



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA



VALENCIA BIOMEDICAL RESEARCH FOUNDATION
CENTRO DE INVESTIGACIÓN PRÍNCIPE FELIPE

Programa de Doctorado en Biotecnología

Tesis doctoral

Obtención y caracterización de perfiles transcripcionales consenso para la comprensión de la progresión y severidad de la lesión medular

Rubén Grillo Risco

Dirigida por:

Dr. Francisco García García

Dra. Victoria Moreno Manzano

Dra. Marta Hidalgo García

Valencia, abril 2025



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA



VALENCIA BIOMEDICAL RESEARCH FOUNDATION
CENTRO DE INVESTIGACIÓN PRÍNCIPE FELIPE

Programa de Doctorado en Biotecnología

Tesis doctoral

**Obtención y caracterización
de perfiles transcripcionales
consenso para la comprensión
de la progresión y severidad
de la lesión medular**

Rubén Grillo Risco

Dirigida por:

Dr. Francisco García García

Dra. Victoria Moreno Manzano

Dra. Marta Hidalgo García

Valencia, abril 2025

Índice

1	Introducción	1
1.1	Lesión medular	2
1.1.1	Etiología y epidemiología	2
1.1.2	La médula espinal en el sistema nervioso	7
1.1.3	Anatomía de la médula espinal	8
1.1.4	La lesión medular en humanos	11
1.1.4.1	Grados de severidad	11
1.1.4.2	Evolución fisiopatológica de la lesión medular: fases temporales	13
1.1.5	La rata como modelo de estudio	18
1.2	Transcriptómica	22
1.2.1	Estrategias de análisis transcriptómicos	24
1.2.1.1	Análisis de expresión diferencial	24
1.2.1.2	Caracterización funcional	24
1.2.2	Metaanálisis de expresión génica	28
2	Justificación y objetivos	31
2.1	Objetivos generales	31
2.2	Objetivos específicos	31
3	Materiales y métodos	33
3.1	Revisión sistemática y selección de estudios	35

3.2	Estandarización de grupos experimentales	35
3.3	Análisis individual de estudios	36
3.3.1	Preprocesamiento	36
3.3.2	Análisis exploratorio	37
3.3.3	Análisis de expresión diferencial	37
3.4	Metaanálisis de expresión génica	38
3.5	Identificación de genes biomarcadores	39
3.6	Caracterización funcional	41
3.6.1	<i>Gene Set Analysis</i>	41
3.6.2	Estimación de la actividad de factores de transcripción . . .	41
3.6.3	Construcción de redes de genes co-expresados	42
3.7	Validación de resultados	43
3.7.1	<i>Dataset</i> transcriptómico de rata	43
3.7.2	<i>Dataset</i> transcriptómico de ratón	44
3.7.3	Validación experimental (qPCR)	45
3.7.4	Análisis de muestras de sangre humana	46
3.7.5	Meta-SCI app	47
4	Resultados	49
4.1	Obtención de perfiles transcripcionales consenso en función de la severidad y la fase de la lesión	49
4.2	Los biomarcadores específicos de severidad y fase identificados permiten una clasificación efectiva de los grupos	54
4.3	Caracterización funcional de las firmas transcripcionales consenso .	65
4.4	Inferencia de la actividad de TF en la progresión de la lesión medular	71
4.5	Clasificación funcional de las redes de genes co-expresados	74
4.6	Los genes biomarcadores de severidad y fase propuestos clasifican correctamente nuevos perfiles transcripcionales de rata	79

4.7	Consistencia en los patrones de expresión temporales entre ratón y rata	81
4.8	Concordancia entre los patrones de expresión obtenidos en el metaanálisis y el análisis qPCR	84
4.9	Los genes biomarcadores específicos de severidad en rata predicen la prognosis de la lesión de los pacientes a partir de muestras de sangre humana	87
4.10	Meta-SCI app - una plataforma interactiva e intuitiva para explotar los resultados	90
5	Discusión	95
6	Conclusiones	107
7	Material suplementario	135

Índice de figuras

Figura 1.	Incidencia anual relativa de lesiones medulares por millón de habitantes	3
Figura 2.	Evolución del número de fallecidos, heridos hospitalizados y lesionados medulares por accidentes de tráfico en España de 2000 a 2009	6
Figura 3.	Organización del sistema nervioso	8
Figura 4.	Organización de los nervios espinales	9
Figura 5.	Anatomía macroscópica de la médula espinal y la vértebra vistas desde el plano transversal	11
Figura 6.	Comparación de la paraplejia/tetraplejia completa e incompleta	13
Figura 7.	Visión general de eventos fisiopatológicos en la lesión primaria y secundaria tras un traumatismo	16
Figura 8.	Comparación de la composición estructural y celular de la médula espinal sana y lesionada	18
Figura 9.	Esquema del dispositivo de contusión mecánico	20
Figura 10.	Resumen de tres tipos de métodos de caracterización funcional	26
Figura 11.	Flujo de trabajo	34
Figura 12.	Diagrama de flujo de la revisión sistemática, siguiendo la declaración PRISMA	50

Figura 13.	Distribución de estudios y muestras en los grupos experimentales	52
Figura 14.	Resumen de resultados del metaanálisis de expresión génica	53
Figura 15.	Heatmap con <i>clustering</i> jerárquico de biomarcadores específicos de severidad	54
Figura 16.	El uso de genes biomarcadores específicos permiten una clasificación efectiva de grupos en base a la severidad o la fase tras la lesión	55
Figura 17.	Red PPI para genes marcadores específicos de severidad <i>down</i> -regulados	57
Figura 18.	Red PPI para genes marcadores específicos de severidad <i>up</i> -regulados	58
Figura 19.	Red PPI para genes marcadores específicos de T1 <i>up</i> -regulados	61
Figura 20.	Red PPI para genes marcadores específicos de T1 <i>down</i> -regulados	62
Figura 21.	Red PPI para genes marcadores específicos de T2	63
Figura 22.	Red PPI para genes marcadores específicos de T4	64
Figura 23.	El <i>Gene Set Analysis</i> revela funciones/rutas desreguladas en base al tiempo y la severidad tras la lesión	66
Figura 24.	<i>Dotplot</i> de una selección de funciones relacionadas con el metabolismo	68
Figura 25.	Heatmap con <i>clustering</i> jerárquico de los genes del término GO:0006695 (proceso de biosíntesis de colesterol)	69
Figura 26.	<i>Dotplot</i> de una selección de funciones relacionadas con las mitocondrias	70
Figura 27.	Resumen del análisis de inferencia de la actividad de los factores de transcripción	73

