



Programa de  
Desarrollo de las  
Ciencias Básicas



FACULTAD DE  
**CIENCIAS**  
UDELAR | [fcien.edu.uy](http://fcien.edu.uy)



Tesis de Maestría en Bioinformática

PEDECIBA Bioinformática

## **Perfiles de pequeños RNAs no codificantes circulantes en mujeres con Lupus Eritematoso Sistémico**

Lic. Sofía Dacosta

Institut Pasteur de Montevideo  
Laboratorio de Genómica Funcional

Orientadores: Dra. María Rosa García, Dra. Tamara Fernández y Dr. Sergio Bianchi

Tribunal: Dra. Natalia Rego, Dra. Leticia Pérez y Dr. Pablo Smircich

Setiembre 2024



# Tabla de contenidos

<b>Glosario</b> .....	<b>5</b>
<b>Resumen</b> .....	<b>6</b>
<b>Introducción</b> .....	<b>7</b>
<i>Sistema inmune y tolerancia inmunológica</i> .....	7
<i>Lupus Eritematoso Sistémico</i> .....	9
Etiopatogenia.....	9
Tratamiento .....	12
Problemas asociados al diagnóstico del LES.....	12
<i>RNAs no codificantes</i> .....	13
<i>RNAs no codificantes en el LES</i> .....	14
microRNAs .....	15
<i>exRNAs</i> .....	17
<i>Biopsias líquidas</i> .....	18
<i>RNA-seq</i> .....	19
<i>Otros métodos de detección de miRNAs utilizados en la búsqueda de biomarcadores del LES</i> .....	20
<b>Objetivo</b> .....	<b>21</b>
<i>Objetivos específicos</i> .....	21
<b>Materiales y métodos</b> .....	<b>22</b>
1. <i>Selección de pacientes y obtención de muestras</i> .....	22
2. <i>Preparación de fracciones de sangre periférica</i> .....	22
2.1 Extracción de RNA .....	22
2.2 Secuenciación de pequeños RNAs.....	22
2. <i>Análisis bioinformático</i> .....	23
3.1 Análisis primario .....	23
3.2 Características generales.....	24
3.3 Búsqueda de miRNAs con potencial biomarcador del LES.....	24
<b>Resultados</b> .....	<b>26</b>
1. <i>Selección de las pacientes</i> .....	26
2. <i>Obtención y secuenciación de bibliotecas de pequeños RNAs</i> .....	28
2.1 Control de calidad de las bibliotecas .....	28
2.2 Resultados de secuenciación.....	29
3. <i>Descripción de las características generales de la población de ncrNAs</i> .....	30
3.1 Composición por Biotipo .....	30
3.2 ncrNAs más representados en las muestras .....	31
3.3 Distribución de tamaño de las secuencias .....	33
3.4 Detección de sesgos de tamaños en las bibliotecas .....	35
4. <i>Búsqueda de exRNAs con potencial de biomarcador de LES</i> .....	37

<b>Discusión .....</b>	<b>41</b>
<b>Conclusiones.....</b>	<b>49</b>
<b>Perspectivas .....</b>	<b>50</b>
<b>Referencias.....</b>	<b>51</b>
<b>Anexo.....</b>	<b>57</b>
<b>Agradecimientos.....</b>	<b>73</b>

# Glosario

## Glosario de términos más utilizados

<b>ANA</b>	Anticuerpos anti-nucleares
<b>circRNA</b>	RNA circular
<b>DNA</b>	Ácido desoxirribonucleico
<b>EA</b>	Enfermedad autoinmune
<b>EV</b>	Vesícula extracelular
<b>exRNA</b>	RNA extracelular
<b>LES</b>	Lupus eritematoso sistémico
<b>lncRNA</b>	RNA no codificante largo
<b>miRNA</b>	Micro RNA
<b>mRNA</b>	RNA mensajero
<b>ncRNA</b>	RNA no codificante
<b>nt</b>	Nucleótidos
<b>PBMC</b>	Célula mononuclear de la sangre periférica
<b>piRNA</b>	RNA asociado a proteínas PIWI
<b>pre-miRNA</b>	microRNA precursor
<b>pri-miRNA</b>	microRNA primario
<b>RISC</b>	Complejo de silenciamiento inducido por RNA
<b>rRNA</b>	RNA ribosomal
<b>RNP</b>	Ribonucleoproteína
<b>siRNA</b>	RNA pequeño de interferencia
<b>snRNA</b>	RNA pequeño nuclear
<b>snoRNA</b>	RNA pequeño nucleolar
<b>tRNA</b>	RNA de transferencia

## Resumen

La pérdida de la autotolerancia por parte del sistema inmune genera respuestas aberrantes contra el propio organismo, lo que está en la base patogénica de las enfermedades autoinmunes (EA). En particular, el Lupus Eritematoso Sistémico (LES) es una enfermedad autoinmune sistémica con alta prevalencia, que presenta problemas diagnósticos importantes y aún no tiene un tratamiento curativo. El transcriptoma humano incluye una amplia gama de ácidos ribonucleicos que no codifican para proteínas (ncRNAs). Estas moléculas, o sus fragmentos, han sido descritos a nivel extracelular como moléculas circulantes en el plasma. Desde entonces, la comunidad científica se ha interesado en la detección de estos ncRNAs en distintos biofluidos, con la finalidad de establecer a estas moléculas como analitos identificables en biopsias líquidas mínimamente invasivas, ya que podrían facilitar el diagnóstico, seguimiento y predicción la evolución de distintas enfermedades. Distintos ncRNAs han sido reportados alterados en la circulación de pacientes con LES, y por lo tanto han sido propuestos como prometedores biomarcadores de la enfermedad a través de su detección en biopsias líquidas. En esta tesis se secuenciaron pequeños RNAs de muestras de plasma de 3 pacientes con LES en actividad, 3 pacientes con LES en remisión y 3 controles sanas. Los resultados obtenidos sugieren la existencia de un perfil de miRNAs asociados a la actividad del LES. De este perfil de miRNAs, planteamos que la detección conjunta de miRNAs cuya expresión está aumentada en las pacientes activas, podría contribuir a mejorar el diagnóstico y seguimiento de la enfermedad. Los resultados de esta tesis son un punto de partida para futuros ensayos en los que se buscará validar estos resultados y continuar explorando el potencial de miRNAs como biomarcadores en biopsias líquidas mínimamente invasivas. Además, observamos una disminución de dos miRNAs – miR-150 y miR-106b – específicos de linfocitos B en las pacientes activas. Este resultado es interesante, ya que sus fluctuaciones ya han sido asociadas a procesos relevantes para la patología del LES, como ser la pérdida de la autotolerancia.

# Introducción

## Sistema inmune y tolerancia inmunológica

El sistema inmune es un sistema fundamental para el mantenimiento de la homeostasis del organismo. Actúa respondiendo frente a patógenos y también frente a otras situaciones de peligro como pueden ser el daño tisular y el crecimiento tumoral. Se trata de una red compleja de células y moléculas que trabajan para reconocer y eliminar amenazas, a través de mecanismos y reacciones coordinadas que se conocen en su conjunto como “respuesta inmune”. Para su comprensión, el sistema inmune puede dividirse en sistema inmune innato y sistema inmune adaptativo. En el caso de la defensa frente a un patógeno, la inmunidad innata proporciona protección inmediata contra la invasión del patógeno, y desencadena a la respuesta adaptativa, que se desarrolla más lentamente y proporciona una defensa especializada contra la infección. El sistema inmune innato cuenta con una serie de receptores, tanto celulares como solubles capaces de identificar motivos conservados en los patógenos y responder a ellos, mientras que el sistema inmune adaptativo se compone de linfocitos con receptores muy diversos y variables para antígenos específicos, y algunos de estos receptores pueden ser secretados dando lugar a los anticuerpos <sup>1</sup>.

Una de las propiedades más notables del sistema inmune es que puede reaccionar ante una enorme variedad de patógenos, pero no reacciona, en condiciones de salud, contra los antígenos propios. Esta falta de respuesta a los antígenos propios, también denominada tolerancia inmunológica, se mantiene a lo largo de toda la vida de un individuo. Los mecanismos moleculares mediante los cuales se generan las especificidades de los receptores de linfocitos no evitan la generación de receptores para antígenos propios, sino que durante el proceso normal de maduración de los linfocitos se generan constantemente linfocitos con la capacidad de reconocer autoantígenos.

Cuando los linfocitos encuentran a su antígeno, es posible cualquiera de varios resultados. Los linfocitos pueden activarse para proliferar y diferenciarse en células efectoras y de memoria, lo que lleva a una respuesta inmune efectora, y se dice que los antígenos que provocan tal respuesta son inmunogénicos. Por otro lado, los linfocitos pueden inactivarse funcionalmente o morir luego de encontrarse con el antígeno, lo que da lugar a la tolerancia. Se dice que los antígenos que inducen tolerancia son tolerogénicos. Normalmente, los patógenos son inmunogénicos y los autoantígenos son tolerogénicos. El destino entre activación y tolerancia de los linfocitos está determinado en gran medida por la naturaleza del antígeno y las señales adicionales presentes cuando el antígeno se muestra al sistema inmune. La tolerancia inmunológica a diferentes autoantígenos puede inducirse cuando los linfocitos en desarrollo encuentran estos antígenos en los órganos linfoides primarios, en un proceso llamado tolerancia central, o cuando los linfocitos maduros encuentran autoantígenos en órganos linfoides periféricos (secundarios) o tejidos periféricos, en un proceso llamado tolerancia periférica. La tolerancia central es un mecanismo de tolerancia únicamente a los antígenos propios que están presentes en los órganos linfoides primarios, es decir, la médula ósea y el timo. Además, la tolerancia central es imperfecta y

algunos linfocitos autorreactivos maduran y están presentes en individuos sanos. En este caso, los mecanismos periféricos pueden impedir la activación de estos linfocitos. Tanto los linfocitos T como los linfocitos B están sujetos a estos mecanismos <sup>1</sup>.

Los principales mecanismos de tolerancia central en las células T que reconocen péptidos propios presentados en el complejo mayor de histocompatibilidad son la apoptosis de las células T inmaduras y la generación de células T reguladoras CD4+ que pasan a los tejidos periféricos. La tolerancia periférica se induce cuando las células T maduras (que escaparon a los mecanismos de tolerancia central) reconocen antígenos propios en los tejidos periféricos, lo que da lugar a 3 posibilidades, el linfocito puede sufrir una inactivación funcional o anergia, puede morir por apoptosis, o puede ser suprimido por células T reguladoras mediante distintos mecanismos. Lo que determina que la respuesta sea tolerogénica es la falta de moléculas coestimuladoras durante la activación del linfocito T autorreactivo, así como un desequilibrio entre la participación de los receptores activadores e inhibidores. El reconocimiento de antígenos propios suele ocurrir en ausencia de estas señales, por el contrario, los patógenos provocan reacciones inmunes innatas, lo que lleva a la expresión de moléculas coestimuladoras que promueven la proliferación de células T y la diferenciación en células efectoras <sup>1</sup>.

Los polisacáridos, lípidos y ácidos nucleicos propios son antígenos que a diferencia de las proteínas no son reconocidos por las células T. Estos antígenos deben inducir tolerancia en los linfocitos B para prevenir la producción de autoanticuerpos que los reconozcan. La tolerancia de los linfocitos B también ocurre a nivel central y periférico. Cuando los linfocitos B inmaduros interactúan fuertemente con antígenos propios en la médula ósea, las células B pueden cambiar su especificidad de receptor, en un proceso conocido como edición del receptor. Mediante la edición del receptor, se reduce la posibilidad de que células B autorreactivas potencialmente dañinas abandonen la médula ósea. Se estima que entre el 25% y el 50% de las células B maduras en un individuo normal pueden haber pasado por el proceso de edición de receptores durante su maduración. Si la edición falla, las células B inmaduras que reconocen fuertemente los antígenos propios reciben señales de muerte y mueren por apoptosis. Algunos autoantígenos, pueden reconocerse más débilmente en la médula ósea. Las células B específicas para estos antígenos sobreviven, pero la expresión del receptor de antígeno se reduce y las células dejan de responder funcionalmente. Las células B que reconocen antígenos propios en la periferia también pueden sufrir apoptosis, o pueden activarse a través de receptores inhibidores, impidiendo así la respuesta efectora. Si las células B reconocen un antígeno proteico pero no reciben ayuda de las células T (porque las células T auxiliares han sido eliminadas o son tolerantes), pueden volverse anérgicas. Mediante estos mecanismos el sistema inmune mantiene la tolerancia, evitando la producción de autoanticuerpos <sup>1</sup>.

Si estos mecanismos fallan, el sistema inmune puede atacar las propias células y tejidos del individuo que pretende defender. Estas reacciones se denominan autoinmunidad y las enfermedades que provocan este tipo de reacciones se conocen como enfermedades autoinmunes (EA). Se piensa que las enfermedades asociadas con la producción de autoanticuerpos, como el lupus eritematoso sistémico, son causadas por una tolerancia defectuosa tanto en los linfocitos B como en las células T colaboradoras<sup>1</sup>.

## Lupus Eritematoso Sistémico

El lupus eritematoso sistémico (LES) es una enfermedad autoinmune compleja con una afectación multisistémica, que se caracteriza por presentar una amplia gama de manifestaciones clínicas (Figura 1). Las presentaciones del LES pueden ocurrir a través de síntomas leves que se resuelven espontáneamente, pero también pueden comprometer órganos vitales, poniendo en riesgo la vida de los pacientes. Se trata de una enfermedad crónica, con una evolución sintomática impredecible en períodos de actividad clínica y períodos relativamente asintomáticos de remisión. Su prevalencia es mayor en mujeres en edad fértil, con un predominio femenino de 9 de cada 10 pacientes <sup>2</sup>. La prevalencia del LES varía entre distintos países y grupos étnicos. Se estima que la prevalencia global de LES y la población afectada en mujeres es de 78,73 (con valores de entre 28,61 a 196,33) por 100.000 personas y 3,04 millones de personas, respectivamente. A nivel regional, la prevalencia de LES en mujeres es de 29,21 (con valores entre 6,04 a 84,23) por 100.000 personas en el sur de América Latina <sup>3</sup>. Los datos disponibles de la prevalencia del LES en Uruguay corresponden al departamento de Paysandú, donde se observó una prevalencia de 57,7 /100.000 habitantes en el año 2021 <sup>4</sup>.

### Síntomas del Lupus Eritematoso Sistémico

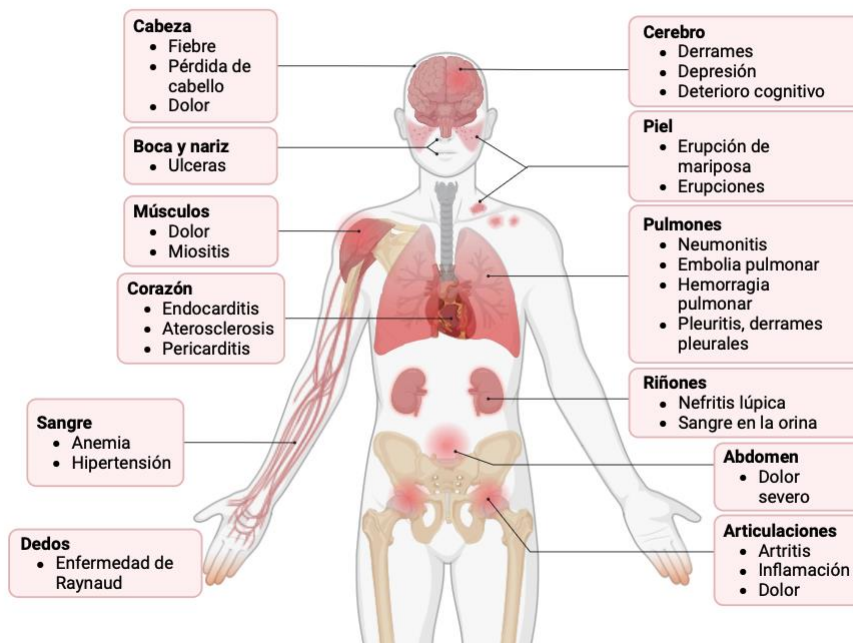


Figura 1. Síntomas del lupus eritematoso sistémico. Imagen creada con BioRender.com.

### Etiopatogenia

La etiología de la enfermedad aún es desconocida. Sin embargo, se han identificado distintos factores de riesgo que pueden contribuir a su desarrollo a través de mecanismos diversos, que conllevan a la producción de anticuerpos contra antígenos nucleares y citoplasmáticos propios, capaces de causar daño en los propios tejidos.

Factores genéticos, epigenéticos, ambientales y relacionados al sexo cromosómico se han asociado a un mayor riesgo de padecer LES. En cuanto a los factores genéticos, alrededor de 200 variantes de riesgo han sido identificadas en el LES <sup>5</sup>. Muchas de ellas corresponden a genes involucrados en la respuesta inmune, y resulta interesante que muchas son compartidas entre distintas EA. También se han descrito patrones epigenéticos alterados en células inmunes de pacientes con LES. La metilación global del DNA se ha visto disminuida en linfocitos de pacientes, lo que se ha correlacionado con la actividad de la enfermedad. Las marcas epigenéticas son dinámicas y pueden verse influenciadas por otros factores, como ser los ambientales <sup>6</sup>. Estudios en gemelos homocigóticos que no presentan concordancia fenotípica en cuanto al desarrollo de la enfermedad, han mostrado que existen cambios en la metilación del DNA y en la expresión en genes relevantes para la patogénesis del LES <sup>7</sup>. La regulación de la expresión mediante microRNAs es otro mecanismo que se ha visto alterado en células del sistema inmune de pacientes con LES, tema que será profundizado más adelante. Dentro de los factores ambientales, la luz ultravioleta, las infecciones con determinados patógenos, particularmente con el virus Epstein-Barr <sup>8</sup>, y la propia flora bacteriana han sido identificados como factores que tienen la capacidad de desencadenar la patología. Se propone que la reactividad cruzada entre autoantígenos y antígenos virales o bacterianos juega un rol en este sentido, y sería capaz de desencadenar y mantener un estado inflamatorio en individuos genéticamente susceptibles <sup>9</sup>. Finalmente, dado que más del 90% de las personas afectadas por el LES son mujeres, el sexo es un factor de riesgo para esta enfermedad, así como para otras patologías autoinmunes. En este sentido influyen factores hormonales, y la presencia de genes del sistema inmune ubicados en el cromosoma X que pueden escapar de la inactivación de dicho cromosoma <sup>10</sup>.

En cuanto a la patogenia del LES, es claro que tanto el sistema inmune innato como el adaptativo se encuentran desregulados, con varios elementos celulares y humorales involucrados. Se han identificado algunas vías clave como impulsoras de la autoinmunidad, como la eliminación deficiente de restos celulares. En esta situación el sistema inmune innato es capaz de reconocer antígenos propios en un contexto inflamatorio, que normalmente no serían visibles para sus receptores, y en consecuencia puede desencadenar la activación de la respuesta adaptativa contra autoantígenos, como ser los ácidos nucleicos propios. La interacción entre células inmunes y sus mediadores también contribuye a la perpetuación de la autoinmunidad y la inflamación crónica. La “firma de interferón”, definida como la sobreexpresión de genes regulados por interferones de tipo 1, es observada con frecuencia en pacientes con LES, y se ha visto implicada en el inicio y progresión de la enfermedad. En la Figura 2 se describen los componentes celulares y humorales que se han visto implicados en la patogénesis del LES <sup>2</sup>.

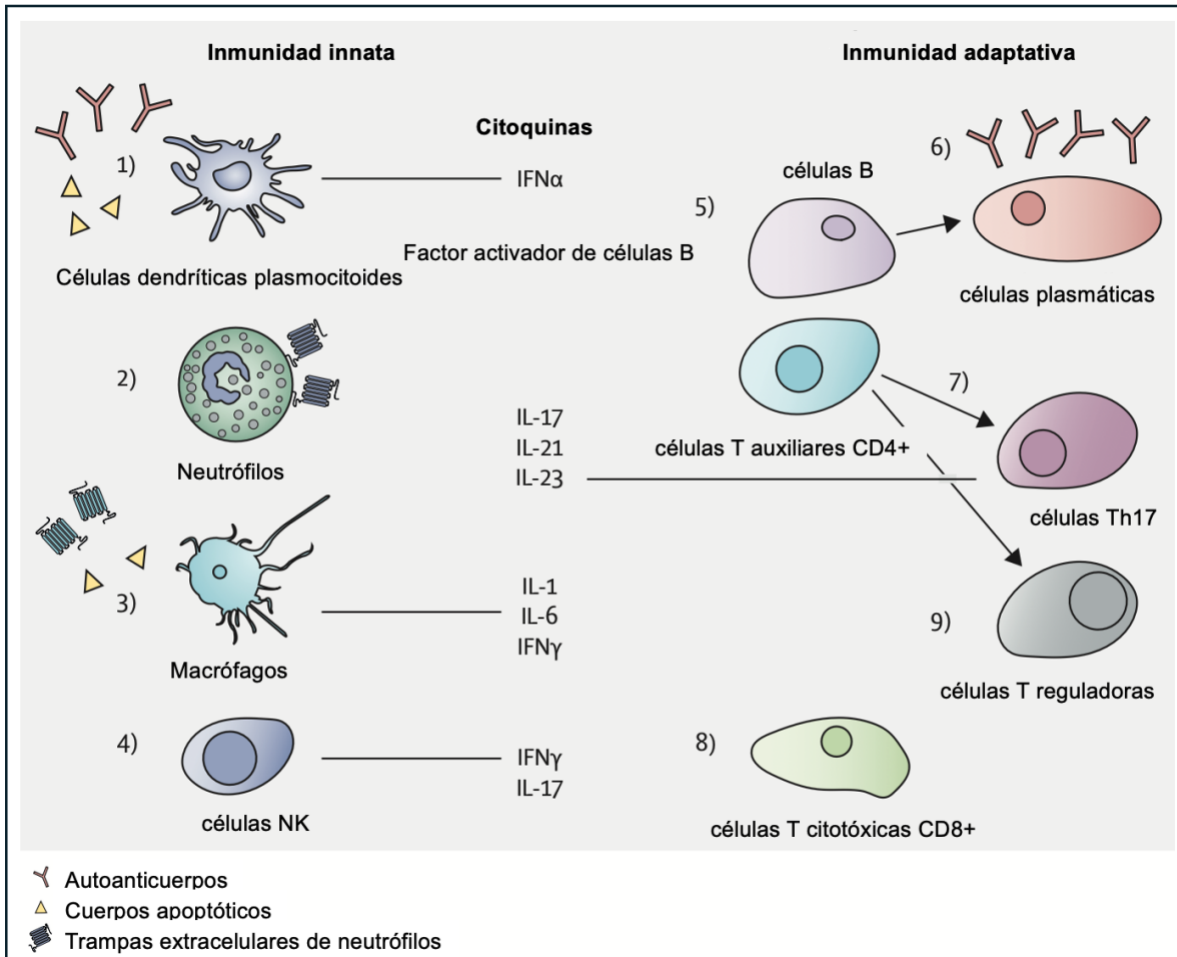


Figura 2. Impulsores celulares y humorales clave de la patogénesis del lupus eritematoso sistémico. Adaptada de Hoi et al., 2024. 1) Las células dendríticas plasmocitoides activadas por DNA y RNA propios dentro de complejos inmunes o cuerpos apoptóticos, son fuentes primarias de interferones tipo I (por ej.,  $IFN\alpha$ ), que amplifican y perpetúan el ciclo de las células autorreactivas y la producción de autoanticuerpos. 2) Los neutrófilos y sus trampas extracelulares son fuentes potentes de ácidos nucleicos, que activan a las células dendríticas plasmocitoides y promueven la activación de linfocitos T y la producción de autoanticuerpos. 3) Los macrófagos contribuyen tanto al inicio como a la progresión del lupus. Defectos en la capacidad de los macrófagos para eliminar las células apoptóticas pueden provocar la acumulación de material apoptótico, que sirve como fuente de autoantígenos. Además, los macrófagos son productores clave de citoquinas proinflamatorias (por ej.,  $IL-1$ ,  $IL-6$  e  $IFN\gamma$ ), actúan como células presentadoras de antígenos y liberan mediadores inflamatorios que causan daño tisular y fibrosis. 4) Las células NK desempeñan un papel importante en la regulación de la actividad y la supervivencia de otras células inmunes y producen citoquinas (p. ej.,  $IFN\gamma$  e  $IL-17$ ). 5) Los linfocitos T auxiliares  $CD4+$  y las células B autorreactivas amplifican la respuesta inmune. Su activación depende de la participación adecuada de las moléculas coestimuladoras. 6) El proceso de maduración de las células plasmáticas secretoras de anticuerpos depende de la proliferación y diferenciación de linfocitos B bajo la influencia de citoquinas. El cambio de clase posterior y la hipermutación

somática son procesos importantes en la maduración de los linfocitos B, que requieren la ayuda de linfocitos T en el centro germinal. 7) Las células Th17, un subconjunto de células T auxiliares CD4+ que producen citoquinas como IL-17 o IL-21, promueven la inflamación y la diferenciación de células plasmáticas secretoras de anticuerpos. Las células Th17 se activan, expanden y mantienen bajo la influencia de la IL-23. 8) Las células T citotóxicas CD8+ pueden mediar en el daño tisular de la piel, los riñones o el cerebro. 9) Las células T reguladoras, un subconjunto de células T CD4+, son esenciales para mantener la tolerancia inmune y prevenir respuestas autoinmunes. Si el contexto favorece que las células T CD4+ se diferencien mayormente a Th17, la población de células T reguladoras se ve disminuida.

## **Tratamiento**

Los objetivos del tratamiento actualmente se centran en el control temprano de la actividad de la enfermedad, la prevención de brotes y la minimización de las comorbilidades y toxicidad relacionada con el tratamiento. Las terapias farmacológicas se seleccionan teniendo en cuenta el nivel de actividad y la gravedad de la enfermedad, los órganos más comprometidos, así como las circunstancias individuales de cada paciente. Los glucocorticoides se utilizan para tratar el LES debido a su capacidad para controlar rápidamente la actividad de la enfermedad e inducir la remisión. Sin embargo, presentan numerosos efectos adversos, especialmente a nivel metabólico, ocular, cerebrovascular, cardiovascular y musculoesquelético, así como un aumento en el riesgo de infecciones. Para pacientes clínicamente activos pero cuyos órganos no están en peligro, se suele lograr la reducción gradual de la dosis de glucocorticoides utilizando inmunosupresores convencionales como el metotrexato. Por otra parte, la hidroxicloroquina es ampliamente utilizada como tratamiento de primera línea del LES por sus beneficios para el control de los síntomas y la prevención de daños en distintos órganos y tejidos con un menor riesgo de efectos secundarios que los glucocorticoides. El tratamiento con anticuerpos monoclonales como belimumab, rituximab o anifrolumab, mejora el control de la actividad en pacientes con LES persistentemente activo que no responden a tratamientos convencionales. Los tratamientos actuales en conjunto con los esfuerzos por promover el diagnóstico temprano de la enfermedad han mejorado la calidad y pronóstico de vida de los pacientes. De todas formas, a la fecha no existe un tratamiento curativo de la enfermedad. Si bien el control temprano de la enfermedad es importante para evitar el daño en órganos y tejidos, el diagnóstico del LES no suele ser sencillo <sup>2</sup>.

