpeedVac, luego de lo cual se resuspendió en ácido fórmico 0.1% y se inyectó en la columna del espectrómetro de masas.

3.2.4 Oligonucleótidos

Además de los mencionados en los capítulos anteriores, se incorporaron nuevos oligonucleótidos y RNAs sintéticos, cuya secuencia se suministra en el Anexo Metodológico. Los mismos fueron adquiridos en Integrated DNA Technologies (IDT, USA).

3.2.5 *UV-melting* y calorimetría de barrido diferencial (DSC)

Para la determinación de la temperatura de fusión de los presuntos dímeros de tRNA (T_m) mediante el monitoreo del hipercromatismo a 260 nm ("*UV-melting*"), se obtuvo la muestra – en concentración suficiente para la formación del dímero – y se colocó en una cubeta de cuarzo de 0.5mL, la cual se posicionó en un espectrofotómetro UV-visible Cary 60 (Agilent), conectado a un baño termostatzado digital. La temperatura del baño se subió desde temperatura ambiente y hasta 95°C, a razón de 2°C / min. La temperatura actual en la cubeta se midió utilizando un termómetro ubicado en una cubeta adyacente, y presente en la misma placa multi-celda. Se registró la absorbancia a 260 nm (corregida respecto del blanco en PBS, y a temperatura ambiente) en función del tiempo, y por tanto, también de la temperatura.

Para los ensayos de DSC, se dejó toda la noche al equipo (MicroCal VP-DSC) ciclando en PBS (referencia: PBS) desde los 25°C – 100°C, a 1°C/min, y hasta completa estabilización del sistema. Llegado a dicho punto, se inyectó la muestra previamente desgaseada (en concentración de 6 μ M), registrándose 3 de los antes mencionados ciclos. El procesamiento de los datos se hizo según especificaciones del fabricante y procedimientos de la Unidad de Biofísica de Proteínas del IPMon. Para la obtención de las gráficas de C_p (capacidad calorífica específica) en función de la temperatura, se utilizó la concentración predicha de dímero, suponiéndola igual a la mitad de la concentración estimada de RNA (y la cual se calculó estimando un hipocromatismo de 15% a partir de los datos de absorbancia a 260 nm de la SEC). Dichas gráficas fueron ajustadas a las líneas de base antes y después del pico mediante el Software Origin 8 y las rutinas pre-cargadas correspondientes al análisis por DSC, calculándose así la T_m , el ΔH calorimétrico (ΔH_{cal}), y el ΔH de Van't Hoff (ΔH_{vh}), mediante un ajuste a un modelo de desnaturalización con múltiples intermediarios.

3.3 RESULTADOS

3.3.1 Sensibilidad de los complejos conteniendo tRNA-h a la RNasaA y Proteinasa K.

Como ya fue descrito en el Capítulo II, la purificación cromatográfica de la fracción S100 del medio condicionado de las células MCF-7 y MCF-10A mostró dos picos en los cuales las mitades de tRNA (de aceptores de Glu y de Gly) pudieron ser detectadas. Denominamos P1 al pico que eluyó primero ($V_e = 10.2$ mL en S75 10/300; correspondiente a lo esperado para una proteína globular de 44kDa) y P2 al que eluyó después ($V_e = 11.8$ mL en S75 10/300; equivalente a 20kDa). Este perfil de elución en dos picos fue independiente del tipo de columna utilizada (Superdex 75 o Superdex 200). Vale aclarar que con ninguna de las dos columnas los picos pudieron ser completamente resueltos.

Los primeros ensayos realizados tuvieron como objeto demostrar que las mitades de tRNA estaban efectivamente formando parte de complejos ribonucleoproteicos. Se esperaba que al tratar la muestra (fracción S100 previo a la inyección) con proteasa, los picos de absorbancia a 260 nm (y picos de tRNA-h por SL-RT-qPCR) desaparecieran, o se corrieran hacia volúmenes de elución mayores (dependiendo que tan lábiles fueran una vez separados de ETBP). Por otra parte, se esperaba que el tratamiento con RNasaA (endo-ribonucleasa pancreática, más activa frente al RNA simple hebra) no produjera mayores cambios, ya que el complejo con proteínas protegería al RNA de la degradación. El experimento final era la combinación de los tratamientos: proteasa → inhibición de la proteasa → RNasa → inhibición de la RNasa → inyección.

Comenzamos probando con la proteasa Tripsina. Pero a pesar de varios intentos (aumentando la concentración, el tiempo de incubación, la alícuota de Tripsina) no logramos observar actividad proteolítica. No solo respecto a los picos P1 y P2, sino a la muestra en general. El medio de cultivo utilizado poseía una formulación en la cual se agregaba como aditivo Transferrina a concentraciones relativamente altas (5µg/mL). En consecuencia, los cromatogramas de la fracción S100 presentaban siempre un alto pico en la absorbancia a 280 nm correspondiente a la Transferrina; que de paso servía como control interno de cada corrida. Pues bien, el tratamiento de la muestra con Tripsina en nada afectaba al cromatograma; ni siquiera el pico de la Transferrina disminuía. Esto nos daba cuentas de que la Tripsina no estaba funcionando. La razón quedó clara al cabo de cierto tiempo, tal como se explicará más adelante.

Decidimos trabajar con una proteasa diferente. Las muestras tratadas con Proteinasa K presentaron una significativa reducción en el pico de la Transferrina respecto del ensayo control (incubado igual tiempo y a igual temperatura, y con adición del inhibidor de proteasas PMSF), lo que valida la actividad proteolítica de la enzima en la muestra. Conforme a lo esperado, se observó una desaparición casi total del pico P1. Sin embargo, el pico P2

permaneció prácticamente inalterado en lo que a su absorbancia a 260 nm se refiere (Figura 29, derecha).

Al tratar las muestras con RNasaA se observó la desaparición total del pico P1 y una reducción muy moderada del pico P2 (Figura 29, centro). Además de inesperado, este comportamiento era difícil de explicar. La sensibilidad del pico P1 a RNasas no excluía, en absoluto, que el RNA en dicho pico estuviese formando complejo con proteínas. El tratamiento con ribonucleasa pancreática bovina (78nM) no es fisiológico, y es esperable que un complejo proteína/RNA exponga parte del RNA al solvente, accesible a las RNasas si su concentración es suficiente. Ya en el Capítulo II se había advertido que la sensibilidad de la fracción S100 a la RNasaA era superior a la de las fracciones vesiculares. Pero entonces ¿cómo explicar la resistencia del pico P2 a las ribonucleasas? Considerando que las mitades de tRNA de Glu y de Gly ya tienen de por sí aproximadamente 10 kDa cada una, la proteína asociada a dichas mitades en el pico P2 (20 kDa) no podría ser demasiado grande, por compacta que fuese la conformación que adoptara el RNA. Y es difícil imaginar una proteína de menos de 20 kDa capaz de “envolver” al RNA de modo tal de volverlo altamente resistente a la degradación endonucleolítica.

Estos ensayos fueron repetidos varias veces y los resultados fueron siempre reproducibles: el tratamiento con RNasaA provocaba la desaparición total del pico P1, mientras que afectaba al pico P2 de una manera parcial (Figura 30, superior). Para corroborar que dichos efectos fuesen producto de la afectación de las mitades de tRNA y no de otros compuestos capaces de absorber a 260 nm, analizamos varias de las fracciones recolectadas por SL-RT-qPCR. Como ya fue descrito, las tRNA-h mostraron un pico definido en las fracciones correspondientes al P1, con grado de amplificación apreciable también en P2. Las mitades de tRNA^{Glu} siempre amplificaron antes que las del tRNA^{Gly} (valores de $2^{-\Delta Cq}$ uno o dos órdenes de magnitud superiores), aunque las diferencias pueden también deberse a la eficiencia de la RT-PCR. Coincidente con lo observado espectrofotométricamente, en las fracciones tratadas con RNasaA el grado de amplificación de las tRNA-h en P1 fue nulo. En cambio, éstas si fueron amplificables en las fracciones correspondientes a P2 (aunque con valores de $2^{-\Delta Cq}$ órdenes de magnitud inferiores a los de las muestras sin tratar).

De todo esto concluimos:

- i) P1 es sensible a Proteinasa K y a RNasaA
- ii) P2 es parcialmente resistente a Proteinasa K y a RNasaA

Luego veremos que los ensayos de sensibilidad a proteasas no fueron válidos, pero esto no lo supimos sino hasta varios meses después.

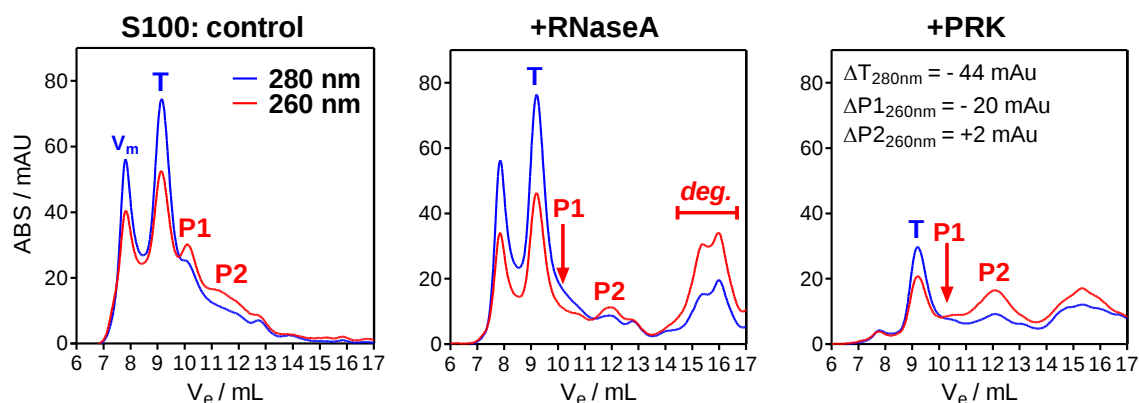


Figura 29: cromatografía de exclusión molecular (Superdex 75 10/300, Amersham; Fase móvil: PBS) de la fracción S100 del medio condicionado de MCF-7. Izquierda: control (muestra incubada a 37°C 30 min. y tratada con DTT 1mM, RNaseOUT 40U, PMSF 5mM). Centro: muestra tratada con 320ng de RNaseA (0.03 Kunitz U), incubada a 37°C 30 min, y luego tratada con DTT 1mM y RNaseOUT 40U. Derecha: muestra tratada con 8µg de Proteinasa K, incubada a 37°C 30 min, y luego tratada con PMSF 5mM. Curva azul: ABS 280nm. Curva roja: ABS 260nm. El volumen de elución (V_e) está medido desde la inyección. V_m : volumen muerto, o excluido. T: Transferrina. P1 y P2: picos donde eluyen las mitades de tRNA analizadas. Deg: posibles productos de degradación.

3.3.2 Intentos de identificar a ETBP mediante espectrometría de masas.

Como primera aproximación, se colectaron las fracciones correspondientes a los picos P1 y P2 y, después de concentrarlas y digerirlas con Tripsina, fueron enviadas para su análisis por espectrometría de masas en un equipo MALDI-TOF. Es importante señalar que, si bien muchos de los estudios proteómicos se realizaron sobre ambos picos, nuestra atención se centró principalmente en el pico P1. Por su tamaño aparente de 44 kDa, parecía claro que dicho pico correspondería a la elución de los complejos ETBP/tRNA-h. Por el contrario, P2 – de tamaño aparente mucho menor – podría originarse por la rotura parcial de los complejos y la consecuente liberación del RNA. Aunque hoy daríamos una interpretación diferente (ver Discusión), esta era nuestra interpretación en el momento.

El análisis por MALDI-TOF arrojó una gran complejidad de péptidos, algunos de los cuales fueron a su vez analizados por fragmentación o espectrometría de masas en tándem (MS/MS). Se pudieron identificar de este modo algunas proteínas: Enolasa 1 o α -Enolasa (ENO1), Calmodulina (CaM) y Transferrina. Ésta última ciertamente esperada, pues de la observación de los cromatogramas se desprende que el pico P1 y la Transferrina no se resuelven completamente. No se pudo identificar, sin embargo, más que un solo péptido por proteína. Esto evidenciaba que un gran número de proteínas se estaban co-purificando en el pico P1, sin ninguna que destacase por sobre las demás. En condiciones de este tipo, la técnica de MALDI-TOF no era la más adecuada.

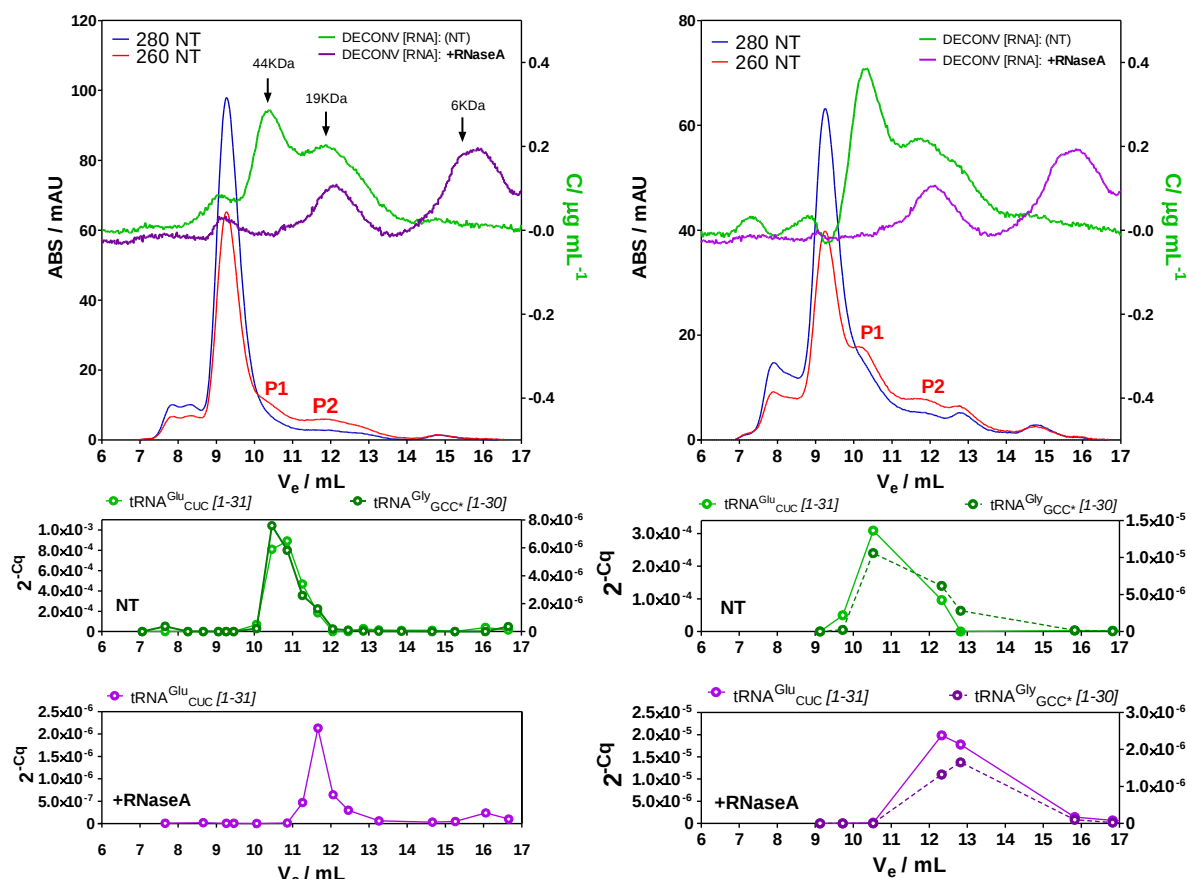


Figura 30: ensayos de sensibilidad de los picos P1 y P2 a la acción de la RNasaA. Se muestran dos réplicas del mismo experimento realizado en días diferentes (izquierda y derecha). Arriba: cromatogramas control (muestras tratadas con DTT 1mM y RnaseOUT 40U e incubadas a 37°C por 30 min; pero sin RNasaA), similares a los mostrados en la Figura 29. En los ejes verticales derechos se muestra la concentración estimada de las mitades de tRNA luego de aplicar una deconvolución de los cromatogramas que permite estimar cuánto de la absorbancia a 260 nm corresponde efectivamente al RNA, para la muestra no tratada (verde) y tratada con RNasaA (violeta). Abajo: resultados de la RT-SL-qPCR para distintas fracciones analizadas de los cromatogramas superiores (valores expresados como $2^{-\Delta Cq}$). Verde claro: fragmentos de 31 nucleótidos del extremo 5' del tRNA^{Glu}_{CUC}. Violeta claro: ídem, pero en la inyección correspondiente a la muestra tratada con RNasaA. Verde oscuro: fragmentos de 30 nucleótidos del extremo 5' del tRNA^{Gly}_{GCC*} (eje vertical derecho). Violeta oscuro: ídem, pero en la inyección correspondiente a la muestra tratada con RNasaA (eje vertical derecho).

Además, la alta cantidad de proteínas co-purificadas en P1 no permitían una aproximación del tipo: “guilty by association” (las tRNA-h eluyen en P1; la proteína detectada eluye en P1; ergo: la proteína es ETBP). Es decir que el problema no consistía en identificar simplemente las proteínas presentes en P1 y P2, sino en manipular el sistema de modo tal que de la comparación entre dos condiciones experimentales surgiera ETBP. Nuestra estrategia fue recurrir al tratamiento con RNasaA como variable: degradando a las tRNA-h en P1 podría afectarse el volumen de elución de ETBP. Las demás proteínas presentes, salvo que también estuviesen unidas a moléculas de RNA, no deberían verse afectadas. Claro que esta aproximación descansaba en una suposición: la unión de las tRNA-h a ETBP debería afectar su radio hidrovolumétrico lo suficiente como para que sus formas unida y no unida eluyesen a volúmenes considerablemente diferentes.

Para obtener información cuantitativa en mezclas con cierta diversidad de proteínas, se decidió recurrir a la espectrometría de masas asociada a un equipo de cromatografía líquida miniaturizada (nano-LC, LTQ-Orbitrap Velos), en colaboración con la Unidad de Bioquímica y Proteómica Analítica del IPMon.

Se aislaron los picos P1 y P2 por SEC³⁴ (con y sin tratamiento con RNasaA), se digirieron con tripsina toda la noche, y se inyectaron en el equipo. En un primer análisis se identificaron 310 proteínas en el pico P1 sin tratar, 263 en P1 luego de tratar con RNasa, 608 en P2 sin tratar, y 614 en P2 luego del tratamiento con RNasa. Como era de esperar, entre las proteínas más abundantes identificadas en P1 se encontraba la Transferrina, que fue tratada como un contaminante. Muchas proteínas desaparecían con el tratamiento con RNasa, otras aparecían, otras bajaban su señal y otras lo aumentaban ¿Cuál era la significancia estadística de todo eso? ¿Cuál sería ETBP? A la primera pregunta se la contestaba fácil: ninguna. Necesitábamos réplicas para empezar a descubrir si alguna proteína *sistemáticamente* disminuía su señal luego del tratamiento con RNasa. Pero con la segunda réplica del ensayo comenzaron los problemas: se detectaron muchas menos proteínas, y cada una con mucho menor número de espectros. Fue así que comenzamos una serie de intentos fallidos para repetir el ensayo, pero cada vez la detección era peor: a veces solo aparecía la Transferrina y la ubicua Queratina. Y no es que fuese un problema técnico (que los hubo); eran las muestras.

Volviendo entonces a la lista del primer ensayo que funcionó de manera razonable, y repasando las proteínas una a una, encontramos una cantidad muy significativa de espectros correspondientes a las proteínas SPINT1 (*Kunitz-type protease inhibitor 1*) e ITIH4 (*inter-alpha-trypsin inhibitor heavy chain*), ambas con localización extracelular descrita ¡Las células MCF-7 estaban secretando inhibidores de Tripsina al medio, parte de los cuales estaban copurificando con las tRNA-h! Esto explicaba lo errático de los resultados, pues las preparaciones con menor cantidad de inhibidor podrían ser al menos parcialmente digeridas y analizadas. Y explicaba también por qué nunca habíamos tenido éxito en digerir la fracción S100 con Tripsina, y sí con Proteinasa K (§3.3.1).

Para solucionar este inconveniente se decidió desnaturalizar todas las proteínas de los picos P1 y P2 previo a su digestión con Tripsina, siguiendo un protocolo de purificación por SDS-PAGE y digestión en gel (ver Materiales y Métodos). El número de proteínas identificadas en P1 no fue mayor por este protocolo (233 y 225 con y sin RNasaA, respectivamente) pero sí se obtuvo un número mucho mayor de espectros por proteína, llegando a casi 1500 para la Transferrina, por ejemplo. Además, los resultados fueron auto-consistentes, obteniéndose una buena correlación entre el número de espectros para cada proteína, con y sin tratamiento con RNasaA ($r = 0.9835$). Esto evidenciaba que ninguna de las proteínas detectadas parecía afectarse a causa del tratamiento con RNasaA. Si bien ahora entendemos que esto era un resultado esperable, en aquella oportunidad fue un poco decepcionante.

Pensamos, errónea pero fortuitamente, que el problema se simplificaría reduciendo la complejidad de la muestra, es decir, aumentando los pasos de purificación de los complejos ETBP/tRNA-h. Se decidió incluir una cromatografía de intercambio aniónico (en adelante

³⁴ Cromatografía de exclusión molecular (*size-exclusion chromatography*).

AEC por *anionic exchange chromatography*) posterior a la SEC. Consideramos que el RNA podría conferirle una carga neta negativa a los complejos ETBP/tRNA-h, que provocase su retención en las columnas. La adición de RNasaA previo a la AEC debería, razonamos, aumentar la cantidad de ETBP colectada en la fracción no unida (FT, por *flow-through*).

Se realizaron una serie de controles accesorios de modo tal de optimizar las condiciones de este nuevo ensayo. Lo primero fue estudiar la estabilidad de los complejos ETBP/tRNA-h frente a la variación de la fuerza iónica. Esto era importante dado que las muestras se inyectarían a una concentración salina tan baja como fuese posible, y la fracción retenida se eluiría posteriormente a alta concentración salina. Existía la posibilidad de que el aumento de la fuerza iónica rompiese los complejos y, por consiguiente, ETBP eluyese antes que el RNA (que es nuestra marca)³⁵.

Se realizó a partir de este punto un nuevo cambio. Dado que la Transferrina contaminante terminaba siendo la proteína mayoritaria en P1, se alteraron las condiciones de cultivo de modo tal de pasar a un medio definido sin Transferrina agregada. Así pues, se comenzaron a cultivar las células MCF-7 en el medio MEGM – el mismo que fue utilizado para las MCF-10A – sin que esto conllevara alteraciones evidentes en la viabilidad celular, tiempo de duplicación o morfología. En la Figura 31 puede observarse el cromatograma de la fracción S100 obtenida mediante estas nuevas condiciones de cultivo, ya sea utilizando una columna analítica Superdex 200 10/300 (A) o Superdex 75 10/300 (D). Nótese cómo la desaparición del pico correspondiente a la Transferrina permite una mejor visualización de los picos P1 y P2, que siguen eluyendo al mismo volumen. De aquí en más, todos los medios condicionados de MCF-7 fueron obtenidos en MEGM.

Volviendo a los ensayos de estabilidad del complejo a distinta fuerza iónica, lo primero que debe decirse es que, cuando la fase móvil de la cromatografía pasó de ser PBS a Tris-Cl 50 mM pH = 7.4; NaCl 150mM, casi todos los picos del cromatograma desaparecieron, y en su lugar aumentó la absorbancia en el volumen muerto o excluido (V_0). Por esta razón se decidió seguir trabajando con base a un *buffer* fosfato (PB: $\text{H}_2\text{PO}_4^-/\text{HPO}_4^{2-}$), alterando la fuerza iónica mediante la concentración de NaCl. A baja concentración salina (50mM de NaCl) se observó una desestabilización de los picos P1 y P2, apareciendo formas de mayor tamaño aparente, con un corrimiento del pico principal equivalente a un incremento de 38 kDa (Figura 31, C). El aumento de la absorbancia a 260 nm llegó incluso a manifestarse en el volumen excluido, que en una columna S200 corresponde a grandes agregados. En cambio, no se observaron diferencias evidentes entre las corridas con PBS y *buffer* PB + 500 mM NaCl (Figura 31, D). De estos resultados concluimos que deberíamos inyectar en la columna de intercambio aniónico la muestra en PBS (aún cuando la fuerza iónica fuese más alta de lo recomendable para el pegado), pero podíamos tener la “tranquilidad” de que el complejo no se rompería al eluir con NaCl (al menos hasta 500mM, y siempre y cuando el comportamiento en la interfase sólido/líquido fuese igual a en solución).

³⁵ No trabajamos con RNA marcado, pero el coeficiente de absortividad molar a 260 nm del RNA es órdenes de magnitud superior al de las proteínas a 280 nm. Supongamos que un complejo proteína/RNA (1:1) pasa por un detector que mide a 280 y 260 nm. La señal del RNA puede ser grande, mientras que la de la proteína asociada puede ser indetectable. Por tanto, la proteína se puede decir que “queda marcada” por el RNA, en la medida en que no se le separe.