Figura 28.	Distribución de frecuencias de las categorías funcionales asociadas a las redes locales formadas por genes co-expresados	75
Figura 29.	Patrones de expresión de las redes locales de lesión moderada clasificadas por categoría funcional	76
Figura 30.	Ejemplos de redes locales de co-expresión	78
Figura 31.	Análisis de <i>clustering</i> del <i>dataset</i> GSE218088 usando biomarcadores específicos de severidad y fase	80
Figura 32.	Comparación de los patrones de expresión del metaanálisis en rata con un <i>dataset</i> de ratón y análisis de <i>clustering</i> usando biomarcadores comunes	82
Figura 33.	Comparación de los patrones de expresión génica de los factores de transcripción y sus genes diana seleccionados para validación experimental	85
Figura 34.	Coeficiente de correlación de Pearson al comparar los valores de logFC de los patrones de expresión obtenidos en la qPCR con los de lesión moderada de rata, lesión severa de rata y el <i>dataset</i> de ratón	86
Figura 35.	Los genes biomarcadores específicos de severidad estratifican muestras humanos y en modelos de LM en ratas	88
Figura 36.	Expresión normalizada de la firma de 12 genes biomarcadores de severidad en el <i>dataset</i> GSE151371 . . .	89
Figura 37.	Visión general de los principales módulos de Meta-SCI app .	91
Figura S1.	Análisis de <i>clustering</i> de las firmas transcripcionales consenso con el top 10 genes específicos de T1	135
Figura S2.	Análisis de <i>clustering</i> de las firmas transcripcionales consenso con el top 10 genes específicos de T2	135

Figura S3. Análisis de <i>clustering</i> de las firmas transcripcionales consenso con el top 10 genes específicos de T3	136
Figura S4. Análisis de <i>clustering</i> de las firmas transcripcionales consenso con el top 10 genes específicos de T2	136
Figura S5. <i>Dotplot</i> de una selección de funciones relacionadas con neurotransmisores	137
Figura S6. <i>Dotplot</i> de una selección de funciones relacionadas con la sinapsis	138
Figura S7. <i>Dotplot</i> de una selección de funciones relacionadas con el desarrollo del sistema nervioso	139
Figura S8. <i>Clusters</i> de genes co-expresados para lesión moderada . . .	140
Figura S9. <i>Clusters</i> de genes co-expresados para lesión severa . . .	141
Figura S10. Patrones de expresión de las redes locales de lesión severa clasificadas por categoría funcional	142

Índice de tablas

Tabla 1. Comparaciones aplicadas para la identificación de biomarcadores específicos	40
Tabla 2. Estudios seleccionados en la revisión sistemática e incluidos en el metaanálisis	51
Tabla 3. Top 10 genes biomarcadores específicos para cada comparación	59

Resumen

La lesión medular espinal (LM) es una condición devastadora que provoca disfunción motora, sensorial y autónoma. Las opciones terapéuticas actuales son limitadas, lo que subraya la necesidad de una mayor comprensión de los mecanismos moleculares asociados a la LM. En este estudio se caracterizaron las diferentes fases y grados de severidad de la LM a nivel génico y funcional, centrándose en la identificación de genes biomarcadores. Nuestra aproximación incluyó una revisión sistemática, análisis transcriptómico individual, metaanálisis de expresión génica y caracterización funcional. Se recopilaron un total de 14 estudios con 273 muestras, lo que permitió identificar genes biomarcadores específicos de severidad y fase, capaces de clasificar de manera precisa los perfiles transcriptómicos. Investigamos la posible transferibilidad de los biomarcadores específicos de severidad y logramos identificar una firma de 12 genes que predice el pronóstico de la lesión a partir de muestras de sangre humana. Además, desarrollamos MetaSCI-app, una aplicación web interactiva diseñada para investigadores, que permite explorar y visualizar todos los resultados generados (<https://metasci-cbl.shinyapps.io/metaSCI>). Así, presentamos una referencia transcriptómica y proporcionamos un marco integral para evaluar la LM considerando las perspectivas de severidad y tiempo, todo integrado en una herramienta de fácil uso.

Resum

La lesió medul·lar espinal (LM) és una condició devastadora que provoca disfunció motora, sensorial i autònoma. Les opcions terapèutiques actuals són limitades, la qual cosa subratlla la necessitat d'una major comprensió dels mecanismes moleculars associats a la LM. En aquest estudi es van caracteritzar les diferents fases i graus de severitat de la LM a nivell gènic i funcional, centrant-se en la identificació de gens biomarcadors. La nostra aproximació va incloure una revisió sistemàtica, anàlisi transcriptòmica individual, metaanàlisi d'expressió gènica i caracterització funcional. Es van recopilar un total de 14 estudis amb 273 mostres, la qual cosa va permetre identificar gens biomarcadors específics de severitat i fase, capaços de classificar de manera precisa els perfils transcriptòmics. Investiguem la possible transferibilitat dels biomarcadors específics de severitat i aconseguim identificar una signatura de 12 gens que prediu el pronòstic de la lesió a partir de mostres de sang humana. A més, desenvolupem MetaSCI-app, una aplicació web interactiva dissenyada per a investigadors, que permet explorar i visualitzar tots els resultats generats (<https://metasci-cbl.shinyapps.io/metasci>). Així, presentem una referència transcriptòmica i proporcionem un marc integral per avaluar la LM considerant les perspectives de severitat i temps, tot integrat en una eina de fàcil ús.

Abstract

Spinal cord injury (SCI) is a devastating condition that leads to motor, sensory, and autonomic dysfunction. Current therapeutic options remain limited, emphasizing the need for a comprehensive understanding of the underlying SCI-associated molecular mechanisms. This study characterized distinct SCI phases and severities at the gene and functional levels, focusing on biomarker gene identification. Our approach involved a systematic review, individual transcriptomic analysis, gene meta-analysis, and functional characterization. We compiled a total of fourteen studies with 273 samples, leading to the identification of severity- and phase-specific biomarker genes that allow the precise classification of transcriptomic profiles. We investigated the potential transferability of severity-specific biomarkers and identified a twelve-gene signature that predicted injury prognosis from human blood samples. We also report the development of MetaSCI-app - an interactive web application designed for researchers - that allows the exploration and visualization of all generated results (<https://metasci-cbl.shinyapps.io/metaSCI>). Overall, we present a transcriptomic reference and provide a comprehensive framework for assessing SCI considering severity and time perspectives, all integrated into a user-friendly tool.

1. Introducción

La lesión medular (LM) es un proceso patológico desencadenado tras un daño en la médula espinal, sea cual sea su origen (traumático o no traumático), que conduce a graves alteraciones en el funcionamiento del sistema nervioso central (SNC) con deterioro, generalmente irreversible, de los sistemas sensoriales y motores, tanto el autónomo como el somático [1].

Las consecuencias de estas alteraciones son devastadoras para la calidad de vida de los pacientes. El grado de deterioro funcional dependerá de la severidad y el nivel donde se produzca la lesión, pudiendo derivar en paraplejia (parálisis de las extremidades inferiores) o tetraplejia (parálisis en las extremidades superiores e inferiores). Los fallos en el sistema nervioso autónomo (SNA) tienen múltiples consecuencias y gran impacto para el paciente, como la pérdida de la sensibilidad y disfunción vesical, intestinal o sexual, entre otras [1].

Actualmente la LM no tiene cura, aunque existen abordajes terapéuticos implantados en la práctica clínica, que permiten mejorar en cierta medida la calidad de vida de los lesionados [1]. Por otro lado, durante las últimas décadas, se ha realizado un gran esfuerzo desde la investigación básica y preclínica, demostrando el éxito de diversas estrategias desde los modelos animales, en algunos casos, implementadas en diversos ensayos clínicos [2–5]. Sin embargo, aún es necesario que se realicen esfuerzos adicionales y coordinados para dilucidar los mecanismos moleculares que subyacen en la propia LM y que contribuirán a mejorar y dirigir de forma más eficiente las estrategias de terapia. Una mejor comprensión de los procesos que regulan las primeras fases de la LM permitiría poder controlar su evolución en los nuevos casos. Así mismo,

también es importante comprender los procesos que persisten desregulados en fases crónicas, con el objetivo de diseñar terapias que permitan una recuperación funcional en personas con una lesión estable.