## **Problemas asociados al diagnóstico del LES**

Uno de los desafíos más comunes e importantes reportados por médicos y pacientes está relacionado con la dificultad diagnóstica. En la mayoría de los pacientes, los síntomas como el cansancio, la falta de apetito y la pérdida de peso, en conjunto con manifestaciones mucocutáneas y musculoesqueléticas representan las primeras manifestaciones clínicas. Estos síntomas pueden incluir fatiga, erupción cutánea, úlceras bucales, alopecia, dolor en las articulaciones y mialgia. Parte del desafío en el diagnóstico del lupus es que no todas las manifestaciones clínicas ocurren simultáneamente y, a menudo, puede transcurrir un tiempo considerable entre las distintas manifestaciones, de manera que normalmente los pacientes

deben someterse a reiteradas intervenciones diagnósticas, al tiempo que la enfermedad evoluciona en períodos de actividad y remisión. Incluso una vez realizado el diagnóstico, sigue siendo difícil predecir qué pacientes progresarán en la enfermedad <sup>2</sup>.

Para facilitar el diagnóstico y seguimiento de la enfermedad, existen herramientas utilizadas por los profesionales de la salud que permiten medir la actividad del LES. En relación a la actividad de la enfermedad, el criterio más utilizado actualmente es el SLEDAI (*Systemic Lupus Erythematosus Disease Activity Index*). Este criterio establece un índice de actividad ponderado compuesto por 24 descriptores clínicos y paraclínicos, cada uno de los cuales tiene un peso diferente, asociado al tipo de manifestación pero no a la gravedad. En teoría, la actividad de la enfermedad puede oscilar entre 0 y 105 puntos, pero cuando este índice se califica adecuadamente es raro encontrar un paciente con una puntuación superior a 20 <sup>11,12</sup>. En 2019 la European League Against Rheumatism (EULAR) en conjunto con el American College of Rheumatology (ACR) definieron nuevos criterios para facilitar el diagnóstico del Lupus Eritematoso Sistémico<sup>13</sup>. En los mismos se incluye como criterios necesarios para el diagnóstico del LES la determinación de anticuerpos anti-nucleares (ANA) positivos con un título mayor o igual a 1:80, ya que la presencia de estos anticuerpos es considerada un sello fundamental del LES. Los ANA se han detectado al comienzo de la enfermedad de prácticamente todos los pacientes. Una vez detectados en su conjunto, para profundizar en el diagnóstico, se evalúa la presencia de anticuerpos anti-DNA doble cadena, anti-nucleosoma, anti-histonas, anti-Sm y anti-RNP (dirigidos contra ribonucleoproteínas nucleares que forman parte del espliceosoma), entre otros. Otros autoanticuerpos encontrados, con menor frecuencia, en pacientes con LES son anti-Ro y anti-La (o anti-SSA y anti-SSB por *Sjögren's Syndrome related antigen A/B* respectivamente) inmunoglobulinas dirigidas contra las proteínas Ro y La asociadas a YRNAs, también compartidos con otras enfermedades autoinmunes <sup>2</sup>.

En las últimas décadas, surgió un gran interés por la identificación de ncRNAs como biomarcadores en distintos biofluidos, con la finalidad de establecer a estas moléculas como analitos identificables en biopsias líquidas mínimamente invasivas. Su estudio puede proporcionar evidencia sobre nuevos mecanismos patogénicos, así como contribuir a la identificación de biomarcadores de utilidad tanto en la terapia como en el diagnóstico de distintas enfermedades, contribuyendo a la detección temprana de la enfermedad, así como a su seguimiento <sup>14</sup>. En el contexto del LES, la identificación de estas moléculas resulta sumamente atractiva, dada la necesidad de mejorar los tiempos de diagnóstico de los pacientes, así como también el seguimiento de la enfermedad. De hecho, distintos ncRNAs han sido reportados alterados en la circulación de pacientes con LES, y por lo tanto han sido propuestos como prometedores biomarcadores de la enfermedad a través de su detección en biopsias líquidas <sup>15-17</sup>.

## **RNAs no codificantes**

Los RNAs no codificantes (ncRNA) son un grupo heterogéneo de transcritos que no son traducidos a proteínas, sino que actúan como reguladores y catalizadores de múltiples procesos celulares (Tabla 1). En los últimos años, con el advenimiento de las técnicas de secuenciación profunda y

de alto rendimiento, se ha puesto énfasis en la caracterización de las poblaciones de ncRNAs tanto en el medio intra como extracelular.

Tabla 1. Principales funciones de los ncRNAs. rRNA: RNA ribosomal; tRNA: RNA de transferencia; snRNA: RNA pequeño nuclear; snoRNA: RNA pequeño nucleolar; lncRNA: RNA largo no codificante; circRNA: RNA circular; piRNAs: RNAs asociados a proteínas PIWI; miRNAs: microRNAs.

ncRNA	Funciones principales
rRNA	Componentes de los ribosomas, directamente involucrados en la síntesis proteica.
tRNA	Decodifican las secuencias de los mRNAs y transfieren aminoácidos a las cadenas peptídicas nacientes durante la traducción.
snRNA	Participan del <i>splicing</i> durante la maduración de los mRNAs.
snoRNA	Participan de la modificación de rRNAs, tRNAs y mRNAs <sup>18</sup> .
lncRNA	Regulan la expresión génica a nivel transcripcional, postranscripcional y epigenético. Actúan como esponjas de miRNAs inhibiendo su función, y pueden servir de andamio para la interacción entre proteínas <sup>14</sup> .
circRNA	Participan de la regulación de la expresión génica al actuar como esponjas de miRNAs, mRNAs y proteínas de unión a RNA <sup>14</sup> .
YRNAs	Son importantes para la replicación del DNA <sup>19</sup> y el control de la calidad de los RNAs <sup>20</sup> .
piRNAs	Actúan como un mecanismo de defensa al silenciar elementos genéticos móviles, como virus y transposones en células germinales.
miRNAs	Participan de la regulación génica a nivel postranscripcional.

Alteraciones en los niveles de distintos ncRNAs han sido asociados con múltiples procesos patológicos. Muchos ncRNAs participan de la regulación de la respuesta inmune y procesos inflamatorios <sup>21</sup>, por lo que han sido foco de estudio en distintas EA complejas, incluyendo el LES. A continuación, se describe brevemente la relación de distintos ncRNAs en el LES.

### RNAs no codificantes en el LES

Distintos ncRNAs han sido asociados con el LES. Particularmente, se conoce que algunos miRNAs, lncRNAs y circRNAs actúan como reguladores epigenéticos y pueden tener incidencia en el desarrollo y progresión de la enfermedad. Múltiples estudios han examinado los niveles y composición de RNA en muestras de sangre periférica (sangre completa, plasma o PBMCs) y medula ósea de pacientes, y han encontrado que la expresión aberrante de ciertos ncRNAs en diferentes células del sistema inmune contribuye a procesos patológicos del LES. Para algunos de estos ncRNAs se ha demostrado que su expresión anormal puede afectar vías de señalización celular, incluidas IFN-I, MAPK y WNT, afectando la activación y diferenciación celular, la producción de citoquinas y la apoptosis <sup>16</sup>. Los circRNAs también han sido caracterizados en

distintas células del sistema inmune de pacientes con LES, observándose que estos también presentan cambios en su expresión respecto a controles sanos, y que estos cambios pueden afectar el funcionamiento de células del sistema inmune <sup>22</sup>. Si bien existen distintos mecanismos por los cuales los lncRNAs y circRNAs ejercen sus funciones, resulta interesante que ambos pueden actuar como esponjas de miRNAs, afectando redes de miRNAs-mRNAs. Los ncRNAs más estudiados en el contexto del LES son probablemente los miRNAs. El foco de este trabajo está puesto en estas moléculas, por lo que a continuación se describe con mayor profundidad sus características y lo que se conoce respecto a su relación con el LES.

### **microRNAs**

Los microRNAs son una clase de RNAs reguladores pequeños no codificantes que están altamente conservados entre diferentes especies y tienen un papel crucial en la regulación de la expresión génica. Actúan inhibiendo la expresión de sus mRNAs blanco a través de un mecanismo dependiente de su secuencia. A su vez, un miRNA puede tener múltiples genes blanco y viceversa, dando lugar a una compleja red de regulación, que se adapta al contexto celular, y que depende de distintos factores, como la disponibilidad de los mRNAs blancos, la localización subcelular, y la actividad de los complejos de silenciamiento inducidos por miRNAs.

Los miRNAs se originan a partir de moléculas precursoras más largas (pri-miRNAs) que son transcritas por la RNA Polimerasa II. Por lo general, son transcritos que se pliegan sobre sí mismos por auto complementariedad, dando lugar a estructuras de tipo horquilla. Los pri-miRNAs son procesados por la endonucleasa nuclear Drosha dando lugar a un pre-miRNA de unos 60 a 120 nt, también con estructura de horquilla. Los pre-miRNAs son exportados al citoplasma donde vuelven a ser procesados, esta vez por la endonucleasa Dicer, dando lugar al miRNA maduro, un dúplex de RNA de 20 a 24 nt con extremos protuyentes. El miRNA resultante es cargado en la proteína Argonauta que retiene una sola de las hebras del dúplex, y en conjunto con otras proteínas forman el complejo de silenciamiento inducido por RNA (RISC, por *RNA-induced silencing complex*). En este complejo los miRNAs guían el silenciamiento génico de mRNAs, al unirse por complementariedad de secuencia a sus mensajeros blanco, dando lugar a inhibición traduccional de los mensajeros y su eventual degradación, en eucariotas <sup>23</sup>.

Alteraciones en los niveles fisiológicos de distintos miRNAs han sido relacionados con diversos procesos patológicos, entre ellos enfermedades cardiovasculares, neurológicas, entre otras. En particular, muchas investigaciones se han centrado en el vínculo entre microRNAs y cáncer, y distintos miRNAs han sido clasificados como OncomiRs o miRNAs supresores de tumor, dependiendo del tipo de mRNA que silencian <sup>24</sup>. Se sabe que muchos miRNAs están involucrados en la maduración y función de las células inmunes y hay evidencia que señala que estos pueden influir en los procesos inmunopatológicos del LES y otras EA <sup>25</sup>.

La lista de microRNAs cuyos niveles han sido encontrados alterados en células inmunes de pacientes con LES es larga. Sin embargo, no para todos ellos la participación en mecanismos patológicos ha sido demostrada. Choi et al. (2023) revisa el conocimiento actual con relación a

miRNAs desregulados en LES, y sostiene que a pesar de las extensas investigaciones realizadas sobre los niveles de miRNAs, su importancia y papel potencial en el diagnóstico del LES siguen sin ser claras. Esto lo atribuyen en parte a resultados contradictorios, a falta de acuerdo en el tipo de muestra a utilizar y también a falta de evidencia directa de la implicancia de los miRNAs en estudio en la patología. A continuación, se describen solamente algunos de los miRNAs que han sido relacionados al LES.

Para muchos miRNAs la relación con el funcionamiento aberrante de las células en las que fueron estudiados no ha sido demostrada, sino inferida mediante la predicción de sus mRNAs blanco, o simplemente fueron reportados cambios en su expresión respecto a controles sanos<sup>17</sup>. Por ejemplo, ha sido descrito que miR-124a, 125a, 222, 125b, 146a y 155 están regulados negativamente en los neutrófilos de pacientes con LES en comparación con los neutrófilos de controles sanos<sup>26</sup>. Sin embargo, el rol de estos miRNAs en la función aberrante de los neutrófilos, en particular durante la formación de trampas extracelulares de neutrófilos, no ha sido evidenciada. Por otra parte, algunos trabajos realizados en linfocitos T de pacientes muestran que ciertos miRNAs sí parecen estar involucrados en la activación alterada de estos linfocitos, como el caso de miR-145 cuya expresión se reportó disminuida al tiempo que su blanco STAT1 (por *signal transducer and activator of transcription 1*) se vio aumentado en estos linfocitos<sup>27</sup>. STAT1 juega un rol muy importante durante la activación de células Th1, que promueven la inflamación mediante la producción de IFN- $\gamma$ . Los niveles de IFN- $\gamma$  han sido reportados aumentados en PBMCs de pacientes con LES, y correlacionan con la actividad de la enfermedad<sup>28</sup>. Distintos miRNAs han sido reportados alterados también en linfocitos B. Por ejemplo, la expresión de miR-29 se vio disminuida en linfocitos B de pacientes al compararla con la de linfocitos B de controles sanos. La disminución de miR-29 conduce al aumento de la expresión de su proteína diana CRKL (por *Crk-like protein*) en estas células, lo que resulta en una mayor secreción de IgG por parte de los linfocitos B secretores de anticuerpos<sup>29</sup>. Otro miRNA cuya expresión ha sido reportada disminuida en linfocitos B es miR-15b. Este miRNA regula la expresión de la Ciclina D3 mediante la degradación de su mRNA. Esta proteína desempeña un papel importante en la proliferación de las células B, y se ha visto sobreexpresada en linfocitos B de pacientes con LES, al tiempo que la expresión de miR15-b se encuentra disminuida. De esta forma, cambios en los niveles de este microRNA podrían tener impacto a nivel del ciclo celular, contribuyendo a las alteraciones de las células B<sup>30</sup>.

Otros trabajos han estudiado la expresión de miRNAs en PBMCs (células mononucleares de sangre periférica) sin mayor distinción del tipo celular. En las PBMCs se incluyen poblaciones heterogéneas de células, tanto de linaje linfoide (linfocitos T, B y células NK) como de linaje mieloide (principalmente monocitos). En estas células, se han identificado distintos miRNAs con expresión alterada respecto a los niveles en controles sanos. Algunos de ellos son miR-16, miR-21, miR-155 que fueron reportados aumentados en PBMCs de pacientes, mientras que la expresión de miR-141 se vio disminuida<sup>31</sup>. En ocasiones los resultados obtenidos por distintos trabajos en PBMCs de pacientes son contradictorios, por ejemplo, la expresión de miR-155 ha sido reportada tanto aumentada<sup>31</sup> como disminuida<sup>32</sup>. La heterogeneidad de la enfermedad representa un desafío en este sentido, y además este efecto podría tener que ver con variaciones en la composición de los distintos tipos celulares en las PBMCs de los pacientes. Lo mismo ocurre

en el caso de los estudios realizados en plasma de pacientes, donde a la heterogeneidad de la enfermedad, se suma la complejidad de este fluido extracelular, compuesto por miRNAs provenientes de distintas células. Por ejemplo, el miRNA-181a ha sido reportado tanto sub<sup>33</sup> como sobrepresado<sup>34</sup> en plasma de pacientes con LES, y lo mismo se ha observado para el miR-223 y el miR-16<sup>15,33</sup>.

Muchos de los miRNAs que se han reportado que tienen una expresión diferencial en pacientes con LES, también han sido reportados alterados en contexto de otras enfermedades. Un claro ejemplo de esto es miR-21, al cual Jenike y Halushka se refieren como “un biomarcador no específico de todas las enfermedades” luego de mostrar que este ha sido propuesto como un biomarcador predictivo o pronóstico de al menos 29 enfermedades diferentes<sup>35</sup>. Los autores adjudican la falta de especificidad a la expresión ubicua de miR-21, sugiriendo que resulta sumamente importante limitar el estudio a aquellos miRNAs que son expresados específicamente por células involucradas en los procesos patológicos de interés, y evitar aquellos miRNAs de expresión ubicua, sobre todo al buscar biomarcadores en el medio extracelular.

## **exRNAs**

Hace varias décadas que se describió que el RNA también se encuentra fuera de las células<sup>36</sup>, dando lugar a los denominados RNAs extracelulares (exRNAs). Estas moléculas son una fuente prometedora de biomarcadores para diferentes enfermedades que pueden ser detectadas en biopsias líquidas mínimamente invasivas. Distintos tipos de RNAs están presentes en el fluido extracelular, y su abundancia puede variar en los diferentes biofluidos<sup>37,38</sup>. Dentro del exRNA se encuentran fragmentos de mRNAs, lncRNAs, rRNAs, YRNAs y tRNAs. Algunas de estas moléculas también pueden encontrarse en su longitud completa como algunos mensajeros de tamaño menor a 1.000 nucleótidos. Los circRNAs y miRNAs también se encuentran en el medio extracelular, así como otros RNAs pequeños no codificantes como los *vault* RNAs, snRNAs, snoRNAs y piRNAs.

La composición de RNAs del medio extracelular no es idéntica a la que se encuentra en el interior de las células, encontrándose distintos RNAs enriquecidos en el medio extracelular<sup>39</sup>. En el plasma, estudios de secuenciación de pequeños RNAs (small-RNA-Seq) han mostrado que los exRNAs se componen mayormente de YRNAs, en particular alrededor del 50 % de las lecturas mapean al RNY4, mientras que el resto de las lecturas mapean a mayormente a miRNAs<sup>38,40</sup>. Las ribonucleasas (RNAses) son proteínas que degradan moléculas de RNA y son muy abundantes en el entorno extracelular, donde degradan rápidamente a todos aquellos RNAs que se encuentren desprotegidos. De forma que, la existencia de exRNA implica su protección frente a la actividad de las ribonucleasas. Esta protección está dada por la compartimentalización del exRNA en distintas estructuras, como lipoproteínas y distintos tipos de vesículas extracelulares (EVs) que pueden provenir de la membrana plasmática celular como los ectosomas o micro vesículas o de compartimentos endosómicos como los endosomas<sup>41</sup>. El exRNA también puede existir por fuera de estas estructuras, y ser protegido por su interacción con proteínas en ribonucleocomplejos<sup>42</sup> o ser protegido por su propia estructura<sup>43</sup>. De hecho, la mayoría del exRNA es extravésicular<sup>39</sup>, y la mayoría de los miRNAs en medios de cultivo celular o que circulan en el plasma sanguíneo

humano se encuentran formando complejos ribonucleoproteicos con proteínas como AGO2 <sup>44,45</sup>, mientras que su abundancia en EVs es muy baja, estimando que para encontrar una copia de un miRNA dado deben contarse cerca de 1000 EVs <sup>46</sup>.

Distintos trabajos han mostrado que los perfiles de exRNAs son distintos a la composición encontrada en las células que los liberan, lo que puede sugerir que las células seleccionan activamente RNAs específicos y los exportan al exterior celular. El empaquetamiento de RNAs específicos en EVs ha sido ampliamente estudiado <sup>47-49</sup>. Pero este no es el único mecanismo capaz de explicar la composición del exRNA. Las células podrían también liberar sus RNAs más abundantes siguiendo gradientes de concentración, o de manera pasiva independiente de las EVs, debido a la muerte celular. Diferentes exRNAs tienen distintos grados de protección frente a la degradación, de forma que el procesamiento extracelular también contribuye a los perfiles de exRNAs<sup>50</sup>. Al inhibir la actividad de las ribonucleasas extracelulares, los perfiles de exRNA cambian dramáticamente, y particularmente reflejan los perfiles intracelulares, lo que sugiere que todo el RNA celular puede liberarse y la acción de ribonucleasas es responsable de su degradación y fragmentación en RNAs más cortos <sup>51</sup>.

La función del exRNA es actualmente un tema de discusión. Varios estudios han reportado la transferencia funcional de miRNAs y mRNAs entre células de mamífero a través de EVs, proponiendo un rol de estas moléculas en la comunicación celular <sup>52-56</sup>. La transferencia de exRNAs también fue estudiada en el microambiente tumoral, donde se propone que cumple funciones importantes para la progresión de la enfermedad <sup>57,58</sup>. A pesar del prometedor rol de los exRNAs en la comunicación celular, el estudio de estos procesos presenta grandes desafíos técnicos. Por ejemplo, resulta difícil saber si es realmente el RNA de las EVs el responsable de los cambios observados en las células receptoras, ya que otras macromoléculas presentes en las EVs, como las proteínas, podrían estar involucradas. En el caso de los miRNAs y mRNAs, aún no está claro cómo estos pueden tener efectos relevantes en las células receptoras dada su baja concentración en las EVs <sup>59</sup>.

Sea para cumplir una función o no, distintos exRNAs son capaces de sobrevivir a la elevada concentración de ribonucleasas presentes en el medio extracelular. Su estabilidad y el acceso relativamente no invasivo a distintas moléculas de exRNA las convierten en una clase interesante de biomarcadores que pueden ser detectados en fluidos extracelulares.

## **Biopsias líquidas**

El término biopsia líquida refiere a cualquier fluido extraído de un organismo con la intención de investigarlo para obtener información médica. A diferencia de las biopsias tradicionales, que requieren la extracción de tejido sólido, las biopsias líquidas se basan en el análisis de componentes moleculares y celulares presentes en fluidos corporales. Las biopsias líquidas ofrecen una alternativa mucho menos invasiva a las biopsias de tejido, y minimizan los riesgos asociados a las intervenciones quirúrgicas de las biopsias convencionales. Al ser más simples y

menos invasivas, podrían facilitar el diagnóstico precoz de la enfermedad, así como permitir un seguimiento continuo que no implique un gran desgaste para los pacientes, y de este modo, una mayor adhesión en el control evolutivo. Existen una gran cantidad de fluidos corporales que pueden utilizarse con este fin: sudor, calostro, leche materna, líquido amniótico, bilis, lágrimas, saliva, plasma, suero, orina, entre otros. También existen una gran cantidad de analitos que pueden ser detectados en estos fluidos como ser: células tumorales circulantes, DNA y RNA libre de células, proteínas, etc.

Este trabajo evalúa el potencial de exRNAs presentes en el plasma a ser utilizados como biomarcadores del LES. Un biomarcador se define como una molécula medible capaz de indicar la presencia de una patología o su grado de avance, o que puede ser utilizado como indicador de susceptibilidad para una enfermedad antes de que esta se desarrolle. El plasma se separa a partir de sangre total mediante centrifugación seriada. Se obtiene una fracción inferior que contiene eritrocitos, seguida de una “capa leucocitaria” que contiene principalmente leucocitos y una capa superior que es el plasma rico en plaquetas (PRP). Cuando se separa el PRP y se vuelve a centrifugar las plaquetas sedimentan y los sobrenadantes pueden retenerse. Diferentes velocidades de centrifugación y cantidad de centrifugaciones seriadas conducen a la generación de plasma pobre en plaquetas (PPP) o libre de plaquetas (PFP) dependiendo del número de plaquetas residuales. Existen distintos métodos que permiten detectar a los exRNAs en las muestras de pacientes.

## **RNA-seq**

La secuenciación masiva ha revolucionado la biología y la medicina en los últimos años, proporcionando una precisión a nivel de base única para la comprensión de las secuencias de ácidos nucleicos, con un rendimiento muy alto, permitiendo la adquisición de una enorme cantidad de datos en poco tiempo. La secuenciación de RNA o RNA-seq es un método ampliamente utilizado para estudiar la expresión de RNAs, descubrir nuevas especies de RNA, y se trata del método más utilizado actualmente para la caracterización del exRNA en biofluidos en búsqueda de biomarcadores de terapia y diagnóstico. RNA-seq comprende un grupo diverso de protocolos únicos. En general, un flujo de trabajo de RNA-seq para pequeños RNAs se puede dividir a grandes rasgos en cuatro pasos: extracción del RNA, preparación de bibliotecas, secuenciación y análisis de datos. Mediante la utilización de distintos protocolos es posible secuenciar RNAs con distintas características, por eso es importante tener presente que los resultados obtenidos mediante ningún método de secuenciación son capaces de reflejar la composición exacta de RNAs de las muestras. En este trabajo se utiliza un protocolo de secuenciación de pequeños RNAs, realizando una selección por tamaño, y además se trata de un método de secuenciación capaz de detectar RNAs cuyos extremos sean 5' monofosfato y 3' OH, favoreciendo la secuenciación de miRNAs y otros pequeños RNAs con estas características. En el caso particular de la búsqueda de biomarcadores del LES, los métodos más comúnmente utilizados han sido la RT-qPCR y los Microarrays.

## **Otros métodos de detección de miRNAs utilizados en la búsqueda de biomarcadores del LES**

La mayoría de los trabajos que buscan caracterizar los perfiles de miRNAs en distintas muestras de pacientes con LES, se basaron en la tecnología de microarrays. Esta tecnología se basa en la hibridación de los miRNAs, o el cDNA generado a partir de estos, con sondas complementarias inmovilizadas en la superficie de un chip. Los miRNAs se marcan con moléculas fluorescentes y se detecta su emisión de fluorescencia unidos en diferentes posiciones en los portaobjetos, pudiéndose evaluar los miRNAs y sus cantidades relativas en la muestra estudiada analizando los datos de la señal de fluorescencia <sup>60</sup>. Esta técnica por lo tanto no permite la detección de nuevas moléculas de RNA, sino que para la detección de un miRNA dado es necesario conocer su secuencia e incluirla en el chip. A diferencia de los microarrays, el RNA-seq no requiere el uso de sondas, haciendo posible la detección de transcritos nuevos o desconocidos. Mediante RNA-seq es posible detectar una mayor cantidad de genes, especialmente en el caso de genes con baja expresión <sup>61</sup>. Otra metodología utilizada por distintos trabajos es la detección de miRNAs mediante Stem Loop RT-qPCR <sup>62</sup>. En este caso también es necesario conocer cuál es el miRNA que se quiere detectar, y la diferencia en la expresión a detectar entre muestras puede ser una limitante. Este método también suele utilizarse para validar resultados obtenidos mediante RNA-seq.

## **Objetivo**

Caracterizar los perfiles de ARNs circulantes en el plasma de pacientes con lupus eritematoso sistémico en actividad y en remisión. Analizar si en estos dos estados de la enfermedad se presentan diferencias cualitativas y/o cuantitativas, así como también las diferencias que puedan tener con los perfiles de individuos sanos. Contribuir a la selección de biomarcadores que faciliten el diagnóstico y seguimiento de la enfermedad.

### **Objetivos específicos**

1. Seleccionar y obtener muestras de sangre periférica de mujeres con diagnóstico confirmado de LES en actividad y en remisión clínica y mujeres clínicamente sanas.
2. Preparar bibliotecas de pequeños RNAs a partir del plasma de las pacientes y controles. Secuenciar las bibliotecas obtenidas.
3. Caracterizar las poblaciones de pequeños RNAs circulantes en las muestras de las tres condiciones.
4. Seleccionar pequeños RNAs con potencial como biomarcadores para su posterior evaluación.

# Materiales y métodos

## 1. Selección de pacientes y obtención de muestras

Se incluyeron en el estudio mujeres con diagnóstico confirmado de Lupus Eritematoso Sistémico atendidas en la Policlínica de Enfermedades Autoinmunes del Hospital Maciel. Se obtuvieron 10 ml de sangre periférica por venopunción en estrictas condiciones de asepsia de 3 pacientes con LES en actividad (A), 3 pacientes con LES en remisión (R) y 3 controles normales (N), en tubos anticoagulados con EDTA. A través de la interacción con las médicas tratantes se accede a datos clínicos y paraclínicos de interés. Se solicitó la firma de consentimiento informado a todas las participantes, y el proyecto cuenta con la aprobación del Comité de Ética del Hospital Maciel.

## 2. Preparación de fracciones de sangre periférica

La sangre periférica extraída a las pacientes y controles normales se procesó en las primeras 2 horas luego de su extracción. El plasma se obtuvo por centrifugación a 1100 g durante 15 minutos a temperatura ambiente y se almacenó a -80 °C con inhibidor de RNAsas hasta su utilización.

### 2.1 Extracción de RNA

A partir de 2 ml de plasma de cada paciente o control se extrajo el RNA circulante utilizando el kit “Quick-cfRNA Serum and Plasma” (ZYMO Research, USA) siguiendo las instrucciones del fabricante. Este kit permite extraer RNAs de un largo  $\geq 17$  nt, incluidos los RNAs unidos a proteínas, exosomales, microRNAs y otros pequeños RNAs. El RNA se cuantificó mediante fluorimetría utilizando el kit “Qubit miRNA Assay”.