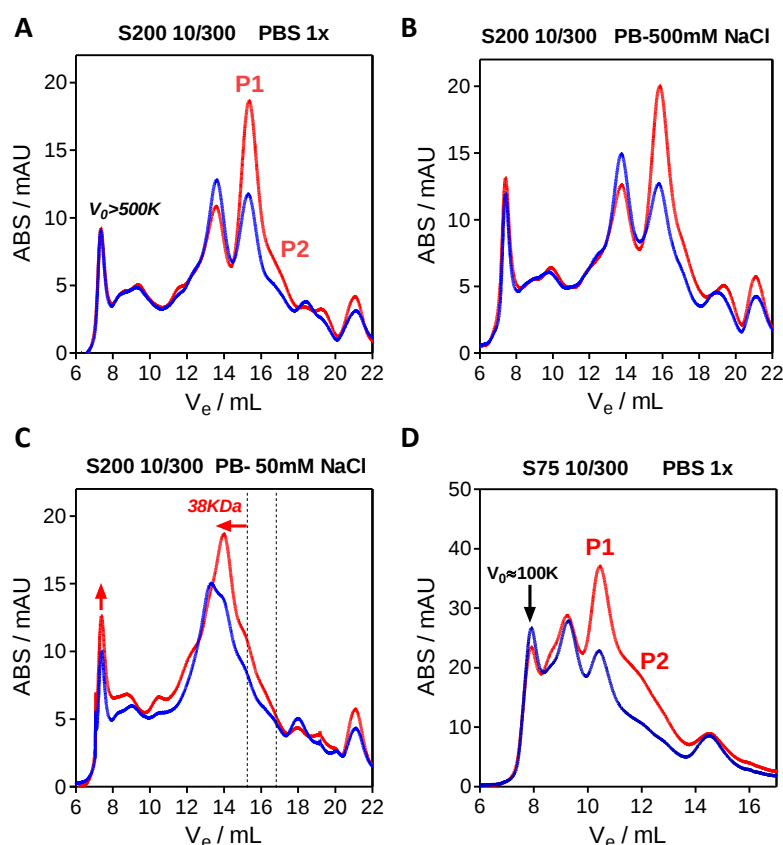


Figura 31: estabilidad de los complejos conteniendo mitades de tRNA a distinta fuerza iónica. Los cromatogramas (SEC; Superdex 200 10/300: A-C; Superdex 75 10/300: D) medidos a 280 (azul) y 260 nm (rojo) corresponden a la fracción S100 del medio condicionado de células MCF-7 utilizando como fase móvil PBS (A,D), buffer fosfato (PB) + 50mM NaCl (C) y buffer fosfato (PB) + 500mM NaCl (B). Las líneas verticales discontinuas marcan la posición de los picos P1 y P2 según el cromatograma (A).

Los cromatogramas (AEC) obtenidos a partir del pico P1 purificado por SEC, con y sin tratamiento posterior con RNasaA, se muestran en la Figura 32. En las muestras no tratadas, la absorbancia a 260 nm en el *flow-through* (FT) fue despreciable; señal de que la mayor parte del RNA quedaba retenido en la columna. Terminada la colección del FT, el pico que aparece de predominante absorbancia a 280 nm corresponde al inhibidor de RNasa murino, agregado tanto a las alícuotas tratadas como a las alícuotas control. Dado que se trata de una proteína ácida ($PI = 4.7$) no sorprende que haya sido retenido en la columna, pero la moderadamente alta concentración salina en el *buffer* de pegado (≈ 137 mM NaCl) termina finalmente provocando su elución. La mayor parte de las moléculas absorbentes a 260 nm eluyen a una alta concentración salina, cuando la conductividad supera los 44 mS/cm, lo que equivale a 400 mM NaCl, aproximadamente. Esto habla de la fuerza de la unión del RNA a la resina, de carga fuertemente positiva (ligandos de amonio cuaternario). Llamamos a este pico colectado P1*. Nótese la ausencia de pico luego del tratamiento con RNasaA. Por el contrario, en este caso, la absorbancia a 260 nm pareció desplazarse hacia el FT.

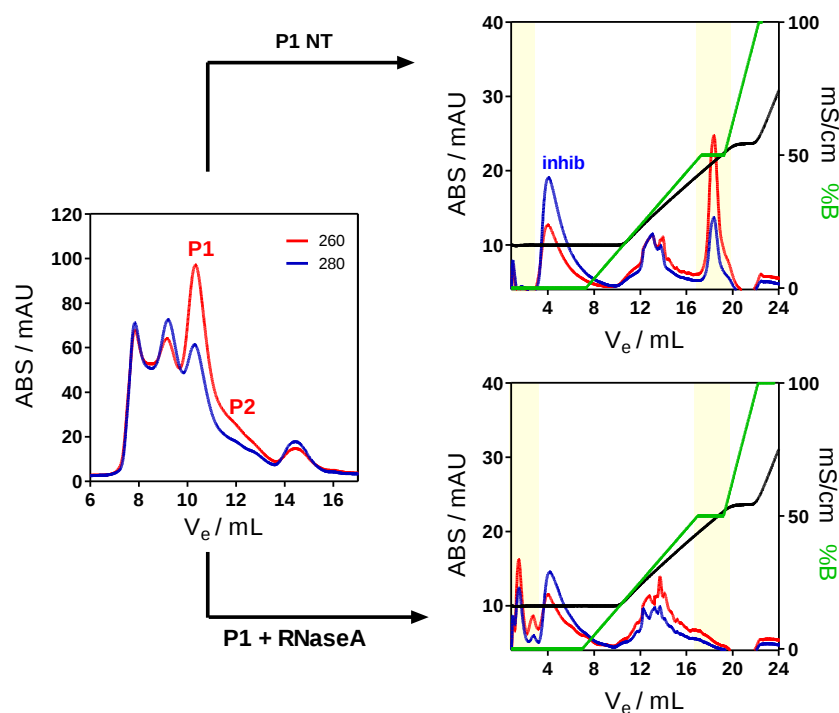


Figura 32: purificación de los complejos conteniendo tRNA-h mediante cromatografía de exclusión molecular (SEC, S75 10/300; izquierda) seguida de cromatografía de intercambio aniónico (AEC, ResourceQ; derecha). El pico P1 colectado de la SEC se dividió en dos alícuotas, una de las cuales fue tratada con RNasaA (gráfico inferior). Los ensayos control (gráfico superior) recibieron DTT 1mM, inhibidor de RNasa murino (inhib), y fueron incubados en las mismas condiciones (30 min 37°C). Las curvas azules y rojas corresponden a la absorbancia a 280 y 260 nm; respectivamente. En AEC, la curva verde corresponde al porcentaje del buffer de elución (PB pH = 7.4; 1M NaCl) y la curva negra a la conductividad iónica a la salida de la columna. Las sombras amarillas señalan las fracciones colectadas de la AEC, correspondientes a la fracción no retenida ("FT"; Vol: 1-3 mL) y a la elución de las tRNA-h ("P1*"; Vol: 17-20 mL; 44.3-50.0 mS/cm).

Al analizar por espectrometría de masas (nLC/MS) el pico P1* se detectaron únicamente tres proteínas glucolíticas: ENO1, TPI1 y PGK1, con 26, 29 y 5 espectros, respectivamente. Lo mismo ocurrió al analizar el pico P2*, lo que implica que no habría proteínas en P2 distintas de las encontradas en P1 y capaces de pegarse a la resina³⁶. Dado que estas proteínas se encontraban también entre las más abundantes en P1, y fueron las mayoritarias también en el FT, podría pensarse en un simple efecto de acción de masas, aunque la retención de PGK1 no es explicable si se considera que tiene un punto isoelectrico de 8.3 (lo que la hace positiva al pH de trabajo).

La cantidad de proteínas identificadas, así como el número de espectros por proteína, fue mucho mayor en el FT, lo que permitió realizar comparaciones entre muestras tratadas y no tratadas con RNasaA previo a la inyección en la AEC (y posteriormente a la SEC). Se realizaron, al respecto, dos ensayos independientes. En el primer caso, la muestra correspondió únicamente al pico P1 aislado de la SEC. Se identificaron 178 y 46 proteínas humanas en las muestras tratadas y no tratadas, respectivamente. La correlación entre el número de espectros para cada proteína fue considerablemente alta ($r = 0.9128$), implicando

³⁶ Por sus pesos moleculares las 3 proteínas deberían eluir en la región de P1 y no en la de P2, pero la baja resolución entre ambos picos provoca un alto grado de solapamiento.

que el tratamiento con RNasaA no produjo mayores variaciones en el perfil de proteínas colectadas en el FT. Sin embargo, las dos proteínas aparentemente más abundantes, TPI1 y ENO1, parecían poseer un número de espectros ligeramente menor del esperable, en las muestras tratadas (Figura 33, A). Esta variación no era, sin embargo, suficientemente grande como para pensar en un efecto específico. El experimento se repitió, pero inyectando a la AEC el total de fracciones correspondientes a los picos P1 y P2 de la SEC. En esta oportunidad, el número de proteínas identificadas en las muestras tratadas y no tratadas fue más parejo (145 vs. 112), pero la correlación de los datos fue, afortunadamente, pobre ($r = 0.7294$). Una única proteína, la ENO1, era la responsable de este desajuste (Figura 33, B). Su exclusión fue suficiente para elevar el coeficiente de correlación a $r = 0.9234$, lo que implica que los datos estaban, por lo demás, bien correlacionados. Esto es importante pues habla de que el efecto de la RNasaA fue específico sobre ENO1, y no sobre las demás proteínas de la muestra. En lugar de los 305 espectros que serían esperables para ENO1 en la muestra tratada, se obtuvieron únicamente 28; pasando de ser la proteína más abundante, a la posición 17. Aunque ambos experimentos no fueron estrictamente réplicas (el tipo de muestra utilizado en cada caso fue diferente), si se los considerase como tales, ENO1 sería la única proteína significativamente alterada por el tratamiento con nucleasa (Figura 33, C).

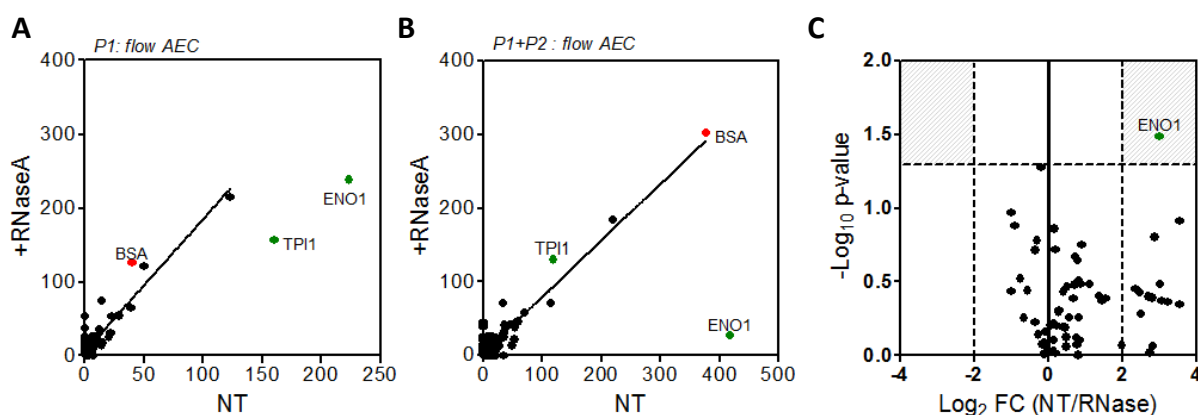


Figura 33: identificación por espectrometría de masas (nano-LC/ESI/LTQ-Orbitrap-Velos) de la proteína ENO1 como candidata a ser ETBP. A) Número de espectros correspondientes a las proteínas presentes en el pico P1 de la SEC y no retenidas (FT: “flow-through”) por la columna de intercambio aniónico (AEC), ya sea con (+RNaseA; eje vertical) o sin (NT; eje horizontal) tratamiento con RNasaA previo a la AEC. Se muestra la recta de mejor ajuste excluyendo a las proteínas TPI1 y ENO1. B) ídem (A), pero inyectando ambos picos P1 y P2 de la SEC, juntos. La recta de mejor ajuste excluye únicamente a ENO1. C) “volcano-plot”, estableciendo la variación entre el número de espectros normalizados para cada proteína en las muestras sin tratar (NT) y tratadas con RNasaA. En sombreado gris se muestran las regiones dónde se ubicarían los puntos significativamente ($P < 0.05$; test t de medidas pareadas a dos colas) enriquecidos o disminuidos (>4 veces) en las muestras tratadas. Los espectros fueron normalizados al número total de espectros obtenidos en cada corrida (y eliminando proteínas contaminantes como las queratinas).

En un principio resultó extraño y enigmático que una enzima glucolítica como la α -Enolasa (ENO1; que cataliza el penúltimo paso de la glucólisis) fuese nuestra principal candidata a ser ETBP. Ciertamente, hubiésemos preferido una proteína con capacidad de unión al RNA reportada. Una breve lectura nos puso al tanto, sin embargo, de la reconocida actividad

“moonlighting”³⁷ que presentan la mayoría de las enzimas glucolíticas. Al día de hoy son varias las enzimas de esta vía cuya capacidad de unión al RNA está ampliamente aceptada (Castello et al., 2015), y la ENO1 es precisamente una de las mejor estudiadas. Más aún, en levaduras, participa del transporte de tRNAs a la mitocondria, lo que la hace ser una validada “tRNA-binding protein” (Entelis et al., 2006). Participa también como una proteína de respuesta al estrés celular, particularmente a la hipoxia, actúa como chaperona, puede unir Plasminógeno en el espacio extracelular, y juega variados roles en cáncer, debido a que participa, como se discutirá más adelante, en la regulación de c-Myc (Ji et al., 2016). Por otro lado, la capacidad de las enzimas glucolíticas de ser secretadas por células tumorales se conoce desde hace varias décadas (Keller et al., 1978).

Aunque estos ensayos deben aún ser reproducidos, el efecto de la RNasaA sobre ENO1 parecería ser potenciado con la inclusión de P2. Otro punto interesante es que, contrario a lo esperable en función de la hipótesis con la que el experimento fue diseñado, ENO1 no aumentó, sino que disminuyó en el FT luego del tratamiento con la nucleasa. Se recordará que lo que se buscaba era una proteína que se retuviera en la columna a causa de su unión al RNA y que, luego de tratar con RNasa, apareciese o aumentase en la fracción no retenida. Se obtuvo aquí un efecto, pero exactamente el contrario. En la discusión se volverá sobre este inesperado resultado.

3.3.3 Los picos P1 y P2 de la SEC son específicos de las tRNA-h.

Por RT-SL-qPCR específicas para las tRNA-h (de Glu y Gly) podíamos saber que los fragmentos de tRNA eluían en P1 y en P2, pero ¿qué hay de los demás pequeños RNAs identificados en el secuenciado profundo de S100? Como se describió en el capítulo II, y si bien las tRNA-h eran la población predominante, otras especies tales como miRNAs y fragmentos de YRNAs también estaban presentes ¿Formarían complejos con las mismas proteínas que los tRNA-h?

La Figura 34 muestra un análisis de distintas fracciones colectadas luego de la purificación cromatográfica de la fracción S100 de las células MCF-7 crecidas en MEGM. Se analizaron las tRNA-h de Glu y de Gly, miR-21-5p (por ser el miRNA más abundante), y los fragmentos 5' del RNA Y4. Los valores de Cq están normalizados (Δ) a $V_e = 14$ mL; no se espera la elución de macromoléculas a partir de dicho volumen. Como puede apreciarse, se observa un pequeño pico para el miR-21-5p y los fragmentos de Y4 a $V_e = 9.2$ mL; un tamaño aparente de 75 kDa que bien podría ser la cola de un pico de mayor tamaño (el siguiente punto analizado está recién a los 105 kDa). Corresponda a lo que corresponda dicho pico [los complejos AGO2/miRNA son de ~100 kDa; (La Rocca et al., 2015)], lo importante a los efectos de esta tesis es que los picos P1 y P2 parecerían ser específicos de las mitades de tRNA.

³⁷ Claro de luna: refiere a la capacidad de una enzima de realizar funciones celulares completamente distintas a la función catalítica por la que dicha enzima es clásicamente reconocida.

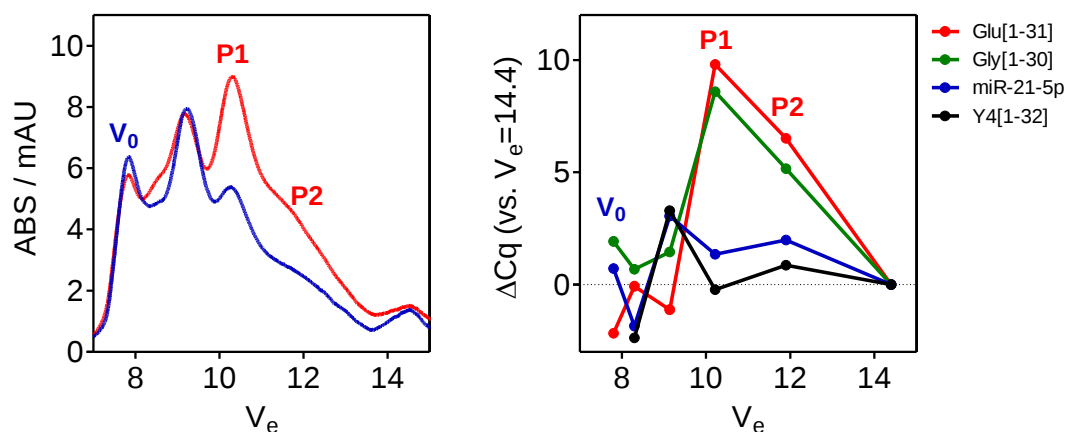


Figura 34: análisis por SL-RT-qPCR (derecha) de los fragmentos 5' del RNA Y4 (negro), del miR-21-5p (azul) y de las mitades 5' del tRNA^{Gly}_{GCC} (1-30; verde) y del tRNA^{Glu}_{CUC} (1-31; rojo). Dichas RT-PCR fueron realizadas sobre fracciones colectadas a partir del cromatograma de la derecha.

3.3.4 Re-evaluación de la sensibilidad de los picos P1/P2 a proteasas: equilibrio entre estados del RNA “desnudo”.

La Figura 29 es representativa de una serie de ensayos que nos llevaron a concluir que P2 era tanto resistente a la RNasaA como a la Proteinasa K (PRK), mientras que P1 era sensible a ambas enzimas, y por tanto *debía* corresponder a un complejo ribonucleoproteico.

Posteriormente realizamos un ensayo de incubación de la fracción S100 por 30 min a 37°C (tiempo y temperatura usados para los ensayos anteriores), pero sin el agregado de ninguna enzima. La única variable fue el agregado o no de inhibidor de RNasa. Se observó que, en ausencia de inhibidor, el resultado era el mismo que el que se obtenía al agregar RNasaA: desaparición del pico P1, y aparición de absorbancia a 260 nm a altos volúmenes de elución (degradación). Concluimos entonces que el pico P1 era tan sensible a las RNasas que una incubación de media hora a 37°C era suficiente para que las nucleasas presentes en la muestra barrieran con dicho pico (al margen de que fueren realmente nucleasas endógenas, o de que se hubiesen agregado accidentalmente durante la preparación de la muestra). El asunto es que los controles en los ensayos de sensibilidad a proteasas se corrían en presencia de PMSF, y no con inhibidor de RNasa. Por consiguiente, los efectos observados podían deberse tanto a la acción de la PRK agregada, como a la acción de RNasas no agregadas conscientemente, pero presentes al fin. Se pensará que esto tendría fácil solución: bastaría con agregar inhibidor de RNasa durante los ensayos con PRK. El problema es que dicho inhibidor es una proteína que actúa por “*binding*”, y al ser degradado por la PRK liberaría las nucleasas secuestradas. Claro que las RNasas también son proteínas, pero al menos la RNasaA mantiene un buen nivel de actividad luego de su digestión con PRK (Rauber et al., 1978). Estos ensayos estaban condenados a no ser concluyentes

Lo que sí parecía claro era la total prescindencia de proteínas del pico P2. El mismo no se veía afectado siquiera después de realizar el tratamiento con PRK en presencia de 0.5% SDS. Es sabido que esta proteasa no se desnaturaliza por la acción del detergente, pero su actividad

aumenta significativamente ($\approx 30\%$) debido a la desnaturalización parcial de las demás proteínas.

Con todo esto en mente se procedió a obtener los picos P1 y P2, aislar el RNA en ellos contenidos mediante el método del Trizol (LS), y re inyectar por separado el RNA puro (Figura 35). Si bien parte de la absorbancia a 260 nm apareció en el volumen excluido (tal vez debido a resuspensión incompleta), llamó la atención que se recuperaron los picos P1 y P2 originales, y a exactamente los mismos volúmenes de elución. Esto demostraba que, contrariamente a lo que hasta entonces era nuestro modelo de trabajo, el tamaño aparente del pico P1 (44 kDa) no se debía a la formación de un complejo entre las tRNA-h (10 kDa) y alguna proteína. La capacidad de formar los picos P1 y P2 residía enteramente en el RNA.

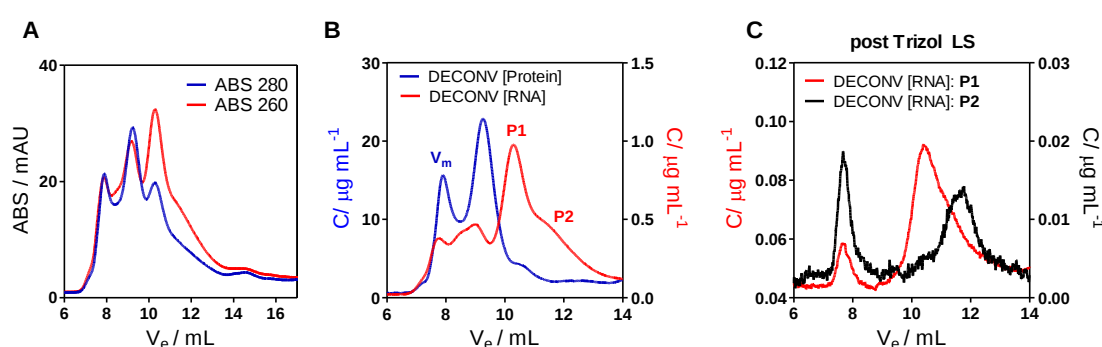


Figura 35: purificación cromatográfica del RNA purificado presente en los picos P1 y P2. A) Cromatograma de partida. B) Deconvolución del cromatograma de partida en la señal correspondiente a las proteínas (utilizando una absorptividad molar a 280 nm genérica) y al RNA (utilizando la absorptividad molar a 260 nm teórica para las mitades del tRNA-Glu). C) Los picos P1 y P2 se aislaron, y el RNA se purificó mediante el agregado de Trizol LS. El RNA puro (50% de la cantidad obtenida) se volvió a inyectar en la misma columna. Se muestra la deconvolución correspondiente al RNA para el pico P1 (rojo) y P2 (negro).

Todo lo anterior llevó a considerar que P1 y P2 fuesen dos estados en equilibrio de las tRNA-h y, dado que P1 corresponde a un tamaño aparente que duplica al de P2, dicho equilibrio podría ser del tipo monómero/dímero. Como primera aproximación se adquirió un oligonucleótido (de DNA) de secuencia complementaria (anti-sentido; as) a la mitad 5' del tRNA-Glu, y marcado en 5' con carboxyfluoresceína (6-FAM). Al inyectar dicho oligo en las columnas y condiciones utilizadas para la fracción S100 (S75 10/300; fase móvil: PBS) se obtuvo un único pico en $V_e = 12.2$ mL, similar al volumen de elución de P2 (Figura 36, B). Dicho oligo puede monitorearse tanto por su absorbancia a 260 nm como por su absorbancia a 494 nm, producto del FAM. Al mezclarlo con un oligo complementario (sentido, s; $V_e = 11.7$ mL) e inyectar la mezcla, los picos individuales desaparecieron y en su lugar apareció un pico de $V_e = 10.5$ mL, que coincide casi exactamente con P1 (Figura 36, C). Dado que se trata de un único pico tanto en el cromatograma a 260 nm como a 494 nm, inferimos que se trata de un dúplex entre ambos oligos (Figura 36, G). Como control se pre-incubó el oligo (as) marcado con FAM con un oligo no complementario (Figura 36, H), y al inyectar la muestra se

obtuvo la envolvente de dos picos no resueltos (nótese la absorbancia a 494 nm únicamente en la mitad de la envolvente a 260 nm), pero sin formación significativa de pico alguno a $V_e = 10.5$ mL. Todo esto en apoyo de la idea de que P1 y P2, por sus volúmenes de elución de ~ 10.5 mL y ~ 12.0 mL, podrían corresponder a un dímero y monómero de RNA, respectivamente. También se probó incubar la fracción S100 del medio condicionado de MCF-7 con el oligo (as), pero no se observó que el oligonucleótido pasase a formar parte de estructuras oligoméricas (Figura 36, E). Sin embargo, como se verá a continuación, los modelos predichos en P1 y en P2 son lo suficientemente estructurados como para que un oligo complementario no sea capaz de formar un híbrido, aun cuando las secuencias lo permitiesen.

