



**Universidad  
Europea**

**Máster en Bioinformática**

**Diferenciación molecular y segregación de  
pacientes con nefritis lúpica basada en la  
respuesta al tratamiento con  
micofenolato de mofetil**

**Autor: Laura Palenzuela Bermejo**

**Tutor: Daniel Toro Domínguez**

**Curso 2023-24**

## **AGRADECIMIENTOS**

Quiero expresar mi más sincero agradecimiento, en primer lugar, a mi tutor, Daniel, por su orientación y apoyo a lo largo de este trabajo. Su disposición para ayudar y su dedicación han sido fundamentales para la realización de este proyecto.

En segundo lugar, a mi compañera de máster Elisabeth, por las ideas compartidas, las dudas resueltas y su apoyo, que han sido de gran ayuda y motivación durante este proceso.

Agradezco también a Luis, a mi familia y a mis amigos, quienes siempre han estado a mi lado brindándome su apoyo incondicional y paciencia en cada uno de mis proyectos. Gracias por acompañarme en este camino.

## RESUMEN

**Introducción:** La nefritis lúpica (NL) es una grave complicación del lupus eritematoso sistémico (LES), donde la acumulación de complejos inmunes en el riñón causa inflamación y pérdida progresiva de la función renal, aumentando la morbilidad y mortalidad de los pacientes. La considerable heterogeneidad clínica y molecular de la NL, junto con la variabilidad en síntomas, pronóstico y respuesta a las terapias, hace crucial una selección temprana y precisa del tratamiento para preservar la función renal.

**Objetivo:** Establecer perfiles moleculares asociados a la respuesta al tratamiento mediante análisis diferencial de expresión génica y enriquecimiento funcional en pacientes con NL tratados con micofenolato de mofetil (MMF), identificando genes diferencialmente expresados (DEGs) y procesos biológicos enriquecidos que distingan a respondedores de no respondedores.

**Material y métodos:** Se realizó un análisis longitudinal de datos clínicos, demográficos y de expresión génica de microarrays, dividido en tres periodos: temprano, al año y de seguimiento desde la biopsia. Se utilizó RStudio para el análisis general y la librería *limma* para la expresión diferencial incluyendo covariables. Los DEGs identificados fueron sometidos a análisis de enriquecimiento funcional con GeneCodis. La respuesta clínica se evaluó mediante uPCR y mPERR.

**Resultados:** El análisis mostró que el periodo de un año desde la biopsia es el que refleja mayores diferencias moleculares entre respondedores y no respondedores. Se identificaron DEGs relevantes en la fisiopatología de NL: CLEC4C, IFIT1 y CD180, junto con el enriquecimiento de distintas rutas: TLRs, vía del interferón tipo I o la respuesta inmune adaptativa. La mayor actividad inmunitaria en no respondedores sugiere una relación con la respuesta limitada al tratamiento.

**Conclusiones:** Los DEGs obtenidos en el análisis y las vías enriquecidas emergen como posibles marcadores y dianas terapéuticas en pacientes resistentes al tratamiento. La validación de estos biomarcadores en cohortes adicionales es clave para evaluar su potencial como posibles objetivos diagnósticos y terapéuticos.

**Palabras clave:** Lupus eritematoso sistémico, nefritis lúpica, micofenolato de mofetil, análisis de expresión diferencial, uPCR, mPERR.

## ABSTRACT

**Introduction:** Lupus nephritis (LN) is a serious complication of systemic lupus erythematosus (SLE), where accumulation of immune complexes in the kidneys causes inflammation and progressive loss of kidney function, increasing morbidity and mortality of the patients. The considerable clinical and molecular heterogeneity of LN, together with variability in symptoms, prognosis and response to therapies, makes early and accurate treatment selection crucial to preserve renal function.

**Objective:** To identify molecular profiles associated with treatment response by differential analysis of gene expression and functional enrichment in patients with NL treated with mycophenolate mofetil (MMF), identifying differential expressed genes (DEGs) and enriched biological processes that may distinguish responders from non-responders.

**Material y methods:** A longitudinal analysis of clinical, demographic and microarray gene expression data was performed, divided into three periods: early, one-year and follow-up since biopsy. RStudio was used for the overall analysis and the *limma* library for differential expression including covariates. The identified DEGs were submitted to functional enrichment analysis with GeneCodis. Clinical response was assessed by uPCR and mPERR.

**Results:** The analysis showed that the one-year period since biopsy reflects the greatest molecular differences between responders and non-responders. Relevant DEGs in the physiopathology of NL were identified: CLEC4C, IFIT1 and CD180, together with enrichment of different pathways: TLRs, type I interferon pathway or adaptive immune response. The higher immune activity in non-responders suggests a relationship with limited response to treatment.

**Conclusions:** The DEGs obtained in the analysis and the enriched pathways emerge as potential markers and therapeutic targets in treatment-resistant patients. Validation of these biomarkers in additional cohorts constitutes a fundamental future line to evaluate their potential as possible diagnostic and therapeutic targets.

**Keywords:** Systemic Lupus Erythematosus, Lupus nephritis, Mycophenolate Mofetil, differential expression analysis, uPCR, mPERR.

## Índice

<b>1.</b>	<b>Introducción .....</b>	<b>6</b>
1.1.	Lupus eritematoso sistémico .....	6
1.2.	Nefritis lúpica .....	6
1.2.1.	Diagnóstico .....	7
1.2.2.	Seguimiento .....	8
1.2.3.	Tratamiento .....	9
1.3.	Aplicación de la bioinformática y las ómicas en el contexto de las enfermedades inmunes 11	
1.4.	Aplicación de la bioinformática en la caracterización molecular de pacientes, diagnóstico y tratamiento en el Lupus.....	12
<b>2.</b>	<b>Hipótesis y Objetivos .....</b>	<b>13</b>
2.1.	Hipótesis.....	13
2.2.	Objetivos .....	13
<b>3.</b>	<b>Metodología .....</b>	<b>14</b>
3.1.	Origen y características de los datos empleados.....	14
3.2.	Software empleado en el trabajo .....	16
3.3.	Procesado y análisis de los de los datos.....	18
3.3.1.	Exploración de los datos clínicos y de expresión .....	18
3.3.2.	Análisis de variables clínicas .....	19
3.3.3.	Procesamiento de los datos de expresión .....	20
3.3.4.	Enriquecimiento funcional e interpretación de resultados .....	22
3.4.	Declaración obligatoria de herramientas de IA .....	22
<b>4.</b>	<b>Resultados .....</b>	<b>23</b>
4.1.	Análisis de variables clínicas .....	23
4.1.1.	Análisis de variables por paciente .....	23
4.1.2.	Análisis de variables por muestra .....	25
4.2.	Expresión diferencial .....	26
4.2.1.	Periodo temprano: hasta 6 meses desde la biopsia. ....	26
4.2.2.	Periodo Año: entre 6 y 19 meses desde la biopsia. ....	29
4.2.3.	Periodo de seguimiento: a partir de 19 meses desde la biopsia.....	31
4.2.4.	Análisis para todas las muestras. ....	33
4.2.5.	Análisis de Componentes Principales de muestras (PCA) .....	36
4.2.6.	Caracterización de Perfiles moleculares .....	39
4.3.	Enriquecimiento funcional.....	41
<b>5.</b>	<b>Discusión.....</b>	<b>49</b>
<b>6.</b>	<b>Conclusiones.....</b>	<b>53</b>
<b>7.</b>	<b>Anexos .....</b>	<b>54</b>
<b>8.</b>	<b>Bibliografía .....</b>	<b>54</b>

## 1. Introducción

### 1.1. *Lupus eritematoso sistémico*

El Lupus eritematoso sistémico (LES) es una enfermedad autoinmune compleja y heterogénea tanto a nivel clínico como molecular, con una etiología aún desconocida. Esta heterogeneidad implica que las manifestaciones del LES pueden variar ampliamente entre pacientes, tanto en los síntomas clínicos como en los procesos inmunológicos y moleculares subyacentes que causan la enfermedad, lo que dificulta tanto el diagnóstico como el tratamiento.(1)

La enfermedad se caracteriza por la producción de autoanticuerpos debida a una disfunción en el sistema inmune. Estos anticuerpos atacan las células del propio cuerpo, formando complejos inmunes (CI) que se depositan en diversos tejidos, lo que provoca inflamación y daño orgánico. Los órganos más comúnmente afectados son los riñones, el sistema nervioso, la piel y articulaciones(1) ,(2).

El LES se manifiesta como una enfermedad intermitente; con alternancia de fases de brotes o recaídas y remisiones. Su incidencia y prevalencia ha aumentado en los últimos años lo que se atribuye fundamentalmente a los avances en el diagnóstico y mayor disponibilidad de datos a nivel global. La incidencia se ve muy influenciada por factores genéticos, socioeconómicos y ambientales siendo la mayor en América del Norte. En concreto la población afroamericana presenta mayor incidencia y mortalidad y, además, responde peor de manera general a los tratamientos con glucocorticoides e inmunosupresores. En cuanto al sexo, las mujeres en edad reproductiva tienen un mayor riesgo por la influencia de factores hormonales en el desarrollo de la enfermedad. (1)

### 1.2. *Nefritis lúpica*

El daño renal ocupa una de las principales complicaciones dentro del LES, pudiendo afectar los riñones en diferentes grados, incluso causar la disfuncionalidad completa de los mismos. Esta sintomatología se conoce como nefritis lúpica (NL) y ocurre en alrededor de un 40% de los pacientes en su mayoría a los 5 años del diagnóstico.(3)(4)

En la nefritis lúpica, los CI, es decir, la combinación de autoanticuerpos y antígenos propios (componentes celulares del propio organismo, como el ADN o proteínas

nucleares), se depositan en los vasos sanguíneos del riñón, especialmente en los glomérulos. Este depósito activa el sistema del complemento, desencadenando una respuesta inflamatoria que atrae células inmunitarias a la zona, agravando el daño renal. La presencia de autoanticuerpos intensifica esta autoinmunidad, lo que perpetúa la inflamación y provoca una respuesta inmune dirigida específicamente contra los componentes del riñón. Este proceso contribuye significativamente a la alta morbilidad y mortalidad de la nefritis lúpica debido a la lesión renal continua y progresiva. (4) Además, el daño renal es a menudo irreversible, formándose lo que se conoce como cicatrices renales.

Las manifestaciones clínicas son variadas: desde proteinuria asintomática hasta el desarrollo de síndromes nefróticos o nefríticos que progresan a insuficiencia renal en etapa terminal (ESRD). En muchos casos, puede ocurrir una NL “silenciosa”, donde no hay síntomas evidentes de afectación renal, pero la biopsia revela cambios histológicos compatibles con NL. En estos casos la nefritis no da la cara hasta que el daño ya se ha producido por lo que es importante un diagnóstico temprano para establecer un tratamiento lo antes posible (3).

#### 1.2.1. Diagnóstico

En pacientes con SLE, es fundamental realizar pruebas periódicas para evaluar la función renal, independientemente de si presentan síntomas clínicos de nefritis. Entre las pruebas se incluyen:(5)

- Determinación de creatinina en suero: La creatinina es un producto de desecho de los músculos. En un organismo sano, los riñones filtran la creatinina de la sangre y la eliminan a través de la orina. La alteración de estos niveles puede indicar que el riñón no está realizando correctamente su función.
- Tasa de Filtración Glomerular Estimada (eGFR): Se utiliza para evaluar la capacidad de filtración renal.
- Analítica de orina que incluye la determinación de los siguientes parámetros:
  - Relación albúmina/creatinina urinaria (uACR).

- Relación proteína/creatinina urinaria (uPCR): mide la gravedad de la proteinuria. La proteinuria se define como la presencia anormal de proteínas en la orina. En condiciones normales el riñón filtra la sangre eliminando productos de desecho, pero retiene las proteínas u otras moléculas relevantes. Sin embargo, cuando hay un daño renal esta capacidad de filtración se reduce y las proteínas pasan a orina en cantidades significativas.
- Sedimento urinario: Se evalúan anomalías como glóbulos rojos dismórficos, leucocitos o cilindros celulares.

Si se detectan valores de proteinuria  $> 0.5$  g/24 horas o uPCR  $> 0.5$  g/g, presencia de hematuria, leucocituria o restos de túbulos renales o bien un deterioro en la función renal, se realiza una biopsia renal para confirmar el diagnóstico de NL.

La NL se clasifica histológicamente en 6 clases definidas en 2003 por *la International Society of Nephrology (ISN)* y *la Renal Pathology Society (RPS)* (6):

- Clase I: No se observan cambios detectables en los glomérulos.
- Clase II: Enfermedad mesangial pura.
- Clase III: Enfermedad proliferativa que afecta a menos del 50% de los glomérulos.
- Clase IV: Enfermedad proliferativa que afecta a más del 50% de los glomérulos.
- Clase V: Cambios membranosos.
- Clase VI: Glomerulonefritis esclerosante avanzada.

### 1.2.2. Seguimiento

Para monitorizar a los pacientes con NL se deben evaluar parámetros renales y extrarrenales. Entre los renales ya hemos hablado sobre eGFR, la proteinuria que se evalúa mediante uPCR o el sedimento urinario. Además de estos hay otro indicador de respuesta que es mPERR (*Modified Primary Efficacy Renal Response*) que se centra en evaluar la función renal en términos de proteinuria y eGFR tras la biopsia. Si la respuesta renal es favorable, definida como uPCR  $\leq 0.7$  g/g y eGFR  $\geq 60$  mL/min/1.73 m<sup>2</sup>, se

considera que el paciente tiene una mejor supervivencia renal a largo plazo y un menor riesgo de insuficiencia renal crónica.(7)

Para el análisis de parámetros tanto renales como extrarrenales, se utiliza un índice que se denomina SLEDAI (*Systemic Lupus Erythematosus Disease Activity Index*). Se trata de una herramienta cuantitativa validada que se utiliza para medir la actividad general de la enfermedad en pacientes con LES. Abarca distintos órganos y sistemas: renal, cutáneo, neurológico, vascular, hematológico, inmunológico y articular.

Mediante una escala de valores que varían desde el 1-8 se asignan puntos a las distintas manifestaciones de la enfermedad, como por ejemplo la proteinuria. Se puede obtener una puntuación global de hasta 105. Cuanto más alto, mayor actividad de la enfermedad. (8)

Otra escala que utilizan los clínicos para monitorizar la actividad de la enfermedad es el PGA (Physician Global Assessment) que se trata de una escala visual que asigna una puntuación de 0 a 3 en función de la gravedad de la enfermedad siendo 3 la mayor.(9)

### 1.2.3. Tratamiento

El tratamiento de la NL está centrado en el control de la inflamación, preservar la función renal y prevenir el daño a largo plazo. Se sigue un enfoque secuencial que varía según la clase de NL y la severidad de la enfermedad.

La guía KDIGO 2024 para LN indica las últimas recomendaciones de fármacos en función de la fase de NL: (10)

El tratamiento de la NL, según las guías de KDIGO 2024, se divide en dos fases principales: terapia de inducción que tiene el objetivo de remitir la inflamación renal activa y la terapia de mantenimiento para prevenir recaídas y mantener la función renal.

#### **Terapia de inducción: Glucocorticoides + inmunosupresor**

En la fase de inducción para los pacientes con clases III/IV activas se utiliza una combinación de glucocorticoide con un inmunosupresor. Las opciones son las siguientes:

- Glucocorticoide: Se administra en primer lugar metilprednisolona intravenosa durante 1-3 días, seguido de prednisona oral que se reduce progresivamente en 3-6 meses.
- Inmunosupresores:
  - Micofenolato de Mofetil (MMF) o análogos: se usa ampliamente como terapia de primera línea durante al menos 6 meses. El MMF es un profármaco del ácido micofenólico (MPA), que inhibe la enzima inosina monofosfato deshidrogenasa (IMPDH). Esta inhibición bloquea la síntesis de purinas, crucial para la proliferación de linfocitos T y B, lo que reduce la respuesta inmunitaria y la producción de autoanticuerpos, como los anti-dsDNA. Además, reduce la producción de citoquinas proinflamatorias, como la IL-2, lo que ayuda a disminuir la inflamación renal. y puede prevenir la fibrosis renal. (11)
  - Ciclofosfamida (CYC): otra opción de primera línea, especialmente en pacientes con enfermedad severa o refractaria. Durante 3 a 6 meses.
  - Azatioprina (AZA): se usa como sustituto del MMF en algunos casos como durante el embarazo o bien cuando el paciente presenta intolerancia al MMF.
  - Belimumab: se trata de un anticuerpo monoclonal que actúa disminuyendo los linfocitos B. Se usa en combinación con la terapia estándar: MMF o CYC.
  - Inhibidores de la Calcineurina (ICN): Tacrolimus o ciclosporina se utilizan en combinación con MMF, especialmente en pacientes que no responden a los tratamientos convencionales o con clases V de NL.

### **Terapia de mantenimiento: glucocorticoide + inmunosupresor en dosis más bajas**

El MMF sigue siendo el tratamiento de elección o bien AZA en combinación de prednisona.

### **Otros fármacos utilizados:**

- Fármacos que actúan sobre el Sistema Renina-Angiotensina: Se recomienda el uso de IECA como enalapril o ramipril, o ARA-II como losartán o valsartán para reducir la proteinuria y controlar la presión arterial, sobre todo se usa en combinación con los otros tratamientos para pacientes clase V.

- Hidroxicloroquina: Se recomienda su uso en pacientes con lupus proliferativo en especial en clase V, debido a sus efectos inmunomoduladores y protectores renales.

La eficacia del tratamiento estándar de la NL con inmunosupresores varía enormemente entre los pacientes, poniendo de manifiesto la alta heterogeneidad molecular de la enfermedad. Entre el 60 y el 70% de los pacientes no logran una remisión a largo plazo ni una respuesta renal completa al año de iniciado el tratamiento. Además, la exposición prolongada a este tipo de terapias estándar, como el MMF o la CYC, conlleva importantes efectos secundarios debido a la toxicidad inducida por estos fármacos (12).