### 2.2 Secuenciación de pequeños RNAs

Se elaboraron bibliotecas para secuenciar los RNAs circulantes en la plataforma MiSeq (Illumina) utilizando el kit “NEBNext small RNA Library Prep Set for Illumina” (New England BioLabs) partiendo de 50 ng de RNA total. Este procedimiento comprende los siguientes pasos: la ligación del adaptador en el extremo 3'-OH de los fragmentos de RNA; la hibridación de un cebador para la retrotranscripción complementario a este adaptador, seguido de la ligación de otro adaptador en el extremo 5'-P; la síntesis de la primera hebra de DNA copia mediante retrotranscripción a partir de los híbridos RNA – cebador de retrotranscripción; y finalmente, la amplificación por PCR con el agregado de secuencias únicas identificadoras (*index*) en uno de los cebadores que permiten identificar a qué muestra corresponden las lecturas de secuenciación (Figura 3). Se purificaron las bibliotecas obtenidas utilizando el kit “Gel and PCR Clean up” (Machery Nagel) utilizando una relación *binding buffer:muestra* de 7:1. Se seleccionó el tamaño de los fragmentos a secuenciar en un gel de poliacrilamida “Novex TBE Gels, 6%” (Thermo). La corrida se realizó durante 1 hora a 120 V. Las bandas de 150 a 200 nucleótidos se recortan y purifican. Estas bandas corresponden a fragmentos de entre 20 y 80 pares de bases sin adaptadores. Se realizó un control del tamaño y de la concentración de cada biblioteca utilizando un chip de DNA de alta

sensibilidad, en un Bioanalyzer 2100 (Agilent). Las bibliotecas se cuantificaron con el kit Qubit “High Sensitivity dsDNA Assay”. Se realizaron pools de 12 nM que fueron diluidos de acuerdo con las indicaciones del fabricante y secuenciados en una corrida *single-end*, durante 50 ciclos utilizando el kit “MiSeq Reagent Kit v2” (Illumina).

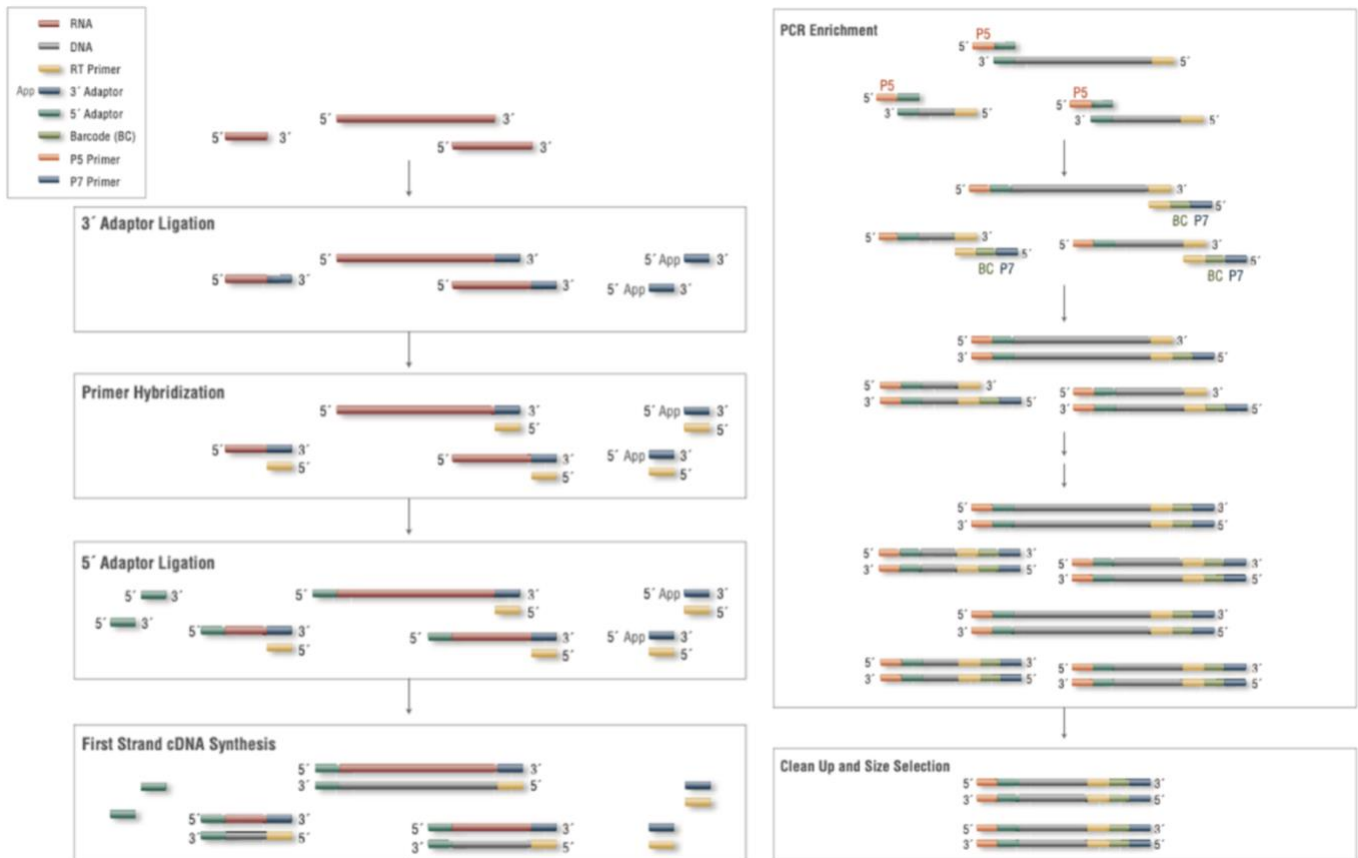


Figura 3. Esquema de la generación de bibliotecas. Figura tomada y modificada del protocolo “NEBNext small RNA Library Prep for Illumina” (NEB).

## 2 Análisis bioinformático

### 3.1 Análisis primario

Las lecturas obtenidas se analizaron utilizando FastQC <sup>63</sup> (V0.12.1) para determinar la cantidad de secuencias obtenidas por muestra y su calidad. Las secuencias adaptadoras se removieron utilizando Cutadapt <sup>64</sup> (V4.9). Solo se analizaron las secuencias con 19 bases o más que contenían un adaptador 3' identificable. De acuerdo al objetivo buscado, estas se mapearon al genoma humano hg38 (GRCh38.p14 (GCA\_000001405.29)) y/o a referencias conteniendo ncRNAs utilizando Bowtie2 <sup>65</sup> (V2.5.3). Para el alineamiento al genoma se utilizaron los parámetros de alineamiento por defecto. Para el alineamiento a la referencia de ncRNAs se modificó el

parámetro -L que determina la longitud de las subcadenas de semillas (-L 8). Las secuencias mapeadas fueron asignadas utilizando featureCounts<sup>66</sup> (V1.6.0.2), contando a nivel de transcrito, (-M, -O, --fraction) permitiendo mapeos múltiples (*multi-mapping reads*) y superposición múltiple (*multi-overlapping features*) asignando una fracción de los reads a cada *feature*. La referencia de ncRNAs utilizada consta de miRNAs maduros (mirBASE), tRNAs (GtRNAdb), lncRNAs, snoRNAs, snRNAs, rRNA, vaultRNAs, YRNA, SRP-RNA (RNACentral), mt rRNA, mt tRNA (Ensembl). La referencia ncRNAs(-Y-t) es la referencia ncRNAs sin incluir YRNAs ni tRNAs.

### 3.2 Características generales

La abundancia de los biotipos se determinó a partir de las tablas de conteo de featureCounts contando a nivel de biotipo, normalizado según el total de secuencias mapeadas para cada muestra. La composición de los distintos YRNAs se obtuvo filtrando los *reads* asignados a YRNAs, normalizando según el total de *reads* mapeados a YRNAs. La composición de los distintos miRNAs se obtuvo filtrando los reads asignados a miRNAs. Los más expresados se determinaron en base al promedio de todas las muestras. La normalización se hizo en base al total de secuencias mapeadas a miRNAs para cada muestra. La abundancia de los 20 transcritos más representados se obtuvo seleccionando los fragmentos más expresados según el promedio por transcrito considerando todas las muestras y se expresaron en log<sub>2</sub> RPM, siendo RPM la cantidad de lecturas asignadas dividiendo su número de recuento absoluto entre la cantidad de lecturas mapeadas y multiplicando por 1 millón.

Para determinar la distribución de tamaños de las lecturas mapeadas, estas se extrajeron del total de los reads utilizando BAM Filter (-mapped=TRUE) y las coordenadas de los features incluidos en la referencia ncRNAs. Para determinar la distribución de tamaños de las lecturas mapeadas a YRNAs o miRNAs, los reads fueron filtrados según hayan mapeado a estos biotipos según sus coordenadas en la referencia de ncRNAs. Las lecturas filtradas se convirtieron a Fasta y se calculó la longitud de cada lectura individual. La abundancia relativa de lecturas con una longitud determinada se expresó como lecturas por millón (RPM).

Para determinar la composición de miRNAs de 21 y 22 nt, los *reads* primero se filtraron por tamaño, seleccionando por un lado los reads de 21 nt y por otro los de 22 nt. Luego fueron alineados con Bowtie2 a una referencia de miRNAs y asignados utilizando featureCounts con los mismos parámetros ya descritos (ver Análisis primario).

### 3.3 Búsqueda de miRNAs con potencial biomarcador del LES

Las tablas de conteo obtenidas se filtraron para mantener solamente aquellos transcritos con 5 o más conteos en todas las muestras. El análisis de expresión diferencial se realizó utilizando DESeq2<sup>67</sup>(v1.20), se consideró expresión diferencial para los transcritos que tuvieran un p-valor ajustado (padj) inferior a 0.05, y un cambio absoluto en el nivel de expresión (log<sub>2</sub>FoldChange) mayor o igual a 1. El análisis de componentes principales se realizó luego de aplicar rlog (blind = TRUE) a los datos de expresión normalizados por DESeq2, utilizando la función prcomp del paquete base de R. La prueba permanova se realizó utilizando la función adonis2 de vegan-package (V2.6-6.1), y la matriz de distancia fue calculada utilizando la distancia de Bray-Curtis. Para la visualización en heatmaps de los genes con expresión diferencial se utilizó la función

heatmap del paquete homónimo de R, a partir de los datos normalizados de DESeq2. La expresión de los genes fue centrada y escalada, y la clusterización jerárquica se realizó sólo a nivel de los genes, permitiendo agrupar genes con perfiles de expresión similares, utilizando la distancia euclidiana y el método de enlace completo.

# Resultados

## 1. Selección de las pacientes

El primer objetivo planteado fue la selección de las pacientes a incluir en el estudio. En colaboración con la Policlínica de Enfermedades Autoinmunes del Hospital Maciel se seleccionaron 3 mujeres con LES en actividad, 3 mujeres con LES en remisión y 3 controles sanas. En la Tabla 1 se muestran los datos clínicos y serológicos de las pacientes, así como su índice de actividad de la enfermedad SLEDAI, y los tratamientos a los que estaban sometidas al momento de la inclusión. En cuanto al fenotipo clínico de las pacientes, todas ellas presentaron manifestaciones mucoso-cutáneas, mientras que solo algunas presentaron las siguientes manifestaciones: articulares (pacientes R1, R2 y A3), musculares (paciente A3), cardíacas (paciente A4), hematológicas (pacientes R1, R2 y R3), y renales (paciente A3). En cuanto al fenotipo serológico, todas las pacientes son positivas para ANAs, y presentan distintos perfiles de autoanticuerpos: anti-DNA (pacientes R1, R2, A2, A3 y A4), anti-Smith (pacientes A2 y A3), anti-Ro (pacientes R2, A2 y A3), anti-La (pacientes R2 y A2), anti-RNP (paciente A3), anti-nucleosoma (pacientes R1 y A3), anti-histona (pacientes R1 y A3). Algunas de las pacientes también presentaron autoanticuerpos anti-fosfolípidos como anti-IL (pacientes R2 y A3) y ACA (anti-cardiolipina) (pacientes A2 y A3). La paciente R3 si bien fue positiva para ANAs, no presentó ninguno de los autoanticuerpos ensayados. La activación del sistema complemento fue detectada para las pacientes R1, R3, A2, A3 y A4. En cuanto a la actividad de la enfermedad, las pacientes en remisión presentaron un puntaje SLEDAI de 0, indicando que no hay actividad de la enfermedad, mientras que las pacientes A2, A3 y A4 presentaron puntajes de 6, 14 y 14 respectivamente, indicando una alta actividad de la enfermedad para las 3 pacientes. Otros indicadores que correlacionan con la actividad de la enfermedad son la velocidad de eritrosedimentación (VES) y la proteína C reactiva (PCR), ambos valores aumentan cuando hay infección o inflamación. Solo la paciente R1 presentó un valor elevado de PCR, y las pacientes A2 y A3 presentaron valores elevado de VES. En cuanto a los tratamientos, las pacientes se encontraban realizando tratamientos con las siguientes drogas al momento de la inclusión: hidroxicloroquina (todas las pacientes), metotrexato (pacientes A2 y A4), azatioprina (pacientes R2 y A4), micofenolato mofetilo (paciente A4), rituximab (pacientes R3 y A4), inmunoglobulina intravenosa (paciente A4), glucocorticoides (pacientes R2, A2, A3, A4) y metilprednisolona (pacientes R3, A2, A3, A4). Una primera observación sobre las características de las pacientes es la gran heterogeneidad que existe entre la sintomatología, los autoanticuerpos que presenta cada una, y algunos marcadores de actividad como la VES y la PCR. Las controles sanas no se sometieron a ningún estudio clínico ni paraclínico.

Tabla 1. Características clínicas de las pacientes con LES. SNP: sistema nervioso periférico); ANA: anticuerpos anti-nucleares; IL: inhibidor lúpico; ACA: anticuerpos anti-cardiolipina; SLEDAI: puntaje obtenido, se considera actividad alta un puntaje  $\geq 6$ ; VES: velocidad de eritrosedimentación; PCR: proteína C reactiva. Los valores normales se indican entre paréntesis en la tabla.

	ID	R1	R2	R3	A2	A3	A4	N1	N2	N3
Fenotipo clínico	Edad	53	38	24	33	20	20	37	27	52
	Actividad	Inactiva	Inactiva	Inactiva	Activa	Activa	Activa	Sana	Sana	Sana
	Cutáneo - mucoso	1	1	1	1	1	1			
	Articular	1	1	0	0	1	0			
	Muscular	0	0	0	0	1	0			
	SNP	0	0	0	0	0	0			
	Cardio	0	0	0	0	0	1			
	Respiratorio	0	0	0	0	0	0			
	Digestivo	0	0	0	0	0	0			
	Hematológico	1	1	1	0	0	0			
	Renal	0	0	0	0	1	0			
	Ginecológico - obstétrico	0	0	0	0	0	0			
	General	0	0	0	0	0	0			
	Plaquetas (150000 - 400000/ $\mu$ l)	210000	289000	184000	407000	232000	394000			
Fenotipo serológico	ANA	1	1	1	1	1	1			
	Título	1280	160	320	2560	1280	640			
	anti-DNA	1	1	0	1	1	1			
	anti-Smith	0	0	0	1	1	0			
	anti-Ro	0	1	0	1	1	0			
	anti-La	0	1	0	1	0	0			
	anti-RNP	0	0	0	0	1	0			
	anti-nucleosoma	1	0	0	0	1	0			
	anti-histona	1	0	0	0	1	0			
	anti-B2GP1	0	0	0	0	0	0			
	Inhibidor lúpico	0	1	0	0	1	0			
	anti-cardiolipina	0	0	0	1	1	0			
	Complemento	1	0	1	1	1	0			
Actividad	SLEDAI	0	0	0	6	14	14			
	Baja	0	0	0	0	0	0			
	Moderada	0	0	0	0	0	0			
	Alta	0	0	0	1	1	1			
	Remisión	1	1	1	0	0	0			
	VES (< 20)	20	13	15	35	76	8			
	PCR (0,5 mg/dl)	0.67	0.24	6.3	0.30	13	0.30			
Tratamiento	Hidroxicloroquina	1	0	1	1	1	1			
	Metotrexato	0	0	0	1	0	1			
	Azatioprina	0	1	0	0	0	1			
	Micofenolato mofetilo	0	0	0	0	0	1			
	Rituximab	0	0	1	0	0	1			
	Inmunoglobulina intravenosa	0	0	0	0	0	1			
	Glucocorticoides	0	1	0	1	1	1			
	Metilprednisolona	0	0	1	1	1	1			

## 2. Obtención y secuenciación de bibliotecas de pequeños RNAs

A partir del RNA purificado del plasma de las 9 muestras se realizaron bibliotecas de pequeños RNAs. Los pequeños RNAs fueron seleccionados mediante electroforesis de poliacrilamida, incluyendo fragmentos correspondientes a miRNAs (de alrededor de 21-22 nt) y mitades de tRNAs y fragmentos de YRNAs (de alrededor de 30-32 nt). Los tamaños esperados para estos RNAs con los adaptadores ligados son de 140 y 150 nt, respectivamente. En la Figura 4 se puede observar que estas bandas no se resuelven claramente, sino que se observó lo que parece ser una única banda a la altura de 150 nt.

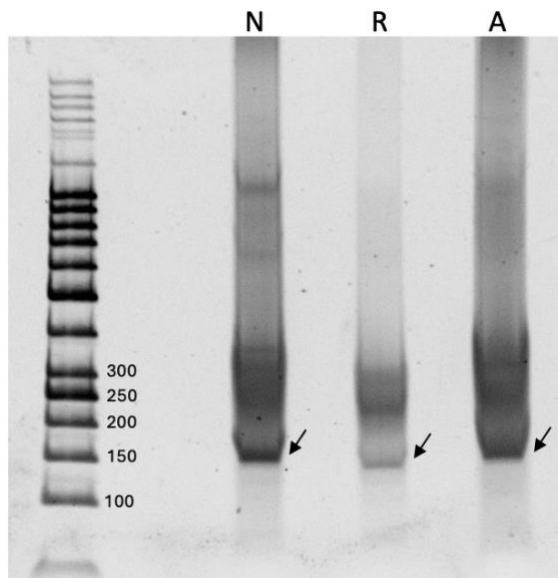


Figura 4. Electroforesis en gel de poliacrilamida al 6% del producto de PCR de las bibliotecas. En el primer carril se encuentra el marcador de peso molecular. Las flechas indican las bandas que fueron cortadas e incluidas en las bibliotecas. El corte se realizó para todas las muestras por igual, incluyendo el fragmento de gel que abarca desde los 150 pb a 200 pb. N: control sana, R: paciente en remisión, A: paciente en actividad.

### 2.1 Control de calidad de las bibliotecas

Previo a la secuenciación, se realizó un control de calidad de las bibliotecas en el analizador de fragmentos Bioanalyzer 2100, utilizando un chip de DNA de alta sensibilidad. Esto permitió conocer con mayor exactitud los tamaños seleccionados para las bibliotecas de cada muestra. Las bibliotecas mostraron un promedio de tamaños de alrededor de 152 a 165 (Figura 5). Las bibliotecas se cuantificaron por fluorometría con el Kit DNA High Sensitivity, Qubit (ver Tabla A1).

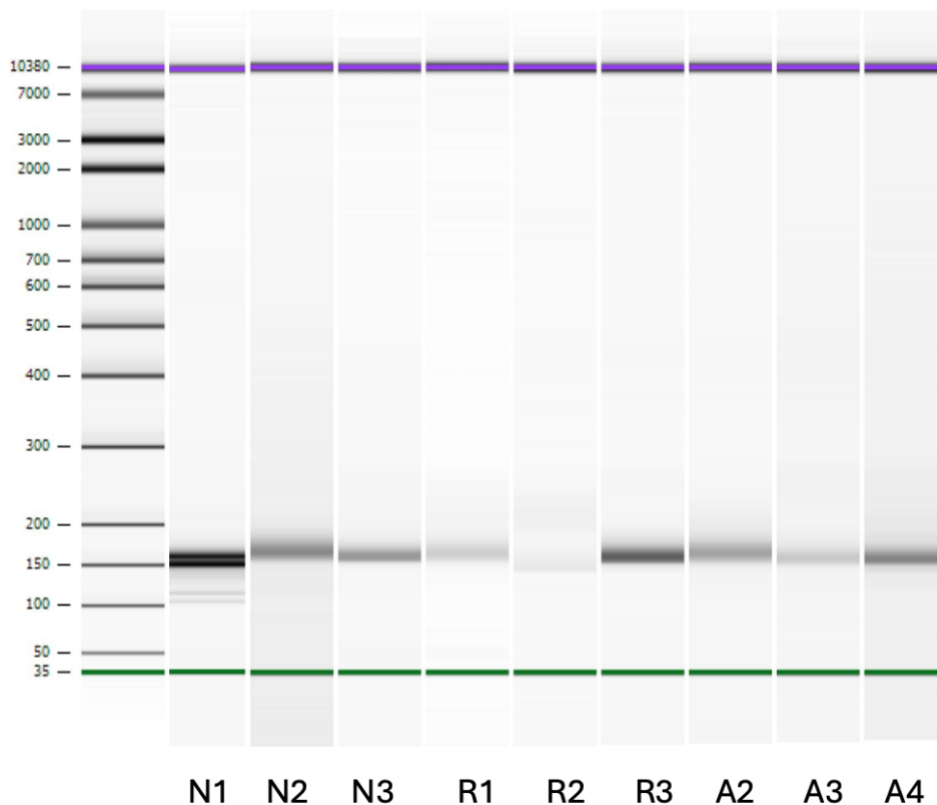


Figura 5. Control de calidad de bibliotecas. Electroferograma de chip de ADN de alta sensibilidad en Bioanalyzer 2100. En el primer carril se encuentra el marcador de peso molecular. Las bandas magenta y verde corresponden al marcador interno superior (10380 pb) e inferior (35 pb) respectivamente.

## 2.2 Resultados de secuenciación

Las bibliotecas fueron secuenciadas en la plataforma MiSeq de Illumina, utilizando un kit de 50 ciclos en una corrida *single-end*. La calidad de las secuencias obtenidas fue buena (más del 99% de las secuencias con Q score > 30), y también se obtuvo un buen porcentaje de alineamiento para todas las muestras (Tabla 2). Las secuencias se alinearon al genoma humano, así como a referencias que contienen únicamente los ncRNAs de interés. Para una descripción inicial de los resultados de secuenciación se utilizó una referencia conteniendo todos los biotipos de ncRNAs mientras que para la búsqueda de pequeños RNAs que puedan servir como biomarcadores de la enfermedad se utilizó una referencia de ncRNAs filtrada sin tRNAs ni YRNAs (ver Materiales y Métodos). El porcentaje de alineamiento al utilizar como referencia el genoma humano fue superior al 96 % para todas las muestras. Al utilizar las referencias de ncRNAs este porcentaje fue apenas más bajo, indicando que gran parte de las secuencias obtenidas corresponden a ncRNAs (Tabla 2). Además, el alineamiento a estas referencias permitió aumentar el porcentaje de secuencias con mapeos únicos. No incluir YRNAs y tRNAs en la referencia aumentó mucho el porcentaje de secuencias con mapeos únicos. Esto se debe a las características de estos genes. El genoma humano contiene múltiples copias de genes para la mayoría de los tRNAs. Estas copias

pueden tener secuencias idénticas o casi idénticas, de forma que es prácticamente imposible distinguir la posición en el genoma de donde provienen. Otro motivo para excluir a los tRNAs fue que los métodos utilizados tanto para la preparación de la muestra como para la secuenciación no favorecen su detección. En el caso de los YRNAs, también se trata de genes que están altamente conservados, especialmente en regiones clave para su función. En la Tabla 2 se muestra la cantidad de secuencias obtenidas y su porcentaje de alineamiento a las distintas referencias utilizadas.

Tabla 2. Secuencias obtenidas y porcentaje de alineamiento a las referencias utilizadas.

Muestra	Secuencias (M)	Genoma (%)	Mapeos únicos (%)	ncRNAs (%)	Mapeos únicos (%)	ncRNAs(-Y-t) (%)	Mapeos únicos (%)
N1	3.9	96	47	93	52	87	74
N2	2.4	99	7	98	20	86	81
N3	4.1	99	10	97	22	86	78
R1	2.7	98	27	95	36	86	79
R2	4.6	98	62	94	67	94	67
R3	3.7	99	9	98	21	87	82
A2	3.2	99	18	96	27	87	77
A3	3.6	99	16	96	26	86	80
A4	3.3	98	16	96	26	84	74

### 3. Descripción de las características generales de la población de ncRNAs

El objetivo principal de esta tesis es la descripción de los perfiles de pequeños RNAs circulantes en las pacientes con LES. Para ello, una vez asignadas las secuencias a la referencia de ncRNAs, se analizaron los datos para describir cómo es la composición de los distintos biotipos de ncRNAs de las muestras, cuáles son los RNAs más representados, y cuáles son los tamaños de estas secuencias.

#### 3.1 Composición por biotipo

En cuanto a la composición por biotipo se puede observar que, salvo para una de las pacientes en remisión y una control, los YRNAs son el biotipo más abundante en el plasma. El segundo biotipo más abundante es el de los miRNAs (ver Figura 6). Sin considerar a la muestra R2, que presentó un comportamiento particular sobre el que se discute más adelante, en promedio las muestras presentaron un 70 % de alineamiento a YRNAs, con un mínimo de 38 % para la muestra N1 y un máximo de 86 % para la muestra N2; y un 24 % de alineamiento a miRNAs con un mínimo de 10 % para la muestra N2 y un máximo de 60 % para la muestra N1.



Figura 6. Abundancias porcentuales de los biotipos de ncRNAs en las muestras. N: control; R: paciente con LES en remisión; A: paciente con LES activo.

### 3.2 ncRNAs más representados en las muestras

El YRNA4 es la secuencia más representada en la mayoría de las muestras, lo que concuerda con lo reportado en la literatura<sup>68</sup>. El plasma de la paciente R2 presentó una marcada disminución de YRNAs, y un cambio en la composición de este biotipo. Si bien el RNY4 es el más abundante, al igual que en el resto de las muestras, presentó una mayor cantidad de secuencias mapeadas a otros YRNAs que el resto de las muestras. Alrededor del 15 % de los reads mapeados a YRNAs en la muestra R2 corresponden al RNY1, y un porcentaje menor a los RNY3 y RNY5 (ver Figura 7).