3.3.5 Posibles conformaciones de las tRNA-h

Los resultados anteriores apuntaban a la presencia de monómeros estructurados en P2 y de dímeros en P1, pero ¿permitiría la secuencia de las tRNA-h esta posibilidad?

Lo primero a considerar es que las mitades 5' de tRNAs presentan ya de por sí una potencial estructura de bucle y horquilla, que es la que corresponde al brazo D de los tRNAs maduros de los cuales provendrían. Sin embargo, dicha estructura es poco probable que se mantenga en solución y a temperatura ambiente, dado que la hélice está formada por 4 pares de Watson y Crick, de los cuales solo uno corresponde a un par GC (en el tRNA^{Gly}_{GCC}). Al realizar una predicción de estructura secundaria con el programa mfold (<http://unafold.rna.albany.edu/>) se obtienen estructuras diferentes, tal como se muestra en la Figura 37. Según mfold y otros programas similares, estas tRNA-h podrían adoptar una estructura de tallo-bucle entre la posición +2 y +17, con 6 apareamientos de Watson y Crick estabilizando la horquilla (2 CG en Gly; 4 CG en Glu, aunque con una base desapareada). En ambos casos, el bucle poseería la secuencia palindrómica UGGU.

De la posición 18 en adelante (es decir, hacia 3') la secuencia permanecería desordenada (Figura 37, B-C). Esto es importante pues dicho segmento podría actuar como elemento de dimerización, tanto para homodímeros Gly-Gly como para heterodímeros Glu-Gly (Figura 37, E-F). En ambos casos, pero especialmente para los homodímeros, la complementariedad es casi perfecta y se daría a lo largo de todo el segmento desordenado, sumando 12 nuevos apareamientos de Watson y Crick (4 CG), apenas interrumpidos por una citosina protruyente ("*base flipping*") en cada hebra. Tal vez lo más interesante de estos dímeros potenciales es que los extremos 5' y 3' – y especialmente este último en Gly-Gly – quedan casi que totalmente "enterrados" en la estructura (la guanina 1 desapareada podría potencialmente plegarse sobre el dímero e interactuar con la guanina 18 mediante apareamientos no Watson y Crick). Lo que predice una natural resistencia a las exonucleasas, que debe de ser probada experimentalmente.

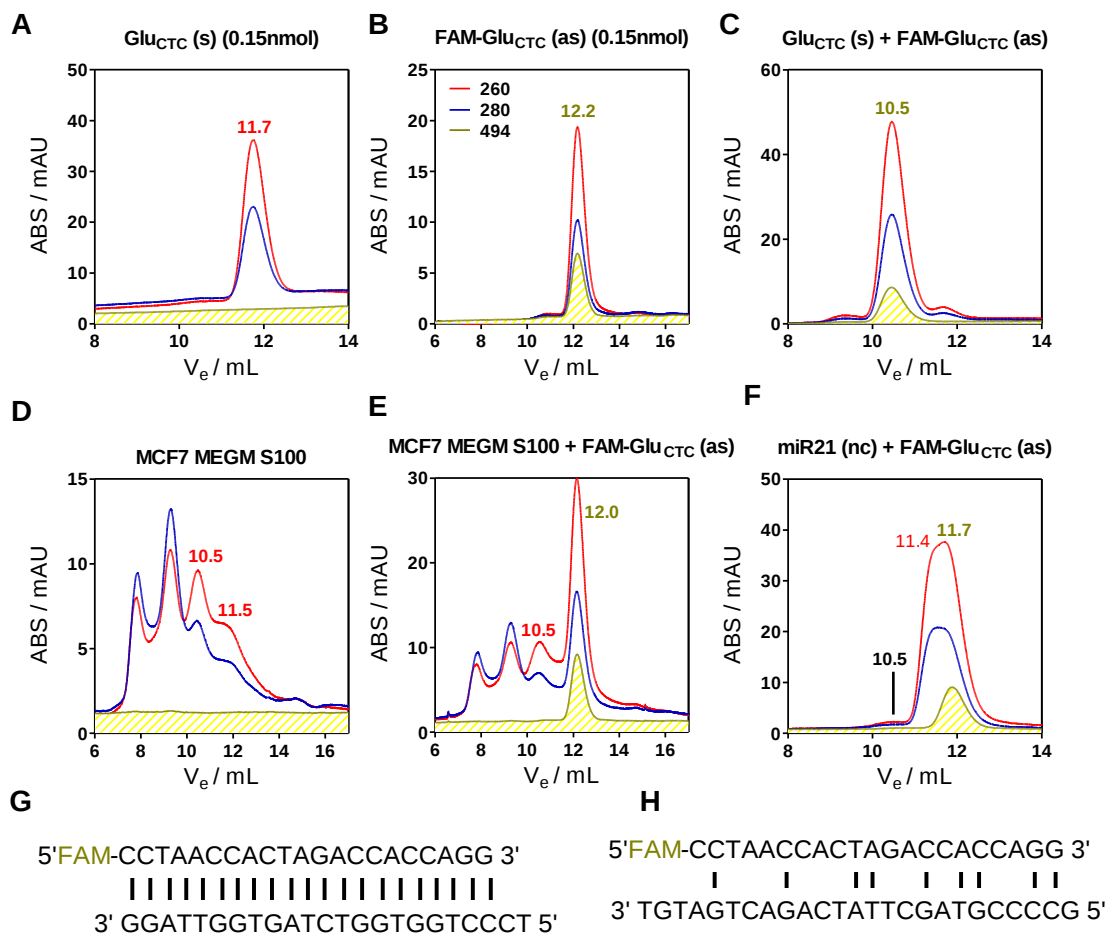


Figura 36: análisis por cromatografía de exclusión molecular de oligonucleótidos en estado de simple y doble hebra. A) inyección de 0.15 nmol del “primer forward” utilizado para las qPCR de las mitades de tRNA-Glu. Su secuencia es sentido (s) a la de dichas mitades. B) inyección de 0.15 nmol de un oligo anti-sentido (as) a las mitades 5’ del tRNA-Glu, y marcado con 6-FAM en 5’. C) inyección de cantidades equimolares de ambos oligos pre-mezclados. D) Purificación de la fracción S100 del medio condicionado de las células MCF-7 crecidas en MEGM. E) inyección de la mezcla entre la fracción S100 y 0.15 nmol del oligo (as) marcado con 6-FAM. F) inyección de la mezcla entre el oligo (as) marcado con 6-FAM y un oligo no complementario (nc) correspondiente al “primer forward” de la qPCR del miR-21-5p. G) Mejor apareamiento posible de los oligos utilizados en (C). H) Mejor apareamiento posible de los oligos utilizados en (F). Las absorbancias a 280 nm y 260 nm se muestran en azul y rojo, respectivamente. La absorbancia a 494 nm se muestra en amarillo.

En la Figura 37 también se hipotetiza sobre posibles bases modificadas que cabría esperar en estas mitades de tRNA, si es que las mismas provienen de sus respectivos tRNA maduros con un patrón de modificación “estándar”. Por estándar se entiende el patrón de modificación descrito para cada especie de tRNA en la base de datos Modomics (<http://modomics.genesilico.pl/>). Lo más relevante es que ninguna de estas bases modificadas (2’-o-metiluridina [Um], N2-metilguanosina [m2G], pseudouridina [Ψ] y dihidrouridina [D]; Figura 37, A-D) afectan la cara Watson y Crick de las bases, por lo que las estructuras propuestas seguirían siendo posibles. Además, tampoco figuran entre las modificaciones que provocan un salto abrupto de las transcriptasas reversas (Cozen et al., 2015; Zheng et al., 2015), lo que es consistente con su eficiente secuenciación y análisis por SL-RT-qPCR.

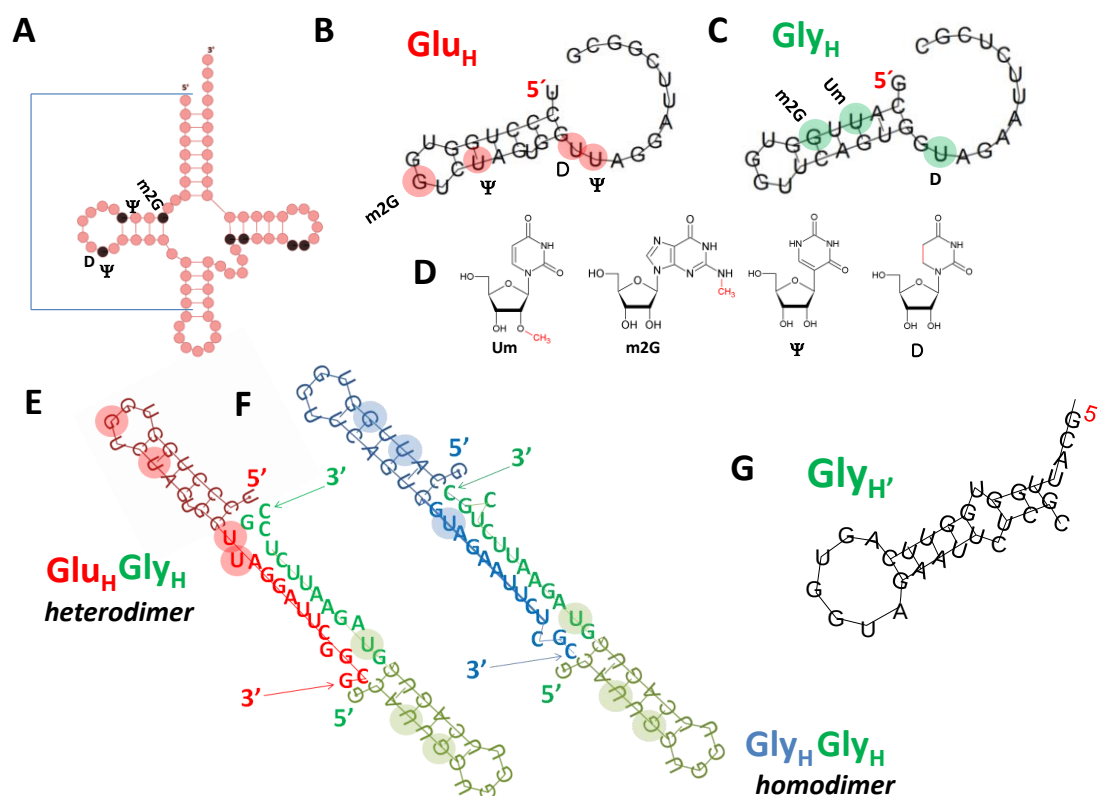


Figura 37: conformaciones e isoformas de las tRNA-h de Glu y de Gly. A) Esquema de un tRNA^{Glu}_{CUC} maduro, con algunas de las bases modificadas más frecuentes en la mitad 5'. Se indica con líneas azules el largo de los fragmentos hallados en la fracción S100 del medio condicionado de MCF-7. B-C) Predicción de estructura secundaria (mfold) de las mitades 5' de tRNA^{Glu}_{CUC} (B) y tRNA^{Gly}_{GCC} (C). Se muestran algunas de las bases que podrían estar modificadas, de acuerdo a lo anotado para el precursor maduro en la base de datos Modomics. D) Estructura de las bases modificadas presentadas en los paneles anteriores, con las modificaciones en rojo. E) Posibles heterodímeros entre mitades 5' de tRNA^{Glu}_{CUC} y tRNA^{Gly}_{GCC}. F) Posibles homodímeros entre mitades 5' de tRNA^{Gly}_{GCC}. G) Estructura secundaria alternativa de las mitades 5' de tRNA^{Gly}_{GCC}.

Al menos para las mitades 5' del tRNA-Gly, otra estructura secundaria es también posible (Figura 37, G). La misma implica una horquilla estabilizada por 8 apareamientos Watson y Crick (interrumpidos por un par UC), con un bucle más grande que también expone UGGU. Dado que en este caso la secuencia 3' desordenada que oficiaría de “interfaz de dimerización” se encuentra apareada con las bases del extremo 5', la estabilidad de esta estructura iría en detrimento de la formación de dímeros. Por lo que cabe hablar de una estructura secundaria pro-dimerización (que llamaremos Glu_H y Gly_H) y de una estructura secundaria anti-dimerización (Gly_{H'}).

3.3.6 Estabilidad de los picos P1 y P2 frente a distintos cationes monovalentes

Culminando el *impasse* teórico de la sección anterior, retomamos aquí el trabajo experimental. Los ácidos nucleicos son a pH fisiológico moléculas poli-aniónicas, por lo que

la formación de estructuras doble hebra enfrenta una considerable repulsión electrostática. La formación de dichas estructuras muchas veces requiere, por lo tanto, de la presencia de cationes que apantallen las cargas negativas presentes en el esqueleto azúcar-fosfato. Como los distintos cationes difieren en carga, radio y solvatación, el plegamiento del RNA en estructuras terciarias definidas es dependiente de la concentración y tipo de cationes presentes (Akat et al., 2014). En la Figura 31 ya se mostró cómo la fuerza iónica era capaz de repercutir en la estabilidad de los picos P1 y P2 de la fracción S100 previo o durante a su análisis cromatográfico.

En esta oportunidad se quiso evaluar el efecto del tipo de catión monovalente (litio, sodio y potasio) a fuerza iónica – casi – constante. Para ello se obtuvo la fracción S100 del medio condicionado de MCF-7, y se procedió a realizar un cambio de matriz (mediante ultrafiltración) a uno de los siguientes tres *buffers*: PB-140mM NaCl, PB-140mM KCl y PB-140mM LiCl. Por haber utilizado sales de sodio, debe considerarse que en cada ensayo hubo también Na^+ a concentración 11.8mM (151.8mM en el primer caso). Una vez alterada la matriz, las muestras fueron inyectadas y corridas utilizando el correspondiente *buffer* como fase móvil.

No se observaron grandes diferencias entre el uso de sales de sodio o de potasio, salvo un pequeño corrimiento de los picos P1 y P2 hacia tamaños menores en este último caso (Figura 38, recuadro izquierdo superior e izquierdo central). Sin embargo, cuando el catión principal fue el litio, se observó la casi total desaparición del pico P2 (Figura 38, recuadros inferiores). Esto quiere decir que, o el litio desestabiliza P2, o la estructura presente en P2 requiere de sodio o de potasio, pero no de litio. Esta observación es de momento una simple curiosidad, pero podría llegar a tener consecuencias funcionales importantes, como se discutirá posteriormente.

3.3.7 Dimerización de las mitades de $\text{tRNA}^{\text{Gly}}_{\text{GCC}}$ sintéticas

Se llegó al punto en el que, más que seguir acumulando evidencias indirectas respecto a la posible formación de dímeros de tRNA-h, lo imperioso era probar la formación de dichos dímeros en el sistema más simple e *in vitro* posible. Para ello se adquirió un oligonucleótido de RNA sin modificación de ningún tipo, de secuencia y largo idénticos a la de las mitades del $\text{tRNA}^{\text{Gly}}_{\text{GCC}}$. Se prefirió dicha especie para los ensayos iniciales por su capacidad potencial de formar homodímeros, haciendo el sistema más simple. El RNA liofilizado se resuspendió en agua libre de nucleasas a una concentración de 100 μM , alicuotándose hasta su uso a -80°C . Las diluciones de trabajo se realizaron en el *buffer* indicado en cada caso (típicamente PBS).

Lo primero a destacar es que, al diluir 1 nmol de RNA en 500 μL de PBS e inyectar (S75 10/300) y correr la SEC en PBS, el oligo fue capaz por sí mismo de generar un cromatograma de distribución bimodal, con dos picos en volúmenes muy aproximados a los de P1 y P2. Ya se sabía que el RNA aislado de los picos P1 y P2 de la fracción S100 reproducía lo observado con la fracción completa (Figura 35). Estos nuevos ensayos confirmaban que la información

necesaria para obtener una estructura capaz de eluir con un tamaño aparente de 44 kDa (P1) residía en la secuencia primaria de las tRNA-h, al menos para las mitades del tRNA-Gly.

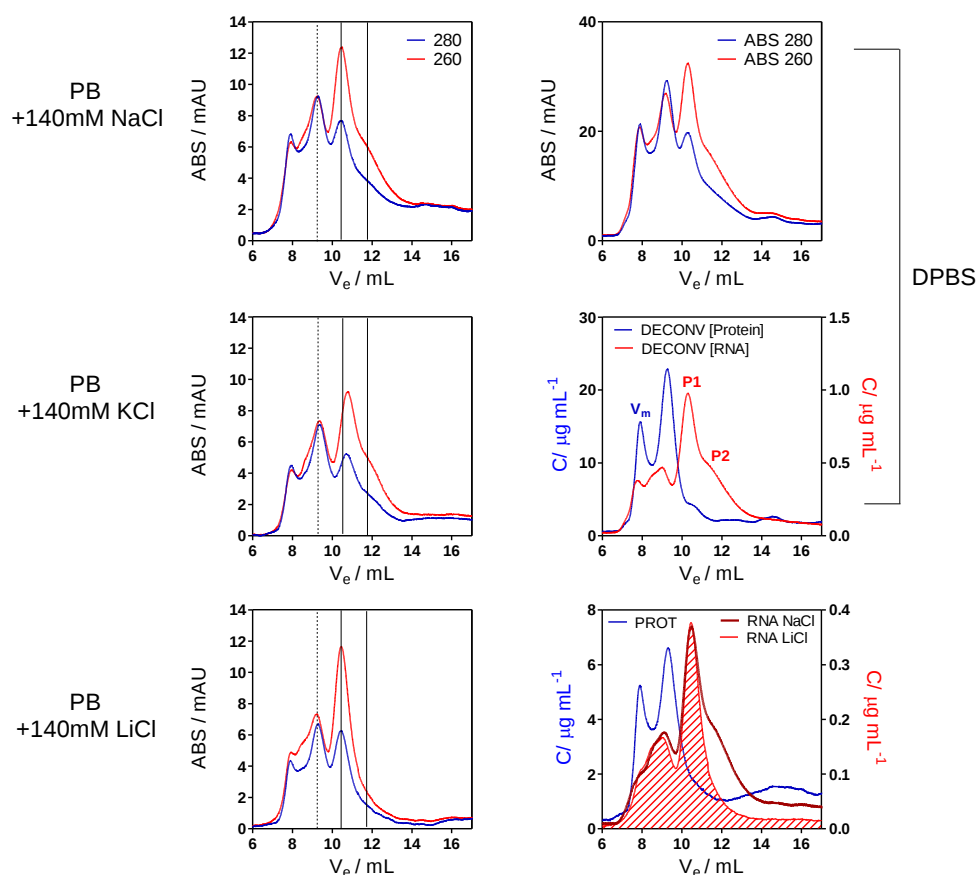


Figura 38: estabilidad de los picos P1 y P2 frente a distintos cationes monovalentes. Izquierda: purificación cromatográfica (S75 10/300) de la fracción S100 del medio condicionado de MCF-7 (medio MEGM), en buffer fosfato (PB; pH = 7.4) conteniendo 140 mM de NaCl (superior), KCl (medio) y LiCl (inferior). Derecha, recuadro superior: ídem, pero utilizando PBS como fase móvil (147mM Na⁺; 4.5mM K⁺). Derecha, recuadro central: deconvolución del cromatograma anterior. Derecha, recuadro inferior: deconvolución del cromatograma superior izquierdo (PB-140mM NaCl) y comparación con el cromatograma inferior izquierdo (PB-140mM LiCl; rayas rojas diagonales).

3.3.7.i Resistencia a RNasaA

Una de las cuestiones que más inquietaban era la resistencia observada del pico P2 a la acción de la RNasaA. Al incubar 1 nmol de oligo con 160 ng (0.015 KunitzU) de RNasaA por 30 min a 37°C (tratamiento similar al utilizado en la sección §3.3.1), se observó una degradación casi total de la muestra (Figura 39, curva negra). Alterando la cantidad de nucleasa agregada y los tiempos de incubación, se logró obtener una degradación parcial y controlada. Esto permitió evidenciar que, conforme la muestra se iba degradando, aparecía un nuevo pico a $V_e = 13$ mL. Dicho pico era de existencia transitoria, y desaparecía al aumentar el tiempo y la

cantidad de enzima agregada (Figura 39). A pesar de ser pico mayoritario (curva violeta), las tRNA-h no se lograban levantar allí por SL-RT-qPCR, mientras que seguían siendo pasibles de amplificación en los volúmenes de elución convencionales (Figura 39, panel inferior). Esto evidencia que el aumento de absorbancia a $V_e = 13$ mL corresponde a formas parcialmente escindidas, y no a un reacomodo estructural de la secuencia completa. Interpretamos la existencia de formas transitorias como una señal de que la acción de la RNasaA no era pareja ni cooperativa a lo largo de estos RNAs (lo que sería esperable de tratarse de un ovillo estadístico), sino que existirían regiones más resistentes a la degradación. Esto es compatible con las estructuras secundarias planteadas en la Figura 37, C, G. Como se verá más adelante – aunque no lo sabíamos en su momento – la concentración de RNA utilizada favorece la presencia del RNA en su forma monomérica más que dimérica (ver sección §3.3.8.iii), lo que se manifiesta en el volumen de elución del pico mayoritario (10.8 mL en vez de 10.5 mL). De cualquier forma, cabe señalar que la notable resistencia del pico P2 a la acción de la RNasaA no pudo ser reproducida aún cuando posteriormente se aumentó la concentración de partida (datos no mostrados). Resultados similares se obtuvieron luego de utilizar la endonucleasa de simple hebra RNase T1.

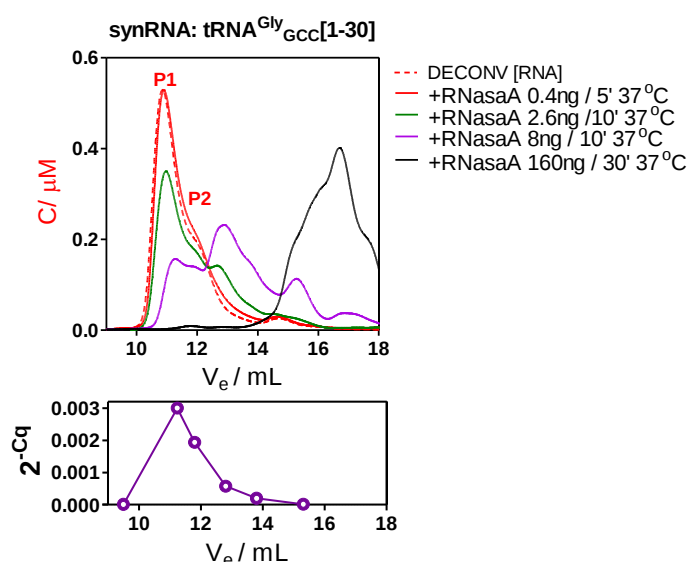


Figura 39: análisis por SEC (S75 10/300 en PBS) de 1 nmol del RNA sintético (Gly_{GCC} 1-30; volumen inyectado = 500 μ L) sin tratar (curva roja punteada) o tratada con RNasaA en distintas condiciones: 0.4 ng y 5 min (curva roja), 2.6 ng y 10 min (curva verde), 8 ng y 10 min (curva violeta) y 32 ng y 30 min (curva negra). Todas las incubaciones se realizaron a 37°C. Se muestra la deconvolución de los cromatogramas registrados a 280 nm y 260 nm, estimando la concentración del RNA. Esto es preferido al cromatograma 260 nm ya que en dicho caso parte de la absorbancia se debe al inhibidor de RNasa agregado para finalizar los tratamientos. Panel inferior: amplificación por SL-RT-qPCR, específica para las mitades 5' del $tRNA^{Gly}_{GCC}$ de 30 nt, de algunas fracciones colectadas correspondientes al cromatograma violeta del panel superior.

3.3.7.ii Estabilidad frente a distintos cationes monovalentes

Se repitió el ensayo realizado en la sección §3.3.6, pero no se observaron diferencias entre el uso de litio, sodio o potasio (resultados no mostrados). Sin embargo, el RNA disuelto en agua y corrido en agua como fase móvil se desplazó casi totalmente al volumen excluido (V_m), sin

remanentes en los volúmenes de elución correspondientes a los picos P1 y P2. Es esperable que el RNA pierda parte o la totalidad de su estructura secundaria en agua pura y que los posibles dímeros no se formen, pero la razón por la cual se convierte en una forma que se excluye no es al momento conocida.

3.3.7.iii Relación entre la concentración de RNA sintético y la presencia de un pico en 10.5 mL (P1'): equilibrios monómero / dímero.

Los oligoribonucleótidos (RNA) tienen un costo considerable, y el principio de austeridad lleva a utilizar las cantidades mínimas posibles. Sin embargo, al comenzar a inyectar cantidades mayores para ensayos que así lo requerían (DSC, sección §3.3.7.iv), se volvió evidente que el volumen de elución de los dos picos observados al analizar la sonda por SEC, así como la relación cuantitativa entre ellos, se modificaba en función del número de moles inyectados (y por ende de la concentración).

Con esto en vista, se inyectaron cantidades crecientes de RNA sintético (0.1, 0.2, 1, 2, 4, 12 y 20 nmol), siempre diluidas en 500 μ L de PBS, en una columna S75 10/300 (SEC), con PBS como fase móvil. Preferimos expresar los resultados en función del número de moles inyectados y no de la concentración, pues esta última varía a lo largo de la SEC y el grado de dilución final no es igual para todos los picos.