### *1.3. Aplicación de la bioinformática y las ómicas en el contexto de las enfermedades inmunes*

En los últimos años, la bioinformática y las ómicas: genómica, transcriptómica, proteómica y metabolómica, han supuesto un avance en el estudio de las enfermedades. Estas tecnologías permiten analizar grandes volúmenes de datos para identificar alteraciones genéticas, perfiles de expresión, proteínas específicas y/o metabolitos relevantes. Esta información es esencial para entender la fisiopatología y la progresión de la enfermedad, así como la respuesta al tratamiento, permitiendo el enfoque hacia la medicina personalizada.(13)

La integración de datos multi-ómicos posibilita la identificación de biomarcadores específicos que pueden mejorar la precisión en el diagnóstico y pronóstico de enfermedades autoinmunes. En enfermedades autoinmunes como el LES o la artritis reumatoide, el uso de RNAseq y análisis proteómico ha revelado vías moleculares y redes de señalización alteradas, lo que ha permitido una comprensión más profunda de los mecanismos inmunológicos y ha facilitado el desarrollo de terapias dirigidas. Lo que pone de manifiesto la importancia de estas tecnologías en el descubrimiento de biomarcadores que pueden ser esenciales para la estratificación de pacientes y la elección de tratamientos más específicos y por tanto que generen mayor tolerabilidad y menor toxicidad.(14)

Estudios recientes en artritis reumatoide han puesto de manifiesto el impacto de estas tecnologías ómicas al identificar biomarcadores y vías específicas que caracterizan las

diferentes fases y estados de actividad de la enfermedad. Por ejemplo, el uso de *Genome-Wide Association Studies* (GWAS) en artritis reumatoide ha revelado loci genéticos asociados con la progresión y severidad de la enfermedad, proporcionando la base para el desarrollo de tratamientos dirigidos según el perfil genético individual.(15)(16)

En artritis reumatoide, además de GWAS, los microarrays de proteínas han resultado ser rápidos, económicos, de alto rendimiento y alta sensibilidad, lo que les posiciona como una tecnología muy útil en la investigación básica. Se han utilizado para detectar autoanticuerpos específicos, que pueden ser utilizados como biomarcadores de diagnóstico y predictores de respuesta al tratamiento. Han permitido caracterizar perfiles moleculares que ayudan a clasificar a los pacientes según su actividad inmunológica y a predecir su respuesta a terapias específicas.(16)

#### *1.4. Aplicación de la bioinformática en la caracterización molecular de pacientes, diagnóstico y tratamiento en el Lupus*

En el contexto del LES, enfermedad que como hemos comentado es muy heterogénea a nivel clínico y molecular y presenta aún una etiología desconocida y tratamientos inespecífico y con múltiples efectos adversos, el uso de tecnologías avanzadas, como los microarrays han facilitado el análisis de datos genómicos, lo que ha permitido profundizar en el conocimiento biológico de la enfermedad, así como la identificación de genes diferencialmente expresados (DEGs)(17) entre pacientes y muestras de individuos sanos o entre enfermos. Estos genes son clave, ya que sus niveles de expresión varían significativamente entre los tejidos sanos y los afectados, lo cual implica que están relacionados con la patología, permitiendo una visión detallada de las alteraciones moleculares detrás de la enfermedad.

El análisis bioinformático de los datos de expresión génica obtenidos permite también identificar vías biológicas específicas que están activas o alteradas en el LES, lo que contribuye a esclarecer los mecanismos que provocan el daño orgánico, en este contexto, en los riñones. Esta información ayuda a detectar biomarcadores clave, como proteínas específicas, que podrían convertirse en dianas terapéuticas para tratamientos

más precisos y que generen menor toxicidad, mejorando así la calidad de vida de los pacientes y promoviendo la medicina personalizada.

Hipotéticamente, mediante el análisis de variaciones en la expresión genética, se puede determinar por qué algunos pacientes responden mejor que otros a tratamientos como el MMF. Esto podría permitir adaptar los tratamientos a las características genéticas de cada paciente.

En este trabajo se aplicarán herramientas bioinformáticas para detectar precisamente qué genes están diferencialmente expresados y qué vías biológicas están activas en una cohorte de pacientes con NL, a diferentes tiempos, comparando respondedores y no respondedores al tratamiento con MMF. Además de este análisis genómico, se incluirá el estudio de diversas variables clínicas asociadas a la respuesta al tratamiento, con el objetivo de comprender las diferencias tanto moleculares como fenotípicas entre los pacientes respondedores y no respondedores.

## 2. Hipótesis y Objetivos

### 2.1. Hipótesis

- La NL presenta una alta heterogeneidad a nivel clínico y molecular, lo que supone diferencias significativas en la respuesta al tratamiento entre los pacientes.
- Existen genes diferencialmente expresados (DEGs) entre pacientes respondedores y no respondedores al tratamiento con MMF, lo que podría permitir establecer perfiles moleculares distintos entre ambos grupos.
- Los genes diferencialmente expresados activan o inhiben vías biológicas específicas que influyen en la progresión de la NL y en la respuesta al tratamiento

### 2.2. Objetivos

- Evaluar la relación entre las variables clínicas, demográficas y otros tratamientos con la respuesta al tratamiento con MMF
- Identificación de DEGs entre dos grupos: respondedores y no respondedores al tratamiento con MMF a distintos tiempos: temprano, al año y de seguimiento tomando como referencia el tiempo desde la biopsia para analizar las diferencias entre periodos.

- Caracterización de procesos biológicos enriquecidos en ambos grupos de pacientes a lo largo de los distintos periodos.

### 3. Metodología

#### 3.1. Origen y características de los datos empleados

##### Cohorte de estudio

Los datos originales de este trabajo provienen de un estudio observacional y longitudinal en el que se reclutaron y siguieron pacientes con NL durante más de dos años bajo el protocolo del estudio SPARE (*Study of Biological Pathways, Disease Activity, and Response Markers in Patients with Systemic Lupus Erythematosus*).

El protocolo fue aprobado por la Junta de Revisión Institucional de la Escuela de Medicina de la Universidad Johns Hopkins, y todos los pacientes con LES que participaron dieron su consentimiento informado.

Los pacientes adultos elegibles tenían entre 18 y 75 años y cumplían con los criterios revisados de clasificación del American College of Rheumatology para el diagnóstico de LES. Los pacientes fueron tratados según la práctica clínica estándar, y se excluyeron aquellos tratados con medicamentos biológicos. En total se seleccionaron 301 pacientes. (18)

Para este trabajo se ha partido de los datos de un estudio posterior: "*Immune and molecular landscape behind nonresponse to Mycophenolate Mofetil and Azathioprine in lupus nephritis therapy*" (18) cuyos datos tienen número de acceso de GEO: GSE224705. En este estudio se realizó un análisis retrospectivo de la cohorte de Hopkins y se seleccionaron los pacientes tratados con MMF, AZA, HC o tratamiento estándar (SOC) y que contaban con al menos dos visitas tras el inicio del tratamiento y se descartaron muestras tratadas con otros inmunosupresores junto con MMF o AZA. (18)

La información clínica fue pseudo-anonimizada e incluyó historia médica, datos demográficos, uso de medicamentos y dosis administradas y títulos de autoanticuerpos. En cada visita, se midió la actividad de la enfermedad mediante el índice SLEDAI versión SELENA y PGA. También se realizaron análisis de orina, anticuerpos anti-dsDNA y

concentraciones plasmáticas de los componentes del complemento C3 y C4(18), además se tomaron también muestras sanguíneas que permitió el análisis de expresión.

Adicionalmente, se registraron datos específicos de las biopsias renales realizadas a los pacientes. Estos datos incluyen el año de diagnóstico de la NL, los años en que se detectó insuficiencia o fallo renal, y las fechas de biopsias en años y meses desde la biopsia. Cada biopsia está acompañada por una clasificación histológica específica, que detalla las clases de nefritis lúpica presentes en cada muestra.

Para este TFM se seleccionaron aquellos pacientes tratados con MMF de dicha cohorte. La evaluación de la respuesta al tratamiento se basó en dos índices: uPCR y mPERR.

- Para uPCR, se consideró una respuesta favorable según las guías de EULAR, la reducción y mantenimiento del índice proteína/creatinina en orina por debajo de 0.5 g/g a los 12 meses del inicio de la terapia. Alternativamente, se aceptó una disminución del 25% a los 3 meses y del 50% a los 6 meses desde el valor inicial.(19)
- Para mPERR, una respuesta favorable requiere de la ausencia de empeoramiento en eGFR, con valores de  $\geq 60$  mL/min/1.73 m<sup>2</sup> o  $\leq 20\%$  por debajo del valor basal.(20)

### **Dataset de expresión**

Los datos de expresión utilizados en este trabajo provienen de microarrays y se encuentran en el repositorio GEO (*Gene Expression Omnibus*) del NCBI con identificador: GPL13158.

La tecnología de microarrays de ADN es una herramienta que permite analizar simultáneamente la expresión de miles de genes. Para analizar la expresión génica con microarrays, primero se extrae ARN de las células de interés y se convierte en ADN complementario (ADNc), el cual se marca con una molécula fluorescente. Esta muestra marcada, denominada "sonda," se aplica sobre el microarray, donde las moléculas de ADNc buscan y se unen (hibridan) con sus secuencias de ADN complementarias (dianas) presentes en el microarray. Tras la hibridación, se mide la fluorescencia en cada diana:

una mayor intensidad de fluorescencia indica una mayor cantidad de hibridación, lo que refleja un nivel más alto de expresión de ese gen específico. (21)

En este trabajo, los datos de expresión génica de los que se parte se habían preprocesado previamente utilizando un flujo estándar para microarrays: La matriz de datos de expresión se generó a partir de archivos FASTQ, que contienen las secuencias de ARN obtenidas de las muestras biológicas (biopsia renal). Estos archivos fueron preprocesados aplicando el método de normalización RMA (Robust Multiarray Average), que es un enfoque comúnmente utilizado en estudios con microarrays (22). Este proceso implica la corrección de valores de fondo para reducir el ruido técnico, minimizando así las variaciones que no reflejan cambios biológicos reales. Después, se aplica un ajuste de cuantil, que iguala las distribuciones de las intensidades entre las distintas muestras, haciendo que las comparaciones sean más consistentes. Por último, se realiza una transformación logarítmica de los datos, lo cual permite una interpretación más clara de la expresión. (21)(22)(23)

### 3.2. *Software empleado en el trabajo*

El código relativo al análisis de los datos descritos se escribió usando el programa Rstudio en su versión 4.3.2. En concreto los scripts se escribieron en formato Rmarkdown que permite combinar código y texto explicativo en un solo documento. Esto facilita tanto la explicación del análisis como la interpretación de los resultados, haciendo el contenido más visual y comprensible. Además del análisis de datos, permite la visualización de gráficas dentro del código lo que facilita la comprensión de este. (24)

Las principales librerías utilizadas fueron:

- **dplyr**: Proporciona funciones para la manipulación eficiente de datos, como filtrado, selección, unión y resumen de datos. (25)
- **readxl**: Permite leer archivos de Excel (formato .xls y .xlsx) y pasarlos a R. (26)
- **readODS**: Utilizado para leer archivos de *OpenDocument Spreadsheet* (ods).(27)
- **tidyr**: Ayuda a organizar los datos en un formato ordenado (tidy), facilitando su manipulación y análisis.(28)

- **biomaRt**: paquete que permite acceder a BioMart, base de datos utilizada para anotar genes. (29)
- **matrixStats**: Permite obtener estadística de datos en filas y columnas de matrices, comúnmente utilizado en análisis de datos de expresión. (30)
- **caret**: proporciona herramientas para preprocesamiento de datos como datos de expresión.(31)
- **limma**: Utilizado para el análisis de datos de microarrays y RNA-seq, en concreto para obtener genes diferencialmente expresados.(32)
- **ggplot2**: Librería para la representación de datos. (33)
- **ggrepel**: Mejora la legibilidad de gráficos evitando que las etiquetas se superpongan entre sí en los gráficos de ggplot2.(34)
- **readr**: Librería utilizada para leer datos de archivos de texto de formatos CSV o TSV.(35)
- **pheatmap**: Librería utilizada para la creación de mapas de calor.(36)
- **RColorBrewer**: Proporciona paletas de colores para usar en gráficos.(37)

Para el análisis de enriquecimiento funcional se utilizó la herramienta web Genecodis. Se trata de una herramienta avanzada de bioinformática diseñada para realizar análisis de enriquecimiento funcional tanto singular (SEA) como modular (MEA). Genecodis se puede utilizar para el análisis de genes, proteínas, miRNAs, factores de transcripción y sitios CpG. En este trabajo, se estudiaron los genes obtenidos a partir del análisis de expresión diferencial. Estos genes se cargaron en la plataforma de Genecodis, donde el programa los procesó para identificar procesos biológicos y funciones moleculares sobrerrepresentadas con esos genes comparándolos con bases de datos de referencia como Gene Ontology (GO). (38)

Los análisis funcionales se emplean para asociar listas de genes con funciones biológicas específicas. Para este propósito, se utilizan pruebas estadísticas, como el test exacto de Fisher, que permite identificar términos funcionales significativamente enriquecidos entre los DEGs obtenidos en expresión diferencial. Este test compara la proporción de genes vinculados a un término funcional particular dentro de la lista de DEGs con su proporción en el conjunto total de genes. De esta manera, se pueden identificar

términos funcionales relevantes que están asociados con procesos biológicos específicos. (39)

Por último, la búsqueda bibliográfica a lo largo de todo el trabajo se llevó a cabo en las bases de datos *PubMed* y *Google Scholar*.

### 3.3. *Procesado y análisis de los de los datos*

#### 3.3.1. Exploración de los datos clínicos y de expresión

En el análisis de datos se realizó una evaluación exhaustiva de varias tablas clínicas, tanto a nivel de paciente como de muestra. Las muestras corresponden a sangre periférica, recolectadas durante diversas visitas en distintos momentos para cada paciente, lo que permitió hacer un seguimiento longitudinal de su evolución clínica. Cada paciente y muestra cuenta con un identificador único en las tablas, facilitando la vinculación entre los datos de las diferentes visitas.

Estas tablas incluyen información importante sobre biomarcadores clínicos, datos demográficos además de indicadores específicos de la actividad del lupus, como el PGA y el SLEDAI, así como los resultados de las biopsias renales. Estos datos resultan esenciales para evaluar la condición clínica de los pacientes y su respuesta al tratamiento.

El análisis se centró en las variables de respuesta mPERR y uPCR ya explicadas. Ambas variables, en formato sí/no, se utilizan para clasificar a los pacientes como respondedores o no respondedores al tratamiento, en función de la mejora renal (mejora de la filtración) y la reducción de uPCR tras el uso MMF.

Durante la limpieza de los datos, se eliminaron columnas sin contenido y datos incompletos. Es relativamente normal encontrar este tipo de situaciones al trabajar con datos generados durante la práctica clínica rutinaria. Este proceso permitió reducir las tablas, manteniendo solo la información más relevante, completa y homogénea para el análisis.

Al finalizar esta etapa, nos quedamos con un total de 106 muestras pertenecientes a 37 pacientes, distribuidas en varias tablas, que serán utilizadas en los análisis estadísticos posteriores. La limpieza de los datos fue esencial para asegurar que el conjunto de datos

final nos permitirá evaluar con precisión la respuesta al tratamiento y los factores clínicos asociados a la evolución de los pacientes.

### 3.3.2. Análisis de variables clínicas

En el análisis de las variables clínicas, por un lado, se estudiaron las variables a nivel de paciente que son variables que no se ven modificadas en las distintas visitas como pueden ser los datos demográficos.

Por otro lado, se llevó a cabo un análisis similar con las variables a nivel de muestras, que fueron obtenidas en diversas visitas de seguimiento como los índices de actividad, los niveles de eGFR o uPCR.

Se realizó análisis de asociación tanto para las variables a nivel de paciente como a nivel de visita para su respuesta a MMF medida con uPCR como mPERR.

#### **Análisis de variables categóricas**

Para las variables categóricas, se utilizaron diferentes pruebas estadísticas dependiendo del tamaño de las categorías, con el objetivo de medir la asociación entre cada variable y la respuesta al fármaco. En los casos en que las categorías eran pequeñas o desiguales, se aplicó el test de Fisher (40), que es adecuado para tablas de contingencia con valores bajos y cuando las frecuencias esperadas por celda son menores a 5. Cuando las frecuencias eran más grandes, se utilizó el test de Chi-cuadrado (41), que es más eficiente cuando las frecuencias esperadas por celda son mayores a 5, lo que permite evaluar la independencia entre las variables de manera robusta. Ambos se realizaron tanto en los datos por paciente como en los datos por muestra, aplicando p-valor de 0.05.

#### **Análisis de variables numéricas**

Para las variables numéricas (como las medidas hematológicas, niveles de colesterol o la tensión arterial), primero se evaluó la normalidad de los datos mediante el test de Shapiro-Wilk (42). Si los datos seguían una distribución normal, se utilizó el t-test(43) para comparar las medias entre los grupos de respuesta (mPERR o uPCR). Si los datos no seguían una distribución normal, se aplicó el test de Wilcoxon (44), una prueba no

paramétrica adecuada para comparar distribuciones entre dos grupos. Todo ello utilizando también p-valor de 0.05

Al final del análisis, se identificaron las variables que mostraron asociaciones significativas con la respuesta al tratamiento.

### 3.3.3. Procesamiento de los datos de expresión

Previo al análisis de expresión diferencial se alinearon los identificadores de las muestras entre la matriz de expresión y los datos clínicos, asegurando que las 106 muestras estuvieran correctamente representadas y sin discrepancias en la nomenclatura.

Posteriormente, se realizó la anotación de sondas o “*probes*” a identificadores únicos de genes “*gene symbol*” utilizando la herramienta *biomaRt*,<sup>(45)</sup> que permite mapear los identificadores de los *probes* a los *symbol* de los genes utilizando la información de base de datos *Ensembl*. Como un mismo gen puede estar representado por varios *probes* en el microarray, se consolidaron estos valores de expresión usando la mediana. Este método fue seleccionado porque es más resistente a los valores atípicos en los datos, ya que, a diferencia de la media, la mediana es menos sensible a valores extremos, lo que permite una representación más robusta de la expresión génica (46). Tras este proceso, la matriz de expresión original, que contenía 54,715 *probes*, se redujo a 22,565 genes. También se eliminó el ruido en los datos identificando y excluyendo genes con varianza cercana a cero mediante el paquete *caret*.

#### 3.3.3.1. Expresión diferencial por periodos con covariables

El análisis de expresión diferencial se llevó a cabo utilizando el paquete *limma* (47), comparando los grupos de pacientes respondedores y no respondedores según las variables uPCR y mPERR. Los p-valores se ajustaron mediante el método de Benjamini-Hochberg (BH) (48) para controlar la tasa de falsos descubrimientos (FDR), considerando significativos aquellos genes con p-valor ajustado  $< 0,05$ .

En este análisis se dividieron los datos en distintos periodos temporales y se añadieron a la función distintas covariables. Los periodos se definieron en base a los meses desde la biopsia renal de los pacientes dividiendo las muestras en tres periodos:

- Temprano (hasta 6 meses post biopsia), para evaluar la respuesta temprana al tratamiento.
- Año (entre 6 y 19 meses después de la biopsia), que refleja los cambios en la respuesta en torno al año.
- Seguimiento (más de 19 meses), que captura los efectos a largo plazo del tratamiento.

Se incorporaron covariables clave: dosis de prednisona, dosis de MMF, actividad según SLEDAI (act) y sexo. Las covariables se añaden al análisis de expresión diferencial para ajustar el modelo y tener en cuenta otros factores clínicos que pueden influir en los resultados, además de la respuesta al tratamiento. Esto le indica al modelo que debe separar el impacto de estas variables adicionales de la respuesta específica al tratamiento, lo que permite identificar con mayor precisión los genes que están verdaderamente asociados con la respuesta.(49)

La actividad se calculó de la siguiente manera:

act = 1 si SLEDAI < 3.

act = 2 si SLEDAI está entre 3 y 7.

act = 3 si SLEDAI > 7.

En cuanto a la consideración de muestras repetidas de un mismo paciente, se decidió optar directamente por un enfoque más restrictivo usando limma a distintos tiempos, sin aplicar modelos mixtos. Este enfoque se eligió porque las muestras de un mismo paciente no se consideran réplicas biológicas; en cambio, representan distintos tiempos y condiciones a lo largo del tratamiento, por lo que se analizaron como datos independientes. Además, al dividir el análisis en periodos temporales, se minimiza el impacto de las muestras repetidas de un mismo paciente en el mismo periodo.

La comparación que realiza la función es de Respondedores contra No respondedores, por tanto, los genes que se obtienen con un LogF<sub>c</sub> positivo o sobreexpresados corresponden a los genes denominados “Up-regulated” o de mayor expresión en respondedores. Por su parte, los genes infraexpresados con valores de LogF<sub>c</sub> negativos

son los “down-regulated”, que son los más diferencialmente expresados en las muestras no respondedoras.