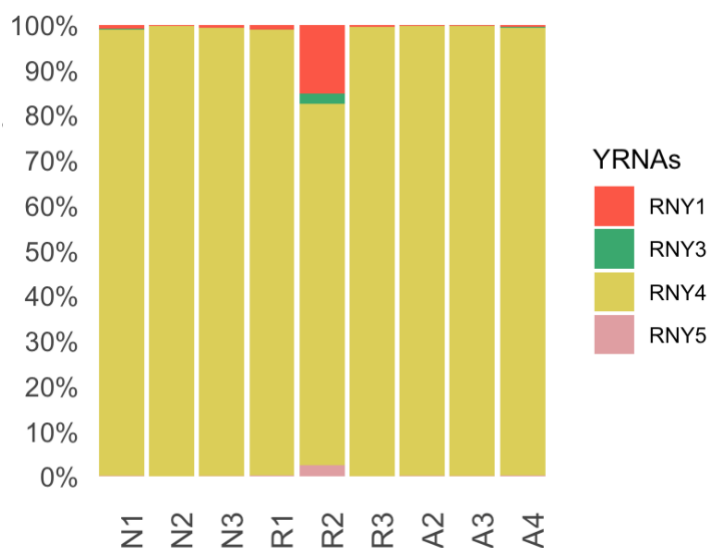


Figura 7. Abundancia porcentual de YRNAs en cada muestra. El 100 % corresponde al total de los reads mapeados a YRNAs. N: control; R: paciente con LES en remisión; A: paciente con LES activo. Las muestras presentan una gran variación en cuanto a la composición de los miRNAs más abundantes. Los microRNAs más representados en las muestras son miR-451a, miR-148a-3p, let-7i-5p y let-7f-5p (ver Figura 8).

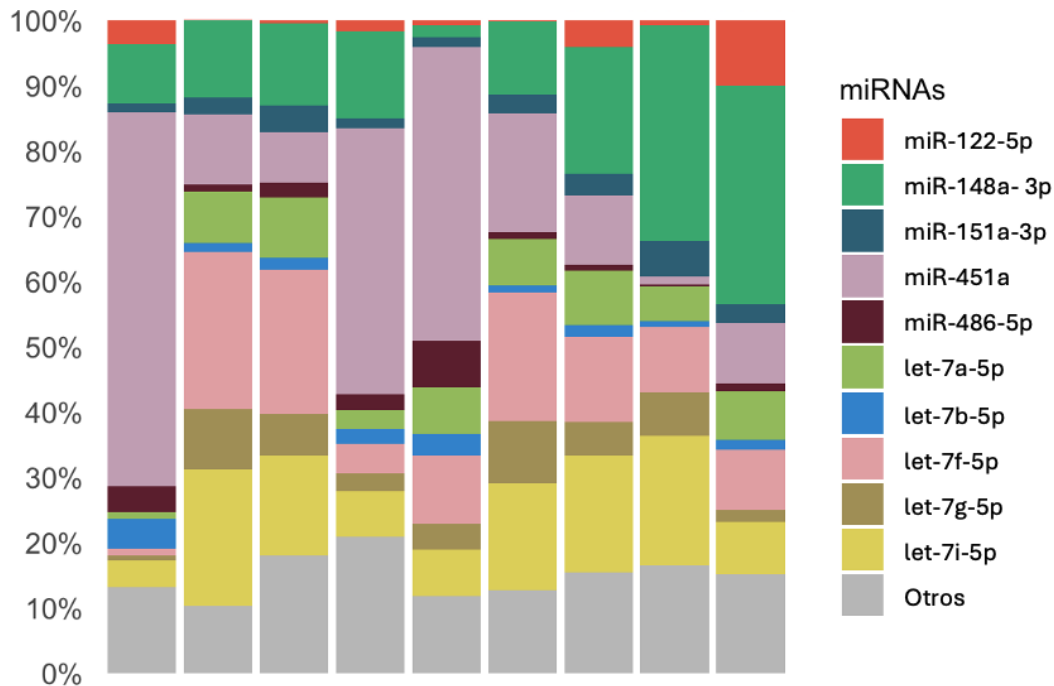


Figura 8. Abundancia porcentual de miRNAs más representados. El 100 % corresponde al total de los reads mapeados a miRNAs. N: control; R: paciente con LES en remisión; A: paciente con LES activo.

El miR-451a es un microRNA expresado por glóbulos rojos. De forma que su abundancia en la circulación está asociada a estas células. Otro miRNA abundante es el miR-486-5p expresado por glóbulos rojos y plaquetas. Los let-7 y miR-151a-3p por el contrario se ha reportado que se expresan en todos los tejidos, si bien esto quizás refleja la presencia de determinados tipos celulares en los tejidos. El miR-122-5p es expresado por hepatocitos. El miR-148a se expresa en distintas células, incluidas células circulantes, como los neutrófilos, pero también es muy expresado por células residentes mastocitos y hepatocitos <sup>69</sup>, que tienen acceso a la circulación.

En la Figura 9 se representan los 20 ncRNAs mayoritarios en el plasma de las muestras de las tres condiciones. El YRNA4 es el ncRNA más abundante, seguido del miR-451a y el resto de los microRNAs. Nuevamente, se evidencia la gran variación entre las muestras, incluso dentro de la misma condición.

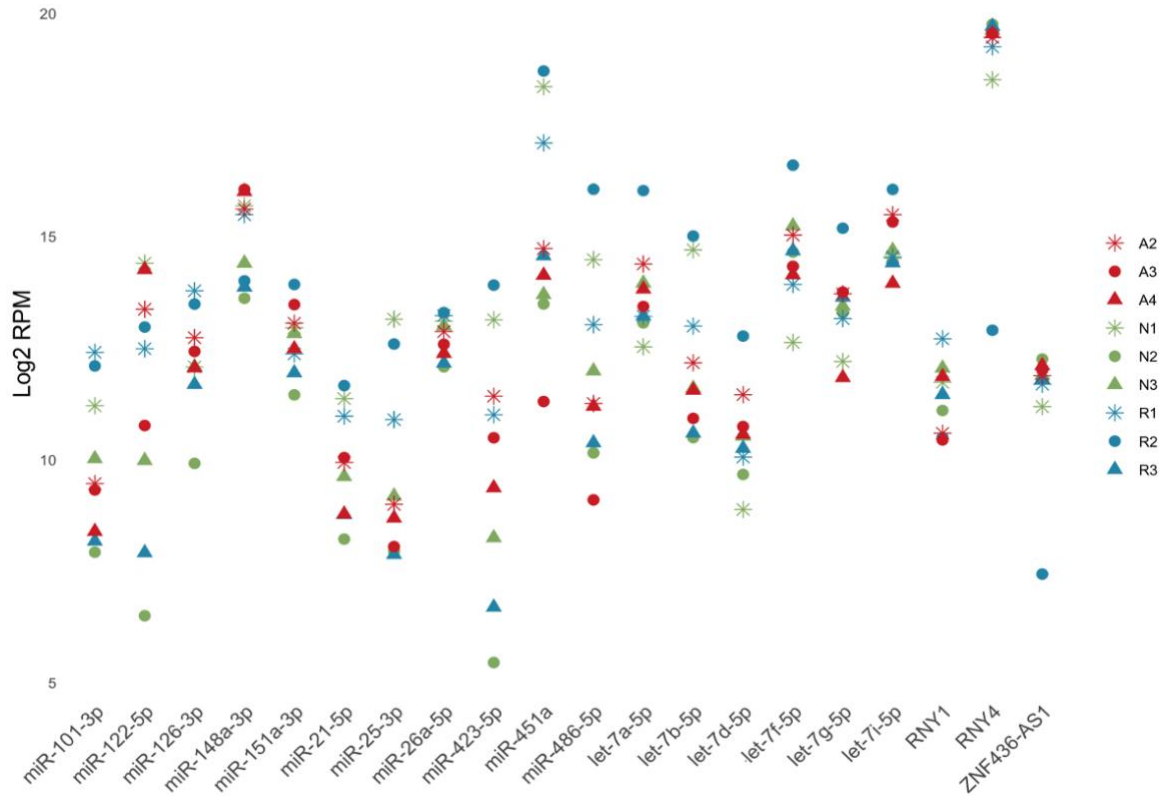


Figura 9. ncRNAs más abundantes expresados en Log2 RPM. N: control; R: paciente con LES en remisión; A: paciente con LES activo.

### 3.3 Distribución de tamaño de las secuencias

Al evaluar la distribución de tamaño de las secuencias se distinguen dos picos, el primero compuesto por secuencias de 20 a 23 nt y el segundo pico compuesto por secuencias de 31 a 33 nt (ver Figura 10). El primer pico de 20 a 22 nt corresponde a secuencias mapeadas a miRNAs (ver Figura 11 A), mientras que el segundo pico de 31 a 33 nt corresponde a secuencias mapeadas a fragmentos de YRNAs (ver Figura 11 B). Para la muestra de la paciente en remisión 2 (R2) se observa que hay una cantidad significativamente menor de secuencias de 31 a 33 nt, lo que coincide con el menor recuento de YRNAs de esta muestra (Figura 6).

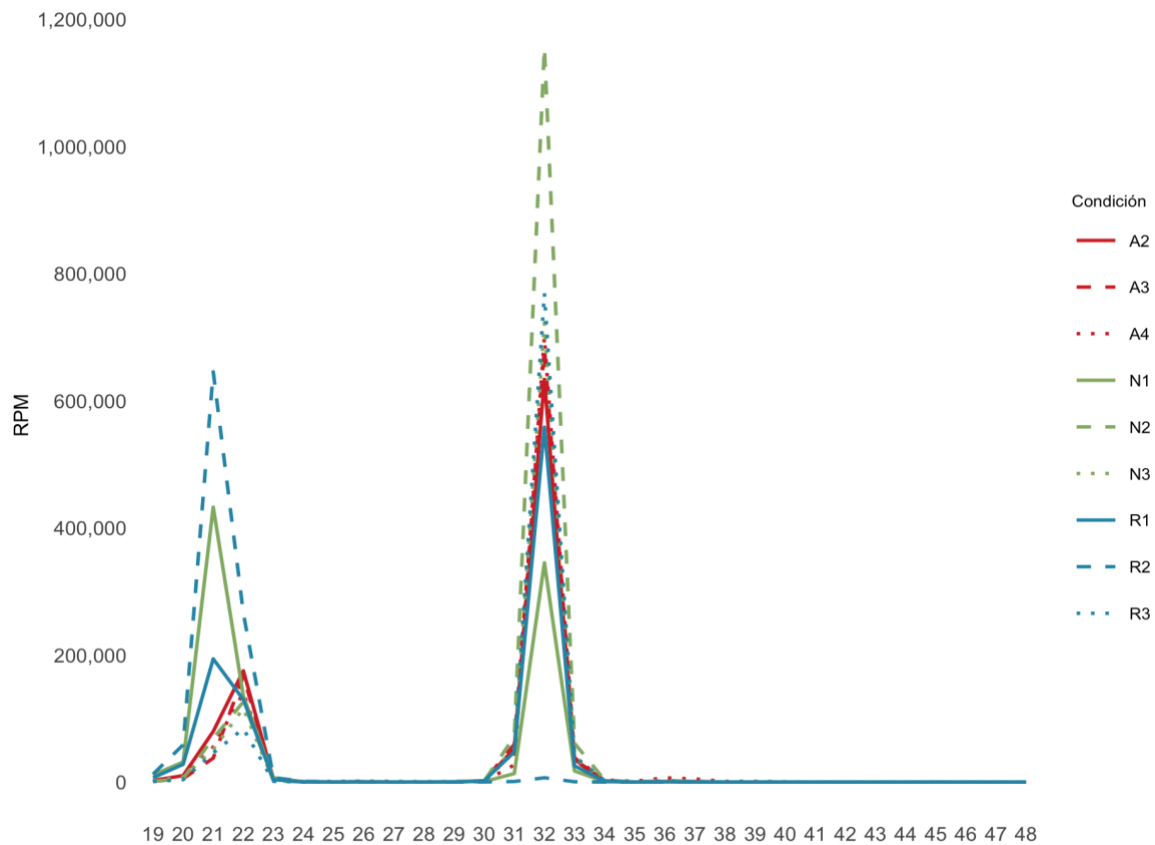


Figura 10. Distribución por tamaño de las secuencias. N: control; R: paciente con LES en remisión; A: paciente con LES activo. RPM: reads por millón.

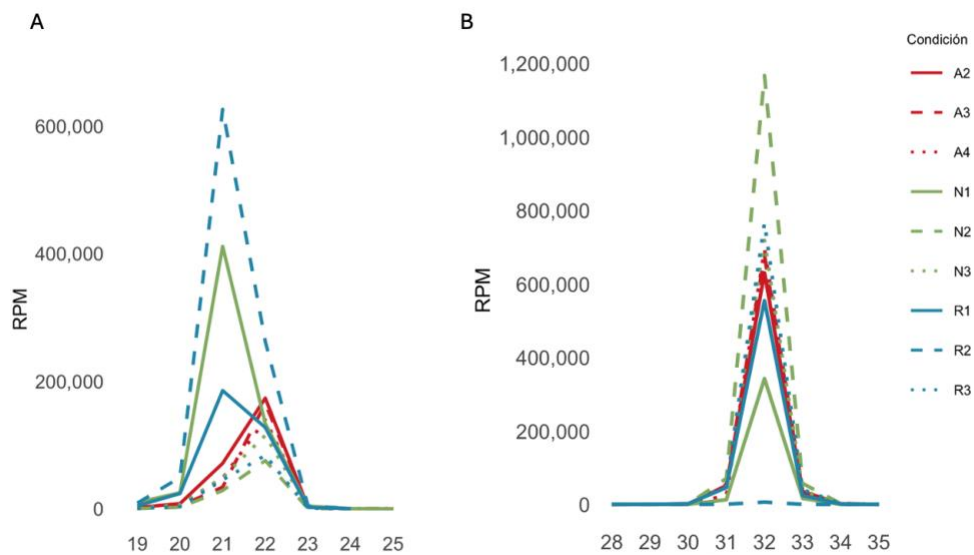


Figura 11. A) Distribución de tamaño de las secuencias mapeadas a miRNAs. B) Distribución de tamaño de las secuencias mapeadas a YRNAs.

### 3.4 Detección de sesgos de tamaños en las bibliotecas

Los fragmentos de las bibliotecas N1 y R2 tienden a ser más cortos que los del resto de las muestras. Esto se puede ver en los resultados de Bioanalyzer (ver tabla A1) y también en la Figura 10, donde se observa que las muestras N1 y sobre todo R2 presentan una menor cantidad de secuencias de 32 nt, acompañado de un aumento de secuencias de 21 nt por sobre las de 22 nt. Se establece una correlación negativa donde las muestras con menor cantidad de secuencias de 32 nt son las que presentan una mayor cantidad de secuencias de 21 nt: R2 que tiene menos fragmentos de 32 es la muestra con más fragmentos de 21, seguida por N1 y R1 (ver Figura 12).

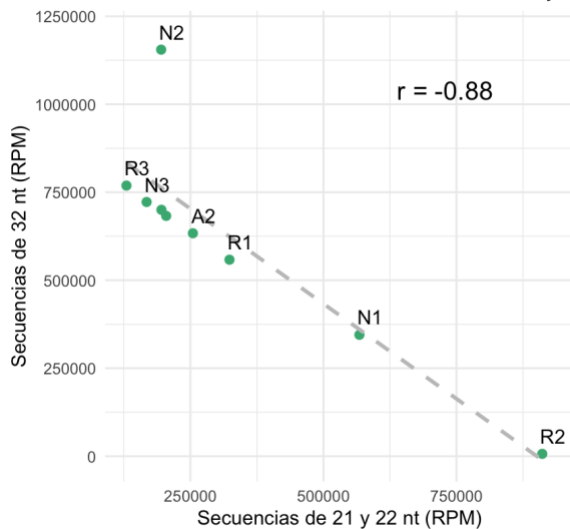
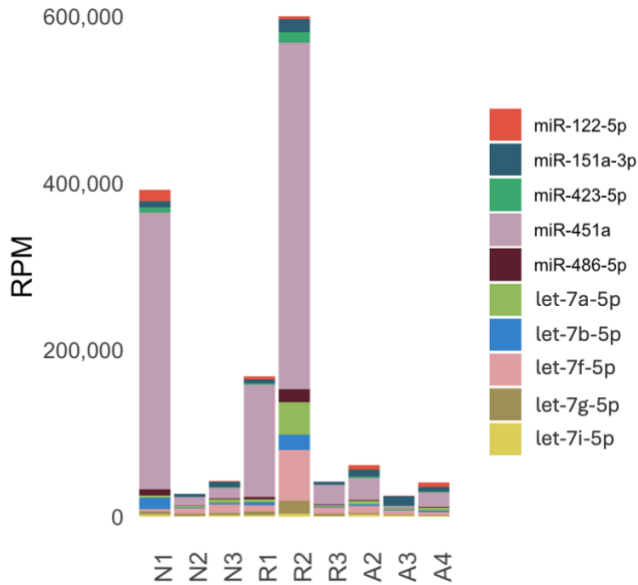


Figura 12. Correlación entre secuencias de 21-22 nt y 32 nt en las muestras. Se utilizó el coeficiente de correlación de Pearson ( $r = -0.88$ ) para evaluar la relación entre las secuencias de 21-22 nt y 32 nt.

Los cambios observados en los tamaños de los fragmentos que componen a las bibliotecas podrían estar reflejando la composición real de pequeños ARNs en las muestras, o bien esta podría deberse a sesgos técnicos, y resulta imposible con estos resultados saber si la composición de tamaños distinta es real o artefactual. Dado que los pequeños ARNs se caracterizan por su tamaño, estas diferencias en tamaños están también acompañadas de diferencias en composición (o viceversa). Al considerar solamente las secuencias mapeadas a miRNAs se puede observar el efecto en las poblaciones de microRNAs. La secuenciación de la población de miRNAs de 21 nt fue favorecida en estas muestras respecto a la población de miRNAs de 22 nt (Figura 11 A), y con ella, la secuenciación de miR-451a en estas muestras (Figura 13), dada la abundancia de la isoforma de 21 nt de este miRNA en la circulación.



A. Figura 13. Secuencias de 21 nt que mapean a miRNAs.

Este efecto también se puede observar al hacer un análisis visual de las secuencias que mapean a los distintos YRNAs, donde resulta claro que se favorece la secuenciación de fragmentos de YRNAs más cortos en las muestras con bibliotecas más pequeñas (ver Figura 14). Este efecto podría deberse a la menor cantidad de secuencias de YRNAs de 33 nt de las muestras N1 y R2 (Figura 11 B), que hace posible la secuenciación de fragmentos de YRNAs menos abundantes.



Figura 14. Visualización de las secuencias alineadas a YRNAs.

#### 4. Búsqueda de exRNAs con potencial de biomarcador de LES

El cuarto objetivo de este trabajo es la selección de pequeños RNAs que puedan servir como biomarcadores de la enfermedad. Para determinar si las muestras de una misma condición presentan un comportamiento común entre sí, pero diferente al de las otras condiciones, se realizó un análisis de componentes principales (Figura 15). Este análisis evidenció un agrupamiento de las muestras correspondientes a las pacientes con lupus activo. Mientras que las pacientes en remisión no se diferencian de las controles sanas. Al repetir este análisis con las 50 secuencias más abundantes en las muestras esta separación se pierde, indicando que la misma no es producto de miRNAs de alta expresión. Por otro lado, se comparó el efecto de incluir o no YRNAs y tRNAs en las referencias, observándose una separación más clara al no considerar los *reads* que mapean a estos biotipos. Al realizar un test Permanova al agrupamiento de las secuencias obtenidas sin YRNAs y tRNAs se obtiene un  $R^2$  (0.34654) que sugiere que una cantidad considerable de la varianza (35%) es explicada por la variable Condición, sin embargo, el p-valor obtenido de 0.113 indica que las diferencias observadas en el PCA entre las condiciones no son estadísticamente significativas.

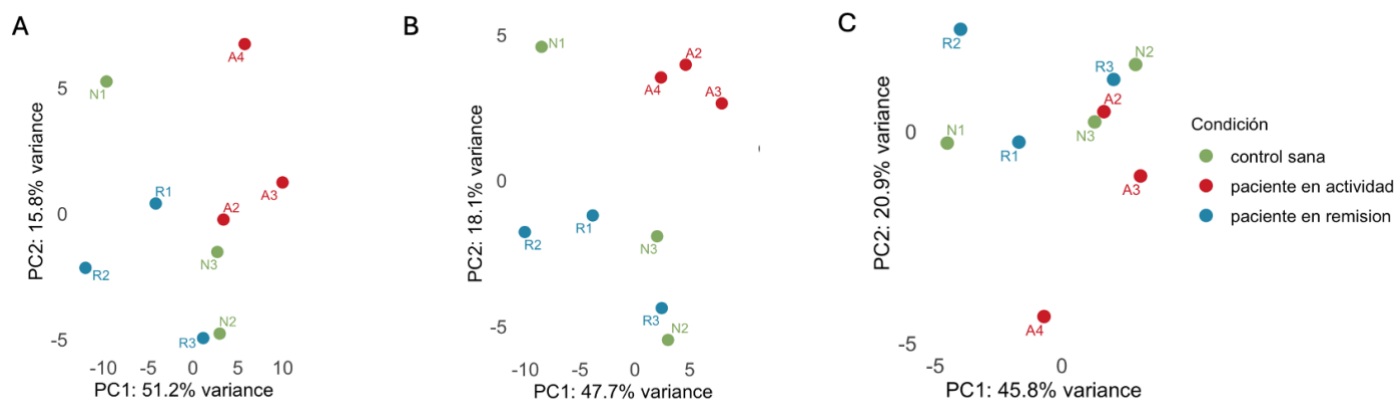


Figura 15. Análisis de componentes principales. A) Considerando todas las secuencias asignadas. B) Considerando todas las secuencias salvo las asignadas a tRNAs e YRNAs. C). Considerando solo el top 50 de secuencias más representadas en las muestras.

Para identificar a los RNAs que contribuyen al agrupamiento de las pacientes activas se realizó un análisis de expresión diferencial, excluyendo las secuencias mapeadas a tRNAs e YRNAs. No se detectó ningún ncRNA diferencialmente expresado entre las pacientes sanas y pacientes en remisión (Figura 16 C). Mientras que las pacientes en actividad si presentan genes con expresión diferencial respecto a las pacientes sanas como a las pacientes en remisión (Figura 16 A y B, respectivamente). En la Tabla A2 se muestran los RPM obtenidos para los transcritos con expresión diferencial reportados.

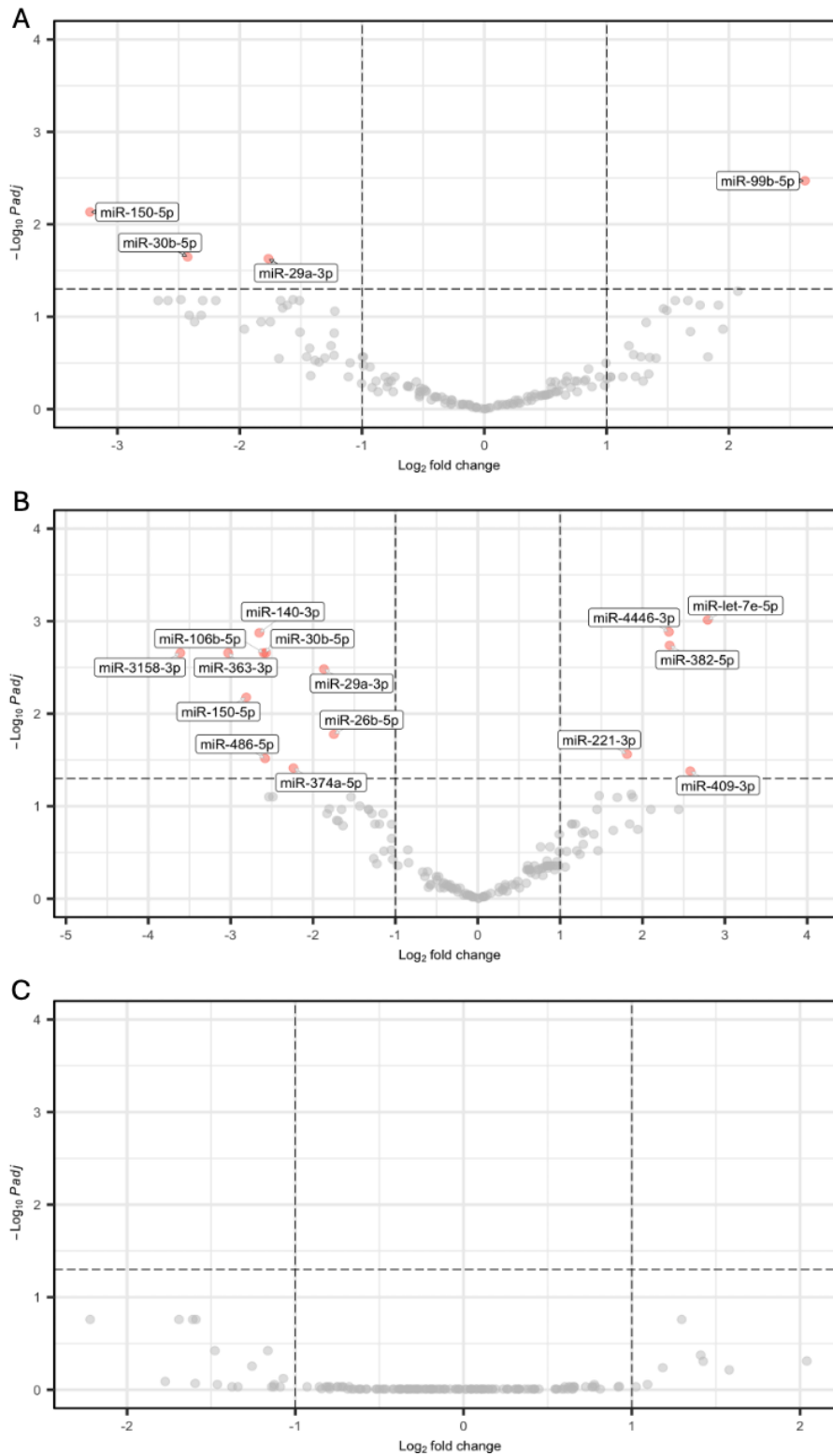


Figura 16. ncRNAs diferencialmente expresados entre las 3 condiciones. A) Pacientes activas respecto controles sanas. B) Pacientes activas respecto pacientes en remisión. C) Pacientes en remisión respecto controles sanas.

La representación en un *heatmap* de los genes con expresión diferencial permite comparar la expresión entre muestras y determinar si muestras de la misma condición comparten la tendencia de la expresión de los genes con expresión diferencial, o si por el contrario se trata de un artefacto introducido por las muestras problemáticas. En la Figura 17 se puede observar que, si bien las muestras N1 y R2 se apartan levemente del comportamiento de su condición, hay una tendencia bastante clara entre grupos. En la Figura 17 se pueden diferenciar 2 grupos de genes, aquellos cuya expresión aumenta en las pacientes con lupus activos y aquellos cuya expresión disminuye en estas pacientes.