Cuando se inyectaron grandes cantidades de RNA sintético (20 nmol), los volúmenes de elución de los dos picos observados, así como la forma de ambos y, en menor medida, la relación entre ellos, coincidieron exactamente con los picos P1 y P2 obtenidos al purificar la fracción S100 del medio condicionado de MCF-7 (P1 = 10.5 mL; P2 = 12.0 mL). Para diferenciar cada caso, nos referiremos a dichos picos como P1' y P2' cuando los mismos correspondan a los oligos sintéticos. Cuando en vez de 20 se inyectaron 12nmol, el máximo volvió a coincidir con P1, pero la relación P2'/P1' aumentó (Figura 40). Al inyectar 4 nmol, también se evidenció un aumento de esta relación, pero comenzó a observarse un desplazamiento del máximo desde P1' hacia volúmenes de elución mayores. Esta tendencia fue sistemática, y cuando las cantidades inyectadas fueron inferiores a 1 nmol, el desplazamiento fue tan grande (V_e del pico más alto > 11 mL) que la absorbancia a $V_e = 10.5$ mL (P1') fue casi cero. Por el contrario, el "pico P2'" (pues no es realmente un pico) se mantuvo invariante en volumen (y al alza en intensidad normalizada). Para 20, 12, 4, 2, 1, 0.2 y 0.1 nmol inyectados, el máximo de absorbancia se ubicó en 10.50, 10.50, 10.61, 10.72, 10.88, 11.22 y 11.45. Para explicar este corrimiento debe considerarse que la absorbancia es aditiva, y que si en un cromatograma se obtienen dos picos no resueltos, entonces lo que se registra es la envolvente de ambos. Si un pico es mucho más grande que el otro, la envolvente tiene un máximo cercano a dicho pico. En cambio, llegado al caso que sean iguales en intensidad, existe una separación a la cual el máximo se sitúa exactamente entre medio. Pero aun considerando esto, lo mostrado en la Figura 40 no puede explicarse invocando a la existencia de dos picos únicos. En las inyecciones de 0.2 y 0.1 nmol (curvas punteadas), la absorbancia en P1' fue prácticamente nula. Si se tratara únicamente de dos picos, el máximo

debería situarse exactamente en P2', ya que el otro pico habría desaparecido. Y sin embargo, se observó un pico en 11.44 mL, con una inflexión en 12.0 mL (P2').

Nuestra interpretación de lo anterior es que existen al menos tres poblaciones: P1' (que atribuimos al dímero de tRNA-h), P2' (monómero, forma Gly_H), y una tercer población con V_e~11.3 mL (P2''). Consideramos que este tercer pico podría corresponder a la conformación alternativa de los monómeros presentada en la Figura 37, G (Gly_H'). Realizando un ajuste a dos curvas gaussianas del cromatograma obtenido al inyectar 0.1 nmol de oligo (Figura 40, B), determinamos la relación de concentración entre ambos picos ([P2'']/[P2'] = 0.8). Las concentraciones fueron estimadas a partir de la integral de las gaussianas, y considerando el coeficiente de absortividad molar a 260 nm predicho para el oligo (297700 M⁻¹cm⁻¹). La relación entre los picos P1' y P2' fue, como se ha dicho, variable con la cantidad de oligo inyectada. Vale aclarar que por la mala resolución entre los picos P2' y P2'', consideramos ambos como un solo pico al no poder realizar un ajuste a tres gaussianas en las concentraciones más altas (Figura 40, C).

Llegado a este punto, realizamos las siguientes racionalizaciones:

- i) Las conformaciones Gly_H y Gly_H' se hallan en equilibrio. Dado que dicho equilibrio sería del tipo $A \leftrightarrow A^*$, la constante de equilibrio viene dada por $[A^*]/[A]$. Dicha relación debería ser independiente de la concentración total. Estimamos el valor de dicha constante: $K_1 = 0.8 = [P2'']/[P2']$.
- ii) El dímero Gly_HGly_H se encuentra en equilibrio con el monómero en su conformación Gly_H. El equilibrio es en este caso del tipo $2A \leftrightarrow A_2$. La constante de equilibrio viene dada por: $K_2 = [A_2]/[A]^2$ y tiene unidades M⁻¹. Desde luego, K₂ es constante; no así la relación $[A_2]/[A]$ que depende de la concentración total (C_T). Por su parte, $C_T = 2[A_2] + [A] + [A']$.

Con las tres ecuaciones planteadas en los puntos anteriores podrían resolverse las incógnitas: [A₂], [A] y [A*]. El problema es que, si bien C_T es una variable conocida (sabemos cuánto inyectamos), y K₁ fue estimada empíricamente, no tenemos conocimiento del valor de K₂. No obstante, ajustando a los valores experimentales de [P1'] y ([P2'] + [P2'']), calculamos que $K_2 = 3 \times 10^6 \text{ M}^{-1}$ (Figura 40, B).

El buen ajuste observado entre los datos y el modelo teórico (aunque uno de los parámetros más importantes fue estimado de los datos) no valida que los picos P1' y P2' correspondan realmente a un equilibrio monómero/dímero. Pero al menos sabemos que el comportamiento observado es compatible con tal hipótesis. Un aspecto relevante, que surge del modelo, es la aparición de un valor de C_T en la cual ocurre un cruzamiento de las curvas de [A₂] y ([A] + [A']), y se sitúa aproximadamente en 1 μM. Por debajo de dicho punto, la concentración de monómero supera a la del dímero, mientras que por encima, la relación dímero/monómero crece con la C_T. Es por tanto esperable que al haber inyectado 0.2 nmol o menos (lo que daría una concentración a la salida de la columna de 0.1 μM) no se haya observado prácticamente ningún pico en P1'.

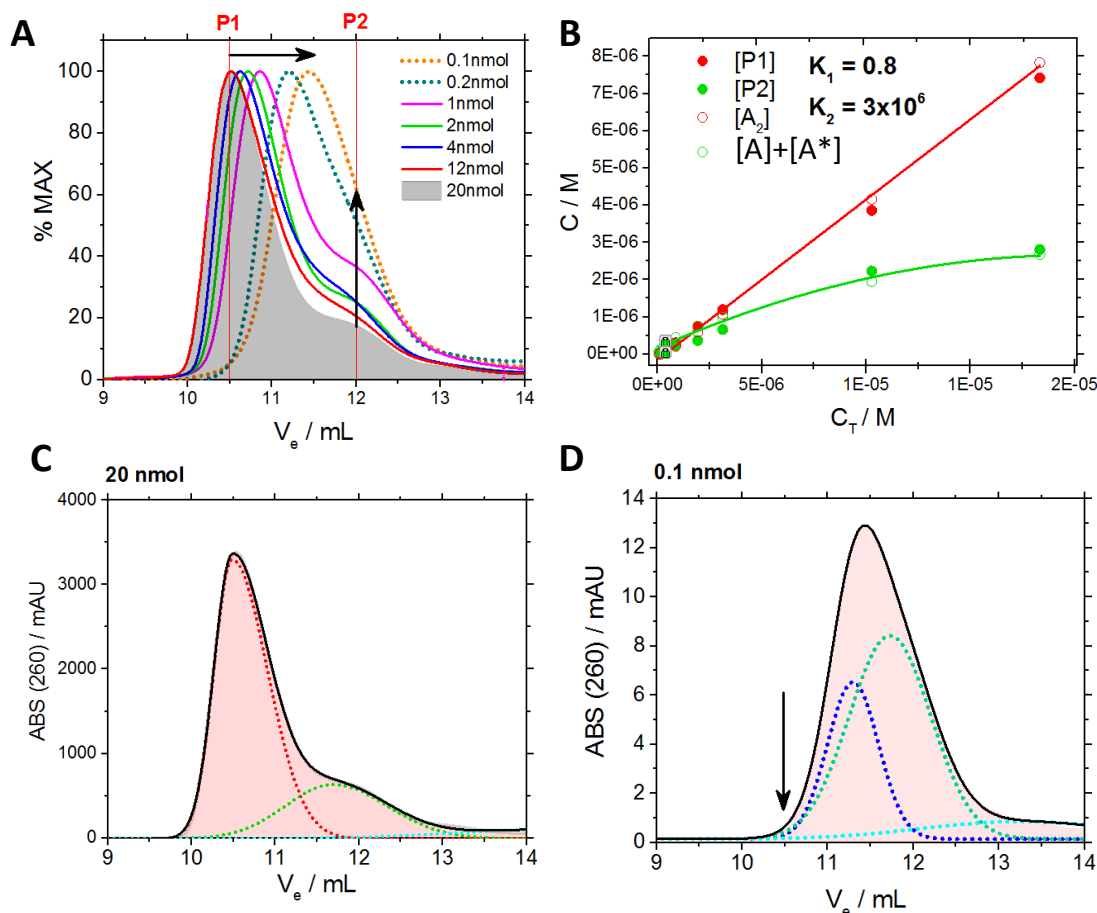


Figura 40: variación de los picos P1' y P2' obtenidos al inyectar el RNA sintético correspondiente a las tRNA-h de Gly, con la cantidad de RNA inyectado. A) Los cromatogramas se normalizaron al valor máximo (100%). El cromatograma (ABS 260 nm) correspondiente a la máxima inyección (20 nmol) se muestra en sombreado. Luego, en orden decreciente: rojo, azul, verde, lila, y las curvas punteadas verde y naranja. Las flechas muestran el desplazamiento del valor máximo hacia volúmenes de elución menores al diluir, así como el aumento del porcentaje de P2' respecto del pico máximo. B) Adecuación entre los valores de concentración en P2' (círculos verdes rellenos) y en P1' (círculos rojos rellenos; se consideró un 15% de hipocromatismo), ambos estimados de la gráfica A según se muestra en C y en D, y los valores teóricos de concentración de monómero (en sus dos estados $[A]$ y $[A^*]$) y dímero $[A_2]$, para un equilibrio monómero/dímero con $K = 3 \times 10^6 M^{-1}$, y para un valor de constante de equilibrio entre las dos conformaciones alternativas del monómero de $K = 0.8$ (círculos abiertos). Se grafica la concentración de cada especie (medida o predicha) en función de la concentración total (C_T) de RNA sintético. Las curvas roja y verde son ajustes a los círculos abiertos. C-D) Deconvolución de los cromatogramas correspondientes a la inyección de 20 nmol (C) y 0.1 nmol (D) en dos gaussianas. A partir de la integral (área) de dichas curvas, se estimaron los valores de $[P1']$ y $[P2']$ utilizados en (B).

3.3.7.iv Determinación de la temperatura de fusión (T_m) de los dímeros de tRNA-h mediante DSC y “UV-melting” (efecto hipercrómico)

Una propiedad de los ácidos nucleicos doble hebra es que su coeficiente de absorptividad molar en el ultravioleta decrece un 10-30% debido al aumento de rigidez de los anillos aromáticos de las bases nitrogenadas, fruto de las interacciones de *stacking* a lo largo de la doble hélice (hipocromatismo). Si se monitorea la absorbancia a 260 nm mientras se aumenta la temperatura de un dímero de DNA o RNA en solución, se observa un aumento más o

menos brusco de la absorbancia en torno a la temperatura de fusión (T_m)³⁸ del dímero. Dicho aumento permite evidenciar el pasaje de formas doble hebra, que absorben menos, a formas simple hebra, que absorben más.

Mediante el empleo de esta técnica (coloquialmente denominada “*UV-melting*”), determinamos la T_m del pico P1' en 46.5°C (Figura 41, izquierda). Un aspecto importante es que, por tratarse de un proceso bi-molecular, el valor de T_m depende de la concentración total de dímero. Dado que no obtuvimos un buen registro de la línea de base a bajas temperaturas, nuestros cálculos del valor de T_m no son suficientemente finos como para pretender establecer tal relación (a partir de la cual podrían calcularse los parámetros termodinámicos ΔH y ΔS mediante la ecuación de Van't Hoff³⁹). Lo que sí fue evidente fue el aumento en el porcentaje de hipercromatismo con el aumento de C_T (concentración total, independiente de la forma), pasando de un 0% a 0.5 μM (datos no mostrados) a un 18% a 4.33 μM . Dicho aumento se explica por el hecho de que el monómero, si bien suponemos tiene una estructura de doble hebra parcial, posee la mayoría de sus bases desapareadas. Cuanto mayor es la concentración total, mayor el porcentaje en forma dimérica, y por tanto mayor la diferencia entre la absorbancia inicial (donde casi todo el RNA es hipocromático) y la absorbancia a altas temperaturas. La T_m del monómero podría también determinarse (pues es esperable un cierto hipocromatismo dado por su estructura de doble hebra parcial), pero para evitar el dímero debería trabajarse a concentraciones bajas. Y a concentraciones bajas (= absorbancias bajas) los efectos provocados por la formación de burbujas conforme aumenta la temperatura impiden la detección de transiciones leves, al menos en las condiciones experimentales utilizadas (nótese cómo por debajo de 2 μM no se obtiene un *plateau* a altas temperaturas, debido precisamente a este fenómeno).

La hibridación de los ácidos nucleicos es además un proceso exotérmico. Recíprocamente, se requiere suministrar energía en forma de calor para separar un RNA de doble hebra. La calorimetría de escaneado diferencial (DSC) mide cuánta energía debe entregarse a una muestra (en comparación con una referencia que contiene únicamente *buffer*) para lograr aumentar la temperatura según una rampa definida (ej: 1°C/min). Al principio toda la energía entregada se utiliza para aumentar la temperatura del solvente. Pero cuando se alcanzan temperaturas cercanas al valor de T_m , la energía entregada se utiliza en la rotura de enlaces de hidrógeno, por lo que debe entregarse más energía que antes para provocar el aumento deseado de temperatura. La capacidad calorífica específica a presión constante (C_p) para un dímero de ácidos nucleicos se hace máxima cuando la temperatura iguala a su T_m , por lo que la técnica de DSC es ideal para estos fines.

El análisis por DSC del pico P1' purificado por SEC arrojó un valor de $T_m = 47.8^\circ C$, similar al obtenido mediante *UV-melting* (Figura 41, derecha). La diferencia de poco más de 1°C se explica por cuestiones de *fitting* de los datos de ambas técnicas, pero también es esperable dado que se utilizaron concentraciones diferentes en cada caso. Otra ventaja de la DSC es que, por tratarse de una técnica calorimétrica, arroja medidas directas de ΔH . Sin entrar en

³⁸ T_m : temperatura a la cual el 50% del DNA o RNA se encuentra en su forma doble hebra, y el restante 50% en forma de hebra simple. O dicho de otro modo, la temperatura a la cual la constante de equilibrio $2A \rightleftharpoons A_2$ se hace igual a uno.

³⁹ $\ln K = -\Delta H/RT + \Delta S/R$ (con T en grados Kelvin)

cuestiones formales que no vienen al caso, vale decir que del área debajo de la curva de C_p vs T se obtiene la medición del ΔH calorimétrico (ΔH_{cal}). Dicho valor depende de la concentración de RNA utilizada, ya que para el cálculo de C_p se divide la capacidad calorífica (que es la medida directa) entre el número de moles de sustancia. Por otro lado, de la forma de la gráfica (y no de su área) se obtiene el parámetro de ΔH de Van't Hoff (ΔH_{vh}), el cual surge de la modelización a un sistema de dos estados (nativo $\rightarrow$ desnaturalizado; ó dímero $\rightarrow$ monómero). Como no depende de los valores de C_p concretos sino de su distribución, el ΔH_{vh} es independiente de la concentración (en nuestro caso, $\Delta H_{vh} = 68$ Kcal/mol). La relación entre ambos parámetros ($\Delta H_{cal}/\Delta H_{vh}$) puede entenderse como el número de unidades de disociación cooperativa por mol de sustancia⁴⁰. Para el caso de un modelo simple de dos estados (monómero/dímero), cada mol de monómero corresponde a 0.5 unidades de disociación. Por otro lado, si lo miramos desde el dímero, cada mol es una unidad de disociación, por lo que el $\Delta H_{cal}/\Delta H_{vh}$ debería valer 1. En nuestro caso, obtuvimos relaciones muy cercanas a 1 al calcular el C_p utilizando la concentración de dímero, y a 0.5 al utilizar las de monómero. Esto valida que la transición que estamos monitoreando es la rotura de los dímeros, y no la desnaturalización de los monómeros parcialmente estructurados. Lo cual es esperable pues a la concentración de trabajo (6 μ M) la mayor parte del RNA debería estar en forma dimérica. Es válido el señalamiento de que nuestro sistema es en realidad un sistema de tres estados: dímero $\rightarrow$ monómero estructurado $\rightarrow$ monómero abierto. De la buena adecuación a un modelo de dos estados se desprende que la fisión del dímero ocurre de manera cooperativa: rotos los apareamientos en la interfaz de dimerización, se desestabiliza la formación de la estructura tallo-bucle de cada monómero, y viceversa.

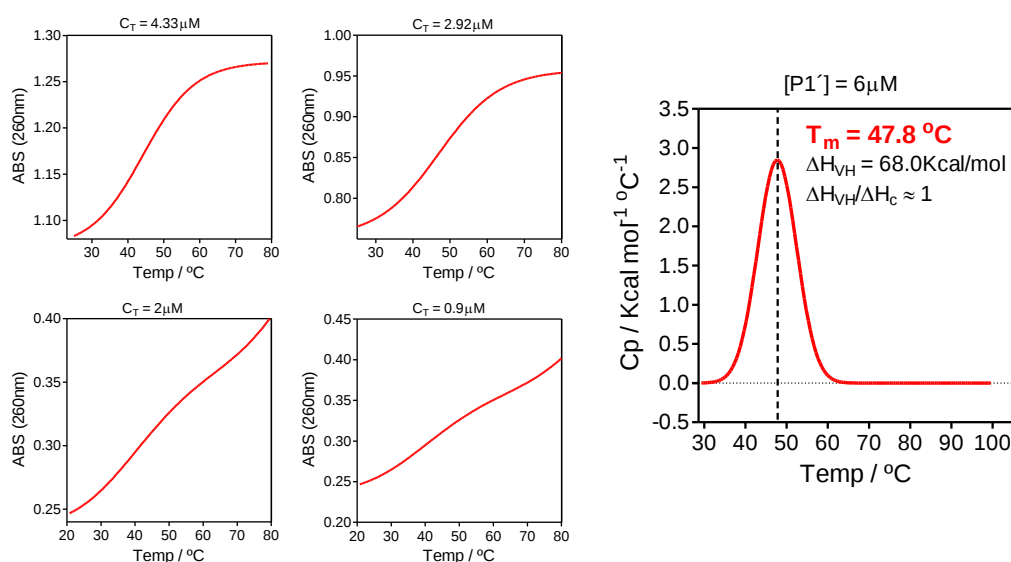


Figura 41: determinación de la T_m de los dímeros de tRNA-h por hipercromatismo en el UV (abs 260 nm); paneles izquierdos, y por calorimetría de escaneado diferencial (derecha). Las T_m calculadas fueron 46.5°C y 47.8°C, respectivamente.

⁴⁰<http://ctrstbio.org.uic.edu/manuals/dscOrigin.pdf>

Importante señalar que el ensayo de DSC se repitió en dos oportunidades con muestras independientes, obteniendo valores muy similares de T_m , ΔH_{vh} y $\Delta H_{cal}/\Delta H_{vh}$. Los valores de T_m y ΔH_{cal} tampoco se alteraron demasiado al realizar tres ciclos de calentamiento y enfriamiento ($20^\circ\text{C} \rightarrow 100^\circ\text{C}$), lo que habla de la reversibilidad del proceso.

La T_m del monómero no se pudo calcular debido a que el DSC, por su sensibilidad, requiere una cantidad importante de muestra. Y una cantidad importante de muestra, en un volumen definido, implica concentraciones grandes. Y por ende formación del dímero. Sería deseable contar en el futuro con un RNA sintético “mutado” en la interfaz de dimerización, de modo tal que pueda formar las estructuras de tallo-bucle pero no dimerizar a altas concentraciones.

3.3.7.v Diferencias entre las mitades de tRNA endógenas y los RNA sintéticos, evaluadas mediante electroforesis nativa

El empleo de las mitades de tRNA-Gly sintéticas permitió evidenciar que la formación de homodímeros es razonable, termodinámica y cinéticamente favorable, y posible aún en ausencia de bases modificadas o de proteínas. No se logró, no obstante, reproducir la resistencia del pico P2 a la degradación mediada por RNasaA, ni provocar la desaparición casi total del pico P2 mediante el simple pasaje de cationes sodio a litio en el solvente. Otra diferencia surge de la purificación de los picos P1 y P2 (endógenos) y de los picos P1' y P2' (RNA sintético), y la corrida electroforética del RNA en condiciones nativas. El RNA sintético migró en los geles nativos de poliacrilamida como una banda de exactamente 30 nucleótidos (estimada a partir de los R_f del marcador de tamaños), independientemente de que proviniese del pico P1' o P2'. Sin embargo, en P2' se advirtió una segunda banda de menor tamaño, que podría corresponder a un producto de degradación parcial o a una isoforma más estructurada y compacta. Curiosamente, dicha banda fue casi inexistente en P1'. Contrario a lo observado con el RNA sintético, donde la formación de dímeros en P1' no pudo reproducirse en las condiciones electroforéticas, el RNA de los picos P1 y P2 purificados corrió, respectivamente, por encima y por debajo de la banda de 30 nucleótidos observada con los ribonucleótidos sintéticos (Figura 42). La “banda” principal en P2 se asemejó a la banda más baja en P2'.

El RNA de P1 y P2 fue también analizado mediante el Agilent Bioanalyzer, utilizando un chip de pequeños RNAs. Dicho instrumento ofrece resultados comparables al de los geles nativos, pero con mucha mayor resolución. Los electroferogramas muestran dos picos principales en P1, de 30 y 60 nucleótidos, que bien podrían corresponder a un equilibrio monómero (30) – dímero (60). En P2, el pico de 60 nucleótidos fue inexistente, y la mayor parte de la población migró con un tamaño aparente menor a 30 nt.

Visto en su conjunto, estos datos refuerzan la posibilidad de que el RNA en P1 se encuentra en forma dimérica, no quedando claro si P2 correspondería al simple monómero, o a una forma más compacta del RNA, como si discutirá posteriormente. Con el RNA sintético se pudo constatar la dimerización en solución, pero dichos dímeros no resistirían las condiciones electroforéticas utilizadas, ni serían suficientes para explicar la gran resistencia a la

degradación observada en las mitades de tRNA extracelulares endógenas. Está claro que queda aún mucho trabajo por delante.

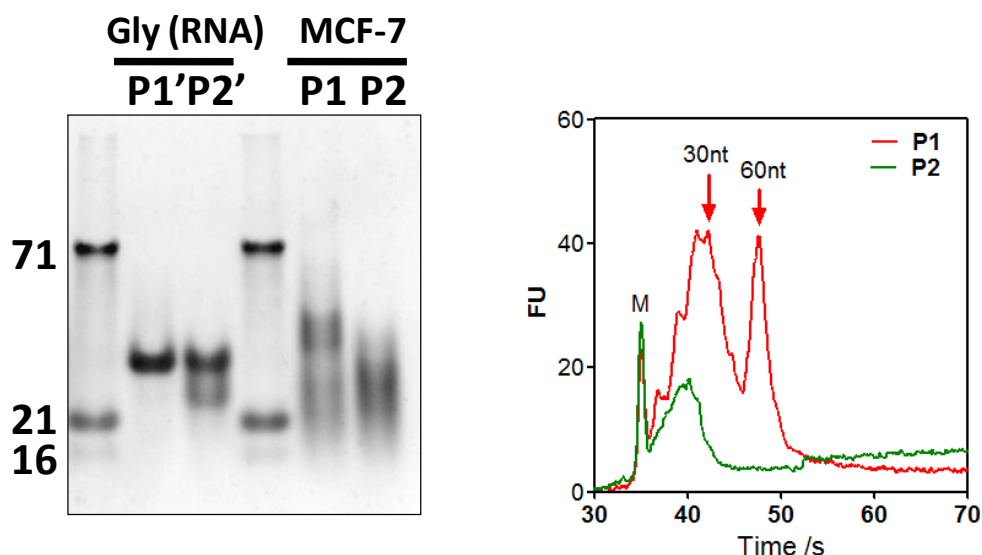


Figura 42: análisis de los picos P1 - P2 (fracción S100 de MCF-7) y P1' - P2' (obtenidos a partir del RNA sintético que mimetiza las mitades de tRNA^{Gly}_{GCC}) por electroforesis nativa en gel de poliacrilamida (20%; izquierda) y mediante el equipo Bioanalyzer (Agilent) utilizando un chip para pequeños RNAs (resolución: 10-800 nt). Las fracciones colectadas por SEC (S75 10/300) fueron sometidas a una extracción del RNA con Trizol previo a su análisis, salvo para los picos P1' y P2' donde la corrida electroforética fue hecha directamente sobre las fracciones colectadas.

3.4 DISCUSIÓN

A diferencia de los capítulos anteriores, esta parte del trabajo permanece abierta, por lo que es prematuro establecer conclusiones definitivas. Aún así, se ha llegado a un grado de desarrollo que permite afirmar algunos conceptos, discutir posibilidades, y elaborar hipótesis. Además, el capítulo de resultados fue presentado en orden cronológico. Entendimos que era la mejor forma de comprender la lógica o la motivación detrás de los experimentos realizados. Se impone, en este momento, trascender la barrera temporal y reinterpretar los resultados con una visión globalizadora.