1.1. Lesión medular

1.1.1. Etiología y epidemiología

La LM es una patología con una etiología muy diversa, la cual se divide en dos grupos principales, la LM traumática y la LM no traumática. Las lesiones de origen traumático abarcan aquellas que se producen por contusión, compresión o transección de la médula espinal, mientras que las de origen no traumático suelen darse a raíz de enfermedades previas, como tumores, infecciones o problemas vasculares, entre otros [6]. Estas consideraciones son importantes a la hora de hablar de la incidencia de la LM, ya que va muy ligada al nivel de desarrollo y estilo de vida de cada país o región geográfica, por lo que el número, la proporción de causas desencadenantes y el perfil sociodemográfico de los pacientes varía mucho en las diferentes partes del mundo. Además, hay que tener en cuenta que dar estimaciones precisas es complicado, ya que no suelen existir registros a nivel nacional, e incluso en países en vías de desarrollo los datos son tan limitados que no se puede obtener información fiable (Figura 1). Al no haber métodos estandarizados entre países para recoger datos epidemiológicos, las estimaciones se extrapolan a partir de zonas geográficas más pequeñas, como regiones administrativas, provincias o incluso hospitales, siendo necesario el uso de estudios poblacionales (prospectivos o retrospectivos) [6–9].

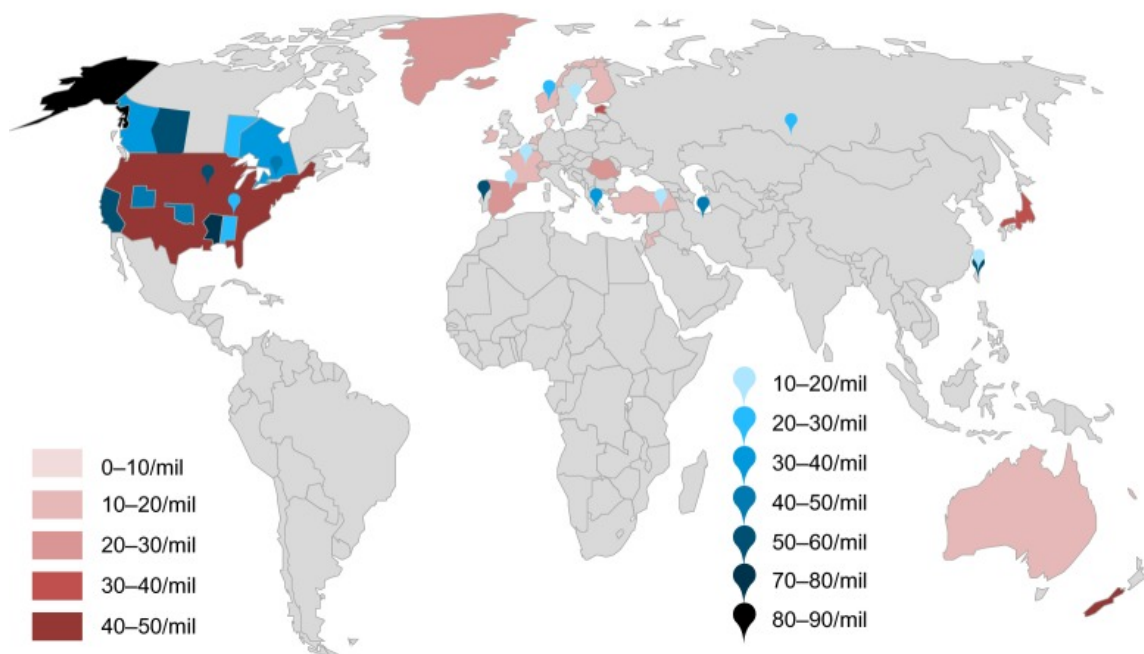


Figura 1. Incidencia anual relativa de lesiones medulares por millón de habitantes. La escala en rojo ilustra la incidencia en países. La escala en azul indica la incidencia en estados/provincias y regiones. En gris se muestran los países con falta de datos epidemiológicos precisos. Figura extraída de Singh et al. [7].

En países desarrollados, la incidencia de la LM traumática fluctúa de los 15 casos por millón en regiones de Europa Occidental o Australia a los 40 casos por millón en Norteamérica. En estos países el perfil del paciente típico son hombres jóvenes cuya causa de la lesión son accidentes de tráfico. La incidencia es de tres a cuatro veces mayor en hombres que en mujeres, mientras que la edad sigue una distribución bimodal, con un máximo de frecuencia entre los 15 y 30 años, y otro a partir de los 65 años [10]. Los accidentes de tráfico son la primera causa de LM traumática, seguida de caídas, eventos violentos y actividades deportivas [6].

Por otro lado, en regiones con conflictos armados como Siria, Yemen, Irak o Afganistán, la principal causa de LM traumática es de origen violento, debido a balas o metralla. En esta misma línea encontramos el estudio realizado por Lugo

et al. [11] en Colombia, en el que el 62 % de los pacientes habían sido lesionados por armas de fuego. Tan solo el 20 % de los participantes iniciales terminaron el estudio, debido a la situación de conflicto del país que les forzaba a desaparecer para poder ocultar su identidad.

Finalmente, las caídas son la causa más frecuente en regiones como el Sureste Asiático, y algunas partes del Mediterráneo oriental, África subsahariana y el Pacífico Occidental, principalmente por la falta de medidas de seguridad en entornos laborales [10]. Aunque las lesiones relacionadas con actividades deportivas suponen una pequeña proporción de las lesiones medulares a nivel global, hay algunas regiones que tienen una mayor incidencia de lesiones asociadas a estos motivos. Por ejemplo, en Hawai el 12 % de las lesiones están relacionadas con el buceo, mientras que en otras regiones de EE.UU. el rango está entre el 3 y el 6 % [12]. Además, Derakhshanrad et al. [13] reportaban que existe una relación entre el nivel de educación y la etiología de la LM traumática. Las actividades deportivas (buceo, gimnasia, wrestling) se vuelven una causa más común de LM traumática a medida que los años de educación aumentan, al mismo tiempo que las lesiones relacionadas con la violencia se vuelven menos comunes.

Las LM no traumáticas también tienen un amplio abanico de causas, habitualmente consecuencia de enfermedades previas. Aunque de nuevo los porcentajes varían mucho de unos estudios a otros, las principales razones son la presencia de tumores, que comprimen la médula espinal, las enfermedades degenerativas derivadas de desórdenes autoinmunes, las infecciones o los desórdenes vasculares. Además, también pueden ser de origen congénito como la espina bífida [14]. Sin embargo, existe un porcentaje de casos donde la causa es desconocida, que suele presentarse en pacientes de edad avanzada y con cuadros clínicos complicados, previos a la lesión, que contribuyen a que la mortalidad en las lesiones no traumáticas sea mayor que en las lesiones

traumáticas. Cabe destacar que en las lesiones no traumáticas la diferencia de incidencia entre hombres y mujeres desaparecen [15, 16].