Los resultados del análisis se representaron mediante: Volcano plots, Heatmaps y la diferenciación de las muestras mediante PCA. Además, se obtuvieron los genes “up” y “down -regulated”

#### 3.3.4. Enriquecimiento funcional e interpretación de resultados

Para el análisis de enriquecimiento funcional, se utilizaron las listas de DEGs correspondientes a cada periodo. Estas listas fueron introducidas en la herramienta GeneCodis para realizar el análisis de enriquecimiento, permitiendo identificar procesos biológicos (BP, *Biological Processes*) asociados con los genes seleccionados.

Se descargaron las tablas de las funciones más enriquecidas a partir de los resultados proporcionados por GeneCodis. Con estos datos, se procedió a representar las funciones en formato de diagramas de barras, destacando las funciones más enriquecidas. Finalmente, se realizó una interpretación de los resultados, agrupando los procesos biológicos más relevantes relacionados con los DEGs obtenidos.

#### 3.4. *Declaración obligatoria de herramientas de IA*

En este trabajo se ha utilizado ChatGPT (50), herramienta desarrollada por OpenAI, como apoyo en diversas áreas: en primer lugar, se utilizó como ayuda en la programación con sugerencias y mejoras de código con objeto de la obtención de resultados relevantes y representación de datos. Se utilizó también como apoyo en la búsqueda bibliográfica y para la obtención de resúmenes de fragmentos de artículos científicos que se consideraron relevantes para la justificación de los resultados obtenidos tras el análisis. Además, se utilizó como apoyo en la redacción y relación de ideas originales, no se recurrió a la herramienta para la creación de contenido original.

Todo el contenido generado se revisó minuciosamente.

## 4. Resultados

### 4.1. Análisis de variables clínicas

Se ha realizado un resumen en formato tabla de los resultados de las variables significativas obtenidas tras el análisis de variables por paciente y por muestra, incluyendo el test estadístico utilizado y los resultados de p-valor. Los datos muestran que algunas variables son significativas únicamente para la respuesta medida con el indicador mPERR o con uPCR, mientras que otras variables muestran significancia en la respuesta con ambos criterios.

#### 4.1.1. Análisis de variables por paciente

Variable	p-valor (mPERR)	p-valor (uPCR)	Tipo de prueba	Respuesta/ No respuesta (MPERR)	Respuesta/ No respuesta (uPCR)
<b>Demográficas:</b>					
Sexo*	-	-	Fisher	F: 70 (72,92%) / 26 (27,08%) M: 8 (80%) / 2 (20%)	F: 68 (70,83%) / 28 (29,17%) M: 8 (80%) / 2 (20%)
Edad de inicio *	-	-	Wilcoxon	25,63/23,57	24,4/26,83
Raza	-	0.0341	Fisher	A: 11 (84,62%) / 2 (15,39%) B: 22 (56,41%) / 17 (43,59%) O: 2 (66,67%) / 1 (33,33%) W: 43 (84,31%) / 8 (15,69%)	A: 11 (84,62%) / 2 (15,39%) B: 16 (41,03%) / 23 (58,97%) O: 2 (66,67%) / 1 (33,33%) W: 47 (92,16%) / 4 (7,84%)
<b>Clínicas:</b>					
Pericarditis	0.0282	-	Fisher	0: 18 (72%) / 6 (50%) 1: 7 (28%) / 6 (50%)	0: 19 (79,17%) / 5 (38,46%) 1: 5 (20,83%) / 8 (61,54%)
Pleuritis	0.0075	0.0323	Fisher	0: 17 (68%) / 3 (25%) 1: 8 (32%) / 9 (75%)	0: 17 (70,83%) / 3 (23,08%) 1: 7 (29,17%) / 10 (76,92%)
Alopecia presente	-	0.0173	Fisher	0: 15 (60%) / 2 (16,67%) 1: 10 (40%) / 10 (83,33%)	0: 13 (54,17%) / 4 (30,77%) 1: 11 (45,83%) / 9 (69,23%)
Trombosis venosa	-	0.0184	Fisher	0: 15 (60%) / 2 (16,67%) 1: 10 (40%) / 10 (83,33%)	0: 13 (54,17%) / 4 (30,77%) 1: 11 (45,83%) / 9 (69,23%)
Síndrome nefrótico	0.0165	-	Fisher	0: 12 (48%) / 4 (33,33%) 1: 13 (52%) / 8 (66,67%)	0: 14 (58,33%) / 2 (15,38%) 1: 10 (41,67%) / 11 (84,62%)
<b>Tratamientos:</b>					
Tratamiento con citoquinas alguna vez	0.0368	-	Fisher	0: 25 (100%) / 9 (75%) 1: 0 (0%) / 3 (25%)	0: 24 (100%) / 10 (76,92%) 1: 0 (0%) / 3 (23,08%)
Tratamiento con prednisona actual	0.0382	0.0128	Fisher	0: 16 (64%) / 2 (16,67%) 1: 9 (36%) / 10 (83,33%)	0: 15 (62,5%) / 3 (23,08%) 1: 9 (37,5%) / 10 (76,92%)
Tratamiento con ASA alguna vez	0.0162	-	Fisher	0: 20 (80%) / 5 (41,67%) 1: 5 (20%) / 7 (58,33%)	0: 19 (79,17%) / 7 (53,85%) 1: 5 (20,83%) / 6 (46,15%)

**Tabla 1. Variables significativas del análisis por paciente.** En la columna de Sexo (M/F), "M" representa al género masculino y "F" al femenino. En la columna de Raza (A, B, O, W), las letras indican los diferentes grupos: A (asiáticos), B (afroamericanos), O (otros) y W (caucásicos). Las variables binarias (0/1) indican la ausencia (0) o presencia (1) de la variable implicada. Los valores en las columnas de respuesta reflejan el número de pacientes

respondedores y no respondedores al tratamiento en cada subgrupo, así como el porcentaje calculado sobre el total de cada categoría.

Las variables demográficas reflejan ciertos patrones con respecto a la respuesta al tratamiento. En cuanto al sexo, no resultó ser significativo en la variación de la respuesta; en este caso los hombres están subrepresentados en la muestra lo que puede poner de manifiesto la menor incidencia de la enfermedad en este grupo. Los datos sugieren que presentan ligeramente una mejor respuesta al tratamiento en comparación con las mujeres, sin embargo, debido a la disparidad en el tamaño de la muestra, esta diferencia no puede considerarse concluyente.

La edad de inicio de la enfermedad, por otro lado, no mostró diferencias significativas entre grupos de respuesta, con una media de inicio en torno a los 25 años en ambos grupos, lo cual sugiere que esta variable no parece influir en la efectividad del tratamiento.

La raza, sin embargo, mostró diferencias; los afroamericanos presentan una mayor prevalencia y severidad del LES, lo que podría influir en su respuesta al tratamiento. Esto sugiere que ciertos grupos étnicos, especialmente los afroamericanos, pueden tener mayor susceptibilidad al daño renal y podrían presentar una respuesta menos favorable al tratamiento inmunosupresor. (1)

Algunas variables clínicas relacionadas con la enfermedad mostraron diferencias entre los grupos de respondedores y no respondedores. La pericarditis está más presente en los no respondedores para la respuesta mPERR. Se trata de una complicación común en LES que refleja actividad inflamatoria sistémica y puede indicar un cuadro clínico más severo (51). Sin embargo, no podemos conocer directamente si existe relación con la respuesta al tratamiento.

Por otro lado, la pleuritis también se observa en mayor medida en los no respondedores para ambos criterios (mPERR y uPCR), lo que podría indicar una relación con una actividad inflamatoria más severa de la enfermedad, aunque no una influencia directa en la respuesta al tratamiento.

La alopecia y la trombosis venosa fueron significativas para la respuesta uPCR y también enriquecidas en el grupo de los no respondedores. La alopecia, incluida en la evaluación SLEDAI, y la trombosis venosa, asociada al síndrome antifosfolípido frecuente en lupus

que aumenta el riesgo de trombosis renal y empeora la proteinuria y función renal (51), podrían reflejar una actividad elevada de la enfermedad, pero a priori no está demostrado que influyan en la respuesta a los tratamientos de NL.

Otro hallazgo significativo para la respuesta mPERR es el síndrome nefrótico, que se caracteriza por la presencia de proteinuria elevada ( $\geq 3000$  mg/24 horas), asociada a una progresión grave de la NL y que sugiere la necesidad de tratamiento inmunosupresor intensivo (52), que, sin embargo, en este caso, presenta enriquecimiento en los respondedores.

Las variables relacionadas con el tratamiento muestran que el uso actual de prednisona, citoquinas y ácido acetil salicílico (ASA) está enriquecido en los no respondedores. Esto sugiere que los pacientes que no responden requieren dosis continuas de prednisona o tratamientos adicionales como citoquinas y ASA, posiblemente debido a una mayor actividad inflamatoria, lo cual puede dificultar la mejoría en su función renal e hipotéticamente su respuesta al tratamiento.

#### 4.1.2. Análisis de variables por muestra

Variable	p-valor mPERR	p-valor uPCR	Tipo de prueba	Media Respondedores mPERR / No respondedores	Media Respondedores uPCR / No respondedores
<b>Variables clínicas</b>					
SLEDAI	0.0465	0.0192	Wilcoxon	3.24 / 3.86	2.96 / 4.53
PGA	0.1199	0.0055	Wilcoxon	0.67 / 1.62	0.62 / 1.67
C3	0.0003	0.0306	Wilcoxon	99.70 / 126.52	98.76 / 128.11
C4	0.0014	0.0192	t-test	17.61 / 23.26	17.18 / 24.21
DNA	0.0042	-	Wilcoxon	69.82 / 9.82	70.74 / 9.46
eGFR	0.0305	-	Wilcoxon	96.47 / 81.61	94.52 / 87.23
<b>Tratamientos</b>					
Dosis de MMF	0.0050	-	Wilcoxon	1690.39 / 2267.86	1583.55 / 2500
Dosis de prednisona	0.0046	-	Wilcoxon	5.26 / 11.25	5.66 / 9.83
Dosis de ASA	0.0019	-	Wilcoxon	30.91 / 63.00	33.93 / 53.07

**Tabla 2. Variables significativas del análisis por muestras.** Comparación de los valores medios de las variables incluidas en la tabla entre respondedores y no respondedores, con significancia estadística evaluada mediante Wilcoxon o t-test. La tabla incluye medidas de actividad y dosis de tratamientos.

En este análisis para muestras, se han seleccionado las variables que resultaron significativas. La tabla sigue el mismo formato que en el análisis por paciente, dividiendo en variables clínicas y de tratamiento. Los resultados presentan el p-valor para cada variable en relación con ambas respuestas, mPERR y uPCR, junto con el tipo de prueba estadística aplicada (Wilcoxon o t-test) y la media de cada grupo de respuesta.

Las puntuaciones de actividad de la enfermedad, medidas por PGA y SLEDAI, son más altas en los no respondedores. La puntuación de PGA resulta significativa para ambas respuestas, lo que sugiere que la falta de respuesta al tratamiento contribuye a un agravamiento de la enfermedad. Por otro lado, SLEDAI muestra significancia únicamente en el contexto de uPCR, y refuerza la conexión entre una alta actividad de la enfermedad y una respuesta reducida, evidenciada a través de la proteinuria.

En cuanto a los niveles de complemento, C3 y C4 resultan significativos para ambas respuestas y son más altos en los no respondedores. Esto sugiere una mayor actividad inmunitaria en los pacientes que no responden bien al tratamiento. Sin embargo, los niveles de anticuerpos anti-dsDNA son significativamente más altos en los respondedores para mPERR.

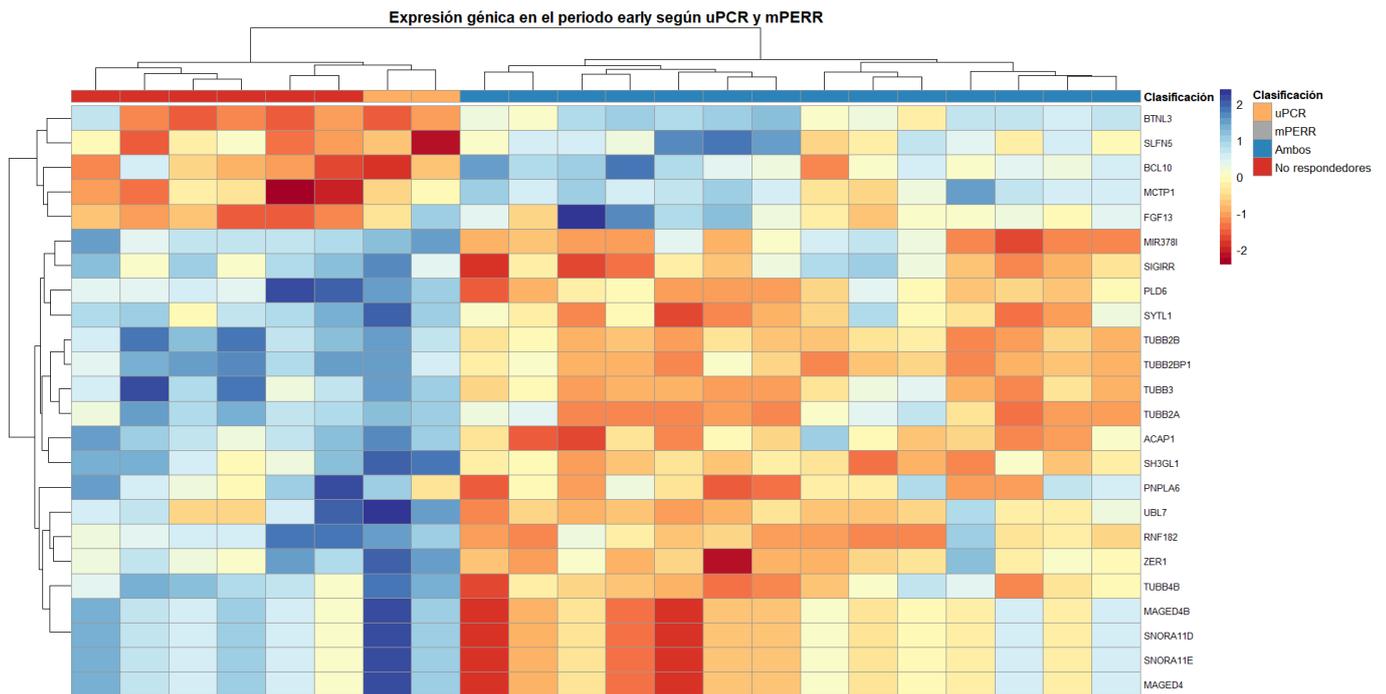
eGFR es significativa solo para mPERR, reflejando que una mejor función renal en términos de filtración se asocia con una respuesta favorable. Este hallazgo tiene sentido ya que el cálculo de mPERR incorpora eGFR en su fórmula.

En cuanto a los tratamientos, las dosis de MMF y prednisona son significativamente más altas en los no respondedores para ambas respuestas, lo que podría indicar la necesidad de dosis más intensivas en pacientes que peor responden al tratamiento. De forma similar, la dosis de ASA es mayor en los no respondedores.

## 4.2. *Expresión diferencial*

### 4.2.1. Periodo temprano: hasta 6 meses desde la biopsia.

En el periodo temprano, se seleccionaron 22 muestras para evaluar la expresión diferencial de genes en respuesta a uPCR y mPERR, obteniéndose 2 genes diferencialmente expresados para uPCR y 23 genes para mPERR, con solo uno en común entre ambas respuestas. Es decir, hay una mayor cantidad de genes asociados específicamente a mPERR en este periodo temprano.



**Figura 1. Heatmap de DEGS para muestras del periodo temprano.** Heatmap de DEGs en muestras del periodo temprano, normalizado por z-scores por fila. Los colores indican niveles de expresión, con tonos azules para mayor expresión y tonos rojos para menor expresión. Las columnas representan las muestras que están etiquetadas según su respuesta uPCR y/o mPERR o la no respuesta

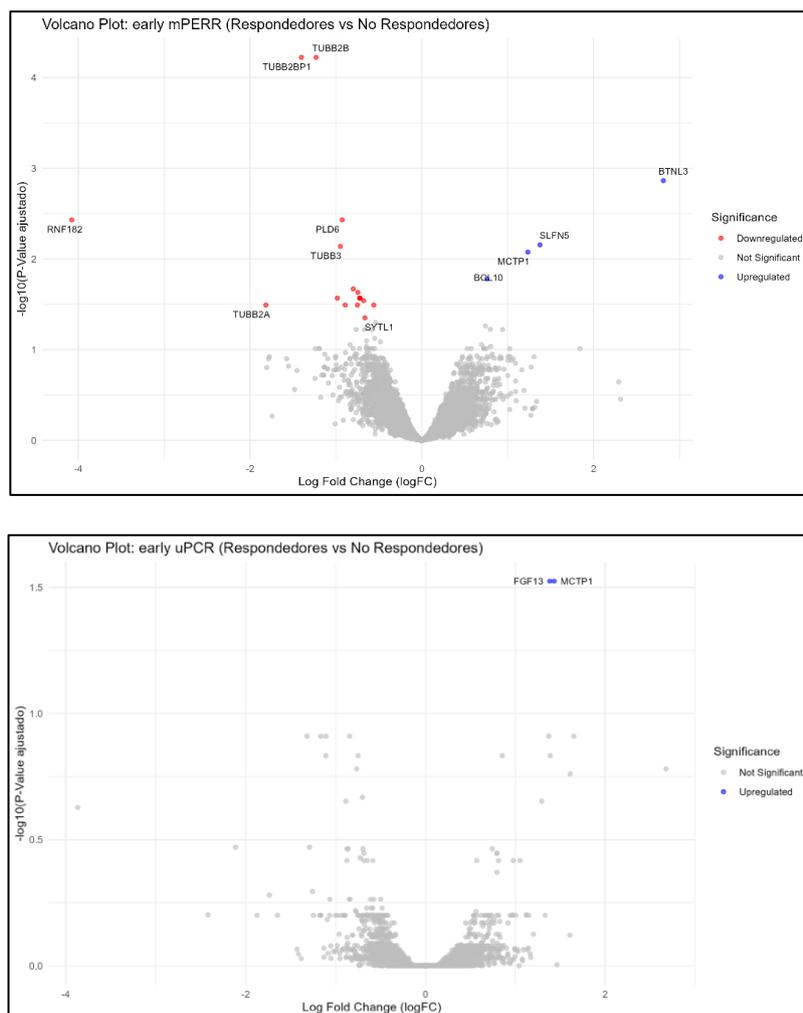
El heatmap muestra la expresión de genes específicos en diferentes muestras, representadas en el eje horizontal, mientras que los genes están etiquetados en el eje vertical. Las muestras se han agrupado de acuerdo con su respuesta clínica en función de uPCR, mPERR, o a ambas respuestas, así como a la ausencia de respuesta. Los colores reflejan la intensidad de la expresión génica dentro de cada muestra, permitiendo observar variaciones de expresión entre grupos de respuesta.

Los dendrogramas agrupan tanto muestras como genes según la similitud de expresión, formando clústeres (grupos de elementos similares); las muestras que responden a una, ambas, o ningún criterio de respuesta tienden a agruparse, reflejando perfiles de expresión génica similares dentro de cada grupo de respuesta.