Cuando comenzamos a trabajar en este capítulo, estábamos convencidos de que las mitades de tRNA halladas en la fracción S100 de varias líneas celulares (y en sangre humana por otros autores) debían ser transportadas y protegidas por proteínas. Sabíamos que podían existir por fuera de las vesículas extracelulares (la prueba definitiva era su no exclusión por SEC), y eso no podía querer decir otra cosa que en complejos ribonucleoproteicos. Así que apostamos a nuestra capacidad de identificar rápidamente la proteína o proteínas asociadas (genéricamente: ETBP)

Dado que las mitades de tRNA tienen un peso molecular menor a 10 kDa (aún cuando se acepten varias modificaciones post-transcripcionales), atribuimos el pico P1 producto de la SEC – cuyo tamaño aparente era 44 kDa – a los supuestos complejos con proteínas. Esto nos llevó a la idea de que la ETBP y los tRNA-h co-purificaban unidas en dicho pico, y que debíamos encontrar la manera de separarlos (lo que provocaría un desplazamiento de ETBP de P1). Allí es donde entró el tratamiento con RNasaA, efectivo para la desaparición de dicho pico.

Lo que no considerábamos, y no lo supimos hasta mucho tiempo después, es que – de existir ETBP – los complejos ribonucleoproteicos se estaban desarmando durante la preparación y análisis de la muestra, probablemente durante la propia cromatografía de exclusión molecular. El análisis con oligodesoxinucleótidos complementarios entre sí permitió evidenciar que tamaños aparentes de 44 kDa podían obtenerse mediante la formación de un dúplex, aún con secuencias cortas. Y finalmente, la inyección de RNA sintético despejó toda duda. Ahora resulta claro que deberíamos haber entrecruzado covalentemente (“*cross-linking*”) los complejos, previo a su preparación y análisis. Pero esto parecía innecesario dado que creíamos que estábamos obteniendo las RNP sin tal recurso. Si aceptamos que la presunta unión ETBP/tRNA-h se estaba rompiendo durante la SEC, ya no habría motivos para suponer que ETBP se hallara en P1. Desde esta óptica, nuestros intentos de “desenmascarar” ETBP como aquella proteína presente en P1 y afectable por RNasa... no tendría sentido. Y efectivamente, no observamos cambios significativos.

Volviendo a la lógica con la que nos manejábamos a la hora de diseñar los experimentos, creímos que los “complejos” podrían retenerse en una cromatografía de intercambio aniónico (AEC), a causa de la carga global negativa conferida por el RNA. Hoy probablemente rechazaríamos la sola idea de realizar este tipo de ensayos, por carencia de fundamento. Si el RNA ya se encuentra dissociado de ETBP a la salida de la SEC, es claro que lo que estaríamos pegando a la columna AEC sería el RNA “desnudo” y no el complejo con proteínas. Ergo, pretender identificar ETBP en la fracción retenida de la AEC era una empresa condenada al fracaso. O no; a veces la ignorancia puede ser más productiva que un parcial conocimiento. En la fracción retenida (P1* y P2*) detectamos con alta confianza algunas enzimas glucolíticas: ENO1, TPI1 y PGK1. La presencia de las mismas, como se dijo, podría explicarse por ser las más abundantes en la muestra, pero el alto punto isoeléctrico de PGK1 impone la invocación de consideraciones adicionales. Al margen de todo esto, los resultados más interesantes se obtuvieron al analizar la fracción no retenida (FT). El modesto efecto producido por la acción de la RNasaA sobre ENO1 fue, al incluir el pico P2 de la SEC, mucho más marcado. Aunque estos ensayos deberían repetirse, la presunta relación entre ENO1 y las tRNA-h parecería depender de la conformación que el RNA adopta en P2.

Pero ¿Cómo explicar que ENO1 casi desaparezca del FT luego del tratamiento con RNasaA? Si el efecto hubiese sido el contrario, tendríamos una explicación mucho más sencilla. Lo cierto es que no podemos responder aún esta pregunta; solo evaluar algunas hipótesis. La disminución de la abundancia relativa de una proteína en el FT puede ser solo consecuencia de un mayor pegado de la misma a la columna. Por tanto, ENO1 debería unirse más luego de tratar con nucleasa. Podría pensarse que el RNA “desnudo” compite con las proteínas de

carga neta negativa por los sitios de unión, y que degradando el RNA las proteínas se unirían con mayor fuerza. Pero de ser así, el efecto debería observarse sobre cualquier proteína negativa. La TPI1 es otra proteína abundante en el pico P1, pero tiene un punto isoeléctrico bastante más bajo que el de ENO1 (5.7 vs 7.0). Cabría esperar que la misma fuese también retenida luego del tratamiento con RNasa, lo que no se observó. El efecto parecería ser, por tanto, específico sobre la α -Enolasa. O que esta tuviese una carga neta mucho más negativa de lo predicho a partir de su punto isoeléctrico. Y esto último podría explicarse a partir de su unión al RNA. En un complejo que tendría que ser resistente a la acción de la RNasaA. Lo que podría explicar por qué el efecto es dependiente de la inclusión del pico P2. Más allá de lo que realmente esté ocurriendo, la vinculación entre las mitades de tRNA y la enzima glucolítica ENO1 es para nosotros sumamente interesante. Sobre todo si se considera que dicha enzima ya ha sido relacionada al transporte de tRNAs en levaduras (Entelis et al., 2006).

Lejos estamos de probar una interacción entre las mitades de tRNA y enzimas de la glucólisis, pero hay evidencia parcial como para justificar un redoble en este esfuerzo. La mayoría de las enzimas que participan en el metabolismo de los carbohidratos presentan el fenómeno denominado “luz de luna” (*protein moonlighting*), que refiere a la capacidad de muchas proteínas de realizar funciones dentro – y fuera – de la célula completamente distintas a aquellas por las cuales reciben su nombre (Huberts and van der Klei, 2010). Y muchas de estas funciones alternativas están relacionadas con la unión de ácidos nucleicos (Castello et al., 2015), lo cual no es del todo sorprendente considerando que la mayoría de ellas reconocen nucleótidos (ATP, ADP) o sustratos fosforilados o fosforilables.

La unión de las mitades de tRNA a ENO1 u otras enzimas glucolíticas (hemos detectado varias en la fracción S100 de MCF-7) podría no ser una interacción fuerte o específica, con un reconocimiento directo de un motivo de secuencia particular, sino una interacción lábil a partir del reconocimiento de una estructura tridimensional determinada. Esto explicaría que los complejos se desarmen durante la SEC (se espera ansiosamente realizar ensayos de entrecruzamiento en el futuro inmediato). Más que una visión clásica, “proteín-centrista” de la biología, podría pensarse en el RNA como una plataforma versátil para el ensamblaje de proteínas. En lugar de pensar que la proteína pliega al RNA para que adopte una conformación determinada, podría pensarse en un RNA naturalmente estructurado, pero de conformación dinámica. Y que en función de este dinamismo estructural se unan o se liberen algunas proteínas.

Habría advertido el lector la sutileza de intentar, en el párrafo anterior, conciliar entre sí las dos mitades inconclusas que conforman el capítulo III de esta tesis. Y sin embargo, no seríamos los primeros en sugerir un rol del RNA como plataforma de anclaje de enzimas glucolíticas. Evidencia estructural reciente (Wu and Minter, 2015) apoya el concepto del “metabolón” propuesto en la década de los 80 para las enzimas del ciclo de Krebs (KC) (Robinson et al., 1987). En vez de actuar separadas y dispersas en la mitocondria, algunas enzimas del KC formarían una gran estructura supramolecular a través de la cual irían pasando los sustratos de cada enzima (productos de la enzima anterior, y adyacente), de un modo análogo a como varias personas, de mano en mano, van bajando los ladrillos de un

camión. Existe evidencia de que algunas enzimas glucolíticas podrían también formar metabolones, en cuya formación el RNA podría estar involucrado. Lo siguiente es traducción textual de las primeras dos frases del resumen de un artículo publicado en 1996: “*El isoelectroenfoco de extractos de células MCF-7 reveló una asociación entre las enzimas glucolíticas GAPDH, PGK1, ENO1 y PK (Piruvato Quinasa). Este complejo entre las enzimas glucolíticas es sensible a RNasa*” (Mazurek et al., 1996). Es decir que dos de las tres proteínas que hallamos retenidas en la AEC podrían formar un complejo supramolecular estabilizado por RNA. Como solemos tener una tendencia a creer más real las cosas que algún artículo plasma en un esquema, se adjunta la Figura 4 de una publicación reciente sobre la capacidad de unión al RNA de las enzimas metabólicas (Castello et al. 2015) (Figura 43).

Aquí enganchamos con la cuestión sobre la posible dimerización de las tRNA-h. Ya sea en el caso de los posibles homodímeros Gly-Gly, o de los supuestos heterodímeros Glu-Gly, las estructuras propuestas son total o casi totalmente simétricas. Dos proteínas podrían reconocer, por ejemplo, la estructura de tallo-bucle (UGGU) de los modelos Glu_H y Gly_H (Figura 37), y de ese modo estar sometidas a una dimerización mediada por mitades de tRNA. Estamos ingresando aquí en terreno demasiado especulativo, pero la posibilidad es sumamente interesante. Más aun considerando que muchas de las enzimas glucolíticas necesitan dimerizar o multimerizar para una catálisis efectiva (ENO1, por ejemplo), y las tRNA-h podrían interferir o favorecer dichos procesos.

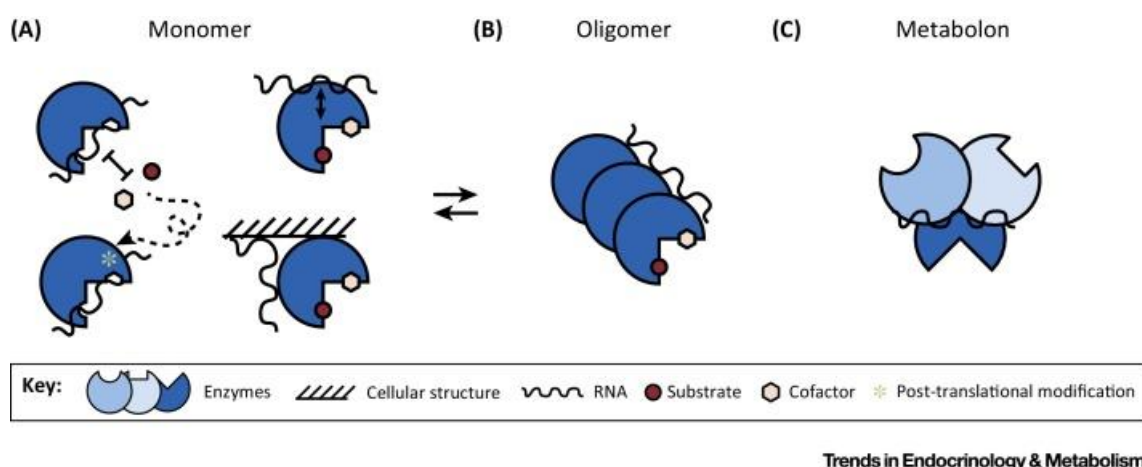


Figura 43: figura y pie de figura extraídos sin modificaciones (salvo por la traducción) de Castello et al. (2015): ¿Será que el RNA es regulado por, o es el regulador de las enzimas metabólicas? Distintos modos de interacción entre el RNA y las enzimas pueden ser previstos. A) El sitio de unión al RNA se solapa con el sitio catalítico o con el bolsillo de unión a ligando (arriba, izquierda), o es distinto de aquel/aquellos (arriba, derecha). La unión al RNA puede ser regulada por modificaciones post-traduccionales mediadas por los metabolitos (abajo, izquierda). El RNA podría también regular la localización de la enzima, por ejemplo, uniéndola a una estructura celular tal como la membrana (abajo, derecha). El RNA puede también servir como plataforma de anclaje para la formación de oligómeros (B), o de complejos multi-enzimáticos en metabolones (C).

Supongamos que en un tiempo futuro no tan lejano llegásemos a probar, fehacientemente, la vinculación de las tRNA-h con una o varias enzimas glucolíticas, o la presencia de un

metabolón ensamblado en torno a dímeros de mitades de tRNA ¿Se tratarían de complejos realmente extracelulares? Aquí podrían entrar en juego algunas de las funciones “luz de luna” de estas enzimas. Pero aun cuando la formación de dichos complejos en el medio extracelular fuera un artefacto ocasionado por la secreción exacerbada de enzimas glucolíticas altamente expresadas por las células tumorales debido al “efecto Warburg”⁴¹, la vinculación entre las tRNA-h extracelulares y el metabolismo energético no dejaría de ser una cuestión interesante.

Volviendo al RNA, y a los datos experimentales: ¿hasta qué punto logramos reproducir con el oligo sintético lo observado en los medios de cultivo condicionados? La formación de dímeros en el pico P1 parecería ser lo más claro. Pero la resistencia a la RNasaA observada en el pico P2 no pudo reproducirse ¿Será que esta reside en la unión de las tRNA-h a proteínas? Es posible, sobre todo por la observación (datos no mostrados) de que el pico P2 purificado de los cultivos celulares se vuelve más sensible a RNasaA cuando el tratamiento se realiza con posterioridad a la SEC. Pero hay otros factores que podrían explicar esta diferencia. En primer lugar, nuestros ensayos experimentales permitieron formar, en el mejor de los casos, dímeros Gly-Gly. Y en el medio extracelular sería previsible la formación de heterodímeros Glu-Gly. Luego está la cuestión de las bases modificadas, que pueden en esto alterar todos los parámetros. Nuestro RNA sintético carece de las mismas.

Con el oligo de RNA, los picos P1' y P2' (haciendo la salvedad de que P2' podría en realidad tratarse de dos picos no resueltos) corresponderían a un par monómero/dímero en equilibrio. Para una predominancia del dímero fueron necesarias concentraciones que superasen los 500 nM, lo que es demasiado alto como para pensar que pudiera esto darse *in vivo*, y especialmente en el medio extracelular. Sin embargo, debe hacerse otra vez la salvedad de que la ausencia de bases modificadas, de las mitades del tRNA^{Glu} y de proteínas que pudieran estabilizar una conformación particular, inhabilitan conclusiones al respecto. Basta ver en los geles nativos donde los picos obtenidos a partir de la purificación cromatográfica del oligo de RNA y del medio condicionado se comportaron completamente distintos. Mientras que en los primeros no se logró mantener un estado de oligomerización en las condiciones electroforéticas, los picos P1 y P2 purificados del medio condicionado sí presentaron un corrimiento por encima y por debajo del tamaño esperado para las secuencias lineales (Figura 42). Por otro lado, mientras que bastaba diluir al oligo de RNA para perder casi totalmente la absorbancia a 260 nm a $V_e = 10.5$ mL (P1'), el pico P1 del medio de cultivo se mantenía como tal, aun cuando se inyectaran a la columna cantidades muy bajas (< 3 mAU).

Lo anterior habla de que, si la constante de dimerización aparente para la sonda sintética fue del orden de $3 \times 10^6 M^{-1}$, la constante aparente con las tRNA-h “biológicas” sería presumiblemente más grande. Pero también es posible que los picos P2 obtenidos con el oligo, y aquellos obtenidos al purificar el medio de cultivo condicionado, no correspondan a las mismas estructuras. Utilizando RNA sintético, el mismo puede estar casi totalmente en la

⁴¹ Efecto observado en la mayoría de las células tumorales, por el cual estas alteran su metabolismo energético y promueven la formación de ATP mediante la glucólisis más que mediante fosforilación oxidativa, aún en condiciones aerobias: Cairns, R.A., Harris, I.S., and Mak, T.W. (2011). Regulation of cancer cell metabolism. *Nat Rev Cancer* 11, 85-95.

forma que suponemos monomérica (P2'). Basta para eso con diluir. Pero nunca podemos tener, por muy alta que sea la concentración, el RNA en forma totalmente dimérica (P1'). La propia ecuación de equilibrio así lo exige. Ambos picos son interdependientes, de modo tal que no puede afectarse uno sin que se afecte el otro. Esto no ocurre, sin embargo, con los picos P1 y P2 del medio condicionado. Se recordará cómo la corrida en PB-LiCl (en lugar de PB-NaCl o PB-KCl) era suficiente para la casi completa desaparición del pico P2, sin afectación del pico P1. El tratamiento con RNasaA provocaba el efecto contrario. Es posible, por tanto, que P2' sea la fracción remanente de monómero en equilibrio con el dímero. Y que P2 sea... alguna otra cosa.

Un reporte reciente mostró como las mitades de tRNA de Ala y de Cys (vinculadas con la inhibición de la traducción vía interacción con la proteína YB-1) adquirirían una conformación particular dentro de la célula, que implicaba la formación de una estructura conocida como “G-cuádruplex” (Ivanov et al., 2014). Los autores mostraron evidencia a favor de la existencia de G-cuádruplex (G4plex) en las mitades de tRNA, pero por lo demás, el reporte publicado en PNAS no ahondó en cuestiones estructurales. Las tRNA-h de Ala y Cys humanas poseen un tracto de varias guaninas consecutivas en su extremo 5', pero esto no es de por sí una cuestión que favorezca la formación de G4plex intramoleculares, para lo cual se requieren cuatro tándems separados de dos o más guaninas. En el mejor de los casos, las tRNA-h de Ala y Cys podrían formar una sola tétrada de guaninas (un G4plex de una sola capa), que es lo que los autores plantearon. Sin embargo, los geles nativos presentados en una de las figuras suplementarias, utilizando sondas sintéticas, muestran claramente la presencia de agregados moleculares, posiblemente dímeros (Figura 44, A-C).

Dado que las estructuras diméricas propuestas en nuestro trabajo exponen bucles de secuencia UGGU, adyacentes además a otros pares GG en cada extremo del dímero, consideramos la posibilidad de que los dímeros se plegaran sobre sí mismos, y que la estructura se cerrara mediante la formación de un G4plex de dos capas intra-dímero e inter-monómero (Figura 44, D). Para la formación de estas estructuras deberían “abrirse” los pares GG/UC mostrados dentro de los rectángulos rojos, lo que conllevaría la pérdida de 10 enlaces de hidrógeno por apareamientos Watson y Crick. Esto se vería compensado por la ganancia de 16 enlaces de hidrógeno conferidos por la formación del G4plex de dos capas (interacciones GG del tipo Watson-Crick/Hoogsteen), así como las interacciones por *stacking* de las cuatro guaninas montadas sobre otras cuatro. Dado que se parte ya de un dímero, no habría por qué considerar grandes impedimentos entrópicos.

¿Por qué nuestro interés en considerar estas estructuras? En primer lugar, porque estamos procurando elucidar las conformaciones más estables que puedan adquirir las tRNA-h; con la esperanza de que sean también las conformaciones biológicamente relevantes, ya sea en el medio extracelular o dentro de la célula. Pero además, la formación de una estructura compacta y estabilizada por G4plex como la mostrada en la Figura 44, D podría explicar la gran resistencia a la RNasaA observada en el pico P2 de la fracción S100 del medio condicionado de MCF-7. Los dímeros, que racionalizamos como la estructura presente en P1, presentan bucles simple hebra a través de los cuales la RNasaA podría comenzar a degradar (un corte bastaría para desensamblarlas estructuras, a 37°C). Pero como dichos bucles

conformarían parte del G4plex en la estructura compacta, ya no habría endonucleasa (ni exonucleasa) de simple hebra capaz de atacar dicha molécula, si existiera. Es sabido que los ácidos nucleicos con G4plex presentan una incrementada resistencia a la degradación por nucleasas (Cao et al., 2006) y parecerían tener una facilidad especial para ingresar a las células (Ivanov et al. 2014).

Fue con todo esto en mente que realizamos los ensayos de estabilidad de los picos P1 y P2 frente a distintos cationes monovalentes: sodio, potasio y litio. Para la estabilización de los G4plex es fundamental la presencia de cationes del grupo I ($K^+ > Na^+ \gg Li^+$) ubicados en el espacio entre los planos y en el eje de simetría de los mismos, capaces de establecer interacciones ión-dipolo con los grupos ceto de las guaninas (Risitano and Fox, 2004). En concordancia con lo anterior, observamos una desaparición casi total del pico P2 al trabajar con un *buffer* sin potasio, bajo en sodio, y rico en litio (para mantener la fuerza iónica). Para nuestra desazón, este resultado no se pudo reproducir con el oligo de RNA sintético. Por el contrario, en estos casos, el pico P2' parecía corresponder más bien al monómero del RNA, con su estructura de doble hebra parcial (Gly_H y Gly_H').

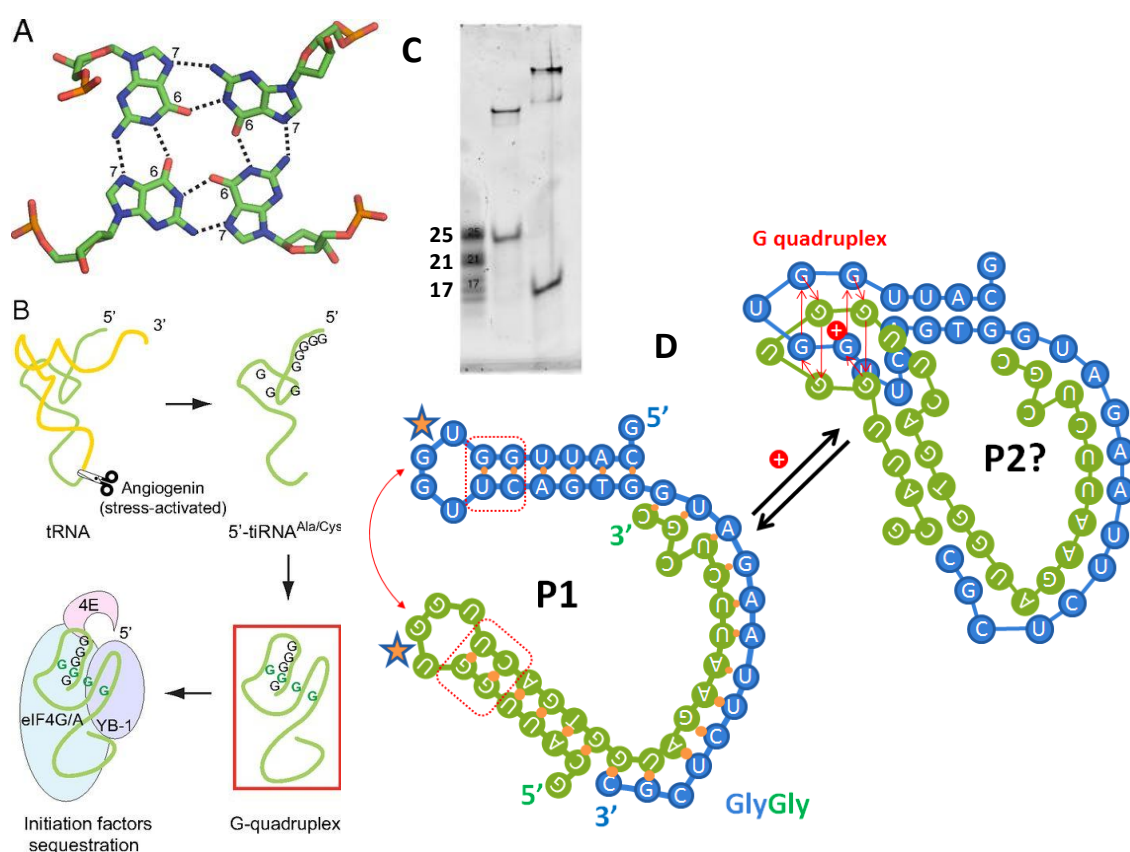


Figura 44: A) Figura del comentario en PNAS (Yang, 2014) sobre el artículo de Ivanov y colaboradores (2014). Se muestra la estructura básica de una tétrada de guaninas. Para la formación de dicha estructura se precisan 4 guaninas provenientes de distintos puntos en el espacio (no consecutivas). Si cada una de dichas guaninas fuese a su vez proseguida por otra/s, se podrían formar varios planos paralelos de tétradas, en lo que se conoce como G-cuádruplex. B) Fracción de la misma figura, sugiriendo la formación de tétradas de guanina de una sola capa entre las tRNA-h en su forma monomérica (las guaninas involucradas se ilustran en verde). C) Gel nativo de poliacrilamida al 20%, presentado en una de las figuras suplementarias del artículo de Ivanov et al. (2014). El carril 1 es un marcador de tamaños (17, 21 y 25 nt). El carril 2 es un

oligoribonucleótido sintético puro (RNA) con la secuencia de las tRNA-h de Ala. El carril 3 es un oligodesoxinucleótido sintético puro (DNA) de igual secuencia (con U en lugar de T). Nótese que ambos son capaces de formar estructuras presuntamente diméricas, a las que los autores no dan mayor importancia. D) Esquema hipotético de una posible estructura terciaria de los dímeros de tRNA-Gly plegados sobre sí mismos y estabilizados mediante la formación de un G-cuádruplex entre las guaninas de los bucles UGGU y las dos guaninas inmediatamente adyacentes en cada hélice. Las estrellas muestran sitios potenciales de corte endonucleolítico.