Si nos centramos en las cifras de España, nos encontramos ante la misma problemática para obtener estimaciones epidemiológicas precisas. Al carecer de un Sistema de Registro Nacional de datos epidemiológicos, es necesario recurrir a estudios poblacionales autonómicos, así como atender a las memorias anuales de los Centros de Referencia Estatales en esta materia, o investigaciones desarrolladas por entidades privadas [17, 18]. De esta manera, el estudio más completo registrado es el de Pérez et al. [17] que evalúa la evolución de incidencia de la LM en España de 2000 a 2009, a partir de datos obtenidos del Registro Nacional de Altas Hospitalarias (Conjunto Mínimo Básico de Datos, CMBDAH). Entre 2000 y 2009 se registraron 10274 ingresos por LM traumática, con una incidencia anual de 23.5 casos por millón.

A comienzos de siglo las lesiones medulares eran debidas principalmente a causas traumáticas, en concreto accidentes de tráfico. Es preciso destacar la diferencia de incidencia entre sexos, afectando 3.4 veces más a hombres que a mujeres. De esta manera, el paciente típico era un varón de edad media de 35 años lesionado en un accidente de tráfico. Sin embargo, en 2009 se observa una reducción del 40.9% de las lesiones por accidentes de tráfico y un aumento en las lesiones causadas por caídas, especialmente entre las personas mayores de 65 años. Este cambio en la incidencia se debió a varias medidas de seguridad vial implementadas durante este período, que incluyeron el Plan de Acción de Medidas Especiales de Seguridad Vial 2004-2005, Plan Estratégico de Seguridad Vial 2005-2008, incremento en la vigilancia y cumplimiento de normas mediante el aumento de radares y controles de alcoholemia, la introducción del carnet por puntos, reforma del Código Penal para la aplicación del código penal a comportamientos de riesgo, pudiendo penalizarse con penas de prisión, además

de una importante inversión en campañas de educación y concienciación de la sociedad (Figura 2) [17].

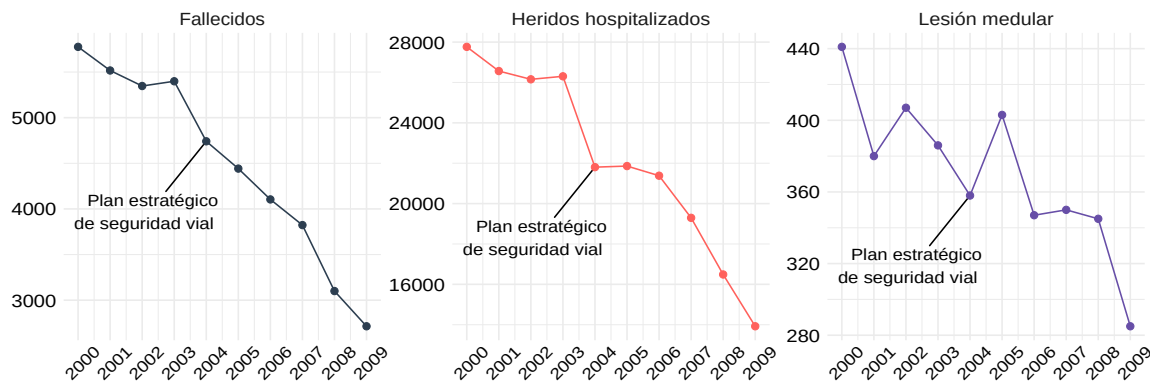


Figura 2. Evolución del número de fallecidos, heridos hospitalizados y lesionados medulares por accidentes de tráfico en España de 2000 a 2009. Los datos de fallecidos y heridos hospitalizados se obtuvieron de la DGT (fecha de consulta 31-05-2024). Las cifras de lesionados medulares por accidentes de tráfico fueron obtenidas de Pérez et al. [17].

Los datos disponibles más recientes indican que en España hay unas 1200 nuevas lesiones medulares cada año con una prevalencia de unas 40 000 personas. Entre las lesiones traumáticas, el 39 % se deben a accidentes de tráfico, mientras que el 55 % son consecuencia de caídas durante actividades deportivas o de ocio, lo que refleja el éxito de las medidas implantadas por el Plan Estratégico de Seguridad Vial a principios de siglo [18, 19].

Desde el punto de vista económico, se estima que en EE.UU. la LM supone un coste aproximado por paciente de 523 000 dólares durante el primer año y de 80 000 dólares en los años posteriores. Sin embargo, estas cifras varían enormemente en función de la región geográfica, la severidad de la lesión o la edad [6, 20]. A este gran impacto económico tanto para el sistema sanitario como para las familias de los pacientes, hay que añadir las enormes consecuencias que suponen estas lesiones a nivel personal y social, lo cual hace que resulte de una mayúscula importancia la investigación de la LM.

1.1.2. La médula espinal en el sistema nervioso

El sistema nervioso se estructura en dos partes, el sistema nervioso central (SNC) y el sistema nervioso periférico (SNP). El SNC lo forman el cerebro y la médula espinal, que se encargan tanto de procesar e integrar toda la información que el individuo percibe del exterior y de los distintos órganos del cuerpo, como de transmitir y ejecutar una respuesta adecuada. Por otro lado, el SNP lo forman ganglios, nervios y terminaciones nerviosas que conducen la información hasta el SNC y desde el SNC. Así pues, desde un punto de vista funcional el SNP tiene un componente sensitivo o aferente, que percibe los estímulos y transmite los impulsos hacia el SNC, y un componente motor o eferente, que lleva los impulsos generados por el cerebro hasta los órganos efectores. Este componente motor se subdivide en el sistema nervioso somático (SNS) y el sistema nervioso autónomo (SNA). El SNS lleva los impulsos generados en el cerebro hasta los músculos esqueléticos, por lo que se encarga de los movimientos voluntarios, mientras que el SNA lleva los impulsos a glándulas, músculo cardíaco o músculo liso, controlando todos los movimientos involuntarios del cuerpo, como el latido del corazón o la vasodilatación en respuesta a cambios de temperatura (Figura 3) [21].

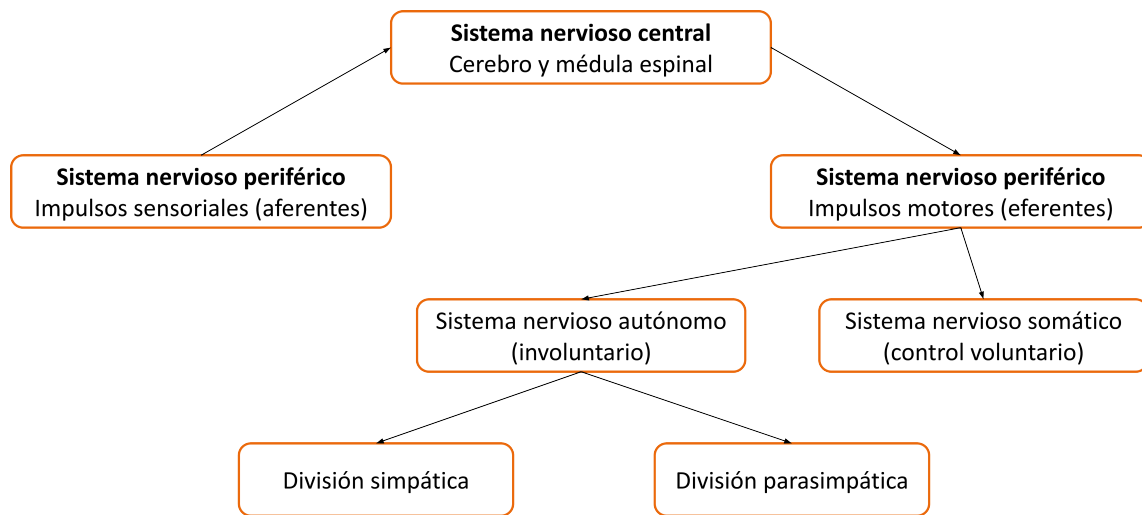


Figura 3. Organización del sistema nervioso. Figura adaptada de Peate y Muralitharan [21].