La agrupación muestra que los pacientes clasificados como "Ambos" y "No respondedores" se separan claramente en clústeres, y además se observa una similitud entre ambas respuestas: mPERR y uPCR. El hecho de que solo haya dos genes diferencialmente expresados en uPCR y más en mPERR puede explicarse observando el

heatmap: algunos pacientes que responden a uPCR se asemejan a los no respondedores, lo que hace que para el criterio de respuesta uPCR se incluyan muestras más heterogéneas. En cambio, mPERR genera una comparación entre grupos más homogéneos permitiendo detectar un mayor número de DEGs debido a la menor variabilidad entre grupos.

**Genes “up” y “down” regulated**



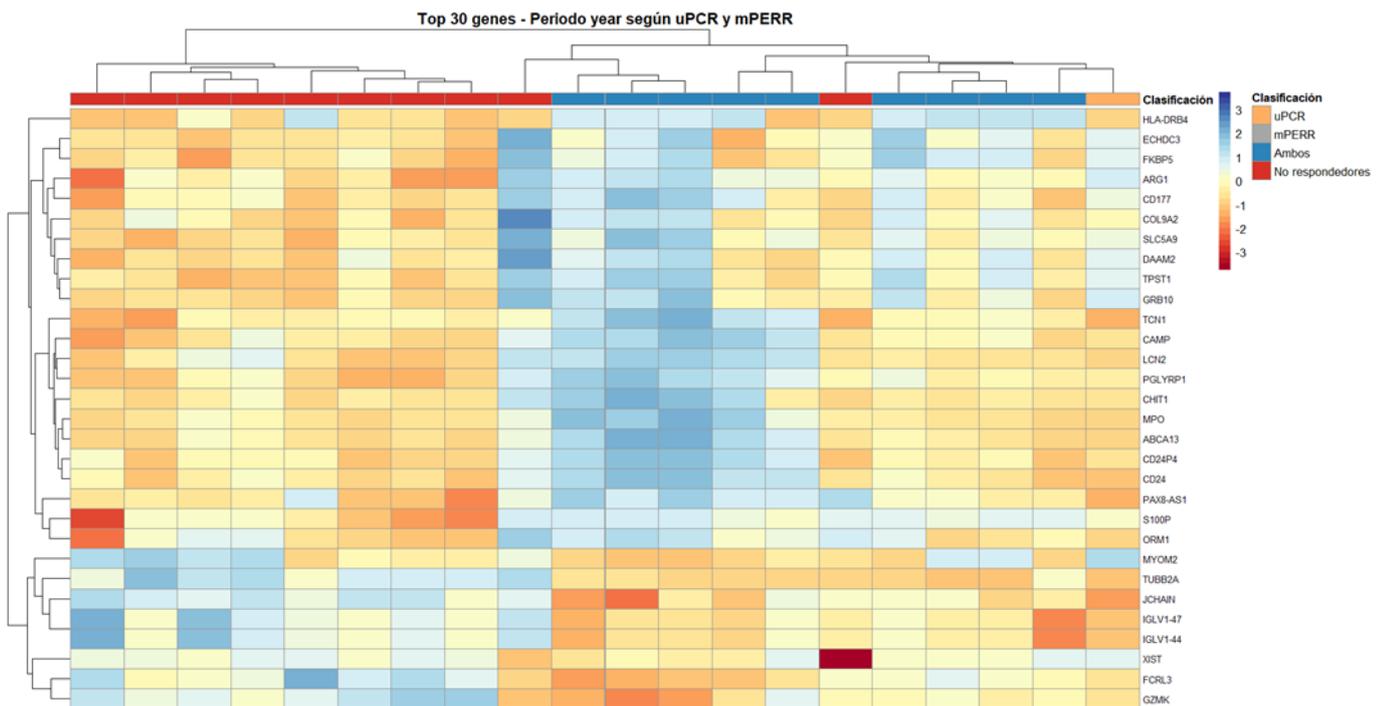
**Figura 2. Volcano plots para genes del periodo temprano.** Volcano plots que muestran los DEGs en respondedores y no respondedores para el periodo temprano según la respuesta a mPERR (arriba) y uPCR (abajo). En el eje X se representa el cambio en la expresión génica medido en logFC y en el eje Y el nivel de significancia estadística de ese cambio (-log10 del p-valor). En azul se representan los genes con sobreexpresión significativa, mientras que en rojo se representan genes con subexpresión significativa. Los puntos en gris corresponden a genes sin cambios significativos en su expresión

En el periodo temprano, el volcano plot para uPCR muestra dos genes diferencialmente expresados, ambos up-regulated, indicando que estos genes están sobreexpresados en respondedores en comparación con los no respondedores. En contraste, el volcano plot

de mPERR presenta una mayor cantidad de genes significativos, con genes tanto up-regulated como down-regulated. Se etiquetan los genes más significativos; con mayor p-valor ajustado.

Esta diferencia entre ambos volcano plots sugiere que mPERR identifica un grupo de genes con patrones de expresión más variados y específicos, ya que detecta tanto genes que aumentan como disminuyen su expresión en los respondedores. Esto podría indicar de nuevo que mPERR permite una comparación más homogénea de grupos, reduciendo la variabilidad dentro de cada grupo y facilitando la detección de un mayor número de DEGs.

#### 4.2.2. Periodo Año: entre 6 y 19 meses desde la biopsia.



**Figura 3. Heatmap de DEGS para muestras del periodo año.** Heatmap de los 30 genes principales de DEGS en el periodo de un año, normalizado por z-scores por fila. Los colores indican niveles de expresión, con tonos azules para mayor expresión y tonos rojos para menor expresión. Las columnas representan las muestras que están etiquetadas según su respuesta uPCR y/o mPERR o la no respuesta

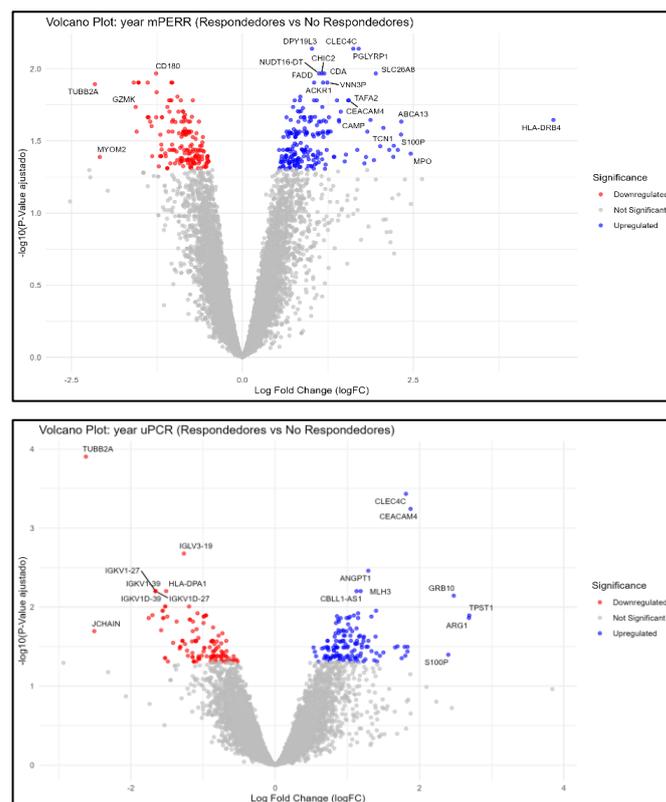
En el periodo de un año desde la biopsia, se analizaron 20 muestras. Se identificaron 273 genes diferencialmente expresados para uPCR y 381 para mPERR, los genes comunes entre las dos respuestas fueron 167.

De estos genes, para la representación gráfica se seleccionaron los 30 genes más representativos en función del p-valor debido a limitaciones de visibilidad en el heatmap.

El heatmap muestra la agrupación de las 20 muestras según sus perfiles de expresión génica. Las 10 muestras clasificadas como no respondedoras forman un clúster definido, aunque una de estas muestras se agrupa junto a los respondedores a ambos criterios, lo que sugiere variabilidad o una posible transición en su perfil de respuesta.

Además, una muestra clasificada como respondedora a uPCR se encuentra cercana a los respondedores a ambos criterios, reflejando una similitud en sus patrones de expresión y una posible superposición en los mecanismos de respuesta.

### Genes “up” y “down-regulated”



**Figura 4. Volcano plots para los genes del periodo año.** Volcano plots para los DEGS en el periodo de un año, comparando respondedores y no respondedores según mPERR (arriba) y uPCR (abajo). En el eje X se representa el cambio en la expresión génica medido en logFC y en el eje Y el nivel de significancia estadística de ese cambio (-log<sub>10</sub> del p-valor). En azul se representan los genes con sobreexpresión significativa, mientras que en rojo se representan genes con subexpresión significativa. Los puntos en gris corresponden a genes sin cambios significativos en su expresión.

En los volcano plots del periodo de año desde la biopsia, se observa una mayor cantidad de genes diferencialmente expresados en comparación con el periodo temprano. El aumento en el número de DEGs implica mayores diferencias entre ambos grupos. Es decir, respondedores y no respondedores son más diferentes al año post tratamiento.

La mayor cantidad de genes significativos para mPERR sugiere una vez más que esta respuesta es más sensible y captura un mayor número de variaciones en la expresión génica. Esto permite una mejor diferenciación entre grupos clínicos, mostrando una respuesta más homogénea y facilitando la identificación de genes clave asociados con la respuesta o la falta de respuesta al tratamiento.

#### 4.2.3. Periodo de seguimiento: a partir de 19 meses desde la biopsia.



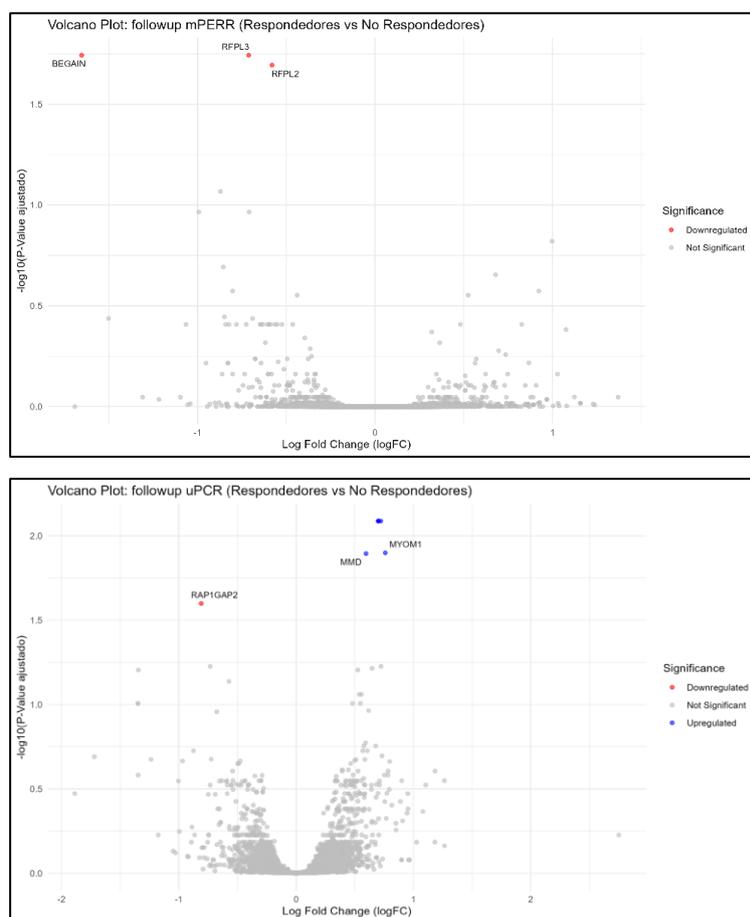
**Figura 5. Heatmap de DEGS para muestras del periodo de seguimiento.** Heatmap de expresión génica en el periodo de seguimiento, normalizado por z-scores por fila. Los colores indican niveles de expresión, con tonos azules para mayor expresión y tonos rojos para menor expresión. Las columnas representan las muestras que están etiquetadas según su respuesta uPCR y/o mPERR o la no respuesta

En el periodo de seguimiento se seleccionaron 64 muestras. Se identificaron 16 genes diferencialmente expresados entre los dos grupos para la respuesta uPCR y 3 para la respuesta mPERR. No se encontraron genes comunes a ambas.

Para este periodo las muestras de las distintas respuestas se entremezclan más. Este patrón refleja que los perfiles de expresión génica no están tan claramente segmentados, lo que podría indicar una respuesta más heterogénea o una menor diferencia en la expresión génica entre los grupos clínicos en este periodo de seguimiento.

Además, la baja cantidad de DEGs en mPERR sugiere que este criterio de respuesta es menos sensible en este periodo.

**Genes “up” y “down-regulated”**



**Figura 6. Volcano plots para los genes del periodo de seguimiento.** Volcano plots que muestran genes diferencialmente expresados en el periodo de seguimiento, comparando respondedores y no respondedores según mPERR (arriba) y uPCR (abajo). En el eje X se representa el cambio en la expresión génica medido en logFC y en el eje Y el nivel de significancia estadística de ese cambio (-log<sub>10</sub> del p-valor). En azul se representan los genes con sobreexpresión significativa, mientras que en rojo se representan genes con subexpresión significativa. Los puntos en gris corresponden a genes sin cambios significativos en su expresión.

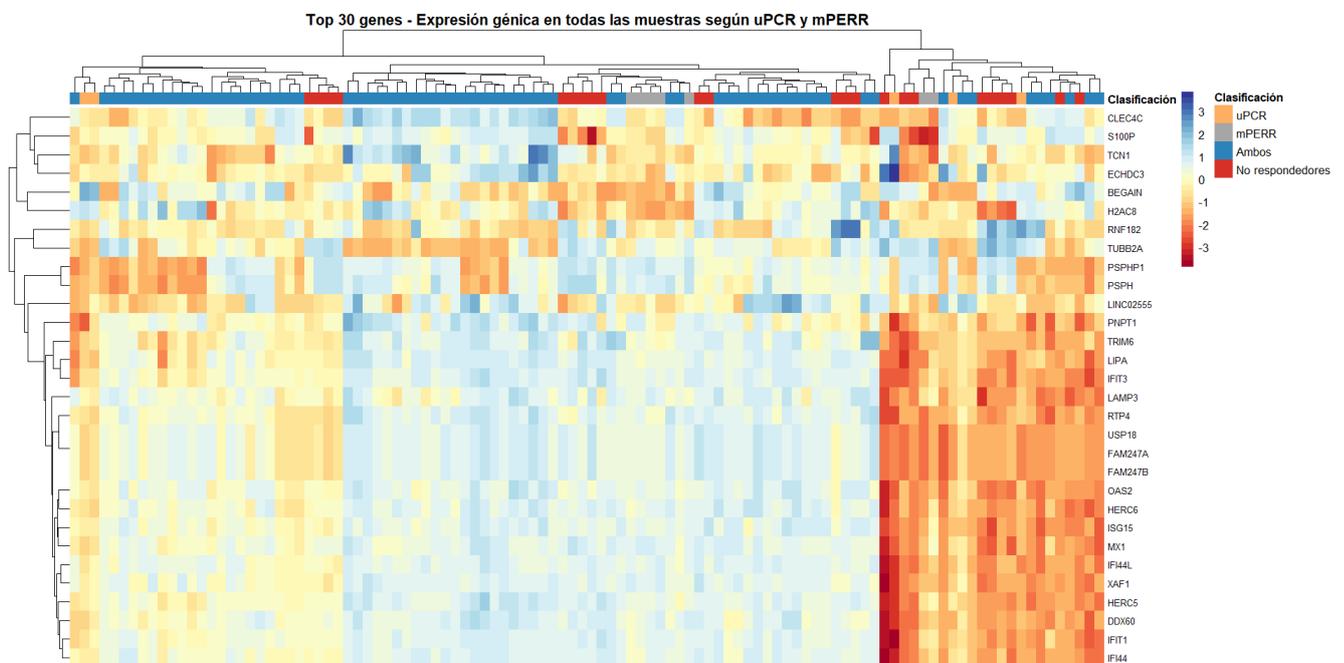
En los volcano plots del periodo de seguimiento, se observa una cantidad muy limitada de genes diferencialmente expresados para ambos criterios de respuesta. En el gráfico

superior de la figura 6, correspondiente a mPERR, sólo se identifican unos pocos genes significativamente sobreexpresados en los no respondedores, mientras que no se observan genes significativos en los respondedores, lo que indica una menor sensibilidad de mPERR en este periodo para detectar diferencias en la expresión génica.

En el gráfico inferior de la figura 6, correspondiente a uPCR, también se detecta un número bajo de genes diferencialmente expresados, con algunos genes sobreexpresados en ambos grupos. Esto sugiere que uPCR capta algo más de variabilidad, aunque la capacidad de ambos criterios para detectar diferencias claras sigue siendo limitada en este periodo.

La escasa cantidad de genes significativos y la similitud en los perfiles de expresión entre respondedores y no respondedores se deben a la heterogeneidad de las muestras en este periodo de seguimiento, que abarca desde 20 meses hasta 10 años. Aunque sería ideal subdividir más los periodos, el número de muestras disponible afectaría a la robustez del análisis. Como resultado, las diferencias en la expresión génica entre los grupos clínicos son menos pronunciadas

#### 4.2.4. Análisis para todas las muestras.

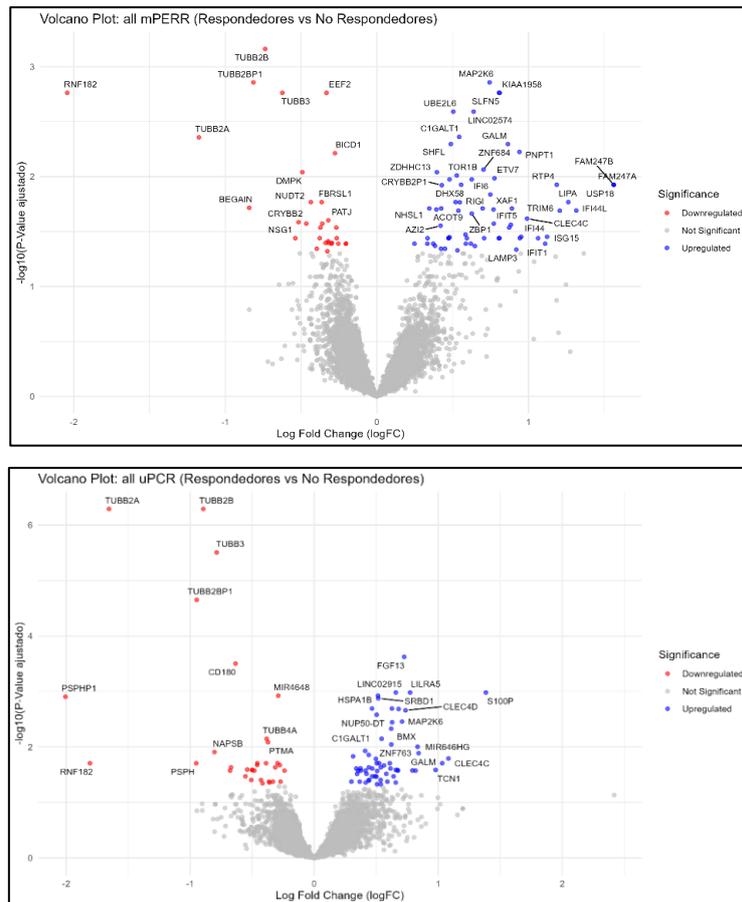


**Figura 7. Heatmap de DEGS para todas las muestras.** Heatmap de los 30 genes principales de los DEGS en todas las muestras, normalizado por z-scores por fila. Los colores indican niveles de expresión, con tonos azules para mayor expresión y tonos rojos para menor expresión. Las columnas representan las muestras que están etiquetadas según su respuesta uPCR y/o mPERR o la no respuesta

Para este análisis se tomaron todas las muestras disponibles: 106. Se realizó el mismo análisis sin incluir covariables e incluyéndose. Sin covariables se obtuvieron: 51 genes diferencialmente expresados para mPERR y 64 para uPCR. Incluyendo las covariables definidas se obtuvieron más genes diferencialmente expresados: 94 para uPCR y 95 para mPERR, lo que sugiere que incluir covariables en el análisis permite detectar con mayor precisión DEGs asociados a la respuesta al tratamiento ya que se indica al modelo que no tenga en cuenta el efecto en la expresión de dichas variables. La similitud en el número encontrado de DEGs para uPCR y mPERR en el análisis general sugiere que ambos criterios capturan, de manera comparable, la variabilidad en la expresión génica relevante para la respuesta clínica, aunque no se detectan los mismos genes.