Si la estructura dimérica compacta estabilizada por G4plex fuese real y posible, la misma no se estaría generando con el oligo de RNA sintético debido a:

- i) Que su existencia es un invento; no es termodinámica, geométrica o cinéticamente posible que tales estructuras se formen.
- ii) La ausencia de proteínas que estabilicen dicha conformación.
- iii) La falta de una enzima que catalice su formación.
- iv) La falta de las tRNA-h de Glu.
- v) La ausencia de bases modificadas.
- vi) La ausencia de cationes divalentes, especialmente Mg^{2+} .

El punto (vi) fue atendido experimentalmente, sin obtenerse resultados diferentes. Tampoco nos animaríamos a descartar de plano la generación de G4plex al utilizar los oligos sintéticos. Se recordará que en los geles nativos se observó, en P2', una segunda banda de migración más rápida que la correspondiente a los monómeros. Sería, por tanto, más correcto afirmar que la formación de estructuras diméricas compactas conteniendo G4plex no ocurre eficientemente con los oligos sintéticos. Al menos con los del tRNA-Gly.

Volviendo al análisis de la cromatografía de intercambio aniónico por espectrometría de masas, seguía pendiente la pregunta de ¿por qué el efecto de la RNasaA sobre ENO1 dependía de la inclusión del pico P2? ¿Qué particularidad tendrían las mitades de tRNA en P2, para dar origen a un reconocimiento específico? Con lo que se ha venido discutiendo, podría pensarse en la presencia de G4plex ¿Sería que ENO1 reconoce este tipo de estructuras? La posibilidad resultó al principio demasiado forzada, pero nada se perdía con leer y conocer mejor a esta versátil enzima. Resulta que el gen *ENO1* posee un codón de iniciación alternativo en la secuencia codificante, que permite la traducción de un producto alternativo: la *Myc promoter binding protein 1* (MBP-1). MBP-1 carece de los 97 aminoácidos N-terminales de ENO1, lo que provoca la pérdida de su actividad catalítica. Sin embargo, es capaz de ingresar al núcleo y actuar como represor en la transcripción del proto-oncogen MYC (más conocido como c-myc). Esto hace de MBP-1 un represor tumoral. El sitio de unión al DNA está mapeado en la región N-terminal de MBP-1, que está también presente en ENO1 (Figura 45). Por su parte, la transcripción de MYC depende, entre otros motivos, de una secuencia *upstream* al promotor y rica en guanina (llamada Pu27, y ubicada en la región de hipersensibilidad a las nucleasas NHE III1, que controla el 85-90% de la transcripción de MYC), la cual puede adoptar una estructura de G4plex (Siddiqui-Jain et al., 2002). Dicha estructura, formada únicamente en una de las hebras, ocurre de forma concomitante a una segunda estructura conocida como “*i-motif*” en la hebra opuesta, y estabilizada por

apareamientos entre citosinas hemiprotonadas. (Brooks and Hurley, 2009). La estabilización de la estructura de G4plex en Pu27, ya sea experimentalmente mediante el agregado de drogas (Siddiqui-Jain et al., 2002), o fisiológicamente a través de proteínas como la Nucleolina (Gonzalez et al., 2009), reprime fuertemente la transcripción de MYC (Figura 45). En 1991 se identificó y clonó MBP-1 como una proteína capaz de unirse *upstream* al promotor de MYC y reprimir su transcripción. Los autores ya habían advertido que dicha proteína poseía una alta homología con ENO1, pero carecía de actividad catalítica (Ray and Miller, 1991). En un congreso de la Asociación Americana para la Investigación en Cáncer (AACR; 2008), se reportó que MBP-1 se une en Pu27, y que tanto MBP-1 como ENO-1 unen oligonucleótidos formadores de G4plex de manera específica⁴². No hay indicios de que dicho trabajo haya sido luego publicado, tal vez porque los resultados fueron incorporados a una patente destinada al tratamiento de pacientes oncológicos (US8410070 B2; concedida y publicada en abril de 2013).

En resumen, si bien estamos en un plano donde predomina la conjetura, tenemos que ENO-1 podría unirse a las mitades de tRNA, cuando estas se encuentran formando dímeros estabilizados por un G4plex (P2). Ya se sabe que ENO-1 es una proteína de unión al RNA, con especificidad por los tRNAs (al menos en levaduras) y, además, es capaz de reconocer y unir ácidos nucleicos formadores de G4plex. Sabemos que en nuestras muestras la proteína detectada fue ENO-1, ya que un alto porcentaje de los péptidos hallados mapean en la región N-terminal de la proteína, que está ausente en MBP-1.

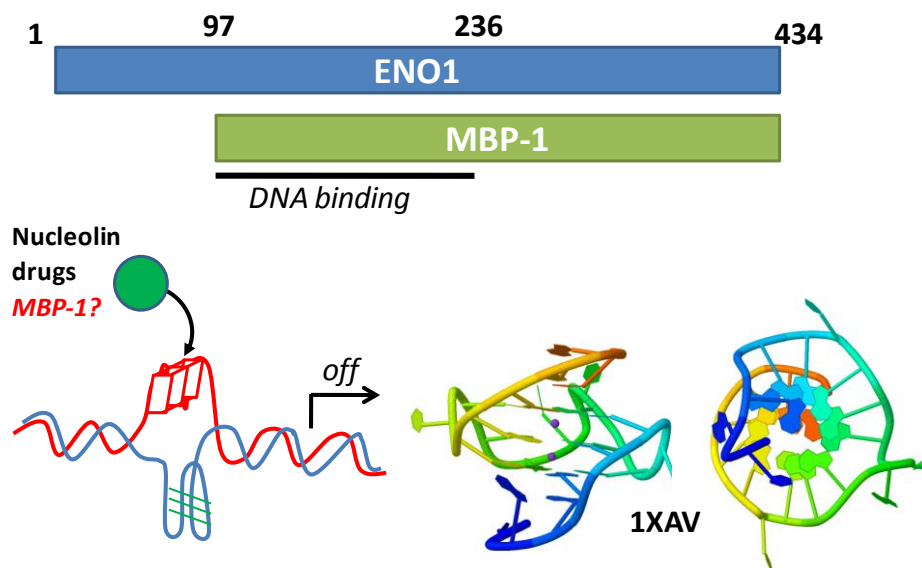


Figura 45: el gen ENO1 codifica para al menos dos proteínas, la alfa-Enolasa (ENO1) y la c-Myc promoter binding protein 1 (MBP-1), la cual se traduce a partir de un codón de iniciación ubicado 97 codones downstream al inicio de ENO1. MBP-1 puede ingresar al núcleo y reprimir la transcripción del famoso oncogén MYC (*c-myc*). La transcripción de MYC se inhibe si se forma una estructura de G cuádruplex (G4plex) en una de las hebras del DNA, y en una región ubicada upstream al promotor del gen. La estructura 1XAV permite evidenciar la formación de un G4plex de tres capas en dicha secuencia.

⁴²The MBP-1 tumor suppressor binds the c-myc promoter quadruplex-forming sequence;
http://cancerres.aacrjournals.org/content/68/9_Supplement/3440.short

Algunas drogas capaces de unir y estabilizar G4plex, así como la proteína Nucleolina, han sido experimentalmente validadas como represores de MYC mediante este mecanismo. Existe evidencia de que MBP-1 podría actuar de la misma manera.

Pues bien dicen que una imagen vale más que mil palabras, la Figura 46 es un intento de resumir algunas de las cuestiones que han sido abordadas en esta discusión. Por el momento, la mayoría de los mecanismos allí representados son especulativos. Pero es una buena hoja de ruta para organizar el trabajo a futuro.

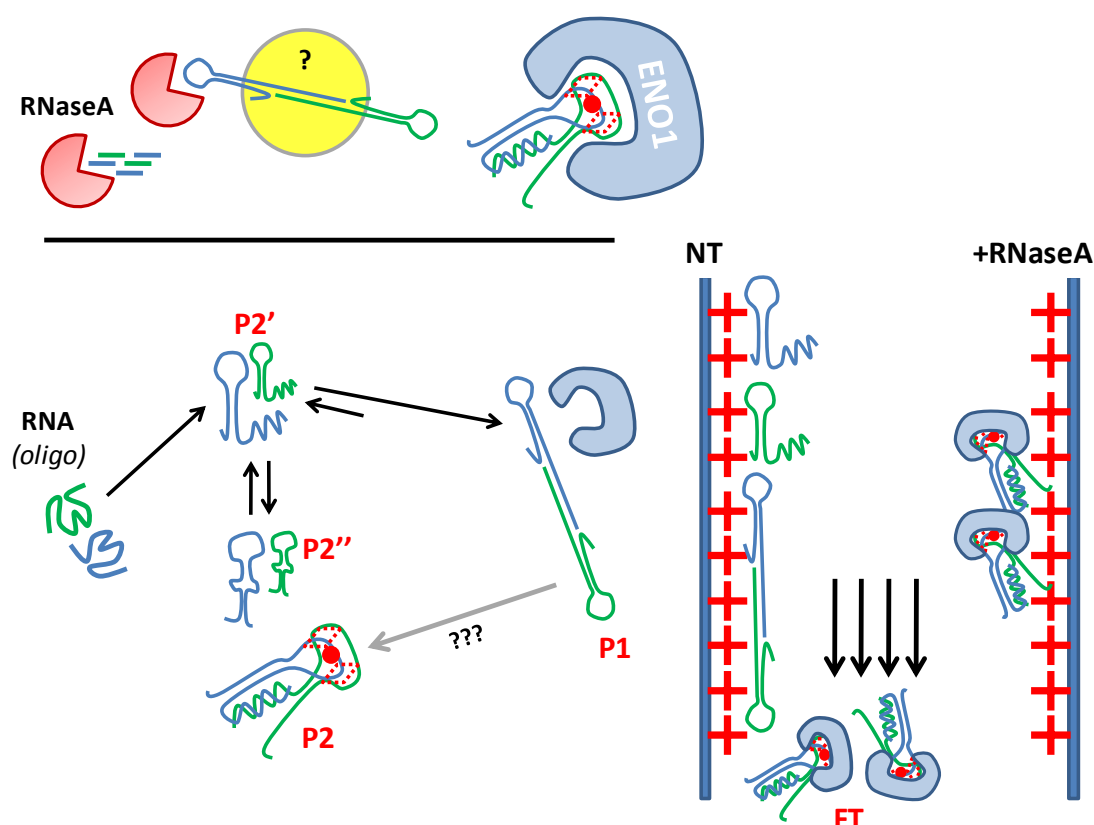


Figura 46: resumen de la discusión del Capítulo III. Izquierda: el RNA sintético (tRNA-Gly) adquiriría en solución dos posibles estructuras, en equilibrio entre sí. Las mismas eluirían en una SEC dando origen a los picos P2' y P2'', casi indistinguibles entre sí. Si la concentración de RNA sintético fuera suficiente, el mismo podría dimerizar, para dar lugar a dímeros Gly-Gly (y posiblemente Glu-Gly) que eluirían dando origen al pico P1', de tamaño aparente 44 kDa. Arriba: en el medio condicionado de MCF-7, las mitades de tRNA estarían en todo momento en su forma dimérica. Pero habría dos posibles conformaciones, una de las cuales implicaría un plegamiento del dímero sobre sí mismo y la estabilización de dicha estructura compacta a través de la formación de un cuádruplex de guanina de dos capas (G4plex). Dicha estructura podría unir específicamente a la enzima glucolítica ENO1 (y/o otras proteínas), que la protegería de la degradación mediada por la RNasaA agregada exógenamente durante los experimentos. Los complejos entre los dímeros de tRNA conteniendo G4plex y ENO1 se disociarían durante la SEC, observándose un pico para el dímero "alargado" (P1), y un segundo pico para el dímero compacto (P2). En muestras previamente tratadas con RNasaA, únicamente se observaría dicho pico. Derecha: al inyectar los picos P1 y P2 purificados de la SEC en una cromatografía de intercambio aniónico, el alto contenido de RNA en las muestras minimizaría la retención de los complejos ENO1/tRNA-h, que aparecerían altamente representados en la fracción no unida (FT). En cambio, en muestras tratadas con RNasaA post-SEC, los complejos (insensibles a la nucleasa) quedarían retenidos, disminuyendo su abundancia relativa en el FT.

3.5 PERSPECTIVAS A CORTO Y MEDIANO PLAZO

Para el cabal cumplimiento del objetivo específico del Capítulo III, la siguiente lista de ensayos será abordada en los próximos meses:

1- Confirmación de la unión entre las tRNA-h y las enzimas glucolíticas, en particular ENO1:

Reproducción del ensayo mostrado en la Figura 33. Entrecruzamiento de las proteínas a los tRNA-h (por distintos métodos) previo a la inyección de la fracción S100 en la columna de exclusión molecular. De observarse un corrimiento en los picos P1 o P2 hacia tamaños mayores: análisis por espectrometría de masas de dichas fracciones, y en las muestras control (sin entrecruzar).

De confirmarse los candidatos: inmunoprecipitación de los mismos a partir del medio condicionado y del lisado celular. Análisis del RNA (previo entrecruzamiento) y de posibles proteínas co-precipitantes. Silenciamiento de dichas proteínas (si la viabilidad celular no se ve afectada) y análisis de las mitades de tRNA en el medio extracelular. Obtención de las proteínas recombinantes, y reconstrucción *in vitro* de los complejos ETBP/tRNA-h utilizando los RNA sintéticos, y nuevos RNAs conteniendo mutaciones en distintos sitios (por ejemplo, alterando el bucle UGGU o la supuesta interfaz de dimerización).

De confirmarse que dichos candidatos se tratan de enzimas glucolíticas: sobre-expresión de las tRNA-h a nivel intracelular y cuantificación de parámetros metabólicos. Dicha sobre-expresión también podría hacerse mediante la inducción de un estrés oxidativo, de modo de que la Angiogenina produzca una mayor síntesis de estas mitades. Alternativamente, silenciamiento de la Angiogenina como modo de manipular hacia abajo la concentración intracelular – y eventualmente extracelular – de las tRNA-h (ya se está trabajando en todos estos puntos, en el marco de la Tesis de Maestría de la Lic. Fabiana Gámbaro).

Ya se dispone, además, del RNA sintético del tRNA-Gly pero marcado con biotina en su extremo 5'. Si la presencia de la biotina no inhibe la formación de dímeros, podría incubarse este RNA con la fracción S100 o con los lisados celulares y, además de inmunoprecipitar las proteínas candidatas, utilizar esferas unidas a estreptavidina para coleccionar el RNA y sus eventuales proteínas asociadas (previo entrecruzamiento). De modo tal de poder arribar al mismo resultado ya sea por la vía del RNA o de las proteínas. Previo a este estudio se realizarían geles de retardo con los extractos, transfiriendo y luego revelando con estreptavidina-HRP.

También se realizará un sencillo ensayo consistente en concentrar la fracción S100 por ultrafiltración, utilizando distintos filtros (a los de 10 kDa que hemos utilizado, sumarle de 50,

100 o 300 kDa). Inyectando los concentrados de cada caso, podría determinarse un estimado del tamaño aparente de los complejos ETBP/tRNA-h, sin necesidad de realizar entrecruzamiento.

2- Resistencia de las tRNA-h a la degradación y capacidad de interaccionar o ingresar a las células.

La resistencia a la RNasaA observada en las tRNA-h provenientes del medio de cultivo no pudo reproducirse totalmente con el RNA sintético. Se estudiará la sensibilidad o resistencia de los dímeros de tRNA-h a exonucleasas, endonucleasas de especificidad más restringida (RNase T1) o a RNasas de doble hebra (RNasa V1). Se adicionará RNA sintético a la fracción S100, o a los picos P1 y P2 purificados por SEC. Posteriormente se agregará RNasaA o las nucleasas que corresponda, de modo de poder estudiar si el efecto protector es dependiente de las proteínas del medio de cultivo condicionado, y en dicho caso, conocer en qué fracción eluyen dicha/s proteína/s.

También está pendiente la repetición de los estudios realizados con el RNA sintético en las mitades de tRNA-Glu, y en una mezcla de Glu y Gly a los efectos de obtener los presuntos heterodímeros. Se adquirirían además RNAs sintéticos “mutados” incapaces de dimerizar.

Queda pendiente la posibilidad de que estas mitades, por sí solas, adquieran una conformación del tipo G4plex, que podría conferirles una gran resistencia y posibilidad de interaccionar con células y/o proteínas de otra manera. Se incubarán las mitades de tRNA-Gly con Mg^{2+} y mayores concentraciones de K^+ a efectos de mejor evaluar esta posibilidad. La presencia de G4plex en las mitades endógenas se estudiará mediante la adición de reactivos químicos que reconocen y responden (ej: fluorescen) ante la presencia de G4plex.

Se estudiará si el RNA sintético tiene, de por sí, la capacidad de ser internalizado por células variadas (ya sea en forma monomérica, dimérica, o de G4plex). Para esto se utilizarán los RNAs sintéticos marcados con biotina, que ya se encuentran en nuestro poder. La captación intracelular se estudiará por microscopía de fluorescencia, utilizando estreptavidina acoplada a un fluorocromo adecuado.

Por último, y aunque es más ambicioso, se procurará obtener la estructura cristalográfica de los dímeros de tRNA-h (Gly-Gly y Glu-Gly), ya sea solos, o en complejos con las proteínas candidatas.

Futuros estudios quedarán supeditados a los resultados obtenidos con todos estos ensayos.

DIFUSIÓN Y PRODUCCIÓN DE ESTA TESIS

Además del presente documento, son productos de este trabajo:

PUBLICACIONES:

Capítulo I:

Tosar, J.P.; Rovira, C.; Naya, H.; Cayota, A. (2014) *Mining of public sequencing datasets supports a non-dietary origin for putative foreign miRNAs: underestimated effects of contamination in NGS.RNA*, v: 20 (6), p.: 754-757

<http://rnajournal.cshlp.org/content/20/6/754.long> (12 citas a marzo de 2016)

Capítulo II:

Tosar, J.P.; Gámbaro, F.; Sanguinetti, J.; Bonilla, B.; Witwer, K.; Cayota, A. (2015) *Assessment of small RNA sorting into different extracellular fractions revealed by high-throughput sequencing of breast cell lines*. *Nucleic Acids Research*, v: 43 (11), p.:5601-5616

<http://nar.oxfordjournals.org/content/43/11/5601> (9 citas a marzo de 2016)

TESIS:

En el marco de esta Tesis, y bajo la co-orientación de su autor, la Lic. Fabiana Gámbaro realizó su Tesina de Grado (Lic. en Bioquímica). Y actualmente realiza su Tesis de Maestría en la temática.

Las Bach. Ximena Doldán y Belén Fernández realizarán su Tesina de Grado (Lic. en Bioquímica) en proyectos no directamente relacionados a esta Tesis de Doctorado, pero sí desprendidos de ella. En el primer caso, el trabajo se encuentra prácticamente terminado, y pronto para su publicación.

PRESENTACIONES EN CONGRESOS:

Resultados parciales de esta Tesis han sido presentados en los Congresos Internacionales: “RNA 2014” (encuentro anual de la *RNA Society*, Québec, Canadá), “ISEV 2015” (encuentro anual de la *International Society for Extracellular Vesicles*, Washington, EE.UU) y “EV-RNA: what is the purpose?”(Utrecht, Holanda). Se presentó un póster en los primeros dos casos, y se realizó una presentación oral en el tercero.

Los resultados también han sido difundidos en el medio local. Se han realizado presentaciones orales y en formato póster en congresos de la Sociedad Uruguaya de Bioquímica y Biología Molecular (SBBM) y de la Sociedad Uruguaya de Biociencias. Algunos trabajos presentados han sido premiados.

COLABORACIONES:

En primer lugar, esta Tesis permitió una colaboración entre el Laboratorio de Genómica Funcional del IPMon, y la Unidad de Bioquímica Analítica del CIN, Facultad de Ciencias, Universidad de la República. Dicha colaboración no fue a efectos de desarrollar el trabajo aquí presentado, sino al desarrollo de una línea de investigación en biosensores amperométricos para la cuantificación de exosomas y otras vesículas extracelulares. Los resultados obtenidos en el marco de dicho proyecto colaborativo han sido muy buenos, y aunque aquí no vienen al caso, vale decir que se ha obtenido un método que permitiría cuantificar exosomas con gran rapidez y bajos costos, obteniendo una respuesta lineal robusta a lo largo de cuatro órdenes de magnitud, y con excelente discriminación entre exosomas y microvesículas. Se está estudiando, actualmente, la protección intelectual de dicho sistema, previo a su publicación. Además, esta vinculación entre el trabajo en biología molecular básica y el desarrollo tecnológico ha sido la base del proyecto presentado a la Universidad de la República para el ingreso del autor al régimen de Dedicación Total (efectivo desde diciembre de 2015).

Esta Tesis ha generado, también, exitosos casos de colaboración internacional. En particular, se ha creado un vínculo académico con el Dr. Kenneth Witwer de la Universidad Johns Hopkins en Baltimore, EE.UU. La colaboración con el Dr. Carlos Rovira (Universidad de Lünd, Suecia) no es en absoluto nueva, pero se destaca. Y se está comenzando un proyecto colaborativo con el Dr. Eric Westhof de la Universidad de Estrasburgo, Francia.

MANUSCRITOS PRÓXIMOS A ENVÍO:

“Sorting mechanisms and heterogeneity in EV RNA content” Este artículo escrito de manera colaborativa e interactiva entre cerca de 40 autores de varias partes del mundo será publicado próximamente en el *Journal of Extracellular Vesicles*. Se trata de un artículo de posición de la *International Society for Extracellular Vesicles*. El orden de los autores no está aún definido.

Tosar, J.P.; Cayota, A. *Protocols for detection of non-vesicular extracellular RNA*. Capítulo a ser publicado en *Methods in Molecular Biology* (invitación aceptada).

Doldán, X.; Fagúndez, P.; Cayota, A.; Laíz, J.; Tosar, J.P.; *Electrochemical sandwich immunosensor for extracellular exosome determination based on surface marker-mediated amplification*.

AGRADECIMIENTOS

Financiación:

El presente trabajo fue financiado específicamente mediante una Beca de Doctorado de la Agencia Nacional de Investigación e Innovación (ANII), y mediante un proyecto “Fondo Clemente Estable, Modalidad III (jóvenes investigadores)”, también de la ANII. Se obtuvo financiación intramuros (IPMon) para algunas partes del trabajo.

El autor realizó este trabajo siendo además docente de la Universidad de la República (grado 2), con Dedicación Total desde noviembre de 2015. Perteneció al Sistema Nacional de Investigadores desde 2011.

Se obtuvo financiación para la asistencia a congresos (Québec, 2014 y Utrecht, 2015) así como para la realización de una pasantía de un mes en la Universidad Johns Hopkins (2014) por parte de la Comisión Sectorial de Investigación Científica (CSIC) de la Universidad de la República.

Reconocimientos académicos:

Además de los miembros del Laboratorio de Genómica Funcional del IPMon, los siguientes investigadores han colaborado *personal y activamente* en alguna de las etapas de realización de este trabajo.

Instituto Pasteur de Montevideo

Unidad de Biología Molecular: Gonzalo Greif, Carlos Robello

Unidad de Bioinformática: Natalia Rego, Hugo Naya

Unidad de Bioquímica y Proteómica Analíticas: Magdalena Gil, Analía Lima, Madelón Portela, Rosario Durán, Carlos Batthyány.

Unidad de Proteínas Recombinantes: Agustín Correa

Unidad de Biofísica de Proteínas: Federico Carrión, Otto Pritsch

Facultad de Ciencias, Universidad de la República

Laboratorio de Biomateriales: Santiago Botassini, Eduardo Mendez

Laboratorios de Físicoquímica Biológica y Enzimología: Lucía Turell, Leonardo Santos, Beatriz Álvarez

Microscopía Electrónica de Transmisión: Gabriela Casanova

Universidad Johns Hopkins, EE.UU: Kenneth Witwer

Universidad de Lünd, Suecia: Carlos Rovira

Universidad de Estrasburgo, Francia: Eric Westhof

Agradecimientos peronsales:

En primer lugar, un agradecimiento a todos aquellos incluidos en la categoría de “reconocimientos académicos”. Porque una cosa es la colaboración técnica, la asistencia experimental, el consejo científico... pero todos o casi todos los mencionados han ido más allá de ese punto y se han realmente interesado por este trabajo, y buscado la mejor manera de llevarlo a cabo. Haciendo preguntas y ayudando a contestarlas.

Lo anterior no quisiera restringirlo únicamente a los colegas antes mencionados. Esa actitud de pro-activo interés en el trabajo del otro (en este caso, del mío), de seguimiento desinteresado (“¿cómo te fue con [tal experimento] que ibas a hacer?”), de “te mando este *paper* que encontré; se me ocurre que te puede interesar”... en definitiva: de aliento y de apoyo... esa actitud, decía, la encontré en buena parte del personal científico – y no solo científico – del IPMon. Creo que es este uno de los mayores activos con que cuenta este Instituto, y lo debe cuidar. Me vienen algunos nombres a la mente a los que quisiera particularmente agradecer, y que son tal vez quienes han mejor encarnado lo dicho en este párrafo. Pero como son varios, y temo olvidarme de alguno, prefiero realizar dichos agradecimientos personalmente.