Como se indicó anteriormente, la médula espinal forma parte del SNC, siendo el mediador entre el cerebro y el resto del cuerpo, ya que recibe en todo momento las señales percibidas, tanto del exterior como del interior, y envía las respuestas que deben realizarse en base a las mismas. Así, es fácil intuir que los daños en la médula espinal tendrán múltiples consecuencias en las funciones del organismo que depende de la gravedad de la lesión y del segmento medular afectado, llegando a poner en riesgo la supervivencia del paciente.

1.1.3. Anatomía de la médula espinal

En humanos, la médula espinal en los adultos mide unos 45 cm de media. A lo largo de la médula espinal se identifican 31 pares de nervios espinales que conectan con diferentes regiones del cuerpo: 8 nervios cervicales, 12 torácicos, 5 lumbares, 5 sacros y un nervio coccígeo (Figura 4). Los nervios espinales son mixtos, es decir, contienen fibras motoras y sensitivas. Los nervios cervicales controlan los músculos y sensibilidad del cuello, la cabeza y parte de las

extremidades superiores. Los nervios torácicos inervan el tronco y los órganos de la cavidad torácica. Tanto los nervios lumbares como los sacros participan en el control del movimiento y la sensibilidad del tren inferior, así como en funciones sexuales, intestinales y vesicales. Finalmente, el nervio coccígeo se asocia con musculatura del coxis y del recto [21, 22]. Esta organización está evolutivamente conservada con ligeras variaciones en roedores [23, 24].

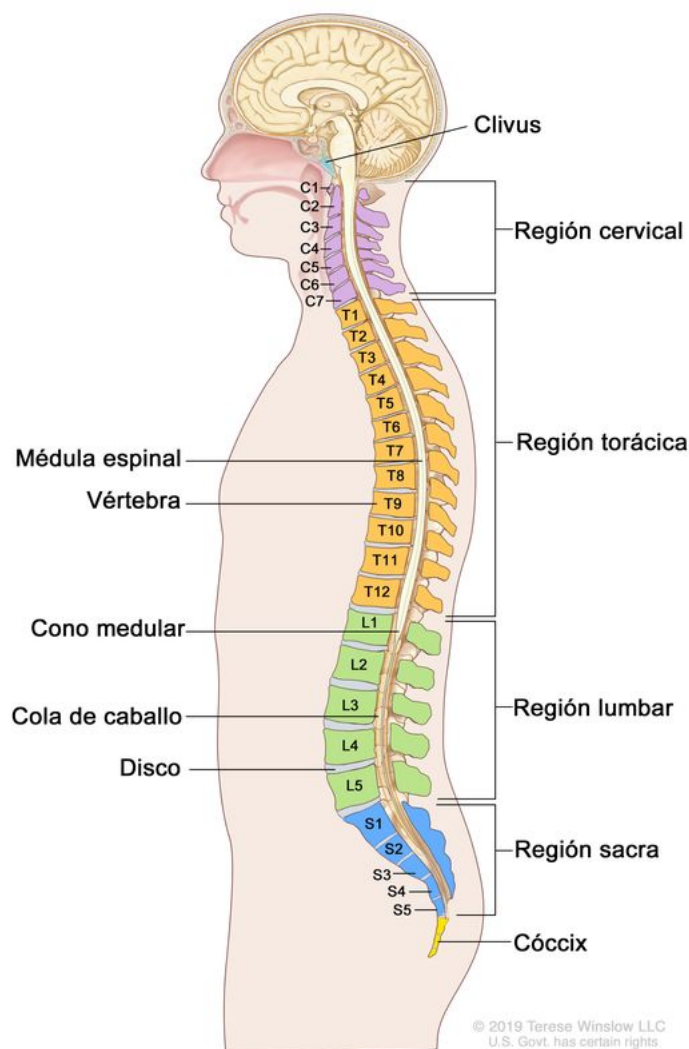


Figura 4. Organización de los nervios espinales. Figura adaptada de [National cancer Institute](#).

Dada la naturaleza no regenerativa del tejido nervioso, es importante para el organismo que esté bien protegido frente a daños. La primera capa de protección es la columna vertebral, que aporta una fuerte defensa rodeando la médula espinal con tejido óseo. Las meninges son la segunda capa de protección, formada por tres membranas que revisten la médula espinal, ubicadas entre la caja ósea y el tejido nervioso: la duramadre (más externa), la aracnoides y la piamadre (más interna). Existe una tercera defensa contra daños mecánicos en el espacio subaracnoideo, entre la aracnoides y la piamadre, el cual contiene líquido cefalorraquídeo, que suspende el tejido.

En un corte transversal de la médula espinal se diferencian dos regiones (Figura 5). Internamente, la sustancia gris con forma de H está formada por dendritas y cuerpos neuronales, axones amielínicos y células gliales (macro- y microglia). La función de la sustancia gris es el procesamiento e integración de los estímulos percibidos. Alrededor de la sustancia gris, se encuentra la sustancia blanca, formada principalmente por axones mielínicos de neuronas, a los que debe su nombre. La mielina es una capa que envuelve los axones para que el impulso nervioso sea más rápido, formada fundamentalmente por lípidos, de ahí que esta región se vea de color blanco. Por ello, la función de la sustancia blanca es mantener la conducción de los impulsos nerviosos hasta y desde el cerebro [22, 25].