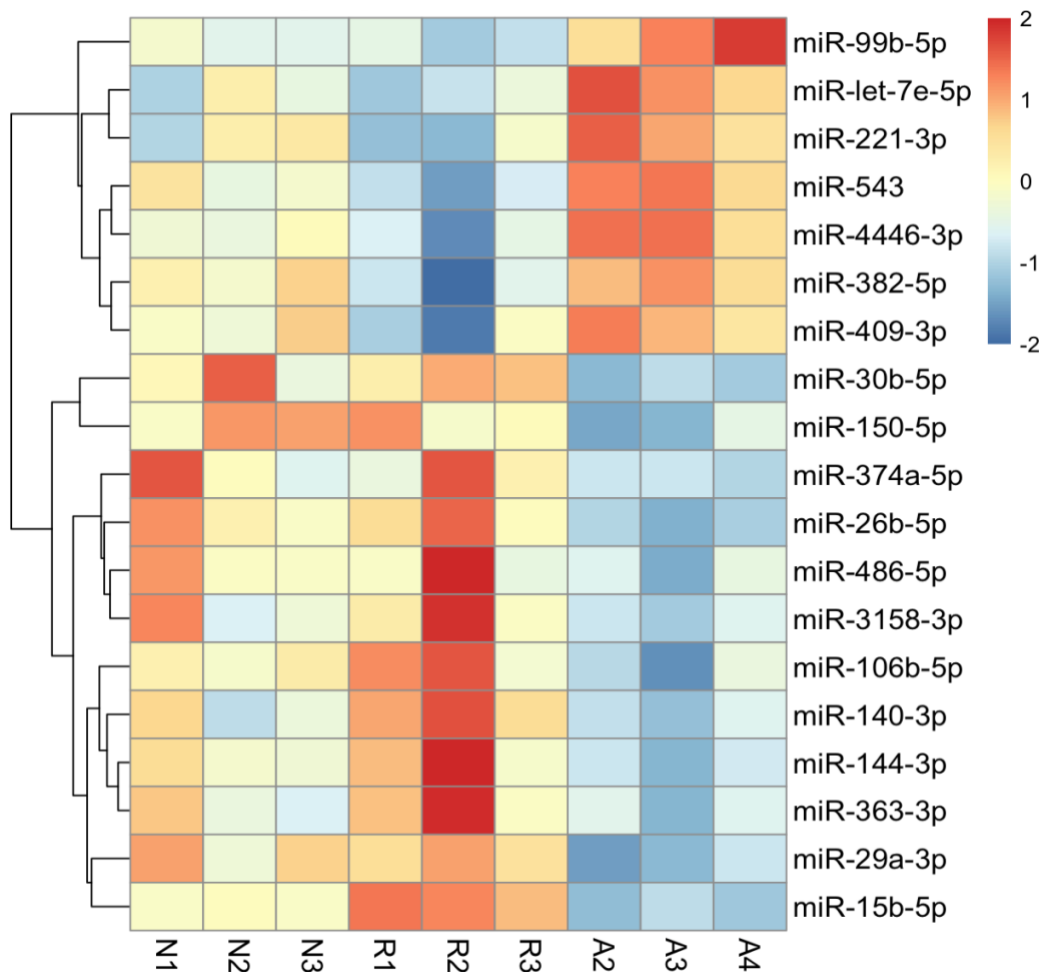


Figura 17. Visualización de la expresión de los genes con expresión diferencial. La clusterización jerárquica de los genes se realizó en base a la distancia euclidiana. N: control; R: paciente con LES en remisión; A: paciente con LES activo. RPM: reads por millón.

A partir del análisis de las secuencias obtenidas para las 3 condiciones se determinó que las pacientes en remisión presentan un perfil de microRNAs circulantes muy similar al de las controles sanas en el sentido de que no se detectaron diferencias estadísticamente significativas. En cambio, las pacientes que están cursando un período de actividad de la enfermedad sí

presentaron un perfil de microRNAs distintivo, caracterizado por una sobreexpresión de miR-99b-5p y una disminución de miR-150-5p, miR-29a-3p y miR-30b-5p respecto a las controles sanas. Mientras que al compararlas con las pacientes en remisión mostraron una sobreexpresión de miR-382-5p, miR-409-3p, miR-221-3p, miR-543, miR-4446-3p y let-7e-5p; y una subexpresión de miR-miR-140-3p, miR-30b-5p, miR-363-3p, 374a-5p, 150-5p, miR-144-3p, miR-26b-5p, miR-106b-5p, miR-3158-3p, miR-486-5p y miR-29a-3p.

En base a estos resultados, la detección conjunta de la firma molecular compuesta por los miRNAs cuya expresión está aumentada en las pacientes activas (Figura 18) podría ser de utilidad para el seguimiento de la enfermedad, de ser validada en un mayor número de pacientes.

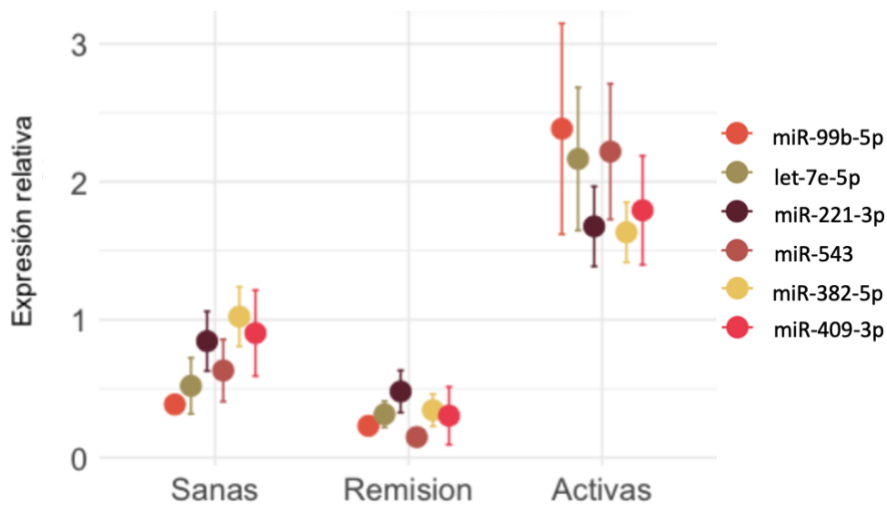


Figura 18. Identificación de una firma molecular de miRNAs sobreexpresados en pacientes con LES activo. La expresión en cada condición normalizada respecto a la expresión promedio de todas las condiciones.

## Discusión

Durante el desarrollo de esta tesis se lograron alcanzar los objetivos planteados mediante la secuenciación de pequeños RNAs de muestras de 6 pacientes con LES y 3 controles sanas. Los resultados obtenidos sirvieron para tener una primera aproximación de los perfiles de miRNAs asociados a la actividad de la enfermedad.

Los resultados obtenidos sugieren que las pacientes que se encuentran cursando un periodo de actividad de la enfermedad presentan un perfil de miRNAs característico que las diferencia de las pacientes en remisión y controles sanas. A pesar de haber contado con un número limitado de muestras, las pacientes activas mostraron un agrupamiento bastante claro al realizar un análisis de componentes principales (Figura 15 B). Al realizar una prueba permanova, se obtiene que el agrupamiento de las pacientes activas no tiene significancia estadística. Esto puede deberse a varias razones. Primero, se cuenta solamente con 3 réplicas de cada condición. Esto limita la sensibilidad del análisis para detectar diferencias sutiles entre las muestras. En segundo lugar, la alta variabilidad intra-grupo puede estar afectando los resultados. Si bien las pacientes activas parecen agruparse visualmente, la variabilidad dentro de cada grupo podría ser suficientemente alta como para diluir la señal de agrupamiento en el análisis estadístico. Esto significa que las diferencias en los perfiles de miRNAs entre las pacientes activas, las pacientes en remisión y las controles sanas podrían no ser suficientemente pronunciadas para superar la dispersión natural de los datos. Además, la falta de significancia estadística podría también indicar la presencia de variables no controladas (*confounding variables*) que influyen en los resultados pero que no fueron consideradas en el diseño del experimento. Estos factores podrían estar introduciendo variabilidad adicional, dificultando la identificación de patrones claros de agrupamiento basados exclusivamente en la condición de la enfermedad. A pesar de que el agrupamiento no haya mostrado significancia estadística, no se descarta la existencia de diferencias que sean biológicamente relevantes entre los grupos. Los resultados observados proporcionan una base para estudios futuros con un mayor número de muestras y un diseño experimental que permita minimizar la variabilidad entre las pacientes así como la variabilidad experimental. A continuación se describen las fuentes de variabilidad que podrían estar influyendo en este análisis y se discuten posibles estrategias para afrontarlas.

Las pacientes con LES presentan una inmensa heterogeneidad, esto además de ser un desafío para el diagnóstico de la enfermedad, es un desafío para su estudio. Las pacientes reclutadas para este trabajo presentaron distintos síntomas, períodos de actividad más o menos graves, y distintos perfiles de autoanticuerpos. La paciente en actividad A3 padece una enfermedad de mayor gravedad ya que presenta una mayor cantidad de manifestaciones clínicas, incluyendo manifestaciones renales. La nefritis lúpica es una de las consecuencias más graves del LES, en la que los riñones se encuentran comprometidos al ser dañados por el sistema inmune, poniendo en riesgo la vida de las pacientes. Existen muchos trabajos que han estudiado la expresión de miRNAs particularmente en el contexto de la nefritis lúpica, asociando múltiples miRNAs a esta condición <sup>71</sup>. Otros esfuerzos para homogeneizar la población en estudio se han realizado separando a las pacientes en distintos grupos según sus perfiles de autoanticuerpos. Los

autoanticuerpos presentes en los pacientes con LES se pueden clasificar en dos grupos principales según estén dirigidos a DNA doble cadena (anti-dsDNA) o a antígenos asociados a RNA (también conocidos como anti-ENA). Hay trabajos que sugieren que los mecanismos inmunorreguladores que conducen a los distintos perfiles de autoanticuerpos son diferentes, pero que la identificación de distintos eventos específicos de la enfermedad a menudo se pasa por alto debido a la gran heterogeneidad de los repertorios de autoanticuerpos de los pacientes. Al clasificar a los pacientes según sus autoanticuerpos, se observó que la presencia de autoanticuerpos anti-ENA y anti-dsDNA se asoció a cambios en los niveles de expresión de receptores de ácidos nucleicos del sistema inmune. La presencia de autoanticuerpos anti-ENA se vio asociada a niveles elevados de TLR7 (sensor de RNA) mientras que la presencia de anticuerpos anti-dsDNA se asoció con niveles elevados de TLR9 (sensor de DNA)<sup>72</sup>. Además, los distintos perfiles de autoanticuerpos se han asociado a cambios en los niveles de distintos miRNAs<sup>73</sup> y mRNAs<sup>74</sup>. Este tipo de estrategias para identificar subgrupos de pacientes con características comunes puede resultar útil para abordar la complejidad de la enfermedad, identificando perfiles de miRNAs asociados a procesos patológicos específicos. Para ello, será necesario reclutar nuevas pacientes y repetir el análisis, evaluando los agrupamientos y perfiles de miRNAs de subgrupos de pacientes que compartan características comunes.

Además de heterogeneidad de las pacientes, hay otros factores capaces de introducir variabilidad en los datos durante la extracción y preparación de las muestras. Las pacientes con lupus activo fueron reclutadas a medida que llegaban a la Policlínica de Enfermedades Autoinmunes del Hospital Maciel. Esto implicó que las muestras de plasma estuvieran almacenadas a -80 °C por períodos de tiempo diferentes, con una diferencia de 11 meses entre la llegada de la muestra de la primer paciente activa a la última. Otro factor importante es el grado de hemólisis de las muestras, que en este trabajo no fue determinado. La hemólisis conlleva a la liberación del contenido de los glóbulos rojos, incluyendo sus miRNAs. Muestras con distintos grados de hemólisis presentarán niveles distintos en los miRNAs más expresados por los glóbulos rojos, y en RNA-Seq, cuando un transcrito aumenta, significa que va a haber otros que disminuyen. Por este motivo la hemólisis es capaz de influir de forma no despreciable en los resultados al alterar las concentraciones de miRNAs. Para considerar esta variable, al momento de incluir nuevas pacientes a este estudio se plantea evaluar el grado de hemólisis de las muestras mediante la determinación de hemoglobina en las preparaciones de plasma. Finalmente, las diferencias en los tamaños de las bibliotecas N1 y R2 también es una fuente de variabilidad adicional que podría estar jugando en contra del agrupamiento de las pacientes activas. Será fundamental no incluir en futuros análisis bibliotecas que presenten este problema.

El análisis de los tamaños de las poblaciones de pequeños ncRNAs es importante, ya que distintos ncRNAs o sus fragmentos se caracterizan por su tamaño. Este análisis, permitió la identificación de diferencias en los tamaños de las bibliotecas N1 y R2, con una mayor representación fragmentos de 21 nt respecto al resto de las muestras. Al observar los resultados de la composición por biotipo (Figura 6) una posible interpretación podría haber sido que la muestra de la paciente en remisión R2 presentaba una disminución de YRNAs en su circulación por motivos biológicos, caracterizada por una menor expresión del RNY4 y una mayor expresión del resto de los YRNAs. Sin embargo, si bien no podemos afirmarlo, pensamos que estamos frente a

un sesgo técnico, y que los fragmentos de alrededor de 33 nt fueron excluidos por motivos técnicos durante la preparación de estas muestras. Las lecturas de RNA-seq se dividen proporcionalmente en función de la abundancia de cada transcrito. Por lo tanto, los RNAs poco abundantes compiten con especies de RNAs muy abundantes. Por lo que una posibilidad es que lo observado se deba a la exclusión de los ncRNAs más abundantes en las muestras de plasma que son los fragmentos 5' del RNY4 de 33 nt. Por lo tanto, para las muestras N1 y R2 se obtuvieron una mayor cantidad de secuencias de especies de menor tamaño que para el resto de las muestras. Los motivos por los cuales las bibliotecas N1 y R2 se componen de fragmentos más pequeños pueden ser varios, y no es posible determinar exactamente qué fue lo que ocurrió. Desde el momento de la toma de muestra hasta su procesamiento se piensa que puede ser el momento más crítico. También la preparación de bibliotecas ocurre en múltiples pasos y en cada uno de estos pasos pueden introducirse distintos sesgos. Durante la preparación de estas muestras los fragmentos de alrededor de 30 nt pueden haber estado expuestos a una mayor fragmentación o degradación, también podrían haber tenido problemas durante la retrotranscripción o incluso haber sido excluidos por error durante la selección de los tamaños de las bibliotecas al momento de cortar el gel. Este resultado reafirma la importancia del control de los tamaños de las bibliotecas de pequeños RNAs, ya que los tamaños de estos y de sus fragmentos derivados determinan qué RNAs serán secuenciadas.

Mediante un análisis de expresión diferencial se identificó un conjunto de miRNAs cuya expresión se encuentra alterada en las pacientes que están cursando un período de actividad (Figura 16). Este perfil está caracterizado por una sobreexpresión de miR-99b-5p y una disminución de miR-150-5p, miR-29a-3p y miR-30b-5p respecto a las controles sanas. Mientras que al compararlas con las pacientes en remisión mostraron una sobreexpresión de miR-382-5p, miR-409-3p, miR-221-3p, miR-543, miR-4446-3p y let-7e-5p; y una subexpresión de miR-miR-140-3p, miR-30b-5p, miR-363-3p, 374a-5p, 150-5p, miR-144-3p, miR-26b-5p, miR-106b-5p, miR-3158-3p, miR-486-5p y miR-29a-3p.

Los niveles de expresión de los transcritos diferencialmente expresados son relativamente bajos para algunos miRNAs (Tabla A2). En pos de mejorar la reproducibilidad y el rigor en la investigación de microRNAs, se recomienda que se respeten niveles basales de expresión, considerando y reportando un umbral mínimo. Para la secuenciación de microRNAs, una regla general es que deben tener valores superiores a 100 RPM <sup>75</sup>. Para varios de los miRNAs reportados en esta tesis, esto no se cumple y por lo tanto hay un riesgo, que es el de estar describiendo algo que simplemente es ruido como un resultado. La gran mayoría de las secuencias obtenidas (alrededor del 70 %) corresponden al RNY4 (Figura 7), y estas resultaron no informativas para este análisis. Por este motivo, se plantea que los miRNAs con bajo recuento de secuencias deben ser considerados como un resultado preliminar, a confirmar mediante la generación de nuevas bibliotecas que no contengan YRNAs, que permitan un secuenciado más profundo de las especies de interés.

A la hora de reflexionar sobre los perfiles de miRNAs obtenidos, es fundamental tener presente que se está trabajando con una muestra que es extracelular, y esto significa que no se conoce qué células son las que expresaron a los RNAs que fueron secuenciados. Además, en el caso del

plasma, se trata de un biofluido circulante al cual distintos tipos celulares tienen acceso. Un aumento en la expresión de un miRNA podría significar que hay un tipo celular, o varios, en los que este miRNA se está expresando más, pero también podría significar que hay un aumento del o los tipos celulares que lo expresan. De forma que los cambios en los perfiles de miRNAs podrían deberse a distintos motivos, habiendo más motivos posibles por tipo celular en el que se exprese dicho miRNA. Por esto, se plantea que a la hora de determinar el potencial de un miRNA como biomarcador del LES es importante considerar qué tipos celulares son los que lo expresan, ya que probablemente sean mejores candidatos aquellos miRNAs que son expresados por tipos celulares específicos, particularmente por células del sistema inmune, ya que de esta forma el miRNA en cuestión podría relacionarse con un proceso patológico puntual, teniendo así un mayor sustento. La expresión no ubicua como requerimiento para que un miRNA sea un buen biomarcador de una enfermedad dada ha sido planteada por otros trabajos <sup>35,76</sup>, y se relaciona muy estrechamente con la especificidad del biomarcador. Utilizando el *Genome Browser* de la Universidad de California, Santa Clara (UCSC) con el *track* de *microRNAome* celulares humanos, es posible acceder a los niveles de expresión de miRNAs en 78 tipos de células primarias. Mediante este recurso, se observó que los microRNAs miR-99b-5p, miR382-5p, miR-409-3p, miR-221-3p, miR-543, let-7e-5p, miR-26b-5p, miR-3158-3p y miR-29a-3p son expresados por múltiples tipos celulares (Figuras A1 – A9), lo que podría poner en duda su potencial como biomarcadores específicos del LES. Sin embargo, se podría plantear que, si bien no se trata de miRNAs específicos de tipos celulares relevantes para la enfermedad, estos miRNAs podrían formar parte, en conjunto, de un perfil de miRNAs asociados al LES. Este enfoque probablemente aporta más poder al estudio que la identificación de un solo miRNA. Cambios en la expresión similares a los observados en este trabajo han sido identificados en el contexto de otras patologías para estos miRNAs, lo que llevó a que sean propuestos como biomarcadores de otras enfermedades. Por ejemplo, un aumento en la expresión de miR-99b en conjunto con miR-409 en exosomas derivados de plasma fue asociado con el desarrollo de cáncer colorrectal, y por lo tanto estos miRNAs fueron propuestos como potenciales biomarcadores para el diagnóstico temprano de este tipo de cáncer <sup>77</sup>. En el caso del miR-221, este miRNA fue propuesto como un biomarcador de melanoma, dado su aumento en la expresión en pacientes que lo padecen <sup>78</sup>. El miR-221 también fue propuesto como un biomarcador de la enfermedad de Parkinson dado su aumento en plasma de pacientes respecto a controles <sup>79</sup>. Estos son algunos ejemplos que muestran que la expresión ubicua de los miRNAs se asocia a una menor especificidad diagnóstica. De todas formas, si al aumentar el número de pacientes este perfil se mantiene, podrían contribuir al diagnóstico y seguimiento de la enfermedad. Sobre todo, pensando en pacientes ya diagnosticadas, donde la especificidad del diagnóstico ya deja de ser un problema.

Dentro del perfil de miRNAs asociados a la actividad del LES, hay algunos miRNAs que tienen una expresión más limitada, y se expresan en un tipo celular o conjunto de células puntuales. Estos son miR-4446, miR-144, miR-106b y miR-150. El miR-4446 se expresa en plaquetas (Figura A10). Una mayor expresión de miR-4446 en pacientes activas resulta llamativa, ya que la actividad de la enfermedad suele asociarse a un menor recuento de plaquetas, conocido como plaquetopenia o trombocitopenia <sup>80</sup>. Este resultado parece ser contraintuitivo en el sentido de que, si los períodos de actividad del LES se asocian a un menor recuento de plaquetas, se esperaría que miR-4446 disminuya en lugar de aumentar en estas pacientes. Lo que sucede es que, a veces, si bien

una paciente puede considerarse en remisión clínica, puede mantener algunas características más asociadas a la actividad de la enfermedad. Al observar las manifestaciones clínicas, se puede ver que todas las pacientes en remisión presentan manifestaciones hematológicas, mientras que las pacientes activas no las presentan (Tabla 1). Dentro de las manifestaciones hematológicas, se considera el recuento de plaquetas, y en este caso las pacientes en remisión mostraron en general un recuento de plaquetas menor que las pacientes activas. De forma que la expresión de miR-4446 mantiene una mayor relación con el fenotipo hematológico de las pacientes que con la actividad de la enfermedad. Si bien lo esperado sería que la plaquetopenia se asocie a la actividad de la enfermedad, la heterogeneidad del LES se vuelve a ver reflejada, y en este caso particular, con un número limitado de muestras, pasa lo contrario. En este trabajo las pacientes con diagnóstico de LES fueron divididas según estén cursando o no un periodo de actividad. La idea detrás de este diseño es poder identificar miRNAs que puedan facilitar el diagnóstico inicial de la enfermedad, al comparar a las pacientes activas respecto a las controles sanas, pero también poder realizar un seguimiento de la evolución de la enfermedad en periodos de actividad y remisión de pacientes que ya han sido diagnosticadas. Al comparar las pacientes en actividad respecto a las controles sanas las diferencias observadas es esperable que estén dadas por la actividad de la enfermedad. Sin embargo, al comparar las pacientes en actividad con las pacientes en remisión hay una mayor complejidad, ya que las diferencias pueden deberse a ambas condiciones, y miR-4446 es un ejemplo de esto. Los cambios observados para el miR-4446, se corresponden con una disminución de este miRNA en pacientes en remisión, debido a un menor recuento de plaquetas, y no mantiene una relación clara con la actividad clínica de la enfermedad para estas pacientes.

El miR-144 es expresado por glóbulos rojos (Figura A11) y se encuentra disminuido en pacientes activas frente a controles sanas y pacientes en remisión. Se ha propuesto que miRNAs expresados por glóbulos rojos no serían adecuados como biomarcadores de cáncer, ya que podrían reflejar fenómenos basados en células sanguíneas en lugar de relacionarse con un tipo específico de cáncer<sup>81</sup>. El mismo razonamiento podría aplicarse al caso del LES, si bien podría plantearse que alguna de las alteraciones de los glóbulos rojos podría tener que ver con la patología, habría que ser cuidadosos con la interpretación de este resultado, y poder confirmar que se debe a una característica de los glóbulos rojos de las pacientes, descartando la influencia de sesgos técnicos, probablemente asociados a la hemólisis, de este resultado.

El miR-106b, es expresado mayormente por linfocitos B CD77 (Figura A12). La molécula CD77 define una vía de maduración de linfocitos B, específica para el centro germinal, y lo expresan linfocitos que están experimentando apoptosis<sup>82</sup>. Este miRNA se encuentra disminuido en las pacientes activas y podría relacionarse con una menor cantidad de células en apoptosis en el centro germinal. El centro germinal es una estructura de los órganos linfoides secundarios, donde operan los mecanismos de tolerancia periféricos. Además, resulta interesante que este miRNA también fue descrito subexpresado respecto a controles en muestras de pacientes con Síndrome de Sjögren<sup>83</sup>. En pacientes con LES fue reportada una disminución del miR-106a respecto a pacientes control<sup>34</sup>. Este miRNA, pertenece a la misma familia que el miR-106b, están evolutivamente relacionados, pero presentan diferencias en su secuencia. Sin embargo, mantienen la misma secuencia semilla (AAAGUGC), lo que indica que podrían compartir sus

mRNAs blanco. La predicción de los mRNA blancos de miR-106b-5p tiene como resultado cientos de posibilidades. Este miRNA resulta interesante como candidato a ser un biomarcador de EA, y resulta interesante confirmar este resultado mediante la inclusión de nuevas pacientes a este estudio.

Finalmente, el miR-150 es expresado por células linfoides (Figura A13), y esto resulta interesante en el contexto de una enfermedad autoinmune. Este miRNA ha sido asociado con diversos procesos patológicos que involucran linfocitos B<sup>84</sup>. La expresión de miR-150 está implicada en la tolerancia del sistema inmune. Uno de los mecanismos de tolerancia es la apoptosis inducida por BCR de las células B inmaduras autorreactivas. Una menor expresión de miR150 en estas células debilita la señalización a través del BCR previniendo la apoptosis de células B autorreactivas, y, por lo tanto, promueve el crecimiento de células B autorreactivas, promoviendo el desarrollo de enfermedades asociadas a linfocitos B, entre ellas, enfermedades autoinmunes<sup>85</sup>. Siguiendo este razonamiento, linfocitos B autorreactivos con una menor expresión de miR-150 tendrían más chances de escapar los mecanismos de tolerancia del sistema inmune, y eventualmente en contextos que así lo favorezcan, conducir a la producción de autoanticuerpos. Uno de los mRNA blancos estudiados es el del factor de transcripción c-Myb, a través del cual se propone que miR-150 controla la diferenciación de linfocitos B, influyendo sobre la respuesta humoral. La deficiencia de miR-150 conduce a la expansión de las células B y a una respuesta humoral aumentada<sup>86</sup>. El miR-150 por lo tanto parece un miRNA interesante a la hora de pensar en un biomarcador del LES. En este trabajo, la expresión de miR-150 fue encontrada disminuida en pacientes activas tanto al comparar con controles sanas como con pacientes en remisión, de forma que es un buen candidato a biomarcador de la actividad del LES. Será importante confirmar este resultado mediante la inclusión de nuevas pacientes a este estudio.

Las especies de RNAs muy abundantes pueden dominar la biblioteca de secuenciación, obstaculizando la detección de transcritos menos abundantes. La abundancia del RNY4 en el plasma puede compararse a lo que ocurre con el rRNA en el interior de las células. El rRNA representa alrededor del 80% del RNA celular, al tiempo que resulta poco informativo para los estudios transcriptómicos que pretenden evaluar la expresión de mRNAs, y por lo tanto el rRNA suele ser depletado para ese tipo de experimentos. Los YRNAs se fragmentan en células en apoptosis, y estos fragmentos se encuentran en múltiples biofluidos, siendo los más abundantes los fragmentos 5' de 30-33 nt. En las bibliotecas generadas en este trabajo los fragmentos del RNY4 llegaron a representar alrededor del 70 % de las lecturas obtenidas para distintas muestras. Estas grandes cantidades de secuencias del RNY4 impactan negativamente en la complejidad de la biblioteca, haciendo que se requiera una secuenciación más profunda para poder secuenciar otras especies de RNA menos abundantes. Las secuencias del RNY4 no fueron consideradas a la hora de realizar la búsqueda de posibles biomarcadores, de forma que el RNA más secuenciado resultó ser no informativo para este trabajo. Para atender este problema, existen métodos que permiten excluir a los YRNAs de las bibliotecas<sup>87</sup> y se plantea utilizarlos a futuro al momento de ampliar el número de pacientes.

Existen numerosas investigaciones sobre microRNAs en el contexto del LES. Al realizar una búsqueda en PubMed con las palabras "*miRNA biomarker SLE*" se obtienen un total de 185

publicaciones, siendo el primer trabajo que propone miRNAs como biomarcadores del LES publicado en 2007<sup>88</sup>. A pesar de las numerosas publicaciones en el tema, en ocasiones la literatura puede ser un poco confusa debido a reportes de resultados contradictorios y la falta de acuerdo en el tipo de muestra a estudiar. Para el estudio de exRNAs, se conoce que el protocolo de preparación del plasma es una fuente importante de variación, ya que la composición de las especies de exRNAs varía con la abundancia de las plaquetas<sup>38</sup>. Otra consideración importante es la elección del tubo de extracción de sangre. Estos tubos tienen un anticoagulante que previene la coagulación de la sangre luego de la extracción. La elección del anticoagulante (EDTA o citrato) e incluso el fabricante influye en los resultados de los componentes del plasma<sup>89</sup>. Otras variables como el tiempo hasta la congelación y la temperatura de almacenamiento también podrían influir en los resultados<sup>90</sup>. Por otro lado, están las diferencias en las técnicas utilizadas para estudiar la expresión de los miRNAs. Incluso habiendo tratado a las muestras de la misma manera, los resultados obtenidos por microarrays y RNA-seq pueden ser distintos, sobre todos para miRNAs de baja expresión. Estas diferencias, dificultan las comparaciones justas entre los resultados de los distintos trabajos, y por lo tanto se plantea que se debe tener cuidado al hacer este tipo de comparaciones. En su reciente revisión Choi et al. reportan alrededor de 90 miRNAs que han sido reportado alterados en el LES por distintos trabajos. De forma que las chances de que los resultados de un trabajo y otro coincidan, simplemente por azar, no son demasiado despreciables. Para avanzar en el estudio de miRNAs en LES sería altamente beneficioso establecer protocolos estandarizados que aseguren condiciones uniformes para el procesamiento de muestras, así como informar el umbral mínimo establecido para la detección de un transcrito, evitando reportes que podrían ser en realidad “ruido”.