Nuestro grupo, el del Laboratorio de Genómica Funcional: un agradecimiento particular. No solo por el trabajo compartido, las discusiones en seminarios, etc. Todos esos aspectos, fundamentales, fueron englobados en la sección de reconocimientos académicos. Aquí quisiera focalizarme en la parte humana. Gracias, compañeros, por bancarme. No es fácil trabajar conmigo, y no voy a enumerar aquí, en lo que se supone que es la alegre coronación de algunos años de trabajo, una larga lista de defectos. Ellos los conocen. Muchas veces han hecho la vista gorda en pos del buen ambiente, y otras veces me han señalado algunas faltas, lo que me ha hecho mucho bien. Gracias. En orden de llegada al laboratorio: María Rosa García, hoy haciendo su *posdoc* en Montpellier, y de quién aprendí mucho; Catalina Guida, que colocó a Chivilcoy en el mapa; Braulio Bonilla, con quién cruzamos los Andes a pie y compartimos infinidad de mates en el balcón, hablando de absolutamente cualquier cosa, hoy estudiante de doctorado en Pittsburgh; Florencia Cabrera-Cabrera, que está comenzando su doctorado en Alemania y que es una persona esencialmente buena; Julia Sanguinetti, una de las valientes que siguen, a punto de terminar su maestría; Fernanda Bangueses, que hizo su Tesina de grado en el laboratorio; y la última incorporación: Fabiana Gámbaro, a quien tuve el gusto de co-tutorar durante su tesis de Licenciatura, y que ahora realiza su maestría en temas muy vinculados a esta tesis. Cabe señalar que Fabiana ha realizado personalmente algunos de los ensayos que aquí se muestran, y trabaja activamente sobre varias de las perspectivas.

Un agradecimiento especial a Alfonso. Me tutoró en la maestría, y luego se animó a continuar con el doctorado. Si recibo finalmente el título de Doctor en Ciencias Biológicas no va a ser por las figuras mostradas en este documento, sino por la aprehensión de una serie de conocimientos y otras cuestiones intangibles que se las debo principalmente a él. Me enseñó una forma de pensar, de proceder, de trabajar... y a escribir sintético e ir al grano. Esta Tesis es una chistosa venganza, como para decir que no me quedé con las ganas. Además, sobre

todo, Alfonso me ha dado libertad. Libertad creativa, paulatina y gradual, mayor en las últimas etapas que en las primeras. No ese “dejar a sus anchas” que no correspondería en un orientador o director de tesis, sino la actitud de algunos buenos padres frente al hijo adolescente que quiere a toda costa darse de frente contra la pared.

Gracias también a los miembros del Tribunal, por haber aceptado evaluar este trabajo. Serán probablemente mis únicos lectores, pero quiero pensar que no escribí esta tesis en función de tan honorable audiencia. Gracias también por los “palos” que puedo intuir me van a dar, y las lecciones que pueda aprender de eso. Con algunos, por cercanía geográfica, hemos intercambiado ideas a lo largo de estos años, y han sido aportes muy importantes y esclarecedores. Seguro que lo seguirán siendo hacia delante.

Del otro lado de la Calle Matajojo, mi otro grupo, el de la Facultad de Ciencias. Durante la realización de este trabajo me desempeñé – y sigo – como Asistente, grado 2, de la Unidad de Bioquímica Analítica del Centro de Investigaciones Nucleares. Nuestra línea de trabajo tiene que ver con los biosensores electroquímicos... lejos del mundo de los RNAs reguladores de la expresión génica. Justo Laíz, responsable de la Unidad, bien podría haber puesto reparos a mi vinculación académica con un laboratorio no estrictamente universitario. O haber objetado mi formación a nivel de posgrado en un campo bastante distinto al del laboratorio en el cual mi cargo residía. Pero no. Justo no solamente toleró, sino que activamente apoyó este deseo mío de formarme en el campo que más me apasiona. Sin perjuicio de mis responsabilidades universitarias, me otorgó la necesaria flexibilidad para poder trabajar en dos frentes distintos, lo cual creo que fue, como superior jerárquico, una apuesta audaz. Con el tiempo terminé actuando de engrudo entre las dos disciplinas, y hoy la línea de inmunosensores electroquímicos para la determinación de exosomas atraviesa su fase productiva. Pero eso no podía adivinarse en su momento. Llegue este agradecimiento al hombre que, sin aparecer en la carátula de esta tesis, ha sido un fundamental e invisible impulsor de la misma. También a todos los demás compañeros con quienes trabajo en la tan querida Facultad de Ciencias, por cuya puerta ingresé en el año 2004 con mucha expectativa, y algo de miedo.

Los amigos. Los de toda la vida, que no saben ni les interesa la diferencia entre purinas y pirimidinas, pero que son la base de todo lo que pasa. Y los de Facultad, que es lo más firme y duradero que me dejó esta casa de estudio. Para que no se ofendan, una mención especial a los llamados “ídolos de la gloria”, un extraño rejunte de la generación 2004, una gemación de la cantera del Cafiaspirina Fútbol Club, con sus miembros hoy dispersos por el mundo, pero más presentes que nunca.

La familia. Mis cuatro hermanos que son lo más grande que hay, aunque nunca se los diga en la cara. Y con sus formas tan particulares de expresar su apoyo constante: “*Papá: este sí que nos vendió humo*” (meses después del día que intenté explicar lo que estaba haciendo, y al comprobar que los informativos no vendrían a buscarme). Mis padres: no tengo palabras para expresar mi agradecimiento. Viví con ellos hasta hace poco. Es verdad que es uno de los pocos métodos para ahorrar unos pesos mientras uno es estudiante de posgrado en ciencias, pero no era esa la (única) razón. La verdad que el obstáculo más difícil a vencer para eso de “independizarme” era lo bien que me lo pasaba en casa. Familia unida, corazón contento. Eso

dice el dicho, ¿no? Y a toda la familia grande, que englobo en la figura de mis dos fantásticas abuelas. Y al Cielo.

Y a Paty. Reescribí este párrafo ya varias veces. No puedo agradecerle por algo en particular. El agradecimiento es general, es total, y obviamente que trasciende la tesis; lo trasciende todo. Sin Paty este documento tendría seguramente el mismo número de páginas y de figuras. Pero no habría amor. En la escritura no habría pasión. Sería una tesis vacía, y frívola, y hueca. Serían tres años vividos en función de un aplauso de cortesía, para despertar al otro día y decir, ¿para qué? Gracias Paty, por hacer que todo tenga sentido.

REFERENCIAS BIBLIOGRÁFICAS

- Akat, K.M., Moore-McGriff, D., Morozov, P., Brown, M., Gogakos, T., Correa Da Rosa, J., Mihailovic, A., Sauer, M., Ji, R., Ramarathnam, A., *et al.* (2014). Comparative RNA-sequencing analysis of myocardial and circulating small RNAs in human heart failure and their utility as biomarkers. *Proc Natl Acad Sci U S A* **111**, 11151-11156.
- Al-Nedawi, K., Meehan, B., Micallef, J., Lhotak, V., May, L., Guha, A., and Rak, J. (2008). Intercellular transfer of the oncogenic receptor EGFRvIII by microvesicles derived from tumour cells. *Nat Cell Biol* **10**, 619-624.
- Anderson, P., and Ivanov, P. (2014). tRNA fragments in human health and disease. *FEBS Lett.*
- Aravin, A., Gaidatzis, D., Pfeffer, S., Lagos-Quintana, M., Landgraf, P., Iovino, N., Morris, P., Brownstein, M.J., Kuramochi-Miyagawa, S., Nakano, T., *et al.* (2006). A novel class of small RNAs bind to MILI protein in mouse testes. *Nature* **442**, 203-207.
- Arroyo, J.D., Chevillet, J.R., Kroh, E.M., Ruf, I.K., Pritchard, C.C., Gibson, D.F., Mitchell, P.S., Bennett, C.F., Pogosova-Agadjanian, E.L., Stirewalt, D.L., *et al.* (2011). Argonaute2 complexes carry a population of circulating microRNAs independent of vesicles in human plasma. *Proc Natl Acad Sci U S A* **108**, 5003-5008.
- Barnes, D., and Sato, G. (1979). Growth of a human mammary tumour cell line in a serum-free medium. *Nature* **281**, 388-389.
- Bissels, U., Wild, S., Tomiuk, S., Holste, A., Hafner, M., Tuschl, T., and Bosio, A. (2009). Absolute quantification of microRNAs by using a universal reference. *Rna* **15**, 2375-2384.
- Blanco, S., Dietmann, S., Flores, J.V., Hussain, S., Kutter, C., Humphreys, P., Lukk, M., Lombard, P., Treps, L., Popis, M., *et al.* (2014). Aberrant methylation of tRNAs links cellular stress to neuro-developmental disorders. *EMBO J* **33**, 2020-2039.
- Borsani, O., Zhu, J., Verslues, P.E., Sunkar, R., and Zhu, J.K. (2005). Endogenous siRNAs derived from a pair of natural cis-antisense transcripts regulate salt tolerance in Arabidopsis. *Cell* **123**, 1279-1291.
- Brennecke, J., Stark, A., Russell, R.B., and Cohen, S.M. (2005). Principles of microRNA-target recognition. *PLoS Biol* **3**, e85.
- Brooks, T.A., and Hurley, L.H. (2009). The role of supercoiling in transcriptional control of MYC and its importance in molecular therapeutics. *Nat Rev Cancer* **9**, 849-861.
- Caby, M.P., Lankar, D., Vincendeau-Scherrer, C., Raposo, G., and Bonnerot, C. (2005). Exosomal-like vesicles are present in human blood plasma. *International immunology* **17**, 879-887.
- Cairns, R.A., Harris, I.S., and Mak, T.W. (2011). Regulation of cancer cell metabolism. *Nat Rev Cancer* **11**, 85-95.
- Cao, Z., Huang, C.C., and Tan, W. (2006). Nuclease resistance of telomere-like oligonucleotides monitored in live cells by fluorescence anisotropy imaging. *Analytical chemistry* **78**, 1478-1484.
- Castello, A., Hentze, M.W., and Preiss, T. (2015). Metabolic Enzymes Enjoying New Partnerships as RNA-Binding Proteins. *Trends in endocrinology and metabolism: TEM* **26**, 746-757.
- Cole, C., Sobala, A., Lu, C., Thatcher, S.R., Bowman, A., Brown, J.W., Green, P.J., Barton, G.J., and Hutvagner, G. (2009). Filtering of deep sequencing data reveals the existence of abundant Dicer-dependent small RNAs derived from tRNAs. *Rna* **15**, 2147-2160.
- Conte, D., Jr., and Mello, C.C. (2003). RNA interference in *Caenorhabditis elegans*. *Curr Protoc Mol Biol Chapter 26*, Unit 26 23.
- Cortez, M.A., Bueso-Ramos, C., Ferdin, J., Lopez-Berestein, G., Sood, A.K., and Calin, G.A. (2011). MicroRNAs in body fluids--the mix of hormones and biomarkers. *Nat Rev Clin Oncol* **8**, 467-477.
- Costa-Silva, B., Aiello, N.M., Ocean, A.J., Singh, S., Zhang, H., Thakur, B.K., Becker, A., Hoshino, A., Mark, M.T., Molina, H., *et al.* (2015). Pancreatic cancer exosomes initiate pre-metastatic niche formation in the liver. *Nat Cell Biol* **17**, 816-826.

- Cozen, A.E., Quartley, E., Holmes, A.D., Hrabeta-Robinson, E., Phizicky, E.M., and Lowe, T.M. (2015). ARM-seq: AlkB-facilitated RNA methylation sequencing reveals a complex landscape of modified tRNA fragments. *Nat Methods* 12, 879-884.
- Chen, C., Ridzon, D.A., Broomer, A.J., Zhou, Z., Lee, D.H., Nguyen, J.T., Barbisin, M., Xu, N.L., Mahuvakar, V.R., Andersen, M.R., *et al.* (2005). Real-time quantification of microRNAs by stem-loop RT-PCR. *Nucleic Acids Res* 33, e179.
- Chen, Q., Yan, M., Cao, Z., Li, X., Zhang, Y., Shi, J., Feng, G.H., Peng, H., Zhang, X., Qian, J., *et al.* (2016). Sperm tsRNAs contribute to intergenerational inheritance of an acquired metabolic disorder. *Science* 351, 397-400.
- Chen, X., Li, Q., Wang, J., Guo, X., Jiang, X., Ren, Z., Weng, C., Sun, G., Wang, X., Liu, Y., *et al.* (2009). Identification and characterization of novel amphioxus microRNAs by Solexa sequencing. *Genome Biol* 10, R78.
- Chen, X., Sim, S., Wurtmann, E.J., Feke, A., and Wolin, S.L. (2014). Bacterial noncoding Y RNAs are widespread and mimic tRNAs. *Rna* 20, 1715-1724.
- Chen, X., Zen, K., and Zhang, C.Y. (2013). Reply to Lack of detectable oral bioavailability of plant microRNAs after feeding in mice. *Nature biotechnology* 31, 967-969.
- Chevillet, J.R., Kang, Q., Ruf, I.K., Briggs, H.A., Vojtech, L.N., Hughes, S.M., Cheng, H.H., Arroyo, J.D., Meredith, E.K., Gallichotte, E.N., *et al.* (2014). Quantitative and stoichiometric analysis of the microRNA content of exosomes. *Proc Natl Acad Sci U S A* 111, 14888-14893.
- Dhahbi, J.M., Spindler, S.R., Atamna, H., Boffelli, D., and Martin, D.I. (2014). Deep Sequencing of Serum Small RNAs Identifies Patterns of 5' tRNA Half and YRNA Fragment Expression Associated with Breast Cancer. *Biomark Cancer* 6, 37-47.
- Dhahbi, J.M., Spindler, S.R., Atamna, H., Boffelli, D., Mote, P., and Martin, D.I. (2013a). 5'-YRNA fragments derived by processing of transcripts from specific YRNA genes and pseudogenes are abundant in human serum and plasma. *Physiol Genomics* 45, 990-998.
- Dhahbi, J.M., Spindler, S.R., Atamna, H., Yamakawa, A., Boffelli, D., Mote, P., and Martin, D.I. (2013b). 5' tRNA halves are present as abundant complexes in serum, concentrated in blood cells, and modulated by aging and calorie restriction. *BMC Genomics* 14, 298.
- Dickinson, B., Zhang, Y., Petrick, J.S., Heck, G., Ivashuta, S., and Marshall, W.S. (2013). Lack of detectable oral bioavailability of plant microRNAs after feeding in mice. *Nature biotechnology* 31, 965-967.
- Duttagupta, R., Jiang, R., Gollub, J., Getts, R.C., and Jones, K.W. (2011). Impact of cellular miRNAs on circulating miRNA biomarker signatures. *PLoS One* 6, e20769.
- Elbarbary, R.A., Takaku, H., Uchiumi, N., Tamiya, H., Abe, M., Takahashi, M., Nishida, H., and Nashimoto, M. (2009). Modulation of gene expression by human cytosolic tRNase Z(L) through 5'-half-tRNA. *PLoS One* 4, e5908.
- Emara, M.M., Ivanov, P., Hickman, T., Dawra, N., Tisdale, S., Kedersha, N., Hu, G.F., and Anderson, P. (2010). Angiogenin-induced tRNA-derived stress-induced RNAs promote stress-induced stress granule assembly. *J Biol Chem* 285, 10959-10968.
- Entelis, N., Brandina, I., Kamenski, P., Krashennikov, I.A., Martin, R.P., and Tarassov, I. (2006). A glycolytic enzyme, enolase, is recruited as a cofactor of tRNA targeting toward mitochondria in *Saccharomyces cerevisiae*. *Genes Dev* 20, 1609-1620.
- Fabian, M.R., and Sonenberg, N. (2012). The mechanics of miRNA-mediated gene silencing: a look under the hood of miRISC. *Nature structural & molecular biology* 19, 586-593.
- Fabris, L., and Calin, G.A. (2016). Circulating free xeno-microRNAs - The new kids on the block. *Molecular oncology* 10, 503-508.
- Faghihi, M.A., Zhang, M., Huang, J., Modarresi, F., Van der Brug, M.P., Nalls, M.A., Cookson, M.R., St-Laurent, G., 3rd, and Wahlestedt, C. (2010). Evidence for natural antisense transcript-mediated inhibition of microRNA function. *Genome Biol* 11, R56.
- Fire, A., Xu, S., Montgomery, M.K., Kostas, S.A., Driver, S.E., and Mello, C.C. (1998). Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* 391, 806-811.

- Fong, M.Y., Zhou, W., Liu, L., Alontaga, A.Y., Chandra, M., Ashby, J., Chow, A., O'Connor, S.T., Li, S., Chin, A.R., *et al.* (2015). Breast-cancer-secreted miR-122 reprograms glucose metabolism in premetastatic niche to promote metastasis. *Nat Cell Biol* 17, 183-194.
- Friedman, R.C., Farh, K.K., Burge, C.B., and Bartel, D.P. (2009). Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res* 19, 92-105.
- Fu, H., Feng, J., Liu, Q., Sun, F., Tie, Y., Zhu, J., Xing, R., Sun, Z., and Zheng, X. (2009). Stress induces tRNA cleavage by angiogenin in mammalian cells. *FEBS Lett* 583, 437-442.
- Garcia-Silva, M.R., Cabrera-Cabrera, F., das Neves, R.F., Souto-Adron, T., de Souza, W., and Cayota, A. (2014a). Gene expression changes induced by *Trypanosoma cruzi* shed microvesicles in mammalian host cells: relevance of tRNA-derived halves. *BioMed research international* 2014, 305239.
- Garcia-Silva, M.R., das Neves, R.F., Cabrera-Cabrera, F., Sanguinetti, J., Medeiros, L.C., Robello, C., Naya, H., Fernandez-Calero, T., Souto-Adron, T., de Souza, W., *et al.* (2014b). Extracellular vesicles shed by *Trypanosoma cruzi* are linked to small RNA pathways, life cycle regulation, and susceptibility to infection of mammalian cells. *Parasitology research* 113, 285-304.
- Garcia-Silva, M.R., Frugier, M., Tosar, J.P., Correa-Dominguez, A., Ronalte-Alves, L., Parodi-Talice, A., Rovira, C., Robello, C., Goldenberg, S., and Cayota, A. (2010). A population of tRNA-derived small RNAs is actively produced in *Trypanosoma cruzi* and recruited to specific cytoplasmic granules. *Molecular and biochemical parasitology* 171, 64-73.
- Garcia-Silva, M.R., Sanguinetti, J., Cabrera-Cabrera, F., Franzen, O., and Cayota, A. (2014c). A particular set of small non-coding RNAs is bound to the distinctive Argonaute protein of *Trypanosoma cruzi*: insights from RNA-interference deficient organisms. *Gene* 538, 379-384.
- Garcia Silva, M.R., Tosar, J.P., Frugier, M., Pantano, S., Bonilla, B., Esteban, L., Serra, E., Rovira, C., Robello, C., and Cayota, A. (2010). Cloning, characterization and subcellular localization of a *Trypanosoma cruzi* argonaute protein defining a new subfamily distinctive of trypanosomatids. *Gene* 466, 26-35.
- Girard, A., Sachidanandam, R., Hannon, G.J., and Carmell, M.A. (2006). A germline-specific class of small RNAs binds mammalian Piwi proteins. *Nature* 442, 199-202.
- Gonzalez, V., Guo, K., Hurley, L., and Sun, D. (2009). Identification and characterization of nucleolin as a c-myc G-quadruplex-binding protein. *J Biol Chem* 284, 23622-23635.
- Goodarzi, H., Liu, X., Nguyen, H.C., Zhang, S., Fish, L., and Tavazoie, S.F. (2015). Endogenous tRNA-Derived Fragments Suppress Breast Cancer Progression via YBX1 Displacement. *Cell* 161, 790-802.
- Gould, S.J., Booth, A.M., and Hildreth, J.E. (2003). The Trojan exosome hypothesis. *Proc Natl Acad Sci U S A* 100, 10592-10597.
- Grivna, S.T., Beyret, E., Wang, Z., and Lin, H. (2006). A novel class of small RNAs in mouse spermatogenic cells. *Genes Dev* 20, 1709-1714.
- Guduric-Fuchs, J., O'Connor, A., Camp, B., O'Neill, C.L., Medina, R.J., and Simpson, D.A. (2012). Selective extracellular vesicle-mediated export of an overlapping set of microRNAs from multiple cell types. *BMC Genomics* 13, 357.
- Haussecker, D., Huang, Y., Lau, A., Parameswaran, P., Fire, A.Z., and Kay, M.A. (2010). Human tRNA-derived small RNAs in the global regulation of RNA silencing. *Rna* 16, 673-695.
- Hindson, C.M., Chevillet, J.R., Briggs, H.A., Gallichotte, E.N., Ruf, I.K., Hindson, B.J., Vessella, R.L., and Tewari, M. (2013). Absolute quantification by droplet digital PCR versus analog real-time PCR. *Nat Methods* 10, 1003-1005.
- Hirschi, K.D., Pruss, G.J., and Vance, V. (2015). Dietary delivery: a new avenue for microRNA therapeutics? *Trends in biotechnology* 33, 431-432.
- Honda, S., Loher, P., Shigematsu, M., Palazzo, J.P., Suzuki, R., Imoto, I., Rigoutsos, I., and Kirino, Y. (2015). Sex hormone-dependent tRNA halves enhance cell proliferation in breast and prostate cancers. *Proc Natl Acad Sci U S A* 112, E3816-3825.
- Hoshino, A., Costa-Silva, B., Shen, T.L., Rodrigues, G., Hashimoto, A., Tesic Mark, M., Molina, H., Kohsaka, S., Di Giannatale, A., Ceder, S., *et al.* (2015). Tumour exosome integrins determine organotropic metastasis. *Nature* 527, 329-335.

- Hou, J., Lin, L., Zhou, W., Wang, Z., Ding, G., Dong, Q., Qin, L., Wu, X., Zheng, Y., Yang, Y., *et al.* (2011). Identification of miRNomes in human liver and hepatocellular carcinoma reveals miR-199a/b-3p as therapeutic target for hepatocellular carcinoma. *Cancer Cell* 19, 232-243.
- Huber, V., Fais, S., Iero, M., Lugini, L., Canese, P., Squarcina, P., Zaccheddu, A., Colone, M., Arancia, G., Gentile, M., *et al.* (2005). Human colorectal cancer cells induce T-cell death through release of proapoptotic microvesicles: role in immune escape. *Gastroenterology* 128, 1796-1804.
- Huberts, D.H., and van der Klei, I.J. (2010). Moonlighting proteins: an intriguing mode of multitasking. *Biochimica et biophysica acta* 1803, 520-525.
- Ivanov, P., Emara, M.M., Villen, J., Gygi, S.P., and Anderson, P. (2011). Angiogenin-induced tRNA fragments inhibit translation initiation. *Mol Cell* 43, 613-623.
- Ivanov, P., O'Day, E., Emara, M.M., Wagner, G., Lieberman, J., and Anderson, P. (2014). G-quadruplex structures contribute to the neuroprotective effects of angiogenin-induced tRNA fragments. *Proc Natl Acad Sci U S A* 111, 18201-18206.
- J. Simpson, R., and Mathivanan, S. (2012). Extracellular Microvesicles: The Need for Internationally Recognised Nomenclature and Stringent Purification Criteria. *Journal of Proteomics & Bioinformatics* 05.
- Ji, H., Chen, M., Greening, D.W., He, W., Rai, A., Zhang, W., and Simpson, R.J. (2014). Deep Sequencing of RNA from Three Different Extracellular Vesicle (EV) Subtypes Released from the Human LIM1863 Colon Cancer Cell Line Uncovers Distinct Mirna-Enrichment Signatures. *PLoS One* 9, e110314.
- Ji, H., Wang, J., Guo, J., Li, Y., Lian, S., Guo, W., and Yang, H. (2016). Progress in the biological function of alpha-enolase. *Animal Nutrition*.
- Jiang, M., Sang, X., and Hong, Z. (2012). Beyond nutrients: food-derived microRNAs provide cross-kingdom regulation. *BioEssays : news and reviews in molecular, cellular and developmental biology* 34, 280-284.
- Jochl, C., Rederstorff, M., Hertel, J., Stadler, P.F., Hofacker, I.L., Schrettl, M., Haas, H., and Huttenhofer, A. (2008). Small ncRNA transcriptome analysis from *Aspergillus fumigatus* suggests a novel mechanism for regulation of protein synthesis. *Nucleic Acids Res* 36, 2677-2689.
- Joyce, C.E., Zhou, X., Xia, J., Ryan, C., Thrash, B., Menter, A., Zhang, W., and Bowcock, A.M. (2011). Deep sequencing of small RNAs from human skin reveals major alterations in the psoriasis miRNAome. *Human molecular genetics* 20, 4025-4040.
- Jung, U., Jiang, X., Kaufmann, S.H., and Patzel, V. (2013). A universal TaqMan-based RT-PCR protocol for cost-efficient detection of small noncoding RNA. *Rna* 19, 1864-1873.
- Kanada, M., Bachmann, M.H., Hardy, J.W., Frimannson, D.O., Bronsart, L., Wang, A., Sylvester, M.D., Schmidt, T.L., Kaspar, R.L., Butte, M.J., *et al.* (2015). Differential fates of biomolecules delivered to target cells via extracellular vesicles. *Proc Natl Acad Sci U S A* 112, E1433-1442.
- Kawaji, H., Nakamura, M., Takahashi, Y., Sandelin, A., Katayama, S., Fukuda, S., Daub, C.O., Kai, C., Kawai, J., Yasuda, J., *et al.* (2008). Hidden layers of human small RNAs. *BMC Genomics* 9, 157.
- Keller, K., Kolbe, H., Lange, K., and Zimmermann, B. (1978). Release of glycolytic enzymes from cultivated tumor cells. *Zeitschrift fur Krebsforschung und klinische Onkologie Cancer research and clinical oncology* 92, 275-286.
- Kirchner, S., and Ignatova, Z. (2015). Emerging roles of tRNA in adaptive translation, signalling dynamics and disease. *Nat Rev Genet* 16, 98-112.
- Kowalski, M.P., and Krude, T. (2015). Functional roles of non-coding Y RNAs. *Int J Biochem Cell Biol* 66, 20-29.
- Kramer, M.F. (2011). Stem-loop RT-qPCR for miRNAs. *Curr Protoc Mol Biol Chapter* 15, Unit 15 10.
- Krawetz, S.A., Kruger, A., Lalancette, C., Tagett, R., Anton, E., Draghici, S., and Diamond, M.P. (2011). A survey of small RNAs in human sperm. *Human reproduction* 26, 3401-3412.