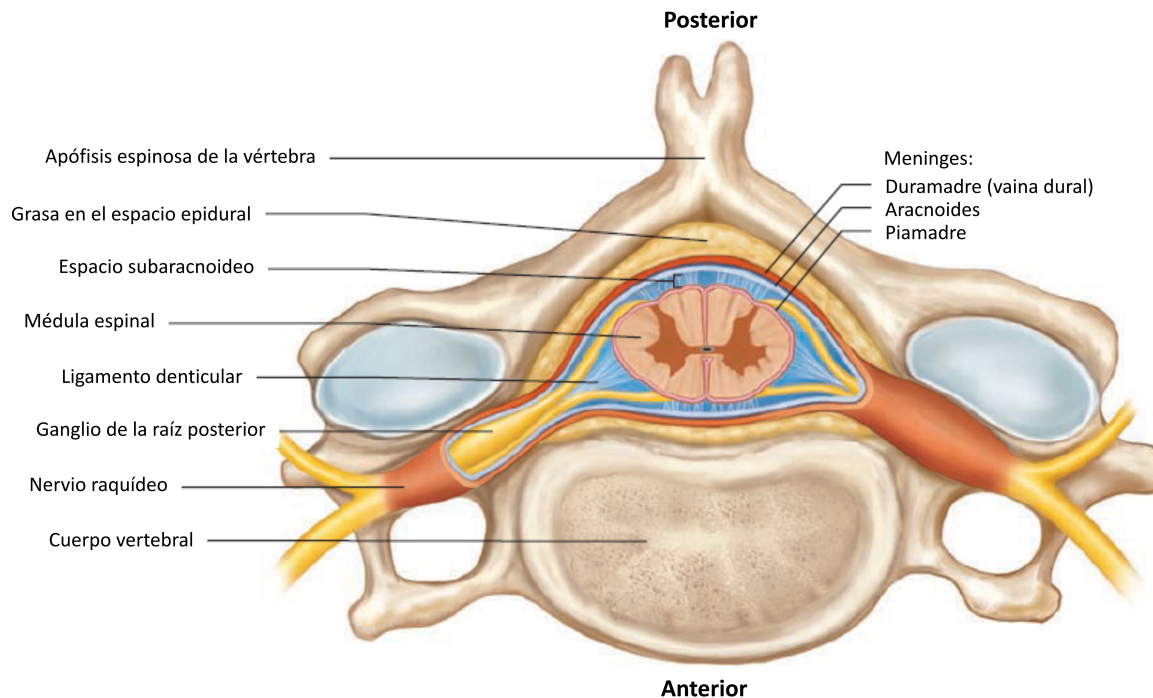


Figura 5. Anatomía macroscópica de la médula espinal y la vértebra vistas desde el plano transversa. Figura adaptada de Saladin [22].

1.1.4. La lesión medular en humanos

1.1.4.1. Grados de severidad

Dependiendo del segmento medular afectado, la LM se puede presentar como paraplejía, donde solo se ven afectadas las funciones sensoriales y/o motoras en las extremidades inferiores, o como tetraplejía, donde se ven afectadas las funciones de las cuatro extremidades. Por otro lado, la LM puede ser clasificada como lesión completa o incompleta (Figura 6). Las lesiones incompletas derivan en diferentes tipos de síndromes dependiendo de las regiones específicas de la médula espinal que se ven afectadas o preservadas (síndrome medular central, anterior, posterior o Brown-Séquard, entre otros). Por ejemplo, en el síndrome

medular central hay una mayor disfunción motora en las extremidades superiores que en las inferiores. Por otro lado, en el síndrome medular anterior se da una parálisis motora completa, pérdida de sensibilidad al dolor y la temperatura pero con preservación del tacto ligero y la propiocepción, mientras que en el síndrome medular posterior sucede lo contrario, hay pérdida de sensibilidad al tacto y propiocepción pero se conserva la función motora y sensibilidad al dolor y la temperatura [6]. La *American Spinal Cord Injury Association* (ASIA) propuso la escala de deterioro AIS (del inglés, *ASIA Impairment Scale*), que es el método más frecuente para clasificar la LM de acuerdo a la severidad y la prognosis, estableciendo una escala de cinco grados (A-E), de la más a la menos severa [26]. Se basa en gran parte en el concepto de preservación aa, de forma que las lesiones incompletas son aquellas en las que se mantiene en algún grado la sensación perineal, la contracción anal voluntaria y/o la flexión del dedo gordo del pie [6].

Grado A (lesión completa). No se conserva ninguna función motora o sensorial por debajo de la lesión, incluyendo los segmentos sacros S4–S5.

Grado B (lesión sensorial incompleta). La función sensorial, pero no la motora, se conserva por debajo del nivel neurológico de la lesión, incluidos los segmentos sacros distales (S4–S5).

Grado C (lesión motora incompleta). Se conserva la función motora por debajo del nivel neurológico de la lesión, y más de la mitad de los músculos clave por debajo del nivel neurológico tienen una fuerza muscular menor de 3. El grado 3 de fuerza muscular implica movimiento activo, rango completo de movimiento y antigraedad.

Grado D (lesión motora incompleta). Se conserva la función motora por debajo del nivel neurológico de la lesión, y al menos la mitad de los músculos clave por

debajo del nivel neurológico tienen una fuerza muscular de 3 o mayor.

Grado E (normal). La función motora y sensorial son normales.

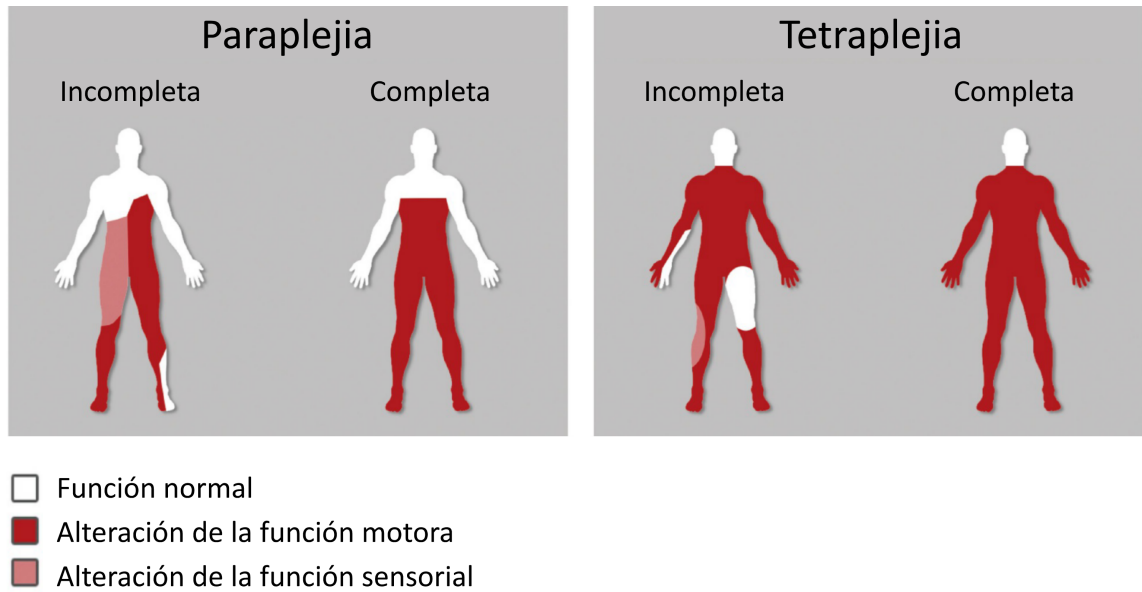


Figura 6. Comparación de la paraplejia/tetraplejia completa e incompleta. Figura adaptada de O'Phelan y De Jesus [27].

1.1.4.2. Evolución fisiopatológica de la lesión medular: fases temporales

Salvando las diferencias anatómicas y funcionales de cada segmento medular, la fisiopatología de la LM a nivel celular y molecular es común. La fisiopatología de la LM tiene dos fases, la lesión primaria y la lesión secundaria. La lesión primaria implica el daño mecánico inicial producido en la médula espinal. Este daño lleva a la rotura de los axones, vasos sanguíneos y membranas celulares, por lo que se interrumpe el impulso nervioso y el flujo sanguíneo, causando shock espinal (pérdida de todas las funciones neurológicas), hipotensión sistémica, vasoespasmo, isquemia (disminución del flujo sanguíneo), desbalance iónico y acumulación de neurotransmisores. Estos eventos desencadenan la lesión

secundaria, que comienza minutos después del trauma inicial y continua durante semanas o meses, causando un daño progresivo del tejido sano que rodea la lesión, aumentando su tamaño [28, 29].