El heatmap muestra los 30 genes más diferencialmente expresados, resaltando cómo las muestras tienden a agruparse según los perfiles de expresión asociados a cada grupo clínico. En particular, las muestras clasificadas como “Respondedoras a ambas” y “No respondedores” presentan patrones de expresión característicos que las agrupan en clústeres específicos, aunque persiste cierta intercalación de muestras de distintos grupos.

**Genes “up” y “down-regulated”**

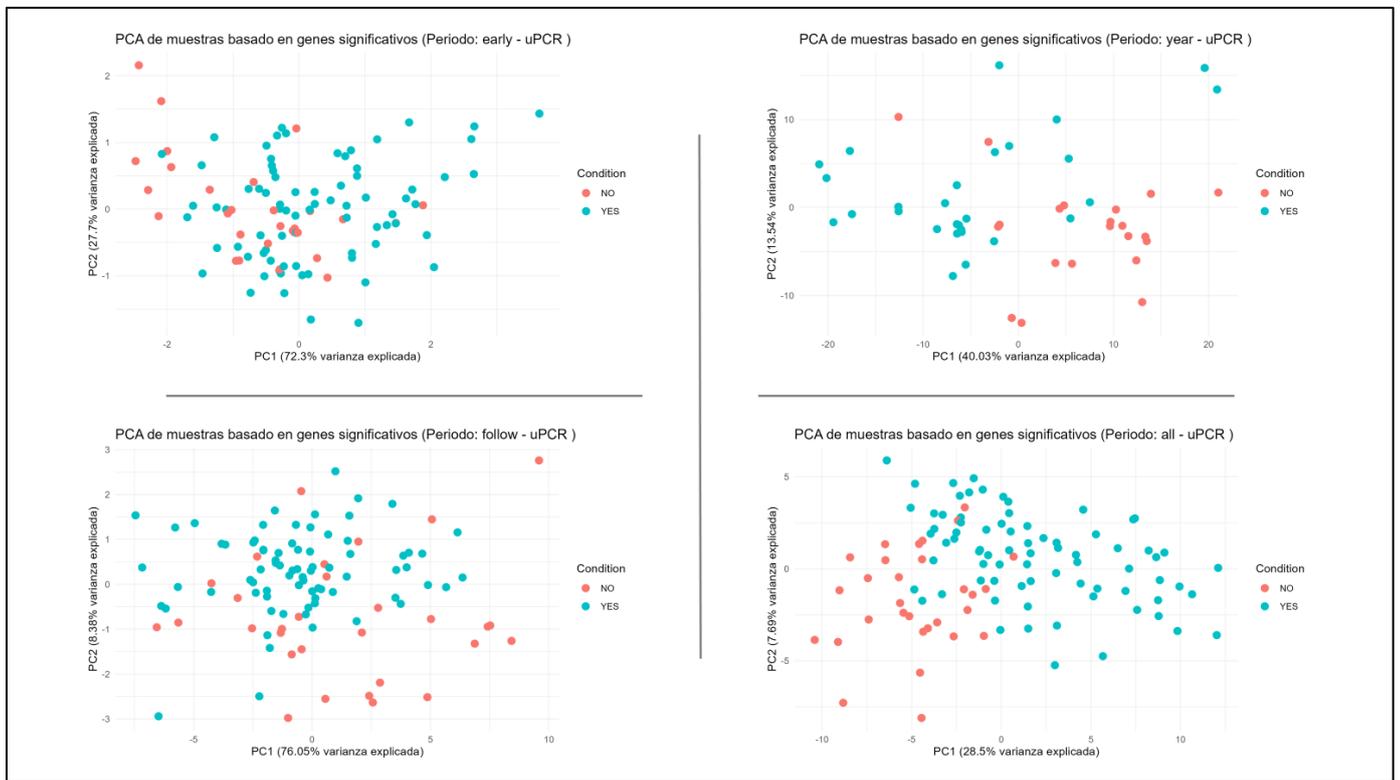


**Figura 8. Volcano plots para los DEGS de todas las muestras.** Volcano plot para los DEGS de todas las muestras, comparando respondedores y no respondedores según mPERR (arriba) y uPCR (abajo). En el eje X se representa el cambio en la expresión génica medido en logFC y en el eje Y el nivel de significancia estadística de ese cambio (-log<sub>10</sub> del p-valor). En azul se representan los genes con sobreexpresión significativa, mientras que en rojo se representan genes con subexpresión significativa. Los puntos en gris corresponden a genes sin cambios significativos en su expresión.

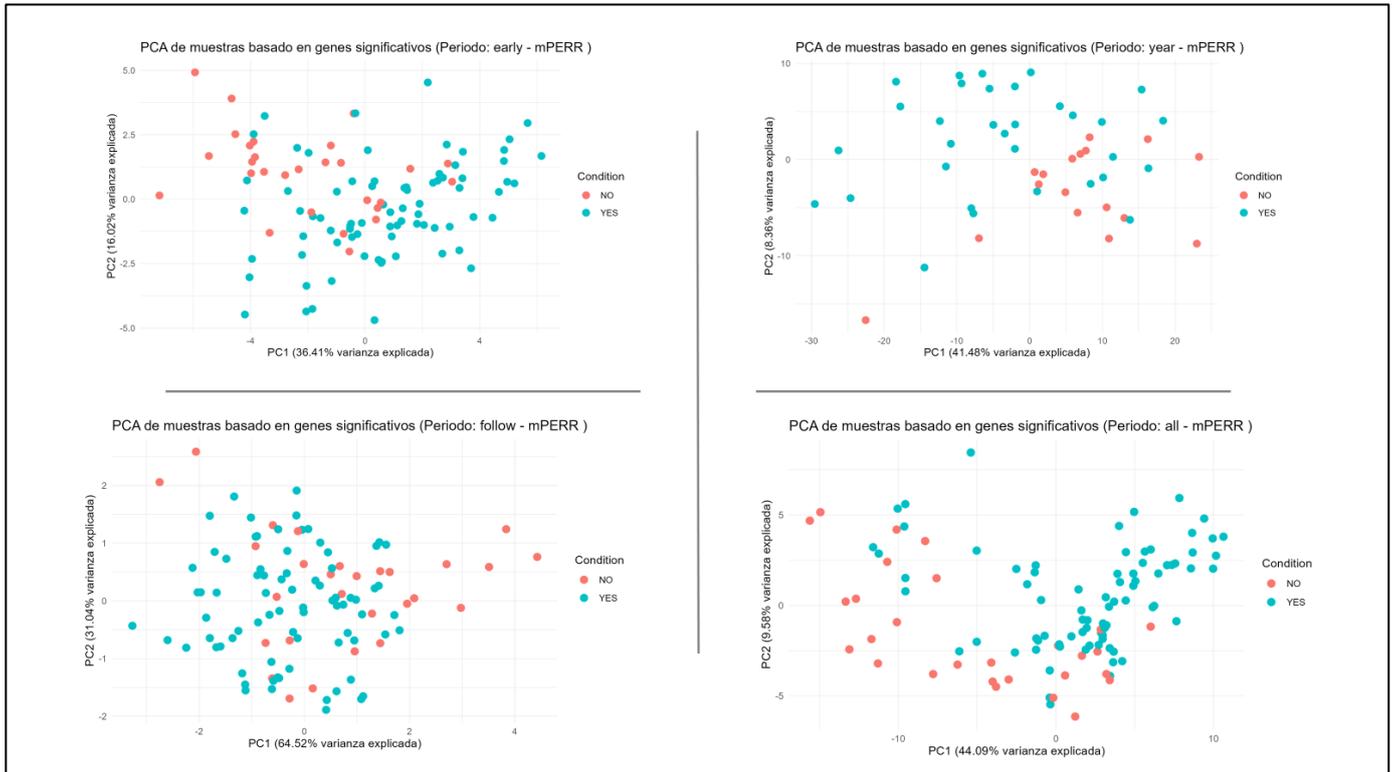
En el análisis que incluye todas las muestras se han obtenido más genes diferencialmente expresados que en los períodos temprano y de seguimiento, sin embargo, tenemos menos que al año, porque abarca una mayor diversidad de tiempos y estados clínicos. Esto permite detectar un número mayor de genes diferencialmente expresados, ya que la inclusión de más muestras refleja una gama más amplia de variabilidad en la expresión génica asociada a la respuesta clínica. Sin embargo, aunque este análisis general ofrece un perfil más completo de la variación genética, no alcanza la especificidad observada en el período de un año, donde las muestras son más homogéneas en términos de tiempo desde la biopsia. Al limitarse a muestras tomadas alrededor del año, el análisis se beneficia de una menor variabilidad temporal.

El volcano plot muestra más cantidad de genes “up-regulated” que “down”.

#### 4.2.5. Análisis de Componentes Principales de muestras (PCA)



**Figura 9. PCA de muestras para la respuesta uPCR en los periodos: temprano, año y seguimiento y todas las muestras.** Resultados del PCA para diferentes periodos en pacientes para la respuesta medida por uPCR, clasificados con las dos respuestas: no respuesta en rojo y respuesta en azul. Las gráficas incluyen los periodos: temprano (arriba izquierda), año (arriba derecha), seguimiento (abajo izquierda) y el análisis de todas las muestras (abajo derecha). Las proporciones de varianza explicada por cada componente principal (PC1 y PC2) se indican en el eje x e y respectivamente.



**Figura 10. PCA de muestras para la respuesta mPERR en los periodos: temprano, año y seguimiento y todas las muestras.** Resultados del PCA para diferentes periodos de seguimiento en pacientes para la respuesta medida por mPERR, clasificados con las dos respuestas: no respuesta en rojo y respuesta en azul. Las gráficas incluyen los periodos: temprano (arriba izquierda), año (arriba derecha), seguimiento (abajo izquierda) y el análisis de todas las muestras (abajo derecha). Las proporciones de varianza explicada por cada componente principal (PC1 y PC2) se indican en el eje x e y respectivamente.

En estos gráficos de PCA, cada punto representa una muestra; en azul las que pertenecen a pacientes respondedores y las que pertenecen a no respondedores en rojo. La distribución de las muestras en cada periodo refleja la similitud en los patrones de expresión génica entre ellas; es decir, las muestras que se agrupan más tienen una expresión génica más parecida.

El porcentaje de variación explicado en cada gráfico indica la proporción de la variabilidad en la expresión génica capturada por los genes significativos en relación con cada periodo y respuesta (mPERR o uPCR).

En el periodo temprano, los gráficos de PCA capturan una variabilidad considerable en la expresión génica: 52.43% para mPERR y 72.3% para uPCR, aunque no hay una separación clara entre respondedores y no respondedores. Esto indica que en esta fase los patrones de expresión génica no son suficientemente distintivos entre los grupos. La mayor cantidad de DEGs en mPERR, observada en los heatmaps y volcano plots, sugiere una mayor sensibilidad de mPERR, aunque sin una clara diferenciación en el PCA.

En el periodo de un año desde la biopsia, los componentes principales capturan una variabilidad significativa de 55.34% en mPERR y 53.57% en uPCR, y comienza a apreciarse una mayor diferenciación entre muestras. La mayor dispersión entre muestras indica una separación más definida entre respondedores y no respondedores, reforzada por los volcano plots, que muestran una cantidad notable de genes diferencialmente expresados.

Para el periodo de seguimiento, los PCA muestran una alta variabilidad significativa: 95.56% en mPERR y 84.98% en uPCR, pero la separación entre respondedores y no respondedores es tenue de nuevo. Esto indica que los perfiles de expresión génica en esta fase de seguimiento son menos específicos, probablemente debido a la heterogeneidad de las muestras (intervalo de 20 meses a 10 años desde la biopsia), que diluye las diferencias entre los grupos clínicos.

En los PCA de todas las muestras, los componentes principales capturan el 53.67% de la variabilidad en mPERR y el 78.98% en uPCR. La combinación de periodos incrementa el número de DEGs y refleja una mayor diversidad de variabilidad genética. Sin embargo, esta inclusión de múltiples periodos también introduce variabilidad temporal, lo que disminuye la especificidad observada en el análisis del periodo de un año, donde los patrones de respuesta eran más definidos.

#### 4.2.6. Caracterización de Perfiles moleculares

Como se recoge en los gráficos previamente comentados es en el periodo de entre 6 y 19 meses desde la biopsia (periodo año) donde se observan las mayores diferencias en la expresión con respecto a la respuesta al tratamiento, por tanto, los genes implicados en dicha respuesta podrían ser objeto de estudio como posibles biomarcadores.

En este periodo y para el análisis realizado incluyendo todas las muestras, el gen **CLEC4C** está sobrerrepresentado en respondedores para ambos criterios de respuestas. Este gen se ha descrito en otros estudios también por estar diferencialmente expresado en el análisis de respuesta al tratamiento en NL (18). CLEC4C codifica un receptor en las células dendríticas plasmocitoides (pDC), que desempeña funciones esenciales en la regulación de la producción de interferón tipo I.(54)

En las enfermedades autoinmunes como el LES, se observa una mayor producción de interferón tipo I, lo que conduce a una activación continua de células inmunes y a la producción de autoanticuerpos, que generan el daño en tejidos y la inflamación.(55). El interferón de tipo I, particularmente IFN- $\alpha$ , es clave en el desarrollo del LES. Es producido principalmente por pDCs y actúa activando la vía JAK-STAT, lo que se traduce en un aumento en la expresión de genes estimulados por IFN (ISGs) relacionado con una exacerbación de la enfermedad. Además, facilita la activación de células B autorreactivas, el cambio de clase de anticuerpos, la diferenciación en células plasmáticas y la sobreexpresión de BAFF, un factor de supervivencia para células B autorreactivas. Todo este proceso contribuye a la desregulación tanto de la inmunidad innata como adaptativa, y algunos estudios sugieren que la inhibición de T1IFN y BAFF podría ser efectiva para reducir la progresión de la enfermedad(56)

La expresión diferencial de CLEC4C en respondedores debido a su acción reguladora sobre este proceso, puede indicar que este gen podría estar contribuyendo a reducir la inflamación y ser indicativo de una mejor respuesta al tratamiento. (54)

En respondedores también se obtuvieron los genes **S100A9** y **S100P** expresados para ambos criterios de respuesta y en el caso de S100P también se obtuvo sobrerrepresentado en respondedores en el análisis de todas las muestras. S100A9 es una proteína proinflamatoria secretada principalmente por neutrófilos que se ha

vinculado con la inflamación activa y daño renal en LES, especialmente en pacientes con anti-dsDNA, por lo que sus niveles podrían ser un buen indicador respuesta al tratamiento (57). **S100P**, por otro lado, muestra correlación con la nefritis lúpica proliferativa, que está asociada a un pronóstico menos favorable de la enfermedad, lo que refuerza su potencial como marcador de actividad renal y su relación con la efectividad de la terapia inmunosupresora.(58).

Los genes **IFIT**: IFIT1, IFIT3, IFIT5 presentan sobreexpresión también en respondedores en el análisis realizado en todas las muestras. Los genes IFIT, inducidos por interferón tipo I, regulan la respuesta inmune y son especialmente relevantes en enfermedades autoinmunes como LES. IFIT1 se ha descrito como un marcador importante en LES, cuya sobreexpresión está relacionada con la actividad de la enfermedad, infiltración de células inmunes, daño renal y cutáneo, lo que sugiere que IFIT1 podría ser un biomarcador útil para el diagnóstico y monitorización, ayudando a evaluar manifestaciones graves, como las renales y cutáneas.(59)

En no respondedores a MMF observamos una sobre regulación del gen **CD180** para ambos criterios de respuesta tanto para el periodo de año desde la biopsia como para el análisis de todas las muestras. CD180, es una molécula asociada a los receptores tipo Toll (TLR) que se expresa principalmente en células inmunes como macrófagos, células dendríticas (DCs) y linfocitos B. Su función principal es regular la activación de estas células, lo cual es relevante en el contexto de enfermedades autoinmunes como el LES. CD180 actúa como regulador de las vías TLR7 y TLR9 (60). Ambos son receptores que detectan fragmentos de ácidos nucleicos, ARN en el caso de TLR7 y ADN no metilado en el caso de TLR9, y su activación en las pDCs estimula la producción de interferón tipo I, que como hemos visto, su sobreproducción en LES desencadena una respuesta autoinmune y la producción de autoanticuerpos, lo que lleva a daño inflamatorio en los tejidos (61). Se ha demostrado que la activación de CD180 inhibe la señalización de TLR7 y TLR9 (60)

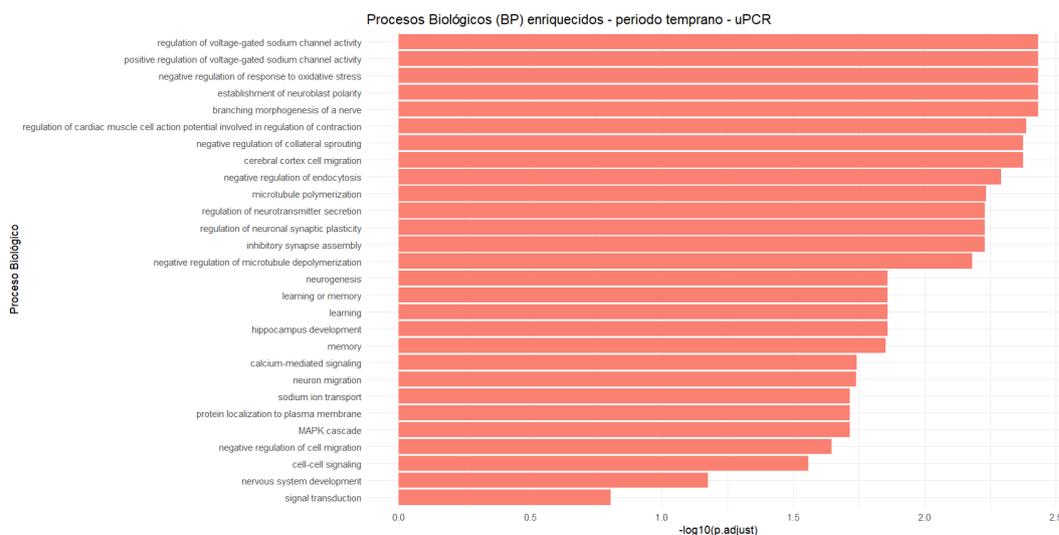
En los no respondedores, y de forma consistente en todos los periodos y en el análisis de todas las muestras para ambos criterios de respuesta, se observa la sobreexpresión de los genes relacionados con **tubulinas**: TUBB2A, TUBB2B, TUBB3, TUBB4. Estos genes

codifican proteínas homólogas fundamentales para la construcción de microtúbulos, los cuales son esenciales en múltiples procesos celulares (62). Sin embargo, no se ha encontrado una relación directa entre la expresión de estos genes y la actividad autoinmune o la respuesta al tratamiento.