Por otro lado, no basta con identificar perfiles de miRNAs con expresión diferencial en pacientes, sino que el conocimiento en este tema podría enriquecerse considerablemente mediante la incorporación de criterios adicionales. Estos criterios ayudarían a refinar y validar las largas listas de miRNAs candidatos a biomarcadores del LES, facilitando la identificación de aquellos con mayor potencial clínico y relevancia diagnóstica. Gran parte de la literatura en el tema se ha limitado a identificar y reportar miRNAs expresados diferencialmente en distintas muestras de pacientes respecto a controles sanos. De forma que hay decenas de miRNAs cuya expresión se reportó aumentada o disminuida en las pacientes, pero sin embargo para la mayoría de ellos no hay evidencia directa sobre su rol en LES, las vías sobre las que podrían estar actuando, e incluso muchas veces no se conoce cuáles son sus mRNAs blanco. Los mRNA blanco suelen predecirse mediante estrategias bioinformáticas y en base a estas predicciones se infieren las vías sobre las que podrían estar actuando. Los miRNAs pueden tener muchos mRNAs blanco, y dependerá del contexto celular la función que puedan ejercer. Además, dado que la predicción de los blancos se realiza en base a las regiones de semilla o *seed* cortas (de 6-8 nt) pueden predecirse muchos blancos simplemente por azar, identificando vías reguladas por un microRNAs in silico, potencialmente sin sentido biológico. Por otro lado, la regulación de la expresión génica mediada por miRNAs también depende los niveles de mRNA blanco, y de la interacción de un complejo de silenciamiento activo con el mismo<sup>75</sup>. Este tipo de estrategias si bien puede resultar útil para luego ir a probar las hipótesis, por sí sola no confirma el rol del miRNA en la patogénesis.

Actualmente, no hay estudios en fase clínica con resultados disponibles que evalúen el uso de un miRNA como biomarcador de LES. Una búsqueda en ClinicalTrials.gov reveló un único estudio clínico titulado “Circulating MicroRNAs in Understanding Pathogenesis of Systemic Lupus Erythematosus” (ClinicalTrials.gov ID: NCT02756546). Este estudio, iniciado en 2014, tiene como objetivo identificar firmas de expresión de miR-181a, miR-196a y miR-21 en el plasma de pacientes con LES, evaluando diferencias en su expresión entre grupos con diferentes puntuaciones SLEDAI. Hasta la fecha, no se han publicado en el sitio resultados. Los 3 miRNAs evaluados en este estudio se expresan prácticamente en todas las células <sup>69</sup> y particularmente miR-21 ha sido sugerido como un biomarcador de casi 30 enfermedades distintas <sup>35</sup>, entonces si bien probablemente están siendo evaluados en base a reportes previos que los encontraron alterados, hay elementos para sospechar que parecerían no ser los mejores candidatos. A pesar del gran volumen de trabajos publicados sobre miRNAs como biomarcadores en LES, ninguno de los miRNAs propuestos ha superado el cuello de botella que representa la validación clínica. Esto subraya la necesidad de más investigaciones y estudios clínicos, pero sobre todo la necesidad de una mayor rigurosidad a la hora de proponer un miRNA como prometedor biomarcador del LES.

## Conclusiones

Durante el desarrollo de esta tesis se estableció una colaboración con la Policlínica de Enfermedades Autoinmunes del Hospital Maciel, que permitió la obtención de muestras de 3 pacientes con LES activo, 3 pacientes con LES en remisión y 3 controles sanas. A partir de estas muestras se logró generar bibliotecas que fueron secuenciadas en la plataforma MiSeq disponible en el IPMon. El análisis de los datos obtenidos mostró que, a pesar de tratarse de un grupo diverso, las pacientes con LES en actividad comparten algunas características en cuanto a la composición de miRNAs en plasma. Estos resultados sugieren la existencia de un perfil de miRNAs asociados a la actividad del LES, y sientan las bases de una línea de trabajo a ser profundizada con la finalidad de confirmar estos resultados en un mayor número de pacientes, así como evaluar el potencial de estos miRNAs como biomarcadores de diagnóstico de la enfermedad y seguimiento de su evolución. Para ello, se propone adoptar un pensamiento crítico, estableciendo criterios que permitan evaluar el potencial de los miRNAs como biomarcadores del LES, considerando qué tan informativos pueden llegar ser en el contexto de la enfermedad. En este sentido se plantea que los miRNAs de expresión ubicua no son los mejores candidatos para brindar especificidad al diagnóstico, así como tampoco lo son, a priori, los miRNAs expresados por glóbulos rojos o plaquetas. Sin embargo, no se descarta que estos puedan contribuir a la anticipación de períodos de actividad en pacientes ya diagnosticadas al ser evaluados en conjunto como firma molecular. Por otro lado, resultan interesantes aquellos miRNAs expresados por tipos celulares involucrados en la patología y que por lo tanto tengan una mayor especificidad. En este sentido se plantea como un resultado interesante la disminución de miR-150 y miR-106b en el plasma de las pacientes cursando un periodo de actividad y se propone confirmar este resultado mediante un aumento en la profundidad de secuenciación de miRNAs por sobre otras especies no informativas. Se trata de dos miRNAs específicos de linfocitos, y sus fluctuaciones han sido asociadas a procesos relevantes para la patología del LES. Finalmente, se plantea la importancia de mejorar la reproducibilidad y el rigor en la investigación de microRNAs en LES. En base al trabajo realizado se identificaron distintos aspectos técnicos a ajustar en pos de mejorar la reproducibilidad de los resultados en siguientes etapas de este estudio.

## Perspectivas

Los resultados de esta tesis constituyen un punto de partida para futuros ensayos en los que se buscará continuar explorando el potencial de miRNAs como biomarcadores en biopsias líquidas mínimamente invasivas. Actualmente, el Laboratorio de Genómica Funcional se encuentra llevando a cabo el proyecto “Identificación y caracterización de pequeños RNAs inmunoreguladores como nuevas vías patogénicas y fuentes de biomarcadores en Lupus Eritematoso Sistémico”. Este proyecto contempla entre sus objetivos darle continuidad al estudio miRNAs en plasma de pacientes con LES activo. En marco de este proyecto, se plantea entonces reclutar un mayor número de pacientes, y continuar obteniendo datos de secuenciación masiva. Contar con una mayor cantidad de réplicas sumará poder estadístico a los resultados, y además permitirá abordar la heterogeneidad de la enfermedad al poder formar distintos subgrupos de pacientes que compartan manifestaciones y perfiles de autoanticuerpos similares. Además, resultará interesante poder realizar un seguimiento de la evolución de las pacientes, accediendo a muestras de las mismas pacientes a medida que sean atendidas en la Policlínica por la reaparición de síntomas y manifestaciones. Para llevar a cabo esto, se ajustarán los protocolos utilizados en esta tesis, incluyendo la determinación del grado de hemólisis de las muestras, la depleción de YRNAs, y un control riguroso del tamaño de las bibliotecas generadas. La hemólisis es un efecto preanalítico evitable en la mayoría de los casos. Su aparición se debe a la técnica de extracción empleada y a las condiciones de transporte y preparación de las muestras. Conocer el grado de hemólisis de las muestras permitirá determinar la calidad de las mismas y en base a esto decidir si pueden ser o no incluidas en el estudio. Esto permitirá tener una *counfounding variable* menos, evitando las fluctuaciones en los recuentos producto de la liberación de miRNAs abundantes de glóbulos rojos. Por otro lado, excluir a los YRNAs de las bibliotecas permitirá un secuenciado más profundo de las especies de interés. Finalmente, otra lección aprendida de este trabajo es la importancia de determinar los tamaños de las bibliotecas. De forma que se planeta continuar realizando un análisis de los tamaños de las especies secuenciadas, y excluyendo del análisis a todas las bibliotecas que presenten sesgos de este tipo.

## Referencias

1. MBBS, A. K. A., PhD, A. H. L. M. & PhD, S. P. M. *Basic Immunology: Functions and Disorders of the Immune System*. (Elsevier, 2019).
2. Hoi, A., Igel, T., Mok, C. C. & Arnaud, L. Systemic lupus erythematosus. *Lancet* **403**, 2326–2338 (2024).
3. Tian, J., Zhang, D., Yao, X., Huang, Y. & Lu, Q. Global epidemiology of systemic lupus erythematosus: a comprehensive systematic analysis and modelling study. *Ann. Rheum. Dis.* **82**, 351–356 (2023).
4. Coppes, E., Lema, M. & Pacello, PhD. MSc. F. Caracterización Clínica y Epidemiológica de pacientes con Lupus en Paysandú. Un estudio descriptivo, observacional. *anfamed.edu.uy* (2021).
5. Guga, S., Wang, Y., Graham, D. C. & Vyse, T. J. A review of genetic risk in systemic lupus erythematosus. *Expert Rev. Clin. Immunol.* **19**, 1247–1258 (2023).
6. Hedrich, C. M., Mbert, K., Rauen, T. & Tsokos, G. C. DNA methylation in systemic lupus erythematosus. *Epigenomics* **9**, 505–525 (2017).
7. Javierre, B. M. *et al.* Changes in the pattern of DNA methylation associate with twin discordance in systemic lupus erythematosus. *Genome Res.* **20**, 170–179 (2010).
8. James, J. A. & Robertson, J. M. Lupus and Epstein-Barr. *Curr. Opin. Rheumatol.* **24**, 383–388 (2012).
9. Greiling, T. M. *et al.* Commensal orthologs of the human autoantigen Ro60 as triggers of autoimmunity in lupus. *Sci. Transl. Med.* **10**, (2018).
10. Ma, Q., Zhou, S.-L., Dang, J., Huo, Z.-H. & Ma, Z.-B. Genes that escape from X-chromosome inactivation and sexual dimorphism of systemic lupus erythematosus. *Yi chuan Hered.* **46**, 18–33 (2024).
11. Tsokos, G. C. & Smolen, C. G. and J. S. *Systemic Lupus Erythematosus. A Companion to Rheumatology*. (2007).
12. Bombardier, C. *et al.* Derivation of the sledai. A disease activity index for lupus patients. *Arthritis Rheum.* **35**, 630–640 (1992).
13. Aringer, M. *et al.* 2019 European League Against Rheumatism/American College of Rheumatology Classification Criteria for Systemic Lupus Erythematosus. *Arthritis Rheumatol.* **71**, 1400–1412 (2019).
14. Nemeth, K., Bayraktar, R., Ferracin, M. & Calin, G. A. Non-coding RNAs in disease: from mechanisms to therapeutics. *Nat. Rev. Genet.* **25**, 211–232 (2024).
15. Wang, H., Peng, W., Ouyang, X., Li, W. & Dai, Y. Circulating microRNAs as candidate biomarkers in patients with systemic lupus erythematosus. *Transl. Res.* **160**, 198–206 (2012).

16. Chen, X., Cheng, Q., Wu, H. & Du, Y. Long non-coding RNAs in systemic lupus erythematosus: New insights into disease pathogenesis and diagnosis. *Scand. J. Immunol.* **95**, e13167 (2022).
17. Choi, D. *et al.* Dysregulated MicroRNAs in the Pathogenesis of Systemic Lupus Erythematosus: A Comprehensive Review. *Int. J. Biol. Sci.* **19**, 2495–2514 (2023).
18. Huang, Z., Du, Y., Wen, J., Lu, B. & Zhao, Y. snoRNAs: functions and mechanisms in biological processes, and roles in tumor pathophysiology. *Cell Death Discov.* **8**, 259 (2022).
19. Christov, C. P., Gardiner, T. J., Szüts, D. & Krude, T. Functional Requirement of Noncoding Y RNAs for Human Chromosomal DNA Replication. *Mol. Cell. Biol.* **26**, 6993–7004 (2006).
20. Fuchs, G., Stein, A. J., Fu, C., Reinisch, K. M. & Wolin, S. L. Structural and biochemical basis for misfolded RNA recognition by the Ro autoantigen. *Nat. Struct. Mol. Biol.* **13**, 1002–1009 (2006).
21. Cheng, Z., Hou, G. & Shen, N. Evolving understandings for the roles of non-coding RNAs in autoimmunity and autoimmune disease. *Journal of Autoimmunity* (2023) doi:10.1016/j.jaut.2022.102948.
22. Liu, H., Zou, Y., Chen, C., Tang, Y. & Guo, J. Current Understanding of Circular RNAs in Systemic Lupus Erythematosus. *Front. Immunol.* **12**, 628872 (2021).
23. PILLAI, R. S. MicroRNA function: Multiple mechanisms for a tiny RNA? *RNA* **11**, 1753–1761 (2005).
24. Esquela-Kerscher, A. & Slack, F. J. Oncomirs — microRNAs with a role in cancer. *Nat. Rev. Cancer* **6**, 259–269 (2006).
25. Zhang, L., Wu, H., Zhao, M., Chang, C. & Lu, Q. Clinical significance of miRNAs in autoimmunity. *J. Autoimmun.* **109**, 102438 (2020).
26. Pérez-Sánchez, C. *et al.* ‘Atherothrombosis-associated microRNAs in Antiphospholipid syndrome and Systemic Lupus Erythematosus patients.’ *Sci. Rep.* **6**, 31375 (2016).
27. Lu, M.-C. *et al.* Decreased microRNA(miR)-145 and increased miR-224 expression in T cells from patients with systemic lupus erythematosus involved in lupus immunopathogenesis. *Clin. Exp. Immunol.* **171**, 91–99 (2012).
28. Karonitsch, T. *et al.* Activation of the interferon- $\gamma$  signaling pathway in systemic lupus erythematosus peripheral blood mononuclear cells. *Arthritis Rheum.* **60**, 1463–1471 (2009).
29. Shi, X. *et al.* Downregulated miR-29a promotes B cell overactivation by upregulating Crk-like protein in systemic lupus erythematosus. *Mol. Med. Rep.* **22**, 841–849 (2020).
30. Ren, D. *et al.* Activation of TLR7 increases CCND3 expression via the downregulation of miR-15b in B cells of systemic lupus erythematosus. *Cell. Mol. Immunol.* **13**, 764–775 (2016).

31. Khoshmirsafa, M. *et al.* Elevated expression of miR-21 and miR-155 in peripheral blood mononuclear cells as potential biomarkers for lupus nephritis. *Int. J. Rheum. Dis.* **22**, 458–467 (2019).
32. Kaga, H. *et al.* Downregulated expression of miR-155, miR-17, and miR-181b, and upregulated expression of activation-induced cytidine deaminase and interferon- $\alpha$  in PBMCs from patients with SLE. *Mod. Rheumatol.* **25**, 865–870 (2015).
33. Zhang, H. *et al.* B Cell-Related Circulating MicroRNAs With the Potential Value of Biomarkers in the Differential Diagnosis, and Distinguishment Between the Disease Activity and Lupus Nephritis for Systemic Lupus Erythematosus. *Front. Immunol.* **9**, 1473 (2018).
34. Carlsen, A. L. *et al.* Circulating MicroRNA Expression Profiles Associated With Systemic Lupus Erythematosus. *Arthritis Rheum.* **65**, 1324–1334 (2013).
35. Jenike, A. E. & Halushka, M. K. miR-21: a non-specific biomarker of all maladies. *Biomark. Res.* **9**, 18 (2021).
36. MANDEL, P. & METAIS, P. Nuclear Acids In Human Blood Plasma. *C. R. des seances Soc. Biol. ses Fil.* **142**, 241–3 (1948).
37. Godoy, P. M. *et al.* Large Differences in Small RNA Composition Between Human Biofluids. *Cell Rep.* **25**, 1346–1358 (2018).
38. Hulstaert, E. *et al.* Charting Extracellular Transcriptomes in The Human Biofluid RNA Atlas. *Cell Rep.* **33**, 108552 (2020).
39. Tosar, J. P. *et al.* Assessment of small RNA sorting into different extracellular fractions revealed by high-throughput sequencing of breast cell lines. *Nucleic Acids Res.* **43**, 5601–5616 (2015).
40. El-Mogy, M. *et al.* Diversity and signature of small RNA in different bodily fluids using next generation sequencing. *BMC Genom.* **19**, 408 (2018).
41. Turchinovich, A., Drapkina, O. & Tonevitsky, A. Transcriptome of Extracellular Vesicles: State-of-the-Art. *Front. Immunol.* **10**, 202 (2019).
42. Yao, J., Wu, D. C., Nottingham, R. M. & Lambowitz, A. M. Identification of protein-protected mRNA fragments and structured excised intron RNAs in human plasma by TGIRT-seq peak calling. *eLife* **9**, e60743 (2020).
43. Tosar, J. P. *et al.* Dimerization confers increased stability to nucleases in 5' halves from glycine and glutamic acid tRNAs. *Nucleic Acids Res.* **46**, 9081–9093 (2018).
44. Arroyo, J. D. *et al.* Argonaute2 complexes carry a population of circulating microRNAs independent of vesicles in human plasma. *Proc. Natl. Acad. Sci.* **108**, 5003–5008 (2011).
45. Turchinovich, A., Weiz, L., Langheinz, A. & Burwinkel, B. Characterization of extracellular circulating microRNA. *Nucleic Acids Res.* **39**, 7223–7233 (2011).

46. Albanese, M. *et al.* MicroRNAs are minor constituents of extracellular vesicles that are rarely delivered to target cells. *PLoS Genet.* **17**, e1009951 (2021).
47. Hinger, S. A. *et al.* Diverse Long RNAs Are Differentially Sorted into Extracellular Vesicles Secreted by Colorectal Cancer Cells. *Cell Rep.* **25**, 715–725.e4 (2018).
48. Cha, D. J. *et al.* KRAS-dependent sorting of miRNA to exosomes. *eLife* **4**, e07197 (2015).
49. Villarroya-Beltri, C., Baixauli, F., Gutiérrez-Vázquez, C., Sánchez-Madrid, F. & Mittelbrunn, M. Sorting it out: Regulation of exosome loading. *Semin. Cancer Biol.* **28**, 3–13 (2014).
50. Tosar, J. P., Witwer, K. & Cayota, A. Revisiting Extracellular RNA Release, Processing, and Function. *Trends Biochem. Sci.* **46**, 438–445 (2021).
51. Tosar, J. P. *et al.* Fragmentation of extracellular ribosomes and tRNAs shapes the extracellular RNAome. *Nucleic Acids Res.* **48**, gkaa674 (2020).
52. Valadi, H. *et al.* Exosome-mediated transfer of mRNAs and microRNAs is a novel mechanism of genetic exchange between cells. *Nat. Cell Biol.* **9**, 654–659 (2007).
53. Ratajczak, J. *et al.* Embryonic stem cell-derived microvesicles reprogram hematopoietic progenitors: evidence for horizontal transfer of mRNA and protein delivery. *Leukemia* **20**, 847–856 (2006).
54. Thomou, T. *et al.* Adipose-derived circulating miRNAs regulate gene expression in other tissues. *Nature* **542**, 450–455 (2017).
55. Montecalvo, A. *et al.* Mechanism of transfer of functional microRNAs between mouse dendritic cells via exosomes. *Blood* **119**, 756–766 (2012).
56. Dosil, S. G. *et al.* Natural killer (NK) cell-derived extracellular-vesicle shuttled microRNAs control T cell responses. *eLife* **11**, e76319 (2022).
57. Skog, J. *et al.* Glioblastoma microvesicles transport RNA and proteins that promote tumour growth and provide diagnostic biomarkers. *Nat. Cell Biol.* **10**, 1470–1476 (2008).
58. Ridder, K. *et al.* Extracellular vesicle-mediated transfer of functional RNA in the tumor microenvironment. *Oncolimmunology* **4**, e1008371 (2015).
59. Gruner, H. N. & McManus, M. T. Examining the evidence for extracellular RNA function in mammals. *Nat. Rev. Genet.* **22**, 448–458 (2021).
60. Li, W. & Ruan, K. MicroRNA detection by microarray. *Anal. Bioanal. Chem.* **394**, 1117–1124 (2009).
61. Wang, C. *et al.* The concordance between RNA-seq and microarray data depends on chemical treatment and transcript abundance. *Nat. Biotechnol.* **32**, 926–932 (2014).
62. Kramer, M. F. Stem-Loop RT-qPCR for miRNAs. *Curr. Protoc. Mol. Biol.* **95**, 15.10.1–15.10.15 (2011).

63. S., A. FastQC A Quality Control tool for High Throughput Sequence Data. *n.d* (2010).
64. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnetJ.* **17**, 10–12 (2011).
65. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
66. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).
67. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
68. Ninomiya, S., Ishikawa, T., Takahashi, M., Seki, M. & Nashimoto, M. Potential physiological roles of the 31/32-nucleotide Y4-RNA fragment in human plasma. *Non-coding RNA Res.* **4**, 135–140 (2019).
69. McCall, M. N. *et al.* Toward the human cellular microRNAome. *Genome Res.* **27**, 1769–1781 (2017).
70. Londin, E. *et al.* Analysis of 13 cell types reveals evidence for the expression of numerous novel primate- and tissue-specific microRNAs. *Proc. Natl. Acad. Sci.* **112**, E1106–E1115 (2015).
71. Roointan, A., Gholaminejad, A., Shojaie, B., Hudkins, K. L. & Gheisari, Y. Candidate MicroRNA Biomarkers in Lupus Nephritis: A Meta-analysis of Profiling Studies in Kidney, Blood and Urine Samples. *Mol. Diagn. Ther.* **27**, 141–158 (2023).
72. Chauhan, S. K., Singh, V. V., Rai, R., Rai, M. & Rai, G. Distinct Autoantibody Profiles in Systemic Lupus Erythematosus Patients are Selectively Associated with TLR7 and TLR9 Upregulation. *J. Clin. Immunol.* **33**, 954–964 (2013).
73. Chauhan, S. K., Singh, V. V., Rai, R., Rai, M. & Rai, G. Differential microRNA Profile and Post-Transcriptional Regulation Exist in Systemic Lupus Erythematosus Patients with Distinct Autoantibody Specificities. *J. Clin. Immunol.* **34**, 491–503 (2014).
74. Rai, R., Chauhan, S. K., Singh, V. V., Rai, M. & Rai, G. RNA-seq Analysis Reveals Unique Transcriptome Signatures in Systemic Lupus Erythematosus Patients with Distinct Autoantibody Specificities. *PLoS ONE* **11**, e0166312 (2016).
75. Witwer, K. W. & Halushka, M. K. Toward the promise of microRNAs – Enhancing reproducibility and rigor in microRNA research. *RNA Biol.* **13**, 1103–1116 (2016).
76. Haider, B. A. *et al.* A Critical Evaluation of microRNA Biomarkers in Non-Neoplastic Disease. *PLoS ONE* **9**, e89565 (2014).
77. Zhang, C. *et al.* Diagnostic and screening potential of plasma exosome miR-99b-5p and its combination with other miRNAs for colorectal cancer. *Oncol. Lett.* **28**, 461 (2024).

78. Li, P., He, Q., Luo, C. & Qian, L. Circulating miR-221 Expression Level and Prognosis of Cutaneous Malignant Melanoma. *Méd. Sci. Monit. : Int. Méd. J. Exp. Clin. Res.* **20**, 2472–2477 (2014).
79. Chen, Q. *et al.* Elevated plasma miR-133b and miR-221-3p as biomarkers for early Parkinson's disease. *Sci. Rep.* **11**, 15268 (2021).
80. Pamuk, O. N. Thrombocytopenia in Patients with Systemic Lupus Erythematosus. *Eur. J. Rheumatol.* **10**, 159–162 (2023).
81. Pritchard, C. C. *et al.* Blood Cell Origin of Circulating MicroRNAs: A Cautionary Note for Cancer Biomarker Studies. *Cancer Prev. Res.* **5**, 492–497 (2012).
82. Mangeney, M., Richard, Y., Coulaud, D., Tursz, T. & Wiels, J. CD77: an antigen of germinal center B cells entering apoptosis. *Eur. J. Immunol.* **21**, 1131–1140 (1991).
83. Kamounah, S., Sembler-Møller, M. L., Nielsen, C. H. & Pedersen, A. M. L. Sjögren's syndrome: novel insights from proteomics and miRNA expression analysis. *Front. Immunol.* **14**, 1183195 (2023).
84. Hu, Y.-Z., Li, Q., Wang, P.-F., Li, X.-P. & Hu, Z.-L. Multiple functions and regulatory network of miR-150 in B lymphocyte-related diseases. *Front. Oncol.* **13**, 1140813 (2023).
85. Kluiver, J. L. & Chen, C.-Z. MicroRNAs regulate B-cell receptor signaling-induced apoptosis. *Genes Immun.* **13**, 239–244 (2012).
86. Xiao, C. *et al.* MiR-150 Controls B Cell Differentiation by Targeting the Transcription Factor c-Myb. *Cell* **131**, 146–159 (2007).
87. Everaert, C. *et al.* Blocking Abundant RNA Transcripts by High-Affinity Oligonucleotides during Transcriptome Library Preparation. *Biol. Proced. Online* **25**, 7 (2023).
88. Dai, Y. *et al.* Microarray analysis of microRNA expression in peripheral blood cells of systemic lupus erythematosus patients. *Lupus* **16**, 939–946 (2007).
89. Consortium, T. exRNAQC *et al.* Performance evaluation of RNA purification kits and blood collection tubes in the Extracellular RNA Quality Control (exRNAQC) study. *bioRxiv* 2021.05.11.442610 (2022) doi:10.1101/2021.05.11.442610.
90. Febbo, P. G. *et al.* Minimum Technical Data Elements for Liquid Biopsy Data Submitted to Public Databases. *Clin. Pharmacol. Ther.* **107**, 730–734 (2020).