Aunque en función del tiempo no hay una clasificación estricta o estandarizada, la fase secundaria puede ser dividida en cuatro fases: aguda (0-3 días), subaguda (4-14 días), fase intermedia o crónica temprana (15 días a 6 meses) y crónica tardía (>6 meses) [6, 30, 31].

A. Fase aguda (0-3 días). La fase aguda está marcada por un conjunto de alteraciones sistémicas y locales, que incluyen alteraciones de los mecanismos vasculares, desequilibrio iónico y una gran respuesta del sistema inmune.

La médula espinal está ampliamente vascularizada, especialmente la materia gris, por lo que al romperse los vasos hay una gran hemorragia que produce edema (retención de líquidos en los tejidos) e isquemia. Además, debido a la rotura de la barrera hematoencefálica, hay una entrada de células inflamatorias y citoquinas pro-inflamatorias, que raramente se ven en el SNC [31].

La combinación del daño mecánico celular con la isquemia hace que se pierda la homeostasis iónica, desregulándose las concentraciones de sodio, potasio, calcio, glutamato y otras moléculas. El glutamato es el principal neurotransmisor excitatorio del SNC, por lo que al liberarse al medio extracelular se produce una excitotoxicidad debido a la activación excesiva de sus receptores. La activación de estos receptores hace que entre calcio a las células y se acumule en el citosol o en diferentes orgánulos, poniendo en marcha mecanismos de apoptosis mediante la activación de fosfolipasas, proteasas y fosforilasas [29–32].

En esta situación de estrés celular, se producen radicales libres que dañan las proteínas, ácidos nucleicos y lípidos. Los radicales libres reaccionan con la

membrana celular mediante peroxidación lipídica, alterando su estructura, por lo que hay un aumento de la permeabilidad de iones, favoreciendo la desregulación de la homeostasis iónica, que finalmente lleva a la inducción de la apoptosis [29–32].

Dada la gravedad de la lesión y el conjunto de daños producido, el sistema inmune responderá de manera muy contundente para intentar contener el daño y restablecer la homeostasis. En esta respuesta inmune, la inflamación juega un papel clave, siendo un proceso de gran complejidad que incluye a múltiples poblaciones celulares y moléculas mediadoras intercelulares. Los mecanismos de inflamación empiezan inmediatamente y algunos persisten hasta varios meses después. Tras el trauma, la microglía residente se activa mientras los linfocitos infiltrados por la rotura de los vasos sanguíneos comienzan a liberar citoquinas y especies reactivas de oxígeno (ROS). Tanto la microglía activada como los linfocitos liberan moléculas pro-inflamatorias como citoquinas IL-6, IL-1, interferones o TFN- α , permitiendo una mayor extravasación de linfocitos y mayor daño tisular. En respuesta a estas moléculas llegan neutrófilos, que aumentan la cantidad de tejido dañado al liberar proteasas y generar ROS, elevándose por tanto la liberación de señales pro-inflamatorias. Los macrófagos se ven atraídos por estas señales y comenzarán a fagocitar el tejido dañado y a liberar moléculas pro-inflamatorias [29–32].

Como se ha visto, en la fase aguda hay una cascada de mecanismos moleculares que son consecuencia del traumatismo inicial y que se retroalimentan, agravando la lesión. La respuesta inflamatoria es muy importante para eliminar los desechos celulares y reparar los tejidos, sin embargo, una respuesta exagerada produce daño del tejido sano y exacerbación de la lesión [29–32]. La Figura 7 muestra de manera esquemática los procesos patológicos que suceden tras una lesión traumática durante la lesión primaria y la fase aguda.

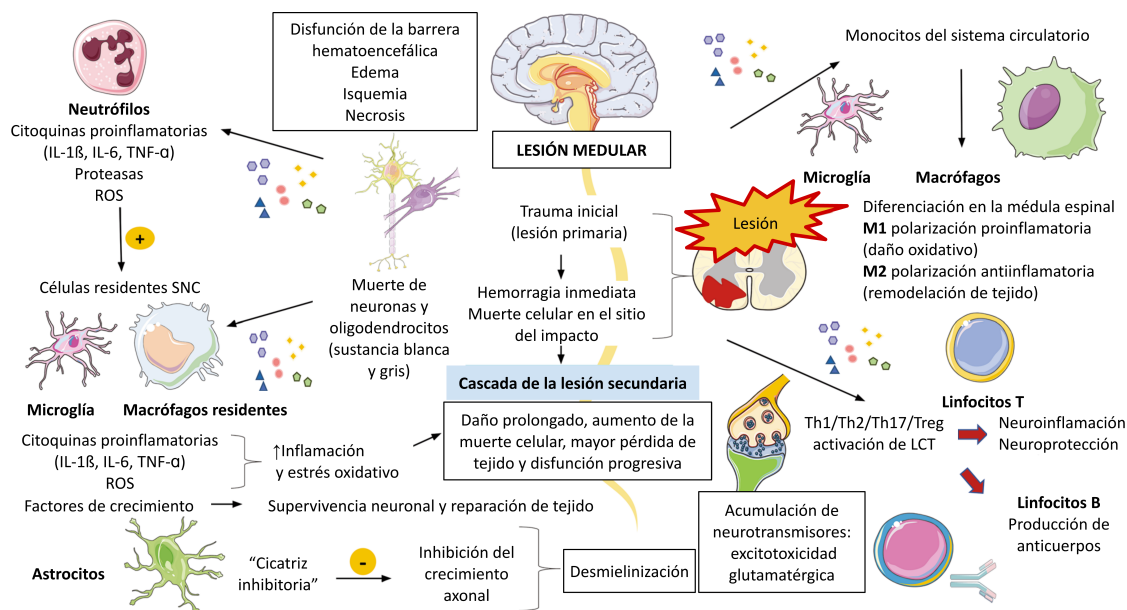


Figura 7. Visión general de eventos fisiopatológicos en la lesión primaria y secundaria tras un traumatismo. Figura adaptada de Ortega et al. [29].

B. Fase subaguda (4-14 días). Los astrocitos son células con numerosas funciones importantes para las neuronas, como soporte, aporte de nutrientes o eliminación de nutrientes. En la fase subaguda los astrocitos comienzan a crecer tanto por hiperplasia como por hipertrofia, es decir, aumentan en número y en tamaño alrededor de la lesión para formar una cicatriz inhibitoria [7]. Esto es beneficioso y perjudicial al mismo tiempo. La formación de la cicatriz permite separar la inflamación de los tejidos sanos para minimizar que se extiendan los daños secundarios descritos en la fase aguda. Sin embargo, en el futuro la cicatriz inhibitoria supondrá una barrera física y química que impedirá la regeneración axonal [6, 33, 34].

Además de los astrocitos, también hay una activación de diferentes poblaciones del sistema inmune, como la microglía residente y macrófagos y linfocitos infiltrados [6]. Tanto microglía como macrófagos pueden polarizar en los fenotipos M1 (pro-inflamatorio) como M2 (pro-regenerativo). Una polarización exacerbada

de M1 crea un ambiente neurotóxico que impide la neuroregeneración y puede contribuir a una extensión de la lesión secundaria. Por otro lado, el fenotipo M2 se asocia con un ambiente anti-inflamatorio que promueve un entorno favorable para la remodelación de tejido, crecimiento isla y la proliferación/diferenciación de progenitores de oligodendrocitos [29]. Por el contrario, rara vez se encuentran neutrófilos en la fase subaguda, por lo que su acción quedaría limitada a la fase aguda [28].