### 4.3. Enriquecimiento funcional

En el análisis de los procesos biológicos enriquecidos con los genes obtenidos tras la expresión diferencial, se identificaron varios procesos relacionados con la NL, vinculados a la respuesta inmune y líneas celulares implicadas. Los gráficos ilustran los principales procesos involucrados, diferenciados para las respuestas uPCR y mPERR.

#### Periodo temprano:



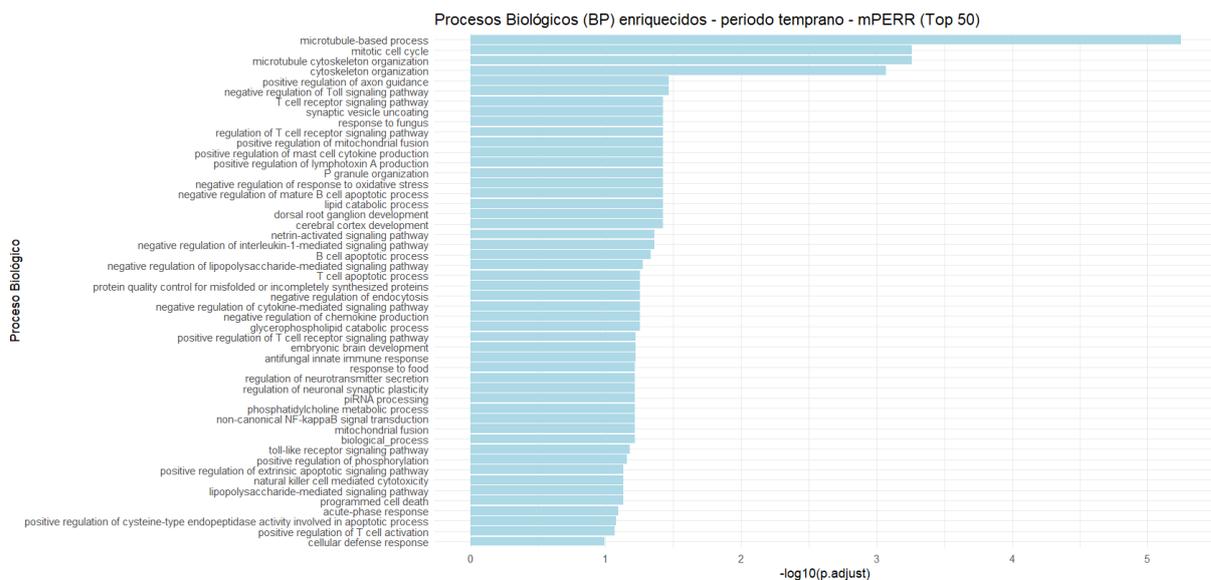
**Figura 11. Procesos biológicos enriquecidos en los DEGs del periodo temprano para uPCR.** Gráfico de barras que muestra los 50 procesos biológicos más enriquecidos durante el periodo temprano con la respuesta medida por uPCR. Se representan los procesos biológicos ordenados de mayor relevancia a menos por su significancia, medida con  $-\log_{10}(p\text{-valor ajustado})$ .

Para el periodo temprano y con la respuesta a uPCR, se observa un enriquecimiento funcional en procesos relacionados con la actividad de canales de sodio activados por voltaje, la regulación negativa de la respuesta al estrés oxidativo y en menor medida la cascada MAPK. Los genes FGF13 y MCTP1, ambos up-regulated, están involucrados.

Los canales de sodio dependientes de voltaje (VGSC) se encuentran en linfocitos y macrófagos. La entrada de sodio a través de estos canales impulsa la activación del intercambiador sodio-calcio. Este aumento en el calcio intracelular es crucial para activar factores de transcripción necesarios para la función de los linfocitos T permitiendo su proliferación y diferenciación. Los bloqueadores de VGSC han demostrado tener efectos inmunomoduladores de los que se pueden beneficiar las enfermedades autoinmunes, pero presentan muchos efectos adversos descritos.(63)

El estrés oxidativo, ocurre cuando hay un desequilibrio entre la producción de especies reactivas de oxígeno y la capacidad del cuerpo para neutralizarlas. La disfunción mitocondrial en los linfocitos T de pacientes con LES resulta en una producción excesiva de especies reactivas, favoreciendo la modificación de antígenos propios y desencadenando la autoinmunidad.(64)

Por otro lado, la cascada de quinasas activadas por mitógenos (MAPK), y específicamente la vía del factor regulado por señales extracelulares (ERK), desempeña un papel significativo en la fisiopatología del LES. Esta vía de señalización está implicada en la activación y proliferación de células T, lo que es crucial para la respuesta inmune.(65)



**Figura 12. Procesos biológicos enriquecidos en los DEGs del periodo temprano para mPERR.** Gráfico de barras que muestra los 50 procesos biológicos más enriquecidos durante el periodo temprano con la respuesta medida por mPERR. Se representan los procesos biológicos ordenados de mayor relevancia a menos por su significancia, medida con  $-\log_{10}(p\text{-valor ajustado})$ .

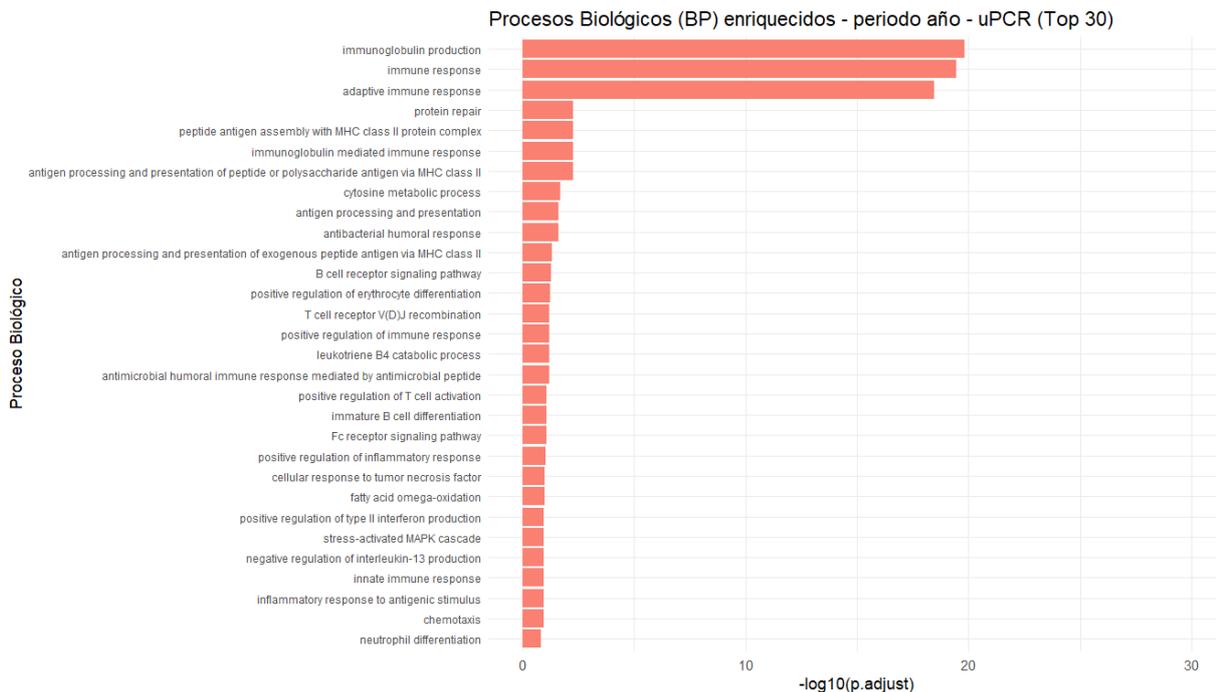
Las funciones más enriquecidas para la respuesta mPERR son las asociadas con la organización de microtúbulos y la estructura del citoesqueleto. Las tubulinas que están sobrerrepresentadas en no respondedores desempeñan dichas funciones fundamentalmente. No se ha descrito una relación directa con la fisiopatología del LES u otras enfermedades autoinmunes similares. (62)

Se observa también un enriquecimiento, en menor medida, en la señalización negativa de receptores Toll . Como se ha explicado anteriormente TLR7 y TLR9, desempeñan un papel crucial en la detección de fragmentos de ácidos nucleicos, lo que desencadena la producción de citocinas inflamatorias, incluidas las que promueven la producción de interferón tipo I que pueden contribuir a la inflamación crónica en NL (61). El gen SIGIRR (down-regulated) está vinculados a este proceso.

En esta etapa también se encuentra enriquecidos los procesos de regulación negativa del estrés oxidativo, en la que también se involucra el gen sobrerrepresentado en respondedores: MTC1.

Otras funciones enriquecidas relacionadas con la respuesta inmune: la vía de señalización del receptor de linfocitos T o la regulación negativa del proceso apoptótico de los linfocitos B en las que están involucradas los genes: BCL10 y BTNL3, sobrerrepresentados en respondedores.

Periodo año:



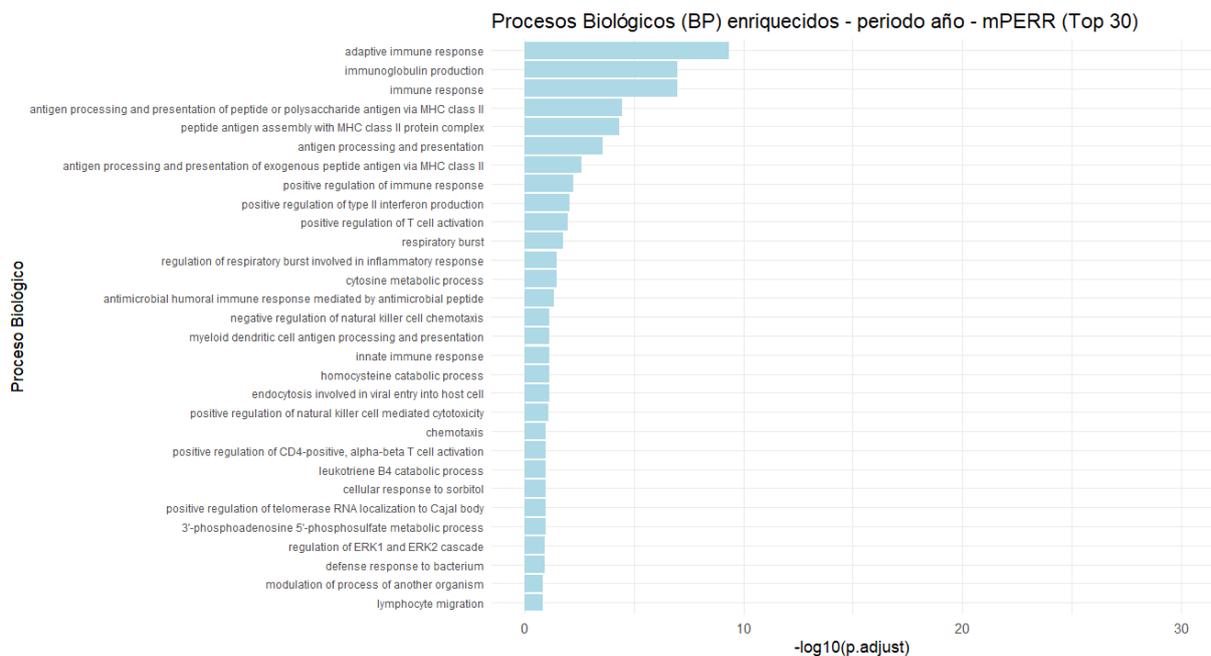
**Figura 13. Procesos biológicos enriquecidos en los DEGs del periodo año para uPCR.** Gráfico de barras que muestra los 30 procesos biológicos más enriquecidos durante el periodo de año desde la biopsia con la respuesta medida por uPCR. Se representan los procesos biológicos ordenados de mayor relevancia a menos por su significancia, medida con  $-\log_{10}(p\text{-valor ajustado})$ .

En el análisis de la respuesta uPCR, se observa un enriquecimiento en la respuesta inmune adaptativa y la respuesta inmune mediada por inmunoglobulinas. En la producción de inmunoglobulinas están implicados genes de la familia de las inmunoglobulinas: IGK, IGL y en la respuesta inmune adaptativa: CLEC4C, genes HLA, IGK, IGL fundamentalmente, la mayor parte de ellos sobrerrepresentados en no respondedores, aunque parece que esta vía está activada en ambos casos.

La respuesta inmune adaptativa es fundamental para la patogénesis del LES. En esta enfermedad dicha respuesta está desregulada, las células T desempeñan un papel importante en este proceso. Además de mostrar una secreción anormal de citoquinas también puede conducir a una activación desregulada de linfocitos B y células dendríticas (Dcs). Los linfocitos B producen autoanticuerpos, presentan autoantígenos y activan células T autorreactivas. Estas interacciones entre linfocitos B y T fomentan la autoinmunidad, llevando a la producción excesiva de autoanticuerpos. Las anomalías en esta respuesta han impulsado el desarrollo de nuevos agentes biológicos para el tratamiento del LES.(66)

La producción de inmunoglobulinas es otra función clave enriquecida en este análisis, con la participación de genes de la familia de las inmunoglobulinas: IGK, IGL. Los linfocitos B desempeñan un papel crucial no solo en la producción de autoanticuerpos (inmunoglobulinas), como los dirigidos contra ADN propio, sino también como células presentadoras de antígenos que colaboran con linfocitos T en la generación de la respuesta autoinmune. Por esta razón, la supresión de la función de los linfocitos B ha sido considerada una estrategia terapéutica potencial en el tratamiento del LES. (67)

El procesamiento y presentación de antígenos por MHC de clase II aparece como una función enriquecida en la que intervienen los genes de la familia HLA (HLA-DRB3, HLA-DPA1, HLA-DRB5, HLA-DQB2). El complejo mayor de histocompatibilidad (MHC) es un conjunto de proteínas en la superficie de las células que presentan fragmentos de antígenos a los linfocitos T CD4+ y que, por tanto, participan en su activación. Este proceso se realiza a través de las células presentadoras de antígenos (APCs) que en las enfermedades autoinmunes como el LES están alteradas: procesan péptidos de origen propio (autólogos) lo que supone un aumento en la presentación de fragmentos de proteínas del propio organismo a los linfocitos T. Esto se traduce en una respuesta inmune inapropiada y la producción de autoanticuerpos, contribuyendo así a la patogenia del LES.(68)



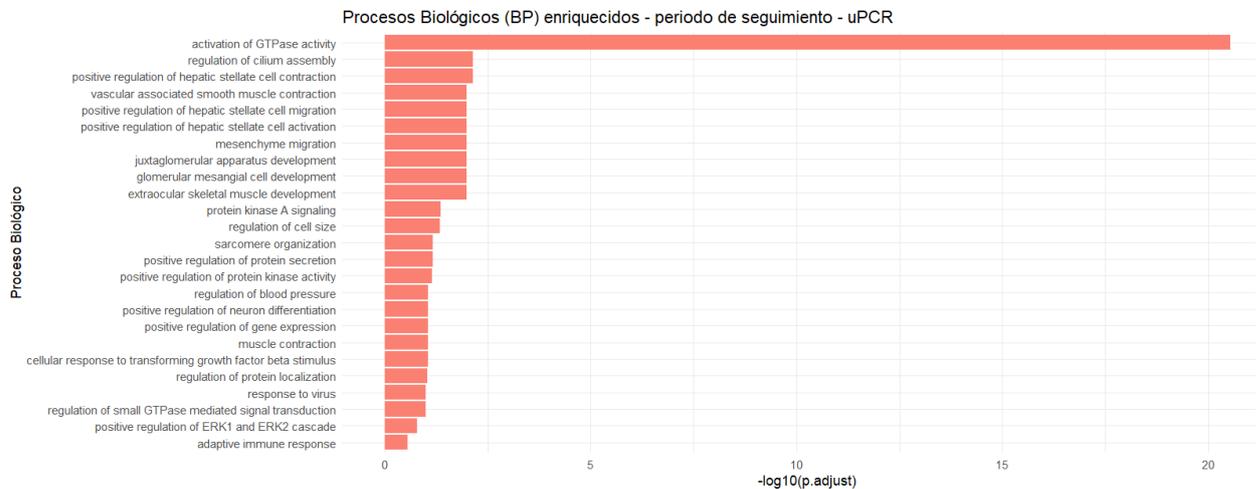
**Figura 14. Procesos biológicos enriquecidos en los DEGs del periodo año para mPERR.** Gráfico de barras que muestra los 30 procesos biológicos más enriquecidos durante el periodo de año desde la biopsia con la respuesta medida por mPERR. Se representan los procesos biológicos ordenados de mayor relevancia a menos por su significancia, medida con  $-\log_{10}(p\text{-valor ajustado})$ .

En el análisis para mPERR de nuevo el proceso más enriquecido fue la respuesta inmune adaptativa, en la que están implicados numerosos genes: genes de la familia, inmunoglobulinas: (IGK-), (IGV-), TNFRSF17 o CLEC4C en su mayoría genes sobrerrepresentados en no respondedores, pero también en respondedores, al igual que en la respuesta medida por uPCR.

Otro de los procesos más enriquecidos como en la respuesta uPCR es la presentación de antígenos, en el cual están involucrados de nuevo genes de la familia HLA, que también están sobreexpresados en no respondedores y la producción de inmunoglobulinas (IGKC, IGKV y IGLV). Esta sobreexpresión en ambos procesos sugiere una sobre activación de la respuesta inmune sobre todo para el grupo de no respondedores.

La vía del interferón II es otro de los procesos enriquecidos para esta respuesta con la participación de genes tanto sobre regulados en respondedores como en no respondedores. El interferón tipo II o IFN- $\gamma$  promueve la activación de linfocitos B autorreactivos y por tanto la producción de autoanticuerpos. (69)

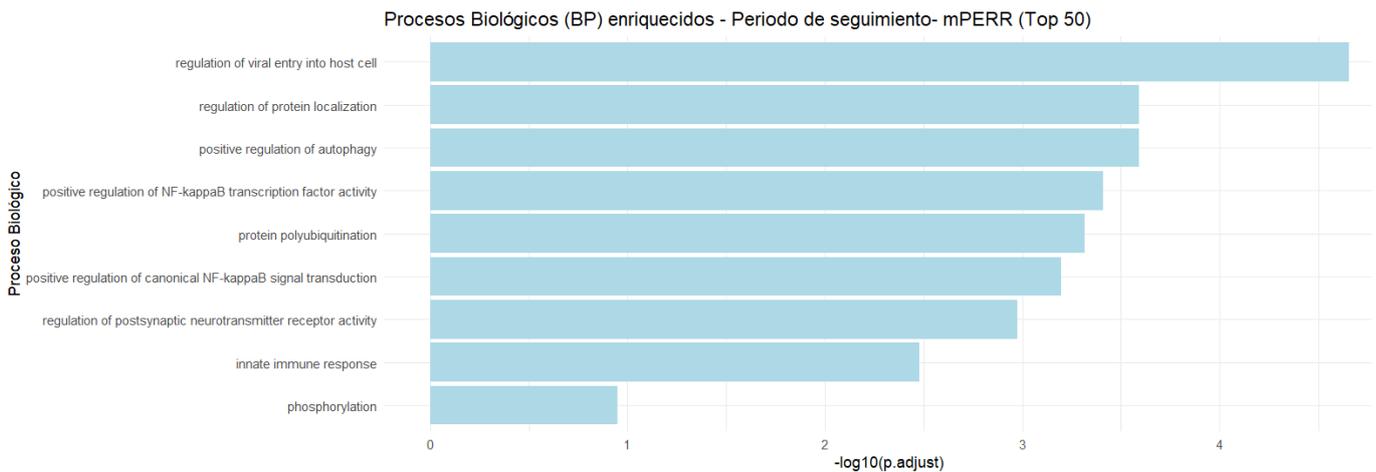
**Periodo de seguimiento:**



**Figura 15. Procesos biológicos enriquecidos en los DEGs de periodo de seguimiento para uPCR.** Gráfico de barras que muestra los procesos biológicos enriquecidos durante el periodo de seguimiento con la respuesta medida por uPCR. Se representan los procesos biológicos ordenados de mayor relevancia a menos por su significancia, medida con  $-\log_{10}(p\text{-valor ajustado})$ .