## Anexo

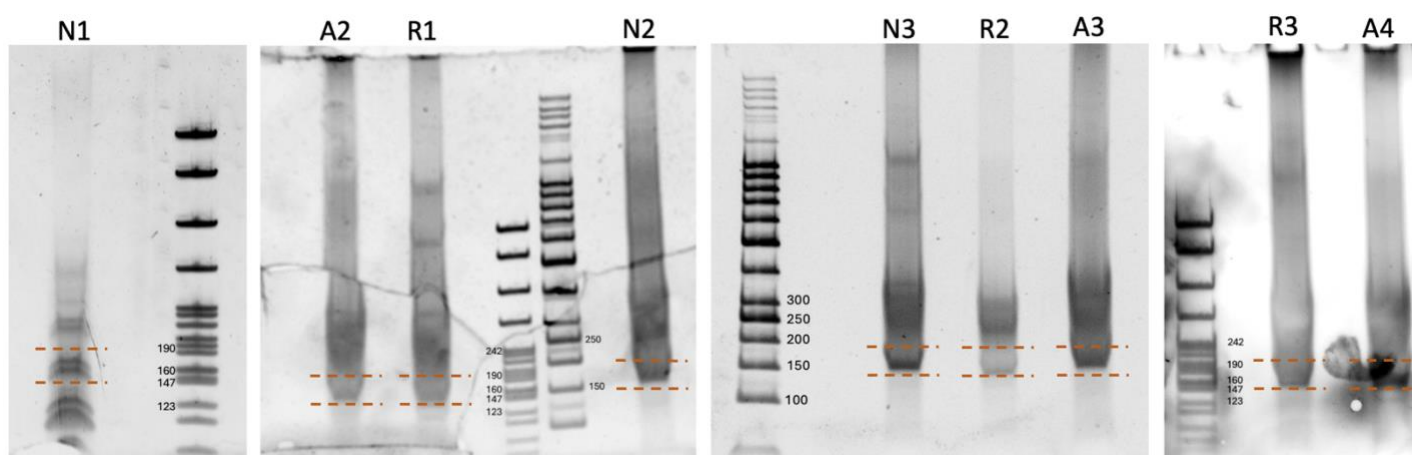


Figura A1. Electroforesis en gel de poliacrilamida al 6%. Se indican con líneas punteadas las bandas incluidas en la biblioteca de cada muestra.

Tabla A1. Concentración nM de las bibliotecas medida por fluorimetría y tamaño promedio obtenido en el analizador de fragmentos.

Muestra	Tamaño promedio (pb)	Concentración (nM)
<b>N1</b>	152	7
<b>N2</b>	166	24
<b>N3</b>	162	23

<b>R1</b>	164	17
<b>R2</b>	155	8
<b>R3</b>	160	25
<b>A2</b>	165	19
<b>A3</b>	158	10
<b>A4</b>	157	20

Tabla A2. Expresión de los transcritos con expresión diferencial, valores en RPM.

	<b>N1</b>	<b>N2</b>	<b>N3</b>	<b>R1</b>	<b>R2</b>	<b>R3</b>	<b>A2</b>	<b>A3</b>	<b>A4</b>
<b>miR-99b-5p</b>	191	21	77	172	73	23	274	554	765
<b>miR-382-5p</b>	28	3	26	15	5	5	31	38	20
<b>miR-409-3p</b>	93	13	132	34	2	31	230	154	87
<b>miR-221-3p</b>	107	45	175	113	114	62	372	268	167
<b>miR-543</b>	54	3	15	11	3	4	85	90	34
<b>miR-4446-3p</b>	35	5	29	33	12	9	82	79	36
<b>let-7e-5p</b>	63	46	84	64	120	45	610	399	212
<b>miR-29a-3p</b>	75	5	38	71	107	16	8	10	12
<b>miR-15b-5p</b>	26	5	16	93	96	16	6	9	6
<b>miR-374a-5p</b>	57	2	5	12	79	5	4	4	3
<b>miR-30b-5p</b>	17	11	7	26	54	10	3	4	3
<b>miR-26b-5p</b>	1482	146	429	1329	2512	229	257	187	201
<b>miR-363-3p</b>	1050	53	135	1423	4470	137	167	53	129
<b>miR-3158-3p</b>	475	7	48	216	1327	33	23	9	27
<b>miR-140-3p</b>	271	10	68	461	822	77	41	25	47
<b>miR-486-5p</b>	24193	1295	4546	9140	67766	1496	2704	615	2641
<b>miR-150-5p</b>	62	37	122	285	82	22	3	5	21
<b>miR-106b-5p</b>	60	7	41	174	263	12	12	4	19
<b>miR-144-3p</b>	316	19	63	591	2239	36	25	2	22



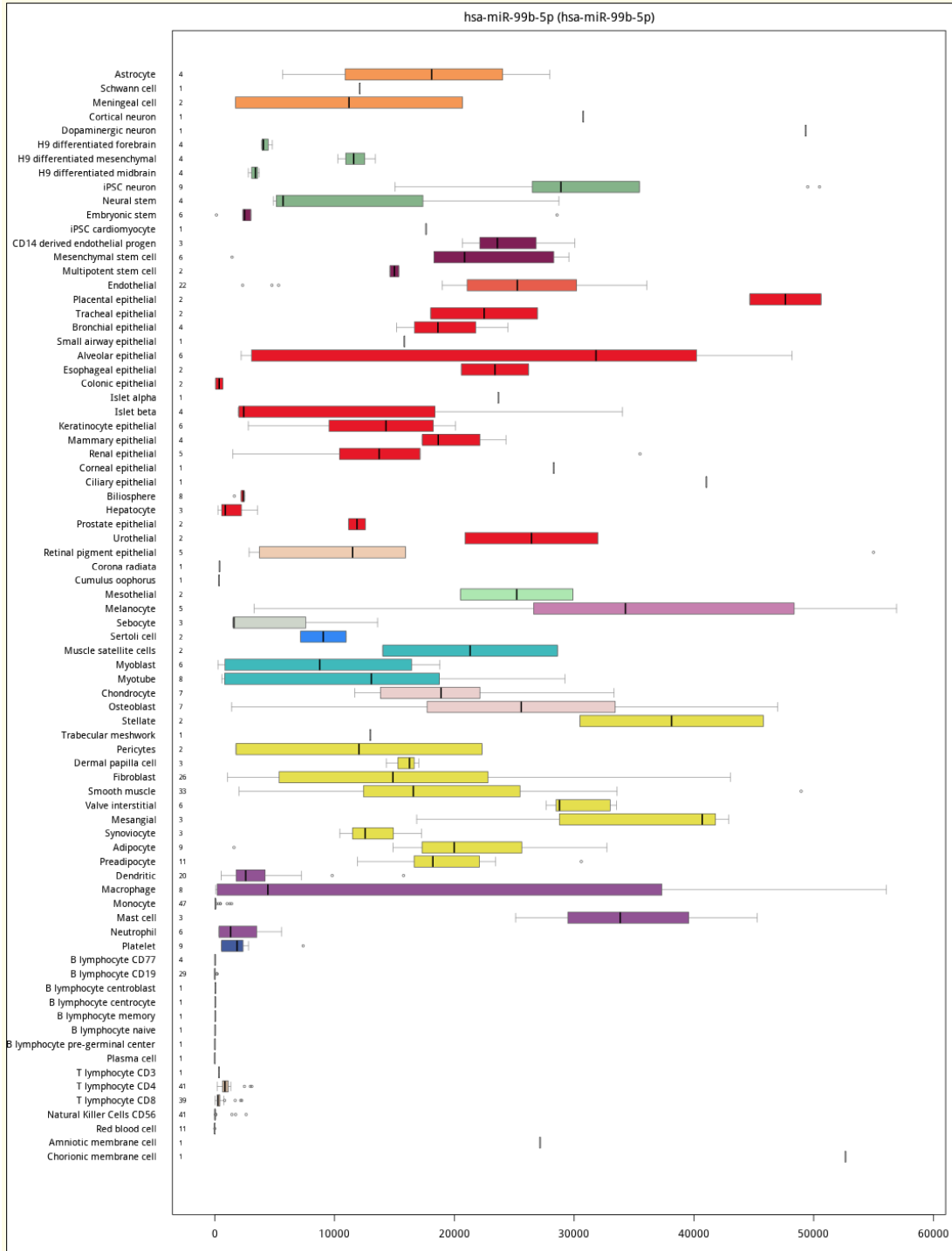


Figura A1. Expresión de miR-99b-5p en 78 tipos de células. Imagen tomada del UCSC Genome Browser.

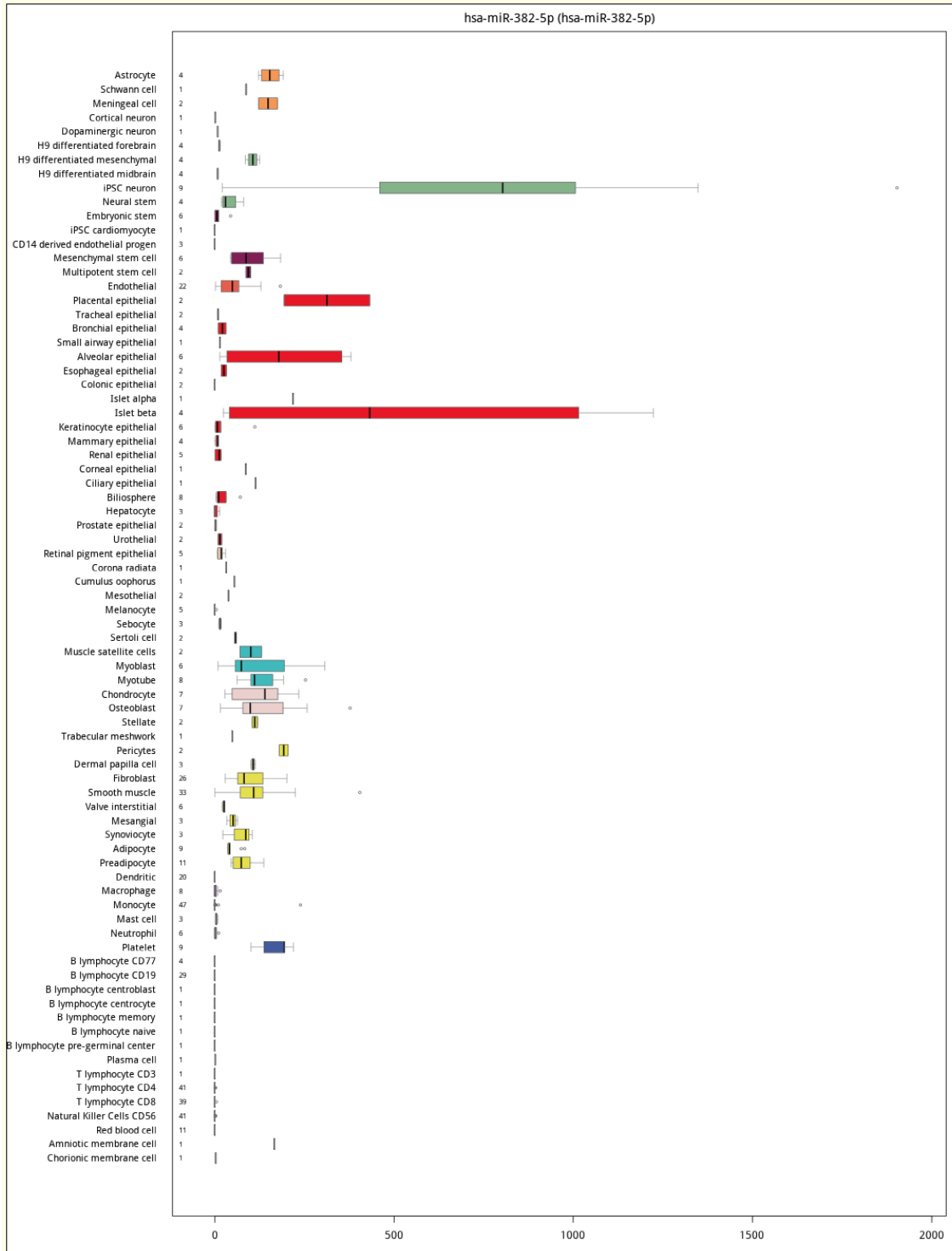


Figura A2. Expresión de miR-382-5p en 78 tipos de células, expresados en valores de RPM. Imagen tomada del UCSC Genome Browser.

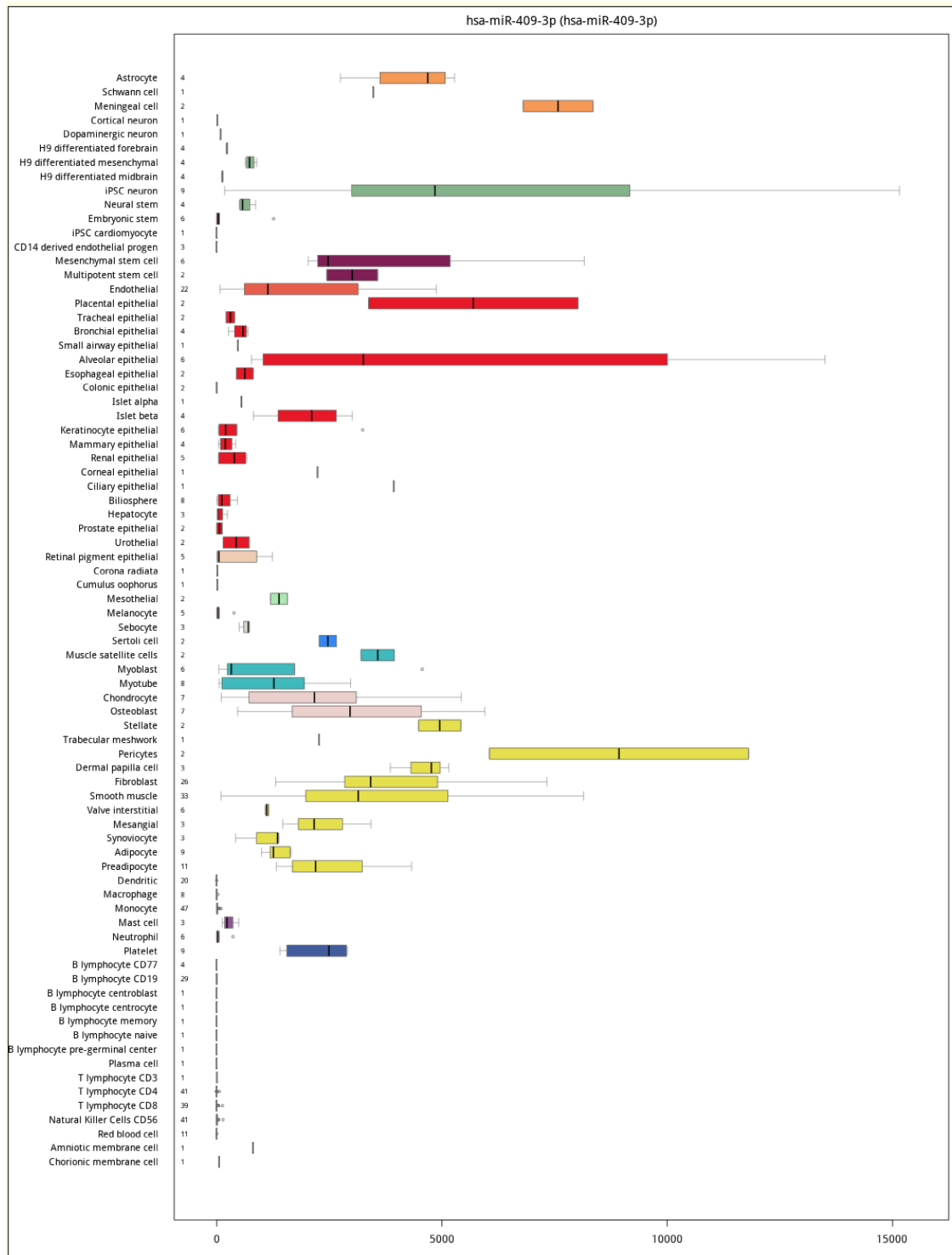


Figura A3. Expresión de miR-409-3p en 78 tipos de células, expresados en valores de RPM. Imagen tomada del UCSC Genome Browser.

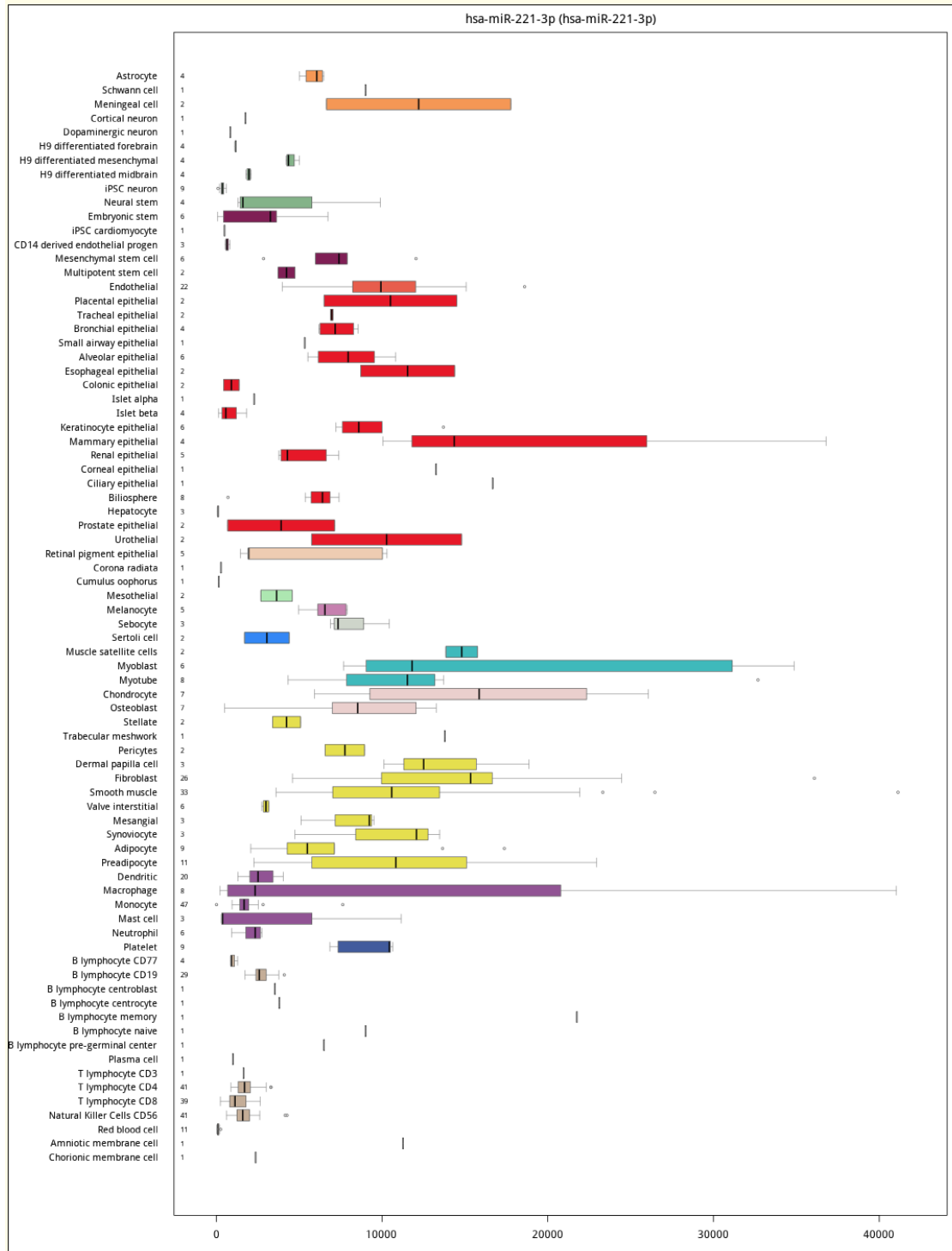


Figura A4. Expresión de miR-221-3p en 78 tipos de células, expresados en valores de RPM. Imagen tomada del UCSC Genome Browser.

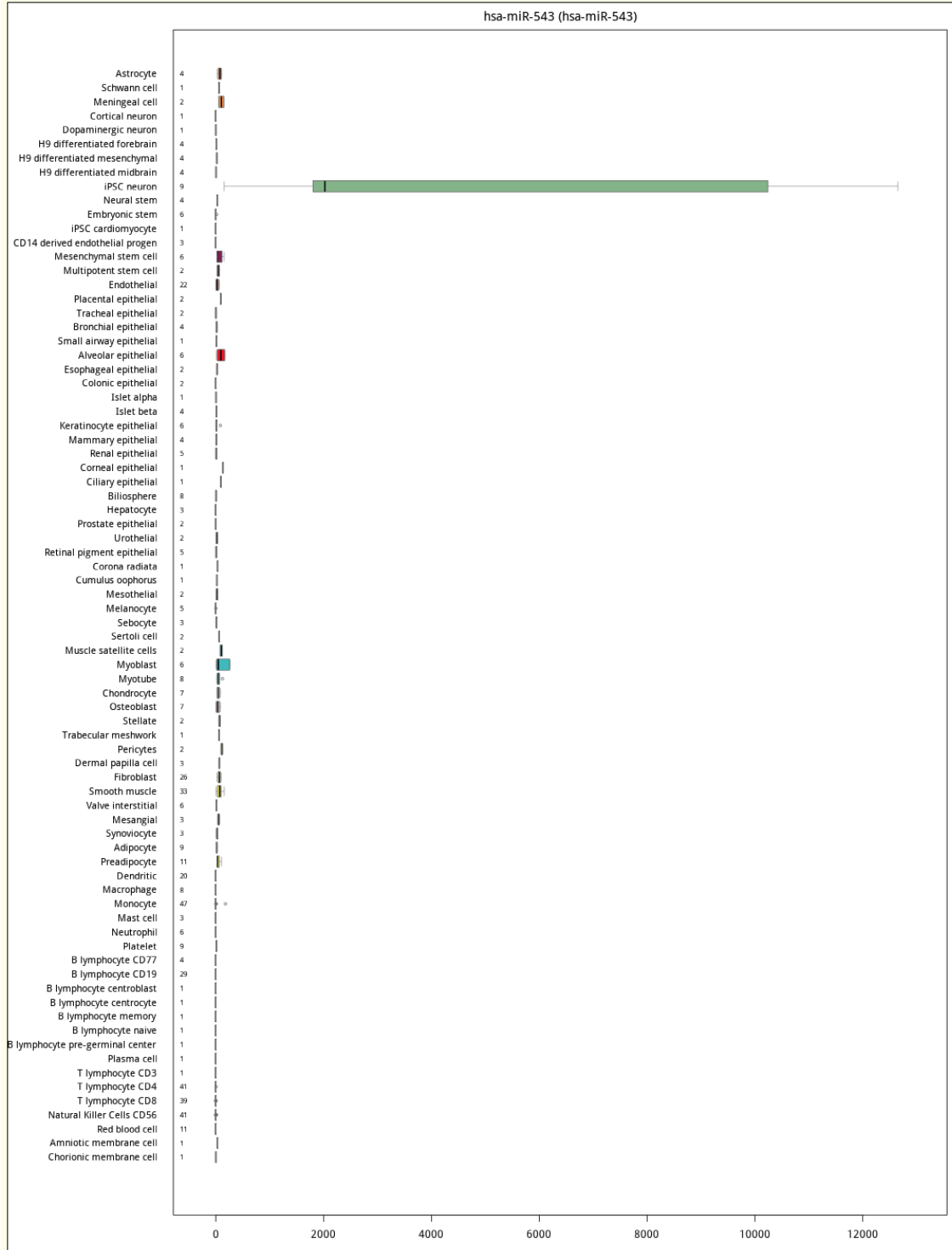


Figura A5. Expresión de miR-543-3p en 78 tipos de células, expresados en valores de RPM. Imagen tomada del UCSC Genome Browser.

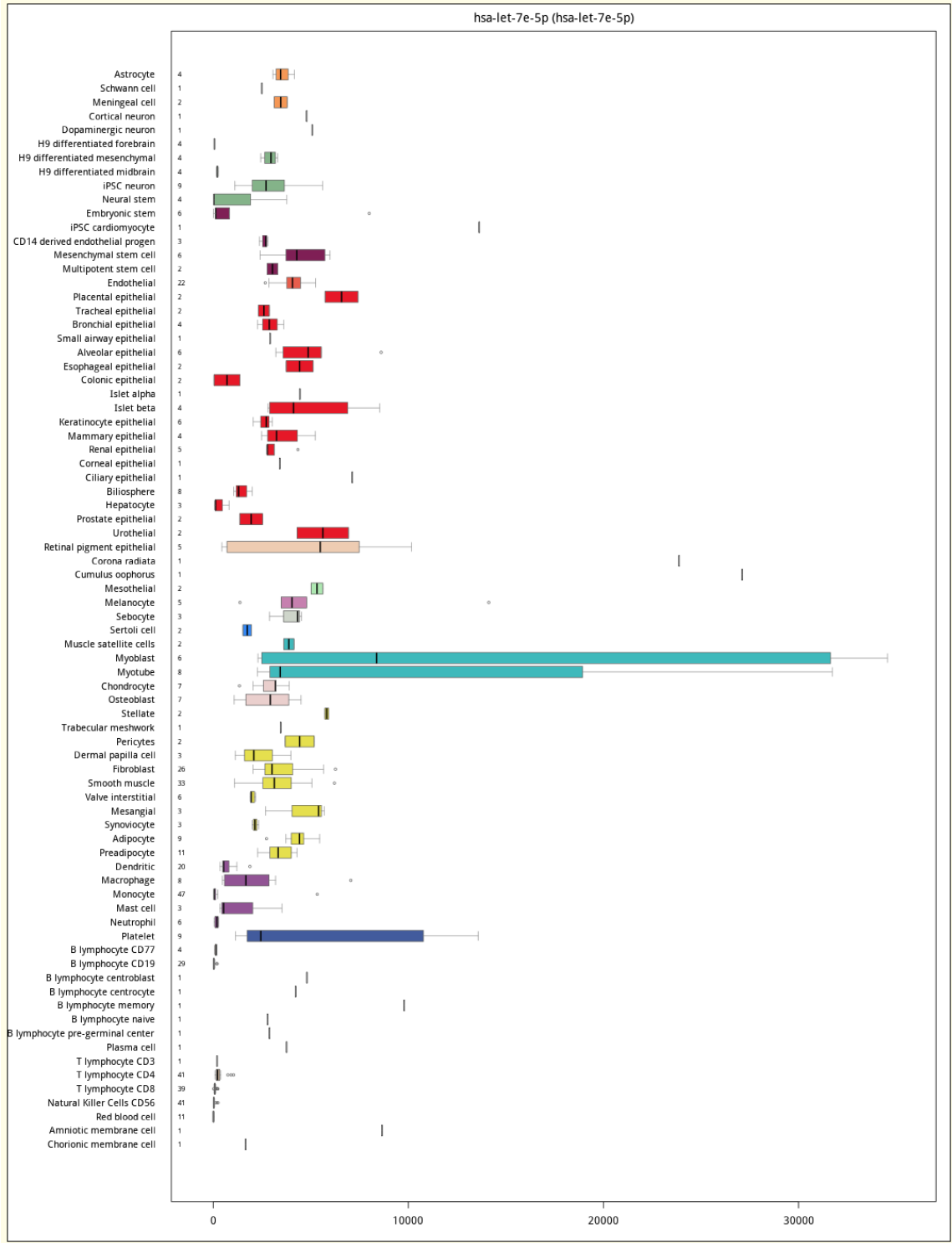


Figura A6. Expresión de let-7e-5p en 78 tipos de células, expresados en valores de RPM. Imagen tomada del UCSC Genome Browser.