- La Rocca, G., Olejniczak, S.H., Gonzalez, A.J., Briskin, D., Vidigal, J.A., Spraggon, L., DeMatteo, R.G., Radler, M.R., Lindsten, T., Ventura, A., *et al.* (2015). In vivo, Argonaute-bound microRNAs exist predominantly in a reservoir of low molecular weight complexes not associated with mRNA. *Proc Natl Acad Sci U S A* **112**, 767-772.
- Lagos-Quintana, M., Rauhut, R., Lendeckel, W., and Tuschl, T. (2001). Identification of novel genes coding for small expressed RNAs. *Science* **294**, 853-858.
- Landgraf, P., Rusu, M., Sheridan, R., Sewer, A., Iovino, N., Aravin, A., Pfeffer, S., Rice, A., Kamphorst, A.O., Landthaler, M., *et al.* (2007). A mammalian microRNA expression atlas based on small RNA library sequencing. *Cell* **129**, 1401-1414.
- Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* **10**, R25.
- Lau, N.C., Lim, L.P., Weinstein, E.G., and Bartel, D.P. (2001). An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*. *Science* **294**, 858-862.
- Lawrie, C.H., Gal, S., Dunlop, H.M., Pushkaran, B., Liggins, A.P., Pulford, K., Banham, A.H., Pezzella, F., Boultonwood, J., Wainscoat, J.S., *et al.* (2008). Detection of elevated levels of tumour-associated microRNAs in serum of patients with diffuse large B-cell lymphoma. *British journal of haematology* **141**, 672-675.
- Lee, F.S., and Vallee, B.L. (1989). Characterization of ribonucleolytic activity of angiogenin towards tRNA. *Biochem Biophys Res Commun* **161**, 121-126.
- Lee, R.C., and Ambros, V. (2001). An extensive class of small RNAs in *Caenorhabditis elegans*. *Science* **294**, 862-864.
- Lee, R.C., Feinbaum, R.L., and Ambros, V. (1993). The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell* **75**, 843-854.
- Lee, S.R., and Collins, K. (2005). Starvation-induced cleavage of the tRNA anticodon loop in *Tetrahymena thermophila*. *J Biol Chem* **280**, 42744-42749.
- Lee, Y.S., Shibata, Y., Malhotra, A., and Dutta, A. (2009). A novel class of small RNAs: tRNA-derived RNA fragments (tRFs). *Genes Dev* **23**, 2639-2649.
- Li, D. (2013). Sanofi-Cell Research outstanding paper award of 2012. *Cell research* **23**, 1241.
- Li, J., Zhang, Y., Li, D., Liu, Y., Chu, D., Jiang, X., Hou, D., Zen, K., and Zhang, C.Y. (2015). Small non-coding RNAs transfer through mammalian placenta and directly regulate fetal gene expression. *Protein & cell* **6**, 391-396.
- Li, S., and Hu, G.F. (2012). Emerging role of angiogenin in stress response and cell survival under adverse conditions. *J Cell Physiol* **227**, 2822-2826.
- Li, Y., Luo, J., Zhou, H., Liao, J.Y., Ma, L.M., Chen, Y.Q., and Qu, L.H. (2008). Stress-induced tRNA-derived RNAs: a novel class of small RNAs in the primitive eukaryote *Giardia lamblia*. *Nucleic Acids Res* **36**, 6048-6055.
- Li, Z., Ender, C., Meister, G., Moore, P.S., Chang, Y., and John, B. (2012). Extensive terminal and asymmetric processing of small RNAs from rRNAs, snoRNAs, snRNAs, and tRNAs. *Nucleic Acids Res* **40**, 6787-6799.
- Liang, H., Huang, L., Cao, J., Zen, K., Chen, X., and Zhang, C.Y. (2012). Regulation of mammalian gene expression by exogenous microRNAs. *Wiley interdisciplinary reviews RNA* **3**, 733-742.
- Liu, W.M., Pang, R.T., Chiu, P.C., Wong, B.P., Lao, K., Lee, K.F., and Yeung, W.S. (2012). Sperm-borne microRNA-34c is required for the first cleavage division in mouse. *Proc Natl Acad Sci U S A* **109**, 490-494.
- Lobb, R.J., Becker, M., Wen, S.W., Wong, C.S., Wiegmanns, A.P., Leimgruber, A., and Moller, A. (2015). Optimized exosome isolation protocol for cell culture supernatant and human plasma. *J Extracell Vesicles* **4**, 27031.
- Lombardi, V.C., Ruscetti, F.W., Das Gupta, J., Pfost, M.A., Hagen, K.S., Peterson, D.L., Ruscetti, S.K., Bagni, R.K., Petrow-Sadowski, C., Gold, B., *et al.* (2009). Detection of an infectious retrovirus, XMRV, in blood cells of patients with chronic fatigue syndrome. *Science* **326**, 585-589.
- Lu, J., Getz, G., Miska, E.A., Alvarez-Saavedra, E., Lamb, J., Peck, D., Sweet-Cordero, A., Ebert, B.L., Mak, R.H., Ferrando, A.A., *et al.* (2005). MicroRNA expression profiles classify human cancers. *Nature* **435**, 834-838.

- Marton, S., Garcia, M.R., Robello, C., Persson, H., Trajtenberg, F., Pritsch, O., Rovira, C., Naya, H., Dighiero, G., and Cayota, A. (2008). Small RNAs analysis in CLL reveals a deregulation of miRNA expression and novel miRNA candidates of putative relevance in CLL pathogenesis. *Leukemia* 22, 330-338.
- Maute, R.L., Schneider, C., Sumazin, P., Holmes, A., Califano, A., Basso, K., and Dalla-Favera, R. (2013). tRNA-derived microRNA modulates proliferation and the DNA damage response and is down-regulated in B cell lymphoma. *Proc Natl Acad Sci U S A* 110, 1404-1409.
- Mazurek, S., Hugo, F., Failing, K., and Eigenbrodt, E. (1996). Studies on associations of glycolytic and glutaminolytic enzymes in MCF-7 cells: role of P36. *J Cell Physiol* 167, 238-250.
- Melo, S.A., Luecke, L.B., Kahlert, C., Fernandez, A.F., Gammon, S.T., Kaye, J., LeBleu, V.S., Mittendorf, E.A., Weitz, J., Rahbari, N., *et al.* (2015). Glypican-1 identifies cancer exosomes and detects early pancreatic cancer. *Nature* 523, 177-182.
- Mittelbrunn, M., Gutierrez-Vazquez, C., Villarroya-Beltri, C., Gonzalez, S., Sanchez-Cabo, F., Gonzalez, M.A., Bernad, A., and Sanchez-Madrid, F. (2011). Unidirectional transfer of microRNA-loaded exosomes from T cells to antigen-presenting cells. *Nature communications* 2, 282.
- Mittelbrunn, M., and Sanchez-Madrid, F. (2012). Intercellular communication: diverse structures for exchange of genetic information. *Nat Rev Mol Cell Biol* 13, 328-335.
- Naccache, S.N., Hackett, J., Jr., Delwart, E.L., and Chiu, C.Y. (2014). Concerns over the origin of NIH-CQV, a novel virus discovered in Chinese patients with seronegative hepatitis. *Proc Natl Acad Sci U S A* 111, E976.
- Nat-Biotechnol (2013). Receptive to replication. *Nature biotechnology* 31, 943.
- Nicolas, F.E., Hall, A.E., Csorba, T., Turnbull, C., and Dalmay, T. (2012). Biogenesis of Y RNA-derived small RNAs is independent of the microRNA pathway. *FEBS Lett* 586, 1226-1230.
- Nilsen, T.W. (2008). Endo-siRNAs: yet another layer of complexity in RNA silencing. *Nature structural & molecular biology* 15, 546-548.
- Nolte-t Hoen, E.N., Buermans, H.P., Waasdorp, M., Stoorvogel, W., Wauben, M.H., and t Hoen, P.A. (2012). Deep sequencing of RNA from immune cell-derived vesicles uncovers the selective incorporation of small non-coding RNA biotypes with potential regulatory functions. *Nucleic Acids Res* 40, 9272-9285.
- Ohara, T., Sakaguchi, Y., Suzuki, T., Ueda, H., Miyauchi, K., and Suzuki, T. (2007). The 3' termini of mouse Piwi-interacting RNAs are 2'-O-methylated. *Nature structural & molecular biology* 14, 349-350.
- Paprotka, T., Delviks-Frankenberry, K.A., Cingoz, O., Martinez, A., Kung, H.J., Tepper, C.G., Hu, W.S., Fivash, M.J., Jr., Coffin, J.M., and Pathak, V.K. (2011). Recombinant origin of the retrovirus XMRV. *Science* 333, 97-101.
- Peinado, H., Aleckovic, M., Lavotshkin, S., Matei, I., Costa-Silva, B., Moreno-Bueno, G., Hergueta-Redondo, M., Williams, C., Garcia-Santos, G., Ghajar, C., *et al.* (2012). Melanoma exosomes educate bone marrow progenitor cells toward a pro-metastatic phenotype through MET. *Nat Med* 18, 883-891.
- Rauber, N.R., Jany, K.D., and Pfeleiderer, G. (1978). Ribonuclease A digestion by proteinase K. *Zeitschrift fur Naturforschung Section C: Biosciences* 33, 660-663.
- Ray, R., and Miller, D.M. (1991). Cloning and characterization of a human c-myc promoter-binding protein. *Mol Cell Biol* 11, 2154-2161.
- Reinhart, B.J., Slack, F.J., Basson, M., Pasquinelli, A.E., Bettinger, J.C., Rougvie, A.E., Horvitz, H.R., and Ruvkun, G. (2000). The 21-nucleotide let-7 RNA regulates developmental timing in *Caenorhabditis elegans*. *Nature* 403, 901-906.
- Risitano, A., and Fox, K.R. (2004). Influence of loop size on the stability of intramolecular DNA quadruplexes. *Nucleic Acids Res* 32, 2598-2606.
- Robinson, J.B., Jr., Inman, L., Sumegi, B., and Srere, P.A. (1987). Further characterization of the Krebs tricarboxylic acid cycle metabolon. *J Biol Chem* 262, 1786-1790.
- Ruegger, S., and Grosshans, H. (2012). MicroRNA turnover: when, how, and why. *Trends Biochem Sci* 37, 436-446.

- Saikia, M., Jobava, R., Parisien, M., Putnam, A., Krokowski, D., Gao, X.H., Guan, B.J., Yuan, Y., Jankowsky, E., Feng, Z., *et al.* (2014). Angiogenin-cleaved tRNA halves interact with cytochrome c, protecting cells from apoptosis during osmotic stress. *Mol Cell Biol* 34, 2450-2463.
- Salter, S.J., Cox, M.J., Turek, E.M., Calus, S.T., Cookson, W.O., Moffatt, M.F., Turner, P., Parkhill, J., Loman, N.J., and Walker, A.W. (2014). Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC biology* 12, 87.
- Schotte, D., Akbari Moqadam, F., Lange-Turenhout, E.A., Chen, C., van Ijcken, W.F., Pieters, R., and den Boer, M.L. (2011). Discovery of new microRNAs by small RNAome deep sequencing in childhood acute lymphoblastic leukemia. *Leukemia* 25, 1389-1399.
- Sharma, U., Conine, C.C., Shea, J.M., Boskovic, A., Derr, A.G., Bing, X.Y., Belleannee, C., Kucukural, A., Serra, R.W., Sun, F., *et al.* (2016). Biogenesis and function of tRNA fragments during sperm maturation and fertilization in mammals. *Science* 351, 391-396.
- Si, M.L., Zhu, S., Wu, H., Lu, Z., Wu, F., and Mo, Y.Y. (2007). miR-21-mediated tumor growth. *Oncogene* 26, 2799-2803.
- Siddiqui-Jain, A., Grand, C.L., Bearss, D.J., and Hurley, L.H. (2002). Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription. *Proc Natl Acad Sci U S A* 99, 11593-11598.
- Siomi, M.C., Sato, K., Pezic, D., and Aravin, A.A. (2011). PIWI-interacting small RNAs: the vanguard of genome defence. *Nat Rev Mol Cell Biol* 12, 246-258.
- Slack, F.J., Basson, M., Liu, Z., Ambros, V., Horvitz, H.R., and Ruvkun, G. (2000). The lin-41 RBCC gene acts in the *C. elegans* heterochronic pathway between the let-7 regulatory RNA and the LIN-29 transcription factor. *Mol Cell* 5, 659-669.
- Snow, J.W., Hale, A.E., Isaacs, S.K., Baggish, A.L., and Chan, S.Y. (2013). Ineffective delivery of diet-derived microRNAs to recipient animal organisms. *RNA Biol* 10, 1107-1116.
- Soule, H.D., Maloney, T.M., Wolman, S.R., Peterson, W.D., Jr., Brenz, R., McGrath, C.M., Russo, J., Pauley, R.J., Jones, R.F., and Brooks, S.C. (1990). Isolation and characterization of a spontaneously immortalized human breast epithelial cell line, MCF-10. *Cancer research* 50, 6075-6086.
- Squadrito, M.L., Baer, C., Burdet, F., Maderna, C., Gilfillan, G.D., Lyle, R., Ibberson, M., and De Palma, M. (2014). Endogenous RNAs modulate microRNA sorting to exosomes and transfer to acceptor cells. *Cell Rep* 8, 1432-1446.
- Thompson, D.M., Lu, C., Green, P.J., and Parker, R. (2008). tRNA cleavage is a conserved response to oxidative stress in eukaryotes. *Rna* 14, 2095-2103.
- Tosar, J.P., Rovira, C., Naya, H., and Cayota, A. (2014). Mining of public sequencing databases supports a non-dietary origin for putative foreign miRNAs: underestimated effects of contamination in NGS. *Rna* 20, 754-757.
- Tosar, J.P.; Gámbaro, F.; Sanguinetti, J.; Bonilla, B.; Witwer, K.; Cayota, A. (2015) Assessment of small RNA sorting into different extracellular fractions revealed by high-throughput sequencing of breast cell lines. *Nucleic Acids Res* 43, 5601-5616
- Turchinovich, A., Weiz, L., Langheinz, A., and Burwinkel, B. (2011). Characterization of extracellular circulating microRNA. *Nucleic Acids Res* 39, 7223-7233.
- Urisman, A., Molinaro, R.J., Fischer, N., Plummer, S.J., Casey, G., Klein, E.A., Malathi, K., Magi-Galluzzi, C., Tubbs, R.R., Ganem, D., *et al.* (2006). Identification of a novel Gammaretrovirus in prostate tumors of patients homozygous for R462Q RNASEL variant. *PLoS pathogens* 2, e25.
- Valadi, H., Ekstrom, K., Bossios, A., Sjostrand, M., Lee, J.J., and Lotvall, J.O. (2007). Exosome-mediated transfer of mRNAs and microRNAs is a novel mechanism of genetic exchange between cells. *Nat Cell Biol* 9, 654-659.
- van Balkom, B.W., Eisele, A.S., Pegtel, D.M., Bervoets, S., and Verhaar, M.C. (2015). Quantitative and qualitative analysis of small RNAs in human endothelial cells and exosomes provides insights into localized RNA processing, degradation and sorting. *J Extracell Vesicles* 4, 26760.

- Van Deun, J., Mestdagh, P., Sormunen, R., Cocquyt, V., Vermaelen, K., Vandesompele, J., Bracke, M., De Wever, O., and Hendrix, A. (2014). The impact of disparate isolation methods for extracellular vesicles on downstream RNA profiling. *J Extracell Vesicles* 3.
- Vickers, K.C., Palmisano, B.T., Shoucri, B.M., Shamburek, R.D., and Remaley, A.T. (2011). MicroRNAs are transported in plasma and delivered to recipient cells by high-density lipoproteins. *Nat Cell Biol* 13, 423-433.
- Villarroya-Beltri, C., Gutierrez-Vazquez, C., Sanchez-Cabo, F., Perez-Hernandez, D., Vazquez, J., Martin-Cofreces, N., Martinez-Herrera, D.J., Pascual-Montano, A., Mittelbrunn, M., and Sanchez-Madrid, F. (2013). Sumoylated hnRNP A2B1 controls the sorting of miRNAs into exosomes through binding to specific motifs. *Nature communications* 4, 2980.
- Vojtech, L., Woo, S., Hughes, S., Levy, C., Ballweber, L., Sauteraud, R.P., Strobl, J., Westerberg, K., Gottardo, R., Tewari, M., *et al.* (2014). Exosomes in human semen carry a distinctive repertoire of small non-coding RNAs with potential regulatory functions. *Nucleic Acids Res* 42, 7290-7304.
- Wang, K., Zhang, S., Weber, J., Baxter, D., and Galas, D.J. (2010). Export of microRNAs and microRNA-protective protein by mammalian cells. *Nucleic Acids Res* 38, 7248-7259.
- Watanabe, T., Takeda, A., Tsukiyama, T., Mise, K., Okuno, T., Sasaki, H., Minami, N., and Imai, H. (2006). Identification and characterization of two novel classes of small RNAs in the mouse germline: retrotransposon-derived siRNAs in oocytes and germline small RNAs in testes. *Genes Dev* 20, 1732-1743.
- Wickramasinghe, N.S., Manavalan, T.T., Dougherty, S.M., Riggs, K.A., Li, Y., and Klinge, C.M. (2009). Estradiol downregulates miR-21 expression and increases miR-21 target gene expression in MCF-7 breast cancer cells. *Nucleic Acids Res* 37, 2584-2595.
- Wightman, B., Ha, I., and Ruvkun, G. (1993). Posttranscriptional regulation of the heterochronic gene *lin-14* by *lin-4* mediates temporal pattern formation in *C. elegans*. *Cell* 75, 855-862.
- Winther, K.S., and Gerdes, K. (2011). Enteric virulence associated protein VapC inhibits translation by cleavage of initiator tRNA. *Proc Natl Acad Sci U S A* 108, 7403-7407.
- Witwer, K.W. (2014). Circulating MicroRNA Biomarker Studies: Pitfalls and Potential Solutions. *Clin Chem*.
- Witwer, K.W., Buzas, E.I., Bemis, L.T., Bora, A., Lasser, C., Lotvall, J., Nolte-'t Hoen, E.N., Piper, M.G., Sivaraman, S., Skog, J., *et al.* (2013a). Standardization of sample collection, isolation and analysis methods in extracellular vesicle research. *J Extracell Vesicles* 2.
- Witwer, K.W., McAleer, M.A., Queen, S.E., and Adams, R.J. (2013b). Real-time quantitative PCR and droplet digital PCR for plant miRNAs in mammalian blood provide little evidence for general uptake of dietary miRNAs: limited evidence for general uptake of dietary plant xenomiRs. *RNA Biol* 10, 1080-1086.
- Witzany, G. (2012). Do we eat gene regulators? *Communicative & integrative biology* 5, 230-232.
- Wu, F., and Minter, S. (2015). Krebs cycle metabolite: structural evidence of substrate channeling revealed by cross-linking and mass spectrometry. *Angewandte Chemie* 54, 1851-1854.
- Wu, X., Somlo, G., Yu, Y., Palomares, M.R., Li, A.X., Zhou, W., Chow, A., Yen, Y., Rossi, J.J., Gao, H., *et al.* (2012). De novo sequencing of circulating miRNAs identifies novel markers predicting clinical outcome of locally advanced breast cancer. *J Transl Med* 10, 42.
- Xu, B., Zhi, N., Hu, G., Wan, Z., Zheng, X., Liu, X., Wong, S., Kajigaya, S., Zhao, K., Mao, Q., *et al.* (2013). Hybrid DNA virus in Chinese patients with seronegative hepatitis discovered by deep sequencing. *Proc Natl Acad Sci U S A* 110, 10264-10269.
- Yamasaki, S., Ivanov, P., Hu, G.F., and Anderson, P. (2009). Angiogenin cleaves tRNA and promotes stress-induced translational repression. *J Cell Biol* 185, 35-42.
- Yang, C., and Robbins, P.D. (2011). The roles of tumor-derived exosomes in cancer pathogenesis. *Clin Dev Immunol* 2011, 842849.
- Zhang, J., Li, S., Li, L., Li, M., Guo, C., Yao, J., and Mi, S. (2015). Exosome and exosomal microRNA: trafficking, sorting, and function. *Genomics, proteomics & bioinformatics* 13, 17-24.

Zhang, L., Hou, D., Chen, X., Li, D., Zhu, L., Zhang, Y., Li, J., Bian, Z., Liang, X., Cai, X., *et al.* (2012a). Exogenous plant MIR168a specifically targets mammalian LDLRAP1: evidence of cross-kingdom regulation by microRNA. *Cell research* 22, 107-126.

Zhang, Y., Wiggins, B.E., Lawrence, C., Petrick, J., Ivashuta, S., and Heck, G. (2012b). Analysis of plant-derived miRNAs in animal small RNA datasets. *BMC Genomics* 13, 381.

Zheng, G., Qin, Y., Clark, W.C., Dai, Q., Yi, C., He, C., Lambowitz, A.M., and Pan, T. (2015). Efficient and quantitative high-throughput tRNA sequencing. *Nat Methods* 12, 835-837.

Zhu, S., Si, M.L., Wu, H., and Mo, Y.Y. (2007). MicroRNA-21 targets the tumor suppressor gene tropomyosin 1 (TPM1). *J Biol Chem* 282, 14328-14336.

Zitvogel, L., Regnault, A., Lozier, A., Wolfers, J., Flament, C., Tenza, D., Ricciardi-Castagnoli, P., Raposo, G., and Amigorena, S. (1998). Eradication of established murine tumors using a novel cell-free vaccine: dendritic cell-derived exosomes. *Nat Med* 4, 594-600.

ANEXO METODOLÓGICO

Sonda TaqMan de hidrólisis universal:

FAM-TCGCACTGGATACG-MGB-NFQ

Reverse primer:

GTGCAGGGTCCGAGGT

SL-RT primer:

GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACXXXXXX.

(X indica bases específicas para cada ensayo).

Bases XXXXXX del SL-RT primer:

miR-122-5p: AAACAC; miR-21-5p [23nt]: GTCAAC; AACAGC; miR-16-5p: CGCCAA;
tRNA^{Gly}_{GCC}[1-30]: GCGAGA; tRNA^{Glu}_{CUC} [1-31]: CGCCGA; tRNA^{Val}_{CAC} [1-23]:
GATAAC

Forward primers:

miR-122-5p: ggtcgcTGGAGTGTGACAATGGT;
miR-21-5p [23nt]: gccccgTAGCTTATCAGACTGATGT;
let-7i-5p: cgtgccTGAGGTAGTAGTTTGTGCT;
miR-16-5p: cgcgctAGCAGCACGTAAAT;
tRNA^{Gly}_{GCC}[1-30]: ccGCATTGGTGGTTCAGTGGTA;
tRNA^{Glu}_{CUC} [1-31]: TCCCTGGTGGTCTAGTGGTTAGG;
tRNA^{Val}_{CAC} [1-23]: ctgcgcGTTTCCGTAGTGTAGTG.

Las bases en minúscula corresponden a los agregados en los extremos 5' a los efectos de aumentar la T_m en el primer ciclo de PCR.

FAM-tRNA^{Glu} AS:

5' FAM-CCTAACCACTAGACCACCAGG 3'

(esqueleto: DNA)

RNA sintético: tRNA-Gly [1-30]:

5' GCAUUGGUGGUUCAGUGGUAGAAUUCUCGC 3'

(esqueleto: RNA)

RNA sintético biotinilado: BIO-tRNA-Gly [1-30]:

5' Biot/GCAUUGGUGGUUCAGUGGUAGAAUUCUCGC 3'

(esqueleto: RNA)