En el periodo de seguimiento, se observa un notable enriquecimiento funcional en la activación de GTPasa asociada con la familia de genes TBC1D3, sobre regulados en respondedores. Las GTPasas, especialmente RhoA, son cruciales en la patogenia del LES ya que regulan funciones como la adhesión celular y la migración. La activación excesiva de la vía RhoA/ROCK se relaciona con una disfunción en las células T y una respuesta inmune desregulada. Además, RhoA está vinculado a la señalización del interferón tipo I. (70)

Otra función enriquecida es la regulación positiva de la cascada de señalización de ERK1 y ERK2 parte de la ya mencionada cascada MAPK, que está relacionada con el gen ACTA2, sobre regulado en respondedores. Este gen también está implicado en otros procesos enriquecidos como la contracción muscular y la regulación de la presión arterial que a priori no tienen una relación directa con la autoinmunidad o la respuesta al tratamiento. Adicionalmente, la respuesta inmune adaptativa se destaca como una función enriquecida también en este periodo, relacionada principalmente con el gen RAP1GAP2, sobre expresado en no respondedores.



**Figura 16. Procesos biológicos enriquecidos en los DEGs de periodo de seguimiento para mPERR.** Gráfico de barras que muestra los procesos biológicos enriquecidos durante el periodo de seguimiento con la respuesta medida por mPERR. Se representan los procesos biológicos ordenados de mayor relevancia a menos por su significancia, medida con  $-\log_{10}(p\text{-valor ajustado})$ .

En el análisis para mPERR observamos entre los procesos más enriquecidos algunos que aparentemente no tendrían relación con la NL: como la regulación de la entrada de virus a células del hospedador.

En esta etapa apenas había DEGs y todos corresponden al grupo de no respondedores. Entre las funciones enriquecidas que podrían tener relación con la enfermedad tenemos: la regulación de la autofagia, la localización de proteínas, regulación positiva de la actividad del factor de transcripción NF-kappaB o la respuesta inmune innata, todas ellas representadas por los genes RFPL2 y 3.

El factor de transcripción nuclear kappa B (NF-κB) es fundamental en la regulación de la respuesta inmune y la inflamación, ya que activa genes determinantes en la supervivencia y proliferación celular. En el LES se ha observado que la activación anormal de NF-κB contribuye a la disfunción de las células T y B, así como a la producción de autoanticuerpos. (71)

## 5. Discusión

Este estudio ha puesto de manifiesto la heterogeneidad clínica y molecular de la enfermedad del LES, en concreto de su complicación en NL y, por tanto, en la respuesta al tratamiento estándar evaluado: MMF. Mediante el análisis de perfiles de expresión génica diferencial y vías biológicas específicas en 106 muestras de 37 pacientes, definidas en periodos temporales por el tiempo desde la biopsia, se identificaron distintos patrones en la respuesta al tratamiento evaluados con los criterios de respuesta definidos: mPERR y uPCR.

El análisis de variables clínicas y demográficas mostró asociaciones significativas entre la fisiopatología de la NL y la respuesta al MMF. En cuanto a las variables demográficas, se confirmó que los pacientes afroamericanos, además de tener una mayor incidencia y severidad de la enfermedad, responden menos al tratamiento con MMF. La edad de inicio de la enfermedad fue similar en respondedores y no respondedores (alrededor de 25 años), y aunque el LES afecta principalmente a mujeres, no se observaron diferencias relevantes de respuesta por sexo, limitadas por la baja representación masculina en la muestra.

En términos clínicos, los índices de actividad de la enfermedad (PGA y SLEDAI) fueron más altos en los no respondedores, indicando una mayor actividad de la enfermedad en ausencia de respuesta. Los no respondedores también presentaron niveles elevados de complemento (C3 y C4), sugiriendo mayor activación inmunológica, sin embargo, los niveles de anticuerpos anti-dsDNA fueron mayores en los respondedores. Además, los pacientes no respondedores requirieron dosis más altas de MMF y prednisona, reflejando la necesidad de un tratamiento intensivo en casos de menor respuesta.

En cuanto al análisis de DEGs, el periodo de un año post-biopsia resultó ser el más informativo, con 273 DEGs asociados a uPCR y 381 a mPERR. Este alto número de DEGs en el periodo de un año indica que las diferencias moleculares entre respondedores y no respondedores se vuelven más evidentes después de un tiempo suficiente de tratamiento. En comparación, el periodo temprano mostró menos DEGs, probablemente debido a que el rango de tiempo es mayor lo que hace que las muestras sean más heterogéneas. Acotar más el periodo sin embargo no era una opción viable

debido al bajo número de muestras: 22. En el periodo de seguimiento, la cantidad de DEGs también fue reducida, lo cual podría explicarse de nuevo por la heterogeneidad temporal de las muestras (que abarcan desde 19 meses hasta 10 años post-biopsia), dificultando una clara segmentación de los pacientes en función de la respuesta y haciendo evidente que no hay una marca clara detrás de la no respuesta o respuesta al tratamiento a lo largo de los años. Al comparar los periodos, mPERR resultó ser el criterio que captura una mayor variabilidad en la expresión génica.

La caracterización de perfiles moleculares tras el análisis de DEGs permitió identificar genes específicos relacionados con la fisiopatología del LES que mostraron diferencias significativas en la respuesta al tratamiento. En los respondedores, destacan CLEC4C, S100A9, S100P y genes de la familia IFIT (IFIT1, IFIT3, IFIT5). La sobreexpresión de CLEC4C, se relaciona con la regulación en la producción de interferón tipo I, por lo que podría tener un papel clave en la actividad de la enfermedad. El interferón I (IFN- $\alpha$ ) es una de las rutas más importantes que explican el mecanismo de la autoinmunidad producida en el LES (56). Se ha demostrado que los pacientes que presentan firmas elevadas de interferón tipo I en estado basal se asocian frecuentemente con una mayor actividad evidenciada por puntuaciones altas en SLEDAI y manifestaciones clínicas más severas como la NL y, además, se ha demostrado que es útil como marcador de NL de clase III/IV, superando a otros marcadores tradicionales como el complemento y los anticuerpos anti-dsDNA.(72)

El gen IFIT1, inducido por interferón tipo I también desempeña un papel importante en la NL especialmente en los casos de clases III y IV donde la expresión de este gen parece estar significativamente aumentada en el glomérulo. Este gen junto a genes proapoptóticos y quimiocinas, promueve el daño en los podocitos, favoreciendo apoptosis e inflamación en el tejido renal.(73). Tanto CLEC4 como IFIT1 podrían ser biomarcadores prometedores y posibles objetivos de nuevos tratamientos.

En los no respondedores, se observó una sobreexpresión significativa de CD180, un gen que regula las vías TLR7 y TLR9, cuya activación estimula la producción de interferón tipo I. La sobreexpresión de CD180 no parece relacionarse con su resistencia al tratamiento ya que su acción reguladora sobre TLR7 y TLR9 inhibe por tanto la producción de

interferón lo que debería disminuir la actividad autoinmune. Modificar la señalización de TLR7 y TLR9 mediante CD180 podría ser una estrategia prometedora para controlar la hiperactivación inmunitaria en pacientes con mayor actividad de la enfermedad o resistentes al tratamiento. Además, en los no respondedores, los genes relacionados con tubulinas como TUBB2A, TUBB2B, TUBB3, o TUBB4 también mostraron sobreexpresión constante en todos los periodos. Las tubulinas son clave para la estructura de los microtúbulos y la funcionalidad de los linfocitos T, esenciales en su activación y capacidad citotóxica. Aunque los estudios sobre tubulinas se han centrado en cáncer, podrían ser una diana terapéutica prometedora en enfermedades autoinmunes como el LES.(74)

Como líneas futuras de este trabajo, los marcadores moleculares identificados en el análisis de DEGs podrían investigarse en cohortes adicionales para confirmar su relación con la respuesta al tratamiento y su impacto en la progresión de la enfermedad. Validar estos genes como biomarcadores ayudaría a mejorar el seguimiento de la enfermedad y permitiría evaluar su viabilidad como posibles dianas terapéuticas.

Marcadores relacionados con el interferón I, ya han demostrado efectividad como dianas terapéuticas en LES. Un ejemplo es el fármaco anifrolumab, que es un anticuerpo monoclonal que inhibe la subunidad 1 del receptor de interferón tipo I. Recientemente ha sido aprobado para tratar el LES de moderado a grave. Este fármaco ha demostrado eficacia en el control de manifestaciones cutáneas en dos ensayos clínicos y ha mostrado buenos resultados en múltiples casos de lupus eritematoso cutáneo (LEC).(75)

En cuanto al análisis de enriquecimiento funcional, en el **periodo temprano**, las vías enriquecidas en respondedores están relacionadas con el estrés oxidativo, la actividad de canales de sodio activados por voltaje y la cascada MAPK (ERK), ambos relacionados con la activación y proliferación de linfocitos. La inhibición de la cascada de MAPK a través de moléculas como MAP4K y DUSP es una estrategia terapéutica novedosa en el tratamiento del LES y en la posible identificación de nuevos biomarcadores (76). Por otro lado, en los no respondedores, destaca la organización de microtúbulos y la señalización de receptores Toll (TLR7 y TLR9), relacionados con la activación en la producción de interferón tipo I y mecanismos inflamatorios, de nuevo esta mayor actividad de las vías

que desencadenan la mayor actividad autoinmune tiene sentido en pacientes que no están respondiendo al tratamiento. Actualmente hay numerosos estudios que posicionan a los TLR como objetivo terapéutico en el LES: uso de antagonistas, inhibidores de proteínas clave de la señalización de TLR, anticuerpos y oligonucleótidos para bloquear esta vía o bien su regulación mediante microRNAs que emerge como una vía prometedora que podría ralentizar la progresión de la enfermedad.(77)

En el **periodo de un año**, el análisis muestra que tanto en respondedores como en no respondedores está enriquecida la respuesta inmune adaptativa y la producción de inmunoglobulinas, vinculadas ambas a la actividad de linfocitos B y T. La mayor parte de los genes implicados en dichas respuestas pertenecen a las familias HLA a IG-K, IG-L. La mayoría sobre expresados en no respondedores.

Además, en este grupo también se observa una mayor activación en la presentación de antígenos (MHC de clase II) en la que, de nuevo, están implicados genes HLA. La sobreexpresión de genes HLA y de inmunoglobulinas sobre todo en no respondedores y la activación de estas vías sugiere que los no respondedores presentan una mayor activación del sistema inmune, lo cual podría explicar que el tratamiento MMF no está surtiendo el efecto esperado.

La vía del interferón II durante este periodo también destaca por estar enriquecida y está relacionada con genes sobrerrepresentados tanto en respondedores como en no respondedores por lo que no hay una asociación clara en la activación de esta vía con la respuesta al tratamiento estudiado. El IFN- $\gamma$  tiene un papel relevante en la autoinmunidad porque activa linfocitos B autorreactivos. Se ha estudiado que la interacción de IFN- $\gamma$  con TLR7 resulta en la activación de factores de transcripción que impulsan la proliferación y diferenciación de linfocitos B, por lo que el bloqueo de ambas podría ser una diana terapéutica interesante. (69)

En el periodo de **seguimiento**, en los respondedores, se observa un enriquecimiento funcional en la activación de GTPasas, en este caso la disfunción de RhoA contribuye a una respuesta inmune desregulada y una disfunción en células T y la activación del interferón I. Inhibir esta vía ha demostrado reducir los anticuerpos anti-dsDNA y mejorar la función de las células T, lo que sugiere que RhoA podría ser un objetivo terapéutico

interesante (70). En no respondedores destaca la activación del factor NF- $\kappa$ B. Su activación anormal supone una mayor expresión de citoquinas proinflamatorias y moduladores inmunes que contribuyen a la patogénesis de la enfermedad: activación de células T y B autorreactivas, y formación de CIs además, se ha descrito la interacción entre NF- $\kappa$ B también en la exacerbación de la autoinmunidad.

## 6. Conclusiones

- La NL muestra una alta heterogeneidad clínica y molecular y su evolución es compleja; alterna fases de brotes o recaídas y remisiones. Estas diferentes fases suponen una distinta activación del sistema inmune lo que dificulta la agrupación de muestras y por tanto la identificación de biomarcadores y la predicción de respuesta a los tratamientos actuales. Por ello es interesante evaluar a la población subdividida por periodos en cuanto a la caracterización de DEGs.
- El periodo que abarca el año desde la biopsia supone el periodo de mayor diferenciación a nivel molecular entre pacientes. Evaluar la respuesta en este periodo podría proporcionar información relevante de si el tratamiento está surgiendo el efecto deseado e hipotéticamente un cambio de estrategia terapéutica si esto no es así. Esto se alinea con la práctica clínica habitual que realiza este seguimiento desde los 6 meses al año desde el establecimiento de un tratamiento y con esta información, se decide si cambiar el mismo en caso de una respuesta no favorable.
- En cuanto a las limitaciones, en este trabajo se cuenta con un número de muestras reducido; sería interesante poder estratificar más a los pacientes porque los periodos temprano y de seguimiento no han mostrado tantas diferencias génicas debido a su heterogeneidad temporal.
- Los DEGs relacionados con el interferón tipo I y los receptores TLR, como CLEC4C, IFIT1 y CD180, muestran un alto potencial como dianas terapéuticas debido a su influencia en la respuesta inmune y su capacidad para modular la actividad autoinmune en NL. Las vías de señalización de TLR, junto con la respuesta inmune adaptativa y la regulación del interferón tipo I, representan mecanismos

clave en la patogénesis de la NL y constituyen prometedoras dianas para el desarrollo de nuevas estrategias terapéuticas.

- El enriquecimiento de la respuesta inmune adaptativa y la producción de inmunoglobulinas en mayor medida en los no respondedores, junto con la sobreexpresión de genes HLA y la mayor actividad en la presentación de antígenos, sugiere una correlación entre la respuesta no favorable al tratamiento con MMF y una mayor actividad del sistema inmune.

## 7. Anexos

Los scripts que incluyen el código para del exploratorio y limpieza de datos, el análisis de variables, la expresión diferencial y el análisis de enriquecimiento funcional están incluidos en el siguiente enlace:

[TFM 22370217 LauraPalenzuelaBermejo](#)

## 8. Bibliografía

- (1) Accapezzato, D., Caccavale, R., Paroli, M. P., Gioia, C., Nguyen, B. L., Spadea, L., & Paroli, M. (2023). Advances in the Pathogenesis and Treatment of Systemic Lupus Erythematosus. *International Journal of Molecular Sciences*, 24(7), Article 7. <https://doi.org/10.3390/ijms24076578>
- (2) Akhil, A., Bansal, R., Anupam, K., Tandon, A., & Bhatnagar, A. (2023). Systemic lupus erythematosus: Latest insight into etiopathogenesis. *Rheumatology International*, 43(8), 1381-1393. <https://doi.org/10.1007/s00296-023-05346-x>
- (3) Gasparotto, M., Gatto, M., Binda, V., Doria, A., & Moroni, G. (2020). Lupus nephritis: Clinical presentations and outcomes in the 21st century. *Rheumatology*, 59(Supplement\_5), v39-v51. <https://doi.org/10.1093/rheumatology/keaa381>
- (4) Alduraibi, F. K., & Tsokos, G. C. (2024). Lupus nephritis biomarkers: A critical review. *International Journal of Molecular Sciences*, 25(2), 805. <https://doi.org/10.3390/ijms25020805>
- (5) Rojas-Rivera, J. E., García-Carro, C., Ávila, A. I., Espino, M., Espinosa, M., Fernández-Juárez, G., Fulladosa, X., Goicoechea, M., Macía, M., Morales, E., Quintana, L. F., & Praga, M. (2023). Diagnosis and treatment of lupus nephritis: A summary of the Consensus Document of the Spanish Group for the Study of Glomerular Diseases (GLOSEN). *Clinical Kidney Journal*, 16(9), 1384. <https://doi.org/10.1093/ckj/sfad055>
- (6) Alforaih, N., Whittall-Garcia, L., & Touma, Z. (2022). A Review of Lupus Nephritis. *The Journal of Applied Laboratory Medicine*, 7(6), 1450-1467. <https://doi.org/10.1093/jalm/jfac036>
- (7) Blenkinsopp, S. C., Fu, Q., Green, Y., Madan, A., Juliao, P., Goldman, D. W., Roth, D. A., & Petri, M. A. (2022). Renal response at 2 years post biopsy to predict long-term renal