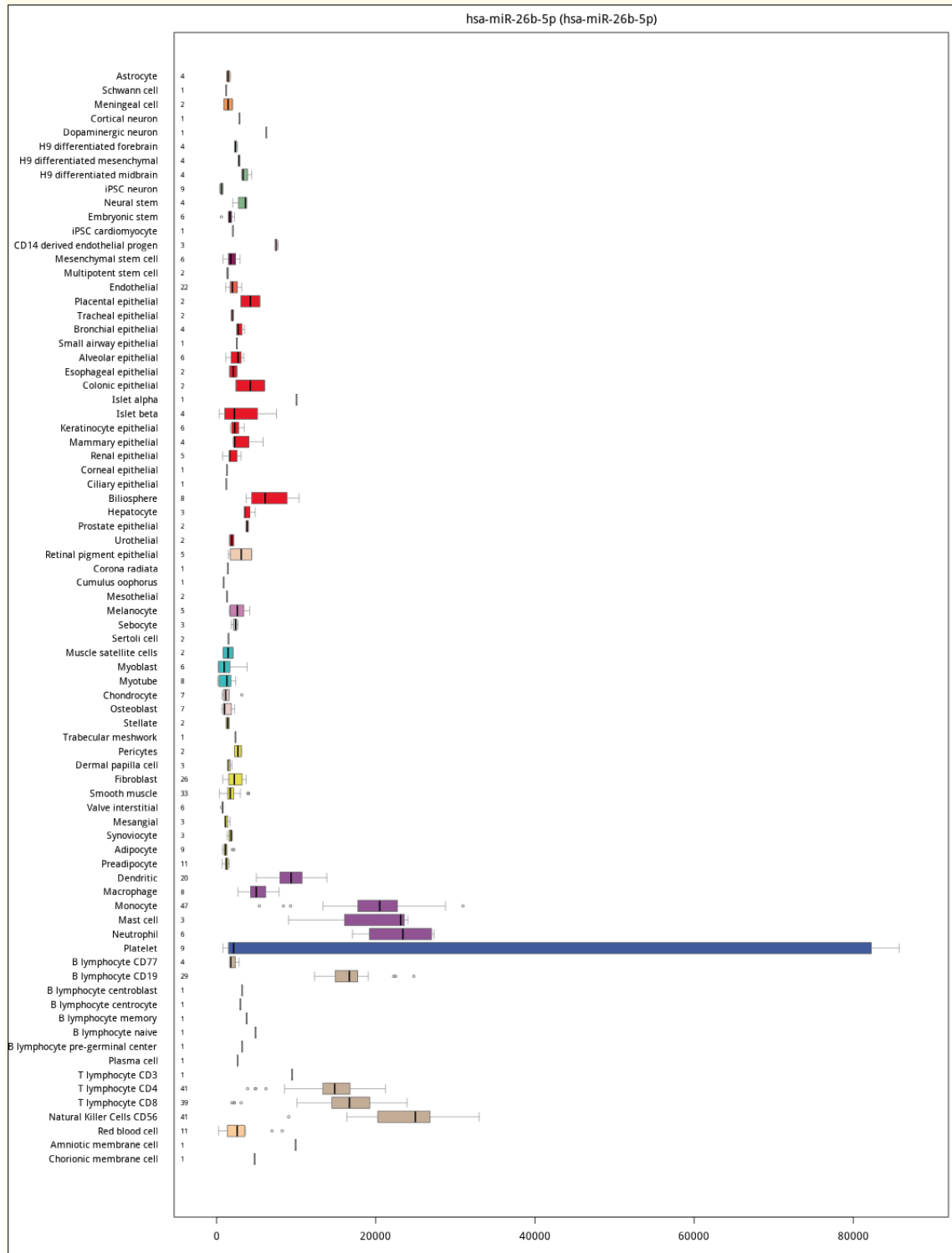


Figura A7. Expresión de miR-26b-5p en 78 tipos de células, expresados en valores de RPM. Imagen tomada del UCSC Genome Browser.

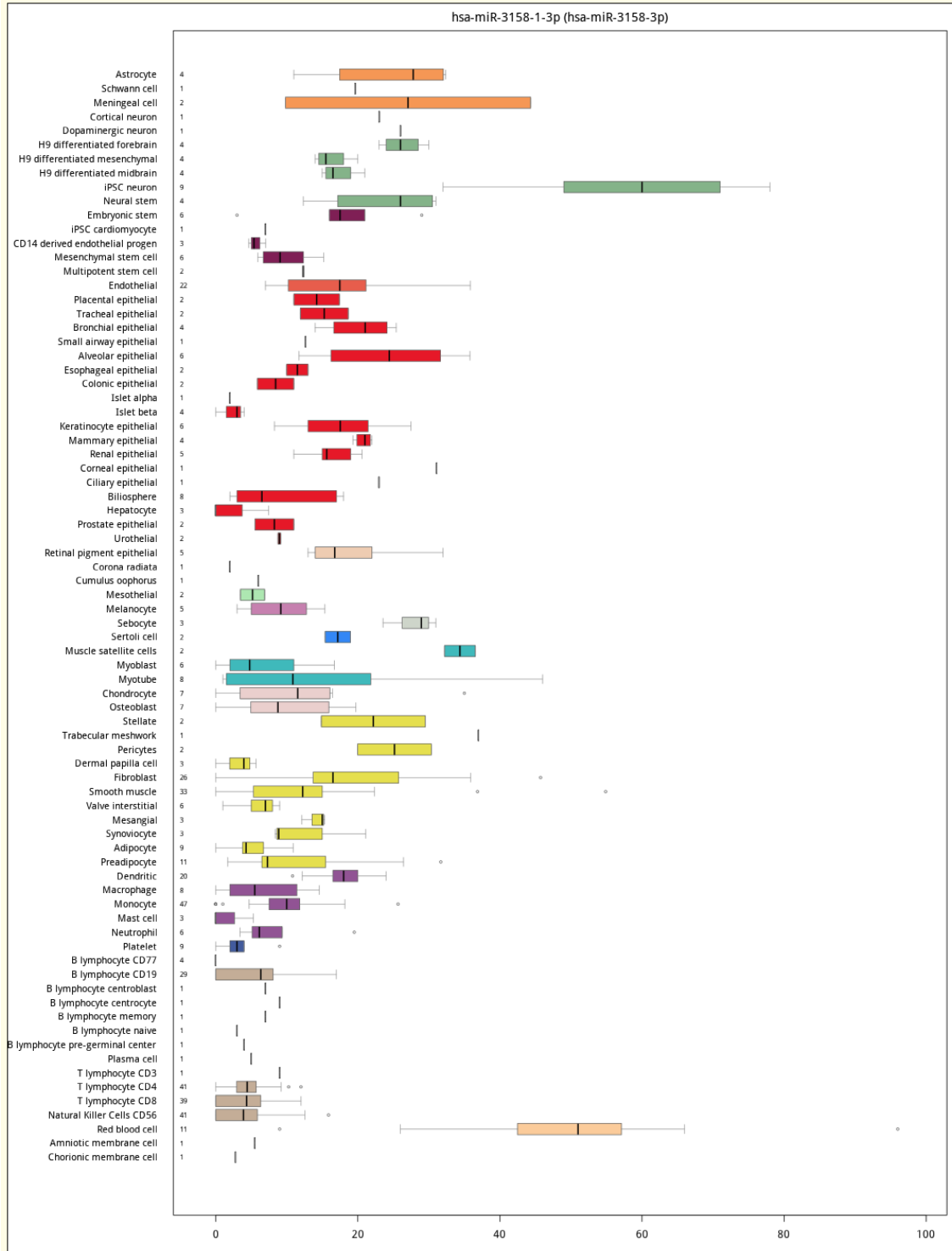


Figura A8. Expresión de miR-3158-3p en 78 tipos de células, expresados en valores de RPM. Imagen tomada del UCSC Genome Browser.

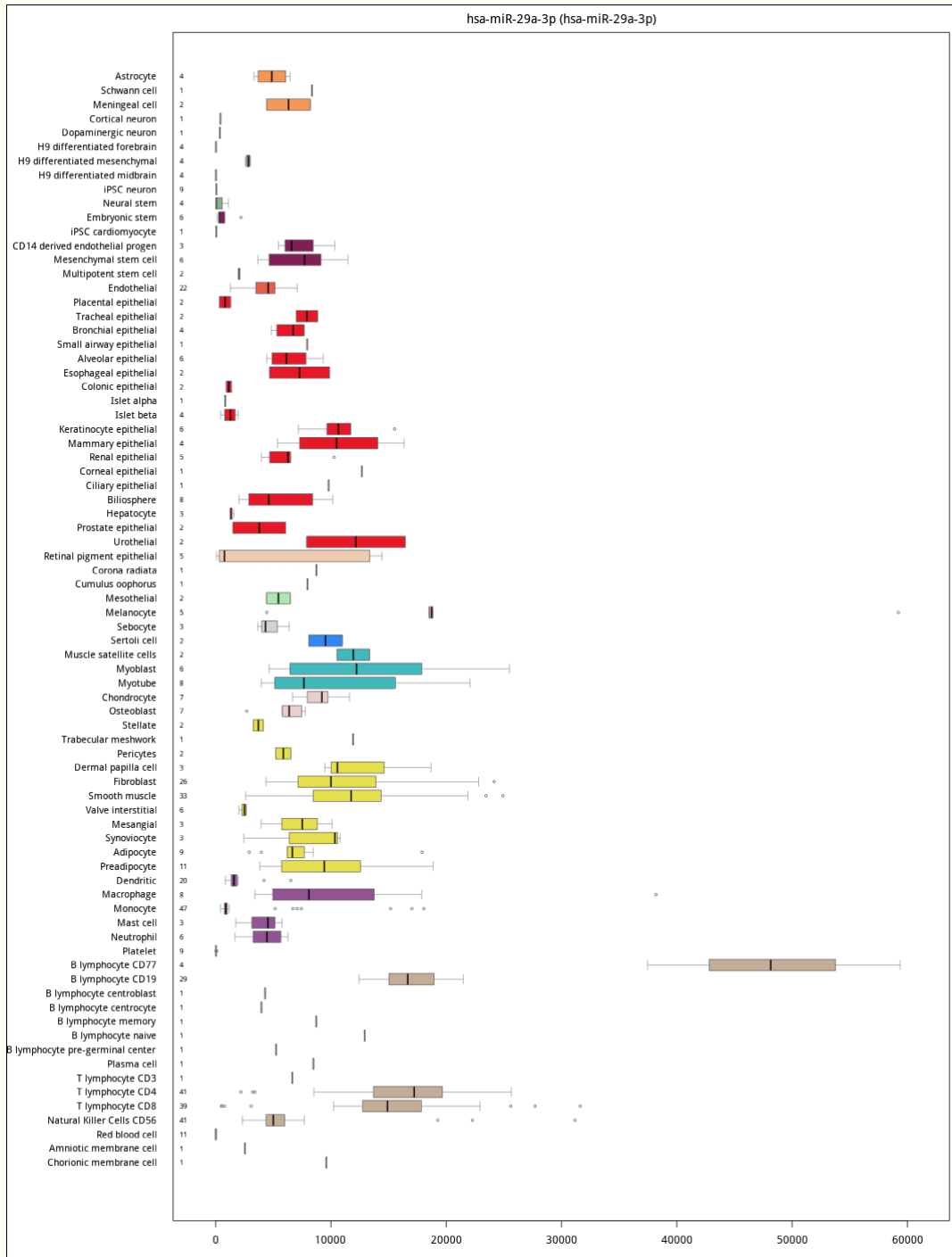


Figura A9. Expresión de miR-29a-3p en 78 tipos de células, expresados en valores de RPM. Imagen tomada del UCSC Genome Browser.

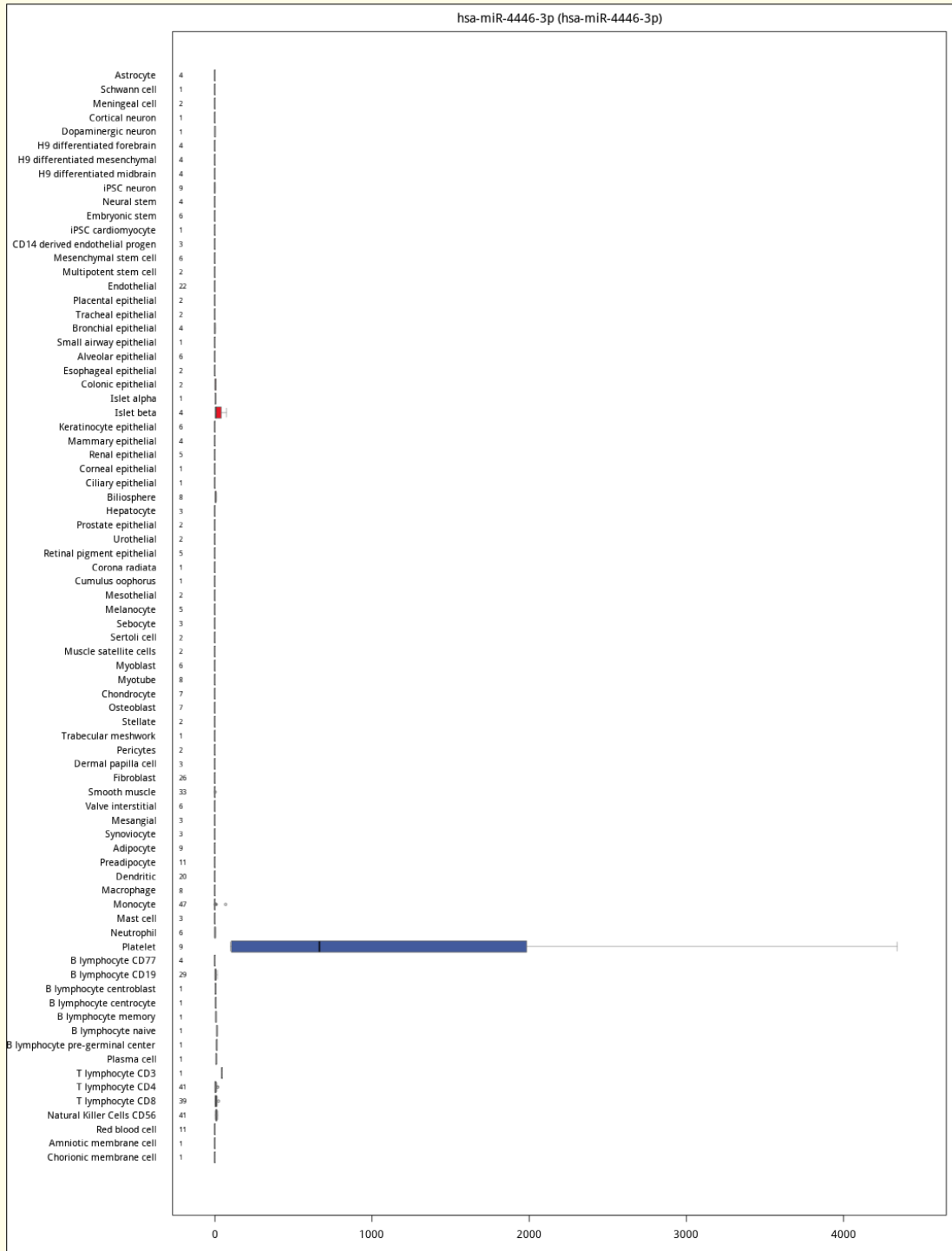


Figura A10. Expresión de miR-4446-3p en 78 tipos de células, expresados en valores de RPM. Imagen tomada del UCSC Genome Browser.

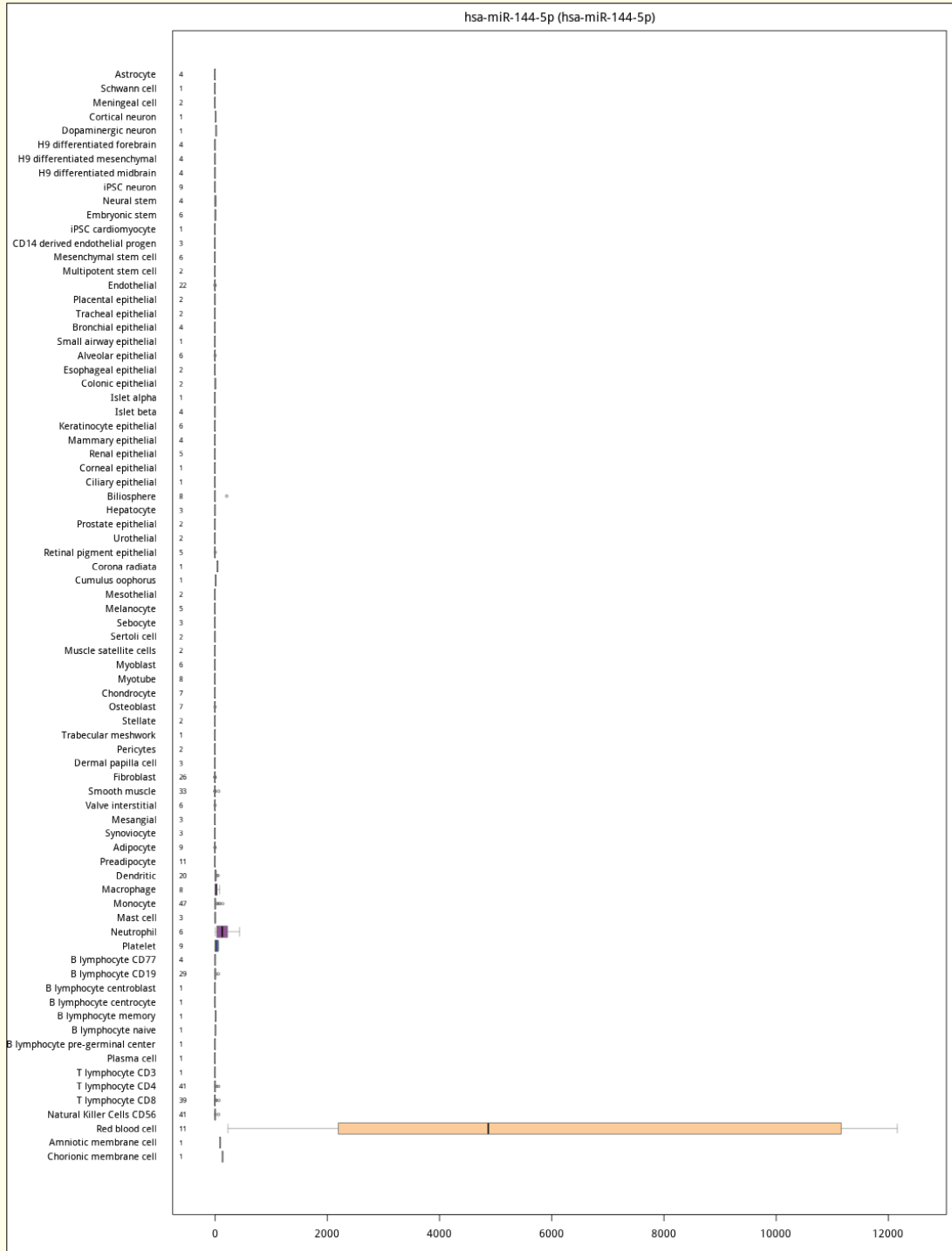


Figura A11. Expresión de miR-144-5p en 78 tipos de células, expresados en valores de RPM. Imagen tomada del UCSC Genome Browser.

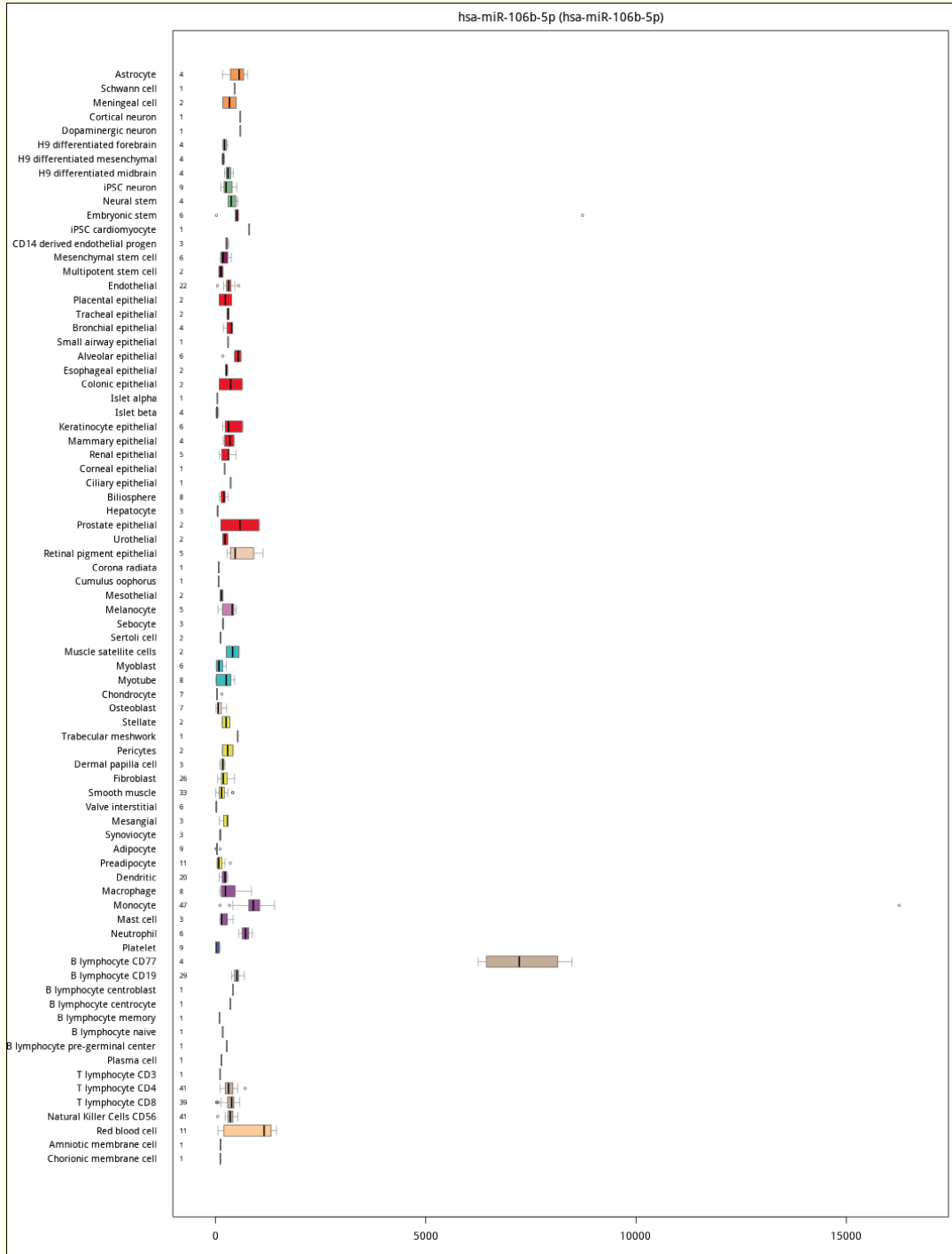


Figura. A12. Expresión de miR-106b-5p en 78 tipos de células, expresados en valores de RPM. Imagen tomada del UCSC Genome Browser.

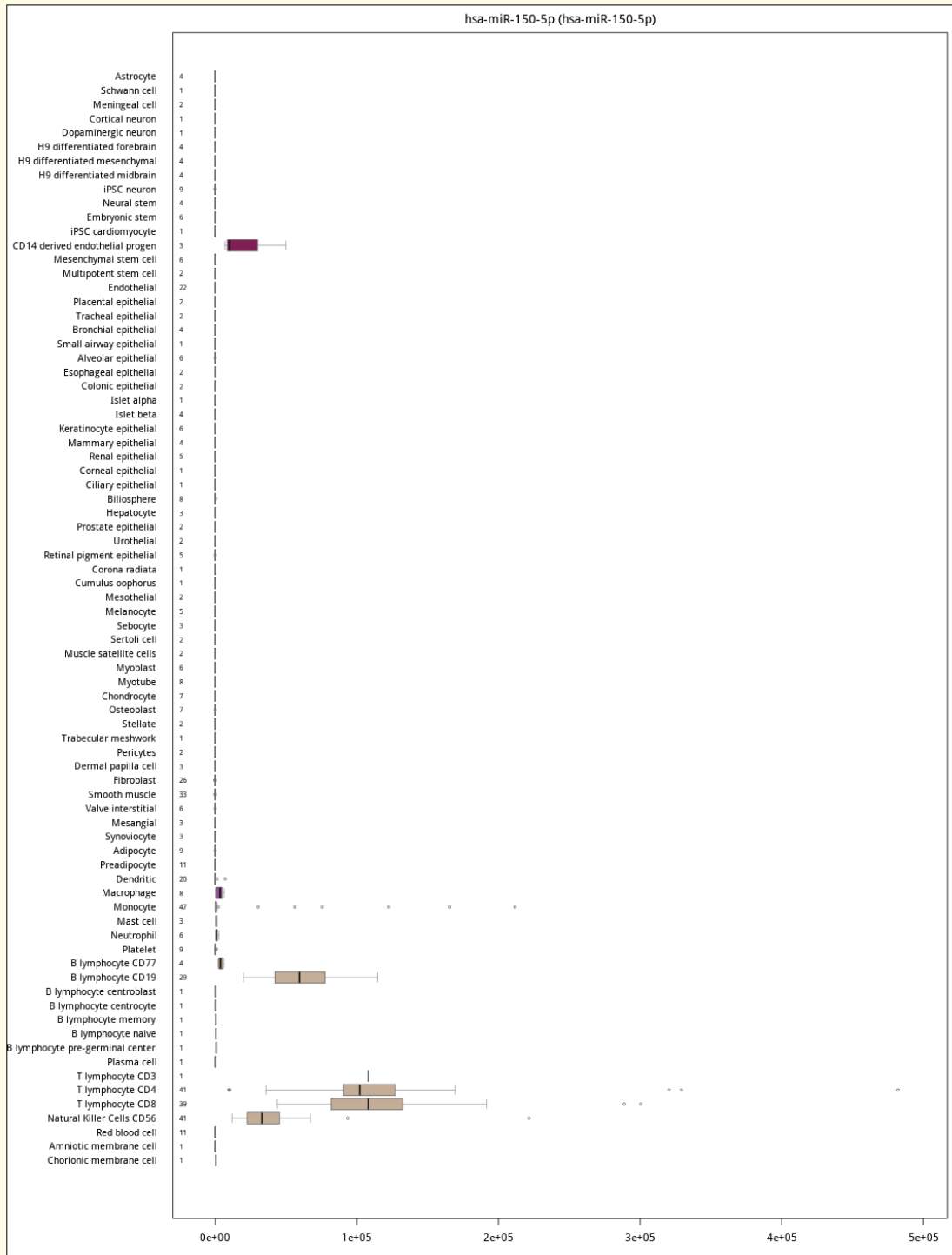


Figura. A13. Expresión de miR-150-5p en 78 tipos de células, expresados en valores de RPM. Imagen tomada del UCSC Genome Browser.

# Agradecimientos

A mis tutores: Dra. Tamara Fernández, Dra. María Rosa García-Silva y Dr. Sergio Bianchi.

A todos los miembros del Laboratorio de Genómica Funcional: Dr. Juan Pablo Tosar, Dra. María Rosa García-Silva, Dr. Sergio Bianchi, Mag. Mauricio Castellano, Mag. Marco Li Calzi, Mag. Valentina Blanco, Lic. Bruno Costa, Lic. Sofía Montenegro, Dr. Pablo Fagúndez y Dr. Alfonso Cayota.

A las médicas encargadas del reclutamiento de las pacientes: Dra. Sandra Consani y Dra. Carolina Díaz.

A los miembros del tribunal: Dra. Natalia Rego, Dra. Leticia Pérez y Dr. Pablo Smircich, por la corrección de esta tesis.

A la Unidad de Biología Molecular del Instituto Pasteur por permitirnos utilizar el secuenciador MiSeq, y al Dr. Gonzalo Grief por su asistencia durante el uso del equipo.

A la ANII y a la CSIC por financiar este trabajo.

Al PEDECIBA y a la Facultad de Ciencias de la Universidad de la República.

Al Institut Pasteur de Montevideo.