- survival in lupus nephritis: a retrospective analysis of the Hopkins Lupus Cohort. *Lupus Science & Medicine*, 9(1), e000598. <https://doi.org/10.1136/lupus-2021-000598>
- (8) Gladman, D. D., Ibañez, D., & Urowitz, M. B. (s/f). Systemic lupus erythematosus disease activity index 2000. *Jrheum.org*. Recuperado el 18 de octubre de 2024, de <https://www.jrheum.org/content/jrheum/29/2/288.full.pdf>
- (9) Chessa, E., Piga, M., Floris, A., Devilliers, H., Cauli, A., & Arnaud, L. (2020). Use of Physician Global Assessment in systemic lupus erythematosus: a systematic review of its psychometric properties. *Rheumatology (Oxford, England)*, 59(12), 3622-3632. <https://doi.org/10.1093/rheumatology/keaa383>
- (10) Rovin, B. H., Ayoub, I. M., Chan, T. M., Liu, Z.-H., Mejía-Vilet, J. M., & Floege, J. (2024). KDIGO 2024 clinical practice guideline for the management of LUPUS NEPHRITIS. *Kidney International*, 105(1), S1-S69. <https://doi.org/10.1016/j.kint.2023.09.002>
- (11) Touma, Z., Gladman, D. D., Urowitz, M. B., Beyene, J., Uleryk, E. M., & Shah, P. S. (2010). Mycophenolate Mofetil for Induction Treatment of Lupus Nephritis: A Systematic Review and Metaanalysis. *The Journal Of Rheumatology*, 38(1), 69-78. <https://doi.org/10.3899/jrheum.100130>
- (12) Furie, R., Rovin, B. H., Houssiau, F., Malvar, A., Teng, Y. O., Contreras, G., Amoura, Z., Yu, X., Mok, C., Santiago, M. B., Saxena, A., Green, Y., Ji, B., Kleoudis, C., Burriss, S. W., Barnett, C., & Roth, D. A. (2020). Two-Year, Randomized, Controlled Trial of Belimumab in Lupus Nephritis. *New England Journal Of Medicine*, 383(12), 1117-1128. <https://doi.org/10.1056/nejmoa2001180>
- (13) Lin, H., & Chu, P. (2024). Special Issue "Bioinformatics Study in Human Diseases: Integration of Omics Data for Personalized Medicine". *International Journal Of Molecular Sciences*, 25(19), 10579. <https://doi.org/10.3390/ijms251910579>
- (14) Robinson, W. H., & Mao, R. (2015b). Biomarkers to guide clinical therapeutics in rheumatology? *Current Opinion In Rheumatology*, 28(2), 168-175. <https://doi.org/10.1097/bor.0000000000000250>
- (15) Gong, X., Su, L., Huang, J., Liu, J., Wang, Q., Luo, X., Yang, G., & Chi, H. (2024). An overview of multi-omics technologies in rheumatoid arthritis: applications in biomarker and pathway discovery. *Frontiers In Immunology*, 15. <https://doi.org/10.3389/fimmu.2024.1381272>
- (16) Li, S., Song, G., Bai, Y., Song, N., Zhao, J., Liu, J., & Hu, C. (2021). Applications of Protein Microarrays in Biomarker Discovery for Autoimmune Diseases. *Frontiers In Immunology*, 12. <https://doi.org/10.3389/fimmu.2021.645632>
- (17) Wang, L., Yang, Z., Yu, H., Lin, W., Wu, R., Yang, H., & Yang, K. (2022). Predicting diagnostic gene expression profiles associated with immune infiltration in patients with lupus nephritis. *Frontiers In Immunology*, 13. <https://doi.org/10.3389/fimmu.2022.839197>
- (18) Toro-Domínguez, D., López-Domínguez, R., Villatoro-García, J. A., Marañón, C., Goldman, D., Petri, M., Carmona-Sáez, P., & Alarcón-Riquelme, M. (2024). Immune and molecular landscape behind non-response to Mycophenolate Mofetil and Azathioprine in lupus nephritis therapy. *Research Square (Research Square)*. <https://doi.org/10.21203/rs.3.rs-3783877/v1>
- (19) Wang, H., Gao, Y., Ma, Y., Cai, F., Huang, X., Lan, L., Ren, P., Wang, Y., Chen, J., & Han, F. (2021). Performance of the 2019 EULAR/ACR systemic lupus erythematosus classification

- criteria in a cohort of patients with biopsy-confirmed lupus nephritis. *Lupus Science & Medicine*, 8(1), e000458. <https://doi.org/10.1136/lupus-2020-000458>
- (20)Blenkinsopp, S. C., Fu, Q., Green, Y., Madan, A., Juliao, P., Goldman, D. W., Roth, D. A., & Petri, M. A. (2022b). Renal response at 2 years post biopsy to predict long-term renal survival in lupus nephritis: a retrospective analysis of the Hopkins Lupus Cohort. *Lupus Science & Medicine*, 9(1), e000598. <https://doi.org/10.1136/lupus-2021-000598>
- (21)Moreno, V., & Solé, X. (2004, 15 febrero). Uso de chips de ADN (microarrays) en medicina: fundamentos técnicos y procedimientos básicos para el análisis estadístico de resultados. *Medicina Clínica*. <https://www.elsevier.es/es-revista-medicina-clinica-2-articulo-uso-chips-adn-microarrays-medicina-13057538>
- (22)Martorell-Marugán, J., López-Domínguez, R., García-Moreno, A., Toro-Domínguez, D., Villatoro-García, J. A., Barturen, G., Martín-Gómez, A., Troule, K., Gómez-López, G., Al-Shahrour, F., González-Rumayor, V., Peña-Chilet, M., Dopazo, J., Sáez-Rodríguez, J., Alarcón-Riquelme, M. E., & Carmona-Sáez, P. (2021b). A comprehensive database for integrated analysis of omics data in autoimmune diseases. *BMC Bioinformatics*, 22(1). <https://doi.org/10.1186/s12859-021-04268-4>
- (23)Irizarry, R. A., Hobbs, B., Collin, F., Beazer-Barclay, Y. D., Antonellis, K. J., Scherf, U., & Speed, T. P. (2003). Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics*, 4(2), 249-264. <https://doi.org/10.1093/biostatistics/4.2.249>
- (24)Introduction. (s.f.). Rstudio.com. Recuperado 20 de octubre de 2024, de <https://rmarkdown.rstudio.com/lesson-1.html>
- (25)A Grammar of Data Manipulation. (s.f.). Recuperado 20 de octubre de 2024, de <https://dplyr.tidyverse.org/>
- (26)Readxl. (s.f.). Tidyverse.org. Recuperado 20 de octubre de 2024, de <https://readxl.tidyverse.org/>
- (27)(S.f.). R-project.org. Recuperado 20 de octubre de 2024, de <https://cran.r-project.org/web/packages/readODS/readODS.pdf>
- (28)Tidy messy data. (s.f.). Recuperado 20 de octubre de 2024, de <https://tidyr.tidyverse.org/>
- (29)BiomaRt. (s.f.). Bioconductor. Recuperado 20 de octubre de 2024, de <https://bioconductor.org/packages/release/bioc/html/biomaRt.html>
- (30)(S.f.). R-project.org. Recuperado 20 de octubre de 2024, de <https://cran.r-project.org/web/packages/matrixStats/matrixStats.pdf>
- (31)Kuhn, M. (2019). The caret Package. <https://topepo.github.io/caret/>
- (32)Limma. (s. f.). Bioconductor. Recuperado 20 de octubre de 2024, de <https://bioconductor.org/packages/release/bioc/html/limma.html>
- (33)Irizarry, R. (2019). ggplot2.Introduction to Data Science. <https://doi.org/10.1201/9780429341830-7>
- (34)(S.f.). R-project.org. Recuperado 20 de octubre de 2024, de <https://cran.r-project.org/web/packages/ggrepel/ggrepel.pdf>
- (35)Readr. (s.f.). Tidyverse.org. Recuperado 20 de octubre de 2024, de <https://readr.tidyverse.org/>
- (36)(S.f.). R-project.org. Recuperado 20 de octubre de 2024, de <https://cran.r-project.org/web/packages/pheatmap/pheatmap.pdf>

- (37)(S. f.). R-project.org. Recuperado 20 de octubre de 2024, de <https://cran.r-project.org/web/packages/RColorBrewer/RColorBrewer.pdf>
- (38)García-Moreno, A.; López-Domínguez, R.; Villatoro-García, J.A.; Ramírez-Mena, A.; Aparicio-Puerta, E.; Hackenberg, M.; Pascual-Montano, A.; Carmona-Saez, P. Functional Enrichment Analysis of Regulatory Elements. *Biomedicines* 2022, 10, 590. <https://doi.org/10.3390/biomedicines10030590>
- (39)Tabas-Madrid, D., Nogales-Cadenas, R., & Pascual-Montano, A. (2012). GeneCodis3: a non-redundant and modular enrichment analysis tool for functional genomics. *Nucleic Acids Research*, 40(W1), W478-W483. <https://doi.org/10.1093/nar/gks402>
- (40)RPubs - Test exacto de Fisher, chi-cuadrado de Pearson, McNemar y Q-Cochran. (s/f). Rpubs.com. Recuperado el 23 de octubre de 2024, de [https://rpubs.com/joaquin\\_ar/220579](https://rpubs.com/joaquin_ar/220579)
- (41)RPubs - Chi-Square in R. (s/f). Rpubs.com. Recuperado el 23 de octubre de 2024, de <https://www.rpubs.com/su-biostatistics/chisquare>
- (42)RPubs - PRUEBA DE SHAPIRO-WILK. (s/f). Rpubs.com. Recuperado el 23 de octubre de 2024, de <https://www.rpubs.com/RubenOrtiz/192363>
- (43)t.test function - RDocumentation. (s/f). Rdocumentation.org. Recuperado el 23 de octubre de 2024, de <https://www.rdocumentation.org/packages/stats/versions/3.6.2/topics/t.test>
- (44)RPubs - Prueba de Wilcoxon. (s/f). Rpubs.com. Recuperado el 23 de octubre de 2024, de <https://rpubs.com/luisrmacias/wilcoxon>
- (45)Durinck, S., Moreau, Y., Kasprzyk, A., Davis, S., De Moor, B., Brazma, A., & Huber, W. (2005). BioMart and Bioconductor: a powerful link between biological databases and microarray data analysis. *Bioinformatics*, 21(16), 3439-3440. <https://doi.org/10.1093/bioinformatics/bti525>
- (46)Bolstad, B., Irizarry, R., Åstrand, M., & Speed, T. (2003). A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics*, 19(2), 185-193. <https://doi.org/10.1093/bioinformatics/19.2.185>
- (47)Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., & Smyth, G. K. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Research*, 43(7), e47. <https://doi.org/10.1093/nar/gkv007>
- (48)Benjamini, Y., & Hochberg, Y. (1995b). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal Of The Royal Statistical Society Series B (Methodological)*, 57(1), 289-300. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>
- (49)Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., & Smyth, G. K. (2015b). limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Research*, 43(7), e47. <https://doi.org/10.1093/nar/gkv007>
- (50)OpenAI. (2023). *ChatGPT* (versión GPT-4, Octubre 2023) [Modelo de lenguaje de inteligencia artificial]. <https://openai.com/chatgpt>
- (51)Management of pericarditis in rheumatic diseases. (s/f). American College of Cardiology. Recuperado el 26 de octubre de 2024, de <https://www.acc.org/Latest-in-Cardiology/Articles/2022/12/05/12/24/Management-of-Pericarditis-in-Rheumatic-Diseases>
- (52)Tektonidou, M. G. (2018). Antiphospholipid Syndrome Nephropathy: From Pathogenesis to Treatment. *Frontiers in immunology*, 9. <https://doi.org/10.3389/fimmu.2018.01181>

- (53) Ben Ariba, Y., Ajili, F., Maïza, M., Louzir, B., & Labidi, J. (2018). Lupus Nephritis: Clinical Characteristics and Prognostic Factors. *Journal of Nephrology & Kidney Diseases*, 2(2), Artículo 1017. Recuperado de [https://www.jsmcentral.org/sm-nephrology-kidney-diseases/fulltext\\_smjnknd-v2-1017.pdf](https://www.jsmcentral.org/sm-nephrology-kidney-diseases/fulltext_smjnknd-v2-1017.pdf)
- (54) Riboldi, E., Daniele, R., Parola, C., Inforzato, A., Arnold, P. L., Bosisio, D., Fremont, D. H., Bastone, A., Colonna, M., & Sozzani, S. (2011). Human C-type Lectin Domain Family 4, Member C (CLEC4C/BDCA-2/CD303) Is a Receptor for Asialo-galactosyl-oligosaccharides. *Journal Of Biological Chemistry*, 286(41), 35329-35333. <https://doi.org/10.1074/jbc.c111.290494>
- (55) Postal, M., Vivaldo, J. F., Fernandez-Ruiz, R., Paredes, J. L., Appenzeller, S., & Niewold, T. B. (2020). Type I interferon in the pathogenesis of systemic lupus erythematosus. *Current Opinion In Immunology*, 67, 87-94. <https://doi.org/10.1016/j.coi.2020.10.014>
- (56) Tanaka, Y., Kusuda, M., & Yamaguchi, Y. (2022). Interferons and systemic lupus erythematosus: Pathogenesis, clinical features, and treatments in interferon-driven disease. *Modern Rheumatology*, 33(5), 857-867. <https://doi.org/10.1093/mr/roac140>
- (57) Tydén, H., Lood, C., Gullstrand, B., Jönsen, A., Ivars, F., Leanderson, T., & Bengtsson, A. A. (2016). Pro-inflammatory S100 proteins are associated with glomerulonephritis and anti-dsDNA antibodies in systemic lupus erythematosus. *Lupus*, 26(2), 139-149. <https://doi.org/10.1177/0961203316655208>
- (58) Elnokeety, M. M., Khalil, N., & Fayed, A. (2023). Serum S100P and C4d proteins as markers for active proliferative lupus nephritis in Egyptian patients with systemic lupus erythematosus. *Ukrainian Journal Of Nephrology And Dialysis*, 4(80), 13-23. [https://doi.org/10.31450/ukrjnd.4\(80\).2023.03](https://doi.org/10.31450/ukrjnd.4(80).2023.03)
- (59) Li, H., Wang, T., Li, B., Huang, T., Hai, Y., Huang, C., & Xiang, W. (2023). Bioinformatic analysis of immune-related transcriptome affected by IFIT1 gene in childhood systemic lupus erythematosus. *Translational Pediatrics*, 12(8), 1517-1526. <https://doi.org/10.21037/tp-23-365>
- (60) Yang, Y., Wang, C., Cheng, P., Zhang, X., Li, X., Hu, Y., Xu, F., Hong, F., Dong, G., & Xiong, H. (2018). CD180 Ligation Inhibits TLR7- and TLR9-Mediated Activation of Macrophages and Dendritic Cells Through the Lyn-SHP-1/2 Axis in Murine Lupus. *Frontiers In Immunology*, 9. <https://doi.org/10.3389/fimmu.2018.02643>
- (61) Wen, L., Zhang, B., Wu, X., Liu, R., Fan, H., Han, L., Zhang, Z., Ma, X., Chu, C., & Shi, X. (2023). Toll-like receptors 7 and 9 regulate the proliferation and differentiation of B cells in systemic lupus erythematosus. *Frontiers In Immunology*, 14. <https://doi.org/10.3389/fimmu.2023.1093208>
- (62) Bittermann, E., Abdelhamed, Z., Liegel, R. P., Menke, C., Timms, A., Beier, D. R., & Stottmann, R. W. (2019). Differential requirements of tubulin genes in mammalian forebrain development. *PLoS Genetics*, 15(8), e1008243. <https://doi.org/10.1371/journal.pgen.1008243>
- (63) Roselli, F., Livrea, P., & Jirillo, E. (2006). Voltage-Gated Sodium Channel Blockers as Immunomodulators. *Recent Patents On CNS Drug Discovery*, 1(1), 83-91. <https://doi.org/10.2174/157488906775245255>

- (64) Perl, A. (2013). Oxidative stress in the pathology and treatment of systemic lupus erythematosus. *Nature Reviews Rheumatology*, 9(11), 674-686. <https://doi.org/10.1038/nrrheum.2013.147>
- (65) Gorelik, G., & Richardson, B. (2009). Key role of ERK pathway signaling in lupus. *Autoimmunity*, 43(1), 17-22. <https://doi.org/10.3109/08916930903374832>
- (66) Pan, L., Lu, M., Wang, J., Xu, M., & Yang, S. (2019). Immunological pathogenesis and treatment of systemic lupus erythematosus. *World Journal Of Pediatrics*, 16(1), 19-30. <https://doi.org/10.1007/s12519-019-00229-3>
- (67) Otón, T., Fernández, L. S., Castro, M. F., & Andreu, J. L. (2011). Terapia biológica dirigida contra los linfocitos B en el lupus eritematoso sistémico. *Seminarios de la Fundación Española de Reumatología*, 12(1), 10-14. <https://doi.org/10.1016/j.semreu.2010.06.003>
- (68) Jurewicz, M. M., & Stern, L. J. (2018). Class II MHC antigen processing in immune tolerance and inflammation. *Immunogenetics*, 71(3), 171-187. <https://doi.org/10.1007/s00251-018-1095-x>
- (69) Chodiseti, S. B., Fike, A. J., Domeier, P. P., Singh, H., Choi, N. M., Corradetti, C., Kawasawa, Y. I., Cooper, T. K., Caricchio, R., & Rahman, Z. S. M. (2020). Type II but Not Type I IFN Signaling Is Indispensable for TLR7-Promoted Development of Autoreactive B Cells and Systemic Autoimmunity. *The Journal Of Immunology*, 204(4), 796-809. <https://doi.org/10.4049/jimmunol.1901175>
- (70) Fan, W., Wei, B., Chen, X., Zhang, Y., Xiao, P., Li, K., Zhang, Y. Q., Huang, J., Leng, L., & Bucala, R. (2023). The RhoA GTPase regulates Type I Interferon Signaling in Systemic lupus erythematosus. *Research Square (Research Square)*. <https://doi.org/10.21203/rs.3.rs-3320841/v1>
- (71) Brightbill, H. D., Suto, E., Blaquiere, N., Ramamoorthi, N., Sujatha-Bhaskar, S., Gogol, E. B., Castanedo, G. M., Jackson, B. T., Kwon, Y. C., Haller, S., Lesch, J., Bents, K., Everett, C., Kohli, P. B., Linge, S., Christian, L., Barrett, K., Jaochico, A., Berezhkovskiy, L. M., . . . Ghilardi, N. (2018). NF- $\kappa$ B inducing kinase is a therapeutic target for systemic lupus erythematosus. *Nature Communications*, 9(1). <https://doi.org/10.1038/s41467-017-02672-0>
- (72) Bruera, S., Chavula, T., Madan, R., & Agarwal, S. K. (2023). Targeting type I interferons in systemic lupus erythematosus. *Frontiers In Pharmacology*, 13. <https://doi.org/10.3389/fphar.2022.1046687>
- (73) Iwamoto, T., Dorschner, J. M., Selvaraj, S., Mezzano, V., Jensen, M. A., Vsetecka, D., Amin, S., Makol, A., Osborn, T., Moder, K., Chowdhary, V. R., Izmirly, P., Belmont, H. M., Clancy, R. M., Buyon, J. P., Wu, M., Loomis, C. A., & Niewold, T. B. (2021). High systemic type I interferon activity is associated with active class III/IV lupus nephritis. *The Journal Of Rheumatology*, 49(4), 388-397. <https://doi.org/10.3899/jrheum.210391>
- (74) Hu, J., Han, C., Zhong, J., Liu, H., Liu, R., Luo, W., Chen, P., & Ling, F. (2021). Dynamic Network Biomarker of Pre-Exhausted CD8+ T Cells Contributed to T Cell Exhaustion in Colorectal Cancer. *Frontiers In Immunology*, 12. <https://doi.org/10.3389/fimmu.2021.691142>
- (75) Martín-Torregrosa, D., Mansilla-Polo, M., & Morgado-Carrasco, D. (2024). Uso del anifrolumab en el lupus eritematoso sistémico, lupus eritematoso cutáneo y otras

dermatosis autoinmunes. Actas dermo-sifiliográficas.

<https://doi.org/10.1016/j.ad.2024.05.024>

(76)Chuang, H., & Tan, T. (2019). MAP4K Family Kinases and DUSP Family Phosphatases in T-Cell Signaling and Systemic Lupus Erythematosus. *Cells*, 8(11), 1433.

<https://doi.org/10.3390/cells8111433>

(77)Wu, Y., Tang, W., & Zuo, J. (2015). Toll-like receptors: potential targets for lupus treatment.

*Acta Pharmacologica Sinica*, 36(12), 1395-1407. <https://doi.org/10.1038/aps.2015.91>

(78)Caster, D. J., & Powell, D. W. (2020, 28 diciembre). Precision Targeting of NF- $\kappa$ B Signaling in Lupus Nephritis. <https://pmc.ncbi.nlm.nih.gov/articles/PMC9245607/>