C. Fase intermedia o crónica temprana (15 días a 6 meses). Este es un periodo de transición hacia la fase crónica que se caracteriza por una maduración de la cicatriz inhibitoria. La disminución de la respuesta inflamatoria da lugar a un intento de remielinización y a reorganizaciones de la matriz extracelular, circuitos neuronales y vasculares, aunque insuficiente para producir una recuperación funcional significativa [6, 35].

D. Fase crónica tardía (>6 meses). Esta fase se extiende a lo largo de la vida de los pacientes y se caracteriza por la maduración y estabilización de la lesión y por tanto, de los déficits funcionales. Puede darse cierta recuperación funcional derivada de mecanismos endógenos de remielinización. Sin embargo, la regeneración resulta ineficaz dada la magnitud de señales inhibitorias, el persistente estado de inflamación y la falta de factores de apoyo en el ambiente celular [6, 28, 36]. En la Figura 8 se representa de manera esquemática una comparación entre la médula espinal sana y la lesionada, en la cual se observan diferentes tipos celulares del sistema inmune, rotura y desmielinización de los axones y la formación de una cicatriz inhibitoria, entre otros procesos.

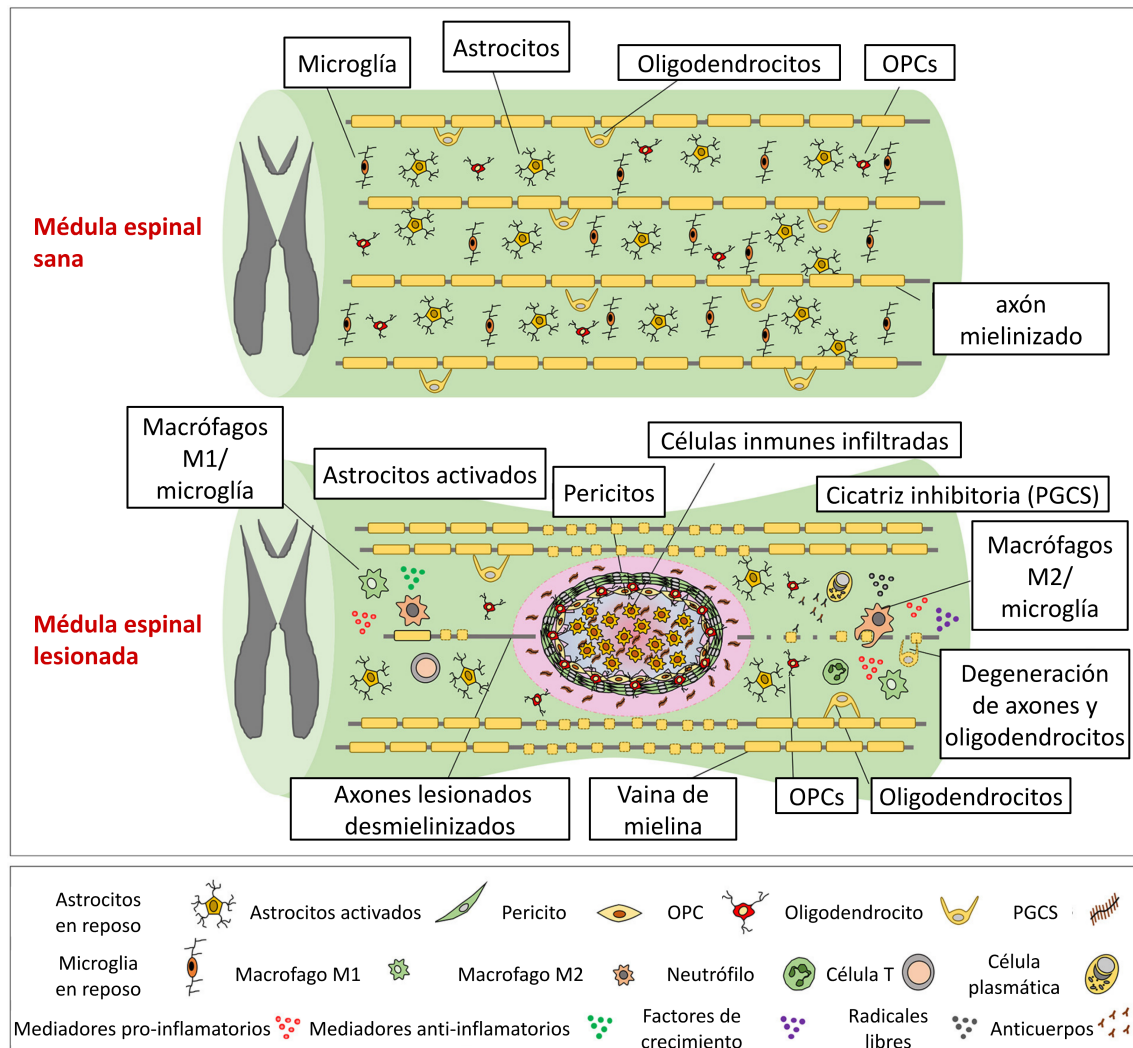


Figura 8. Comparación de la composición estructural y celular de la médula espinal sana y lesionada. OPCs, *Oligodendrocyte progenitor cells*; PGCS, proteoglicanos de condroitin sulfato. Figura adaptada de Alizadeh et al. [28].

1.1.5. La rata como modelo de estudio

Los modelos animales en LM históricamente han incluido gatos, perros, cerdos, primates no humanos, ratones y ratas [32, 37]. Sin embargo, la rata se ha consolidado como el modelo experimental más utilizado en la investigación básica de la LM, habiendo contribuido enormemente al entendimiento de su fisiopatología

y al desarrollo de aproximaciones terapéuticas. El uso de ratas tiene múltiples ventajas como su fácil manipulación, bajo coste y la existencia de protocolos bien establecidos y estandarizados para realizar las lesiones y para el estudio funcional motor y sensitivo [37, 38].

Por otro lado, la anatomía y las respuestas fisiopatológicas a la LM de las ratas se asemejan a la humana. Por ejemplo, al igual que los humanos las ratas desarrollan cavidades quísticas en la lesión, lo que no ocurre en ratones. La mayoría de lesiones vistas en humanos pueden ser replicadas en ratas, incluyendo LM completas e incompletas a diferentes niveles [37]. Existen diferentes modelos experimentales que permiten generar tipos de lesión clínicamente relevantes, como la lesión por contusión, compresión, hemisección o transección, entre otros [39].

Las lesiones por contusión son muy utilizadas ya que modelan con bastante fidelidad las lesiones producidas por impactos en humanos. Se realizan con un dispositivo que deja caer un peso. De esta manera, se pueden originar lesiones de mayor o menor severidad ajustando la fuerza con la que impacta el peso en la columna Figura 9. Las lesiones por compresión también imitan la patología observada en humanos. Se realizan con pinzas o clips, generando lesiones más severas prolongando el tiempo o la fuerza de compresión.

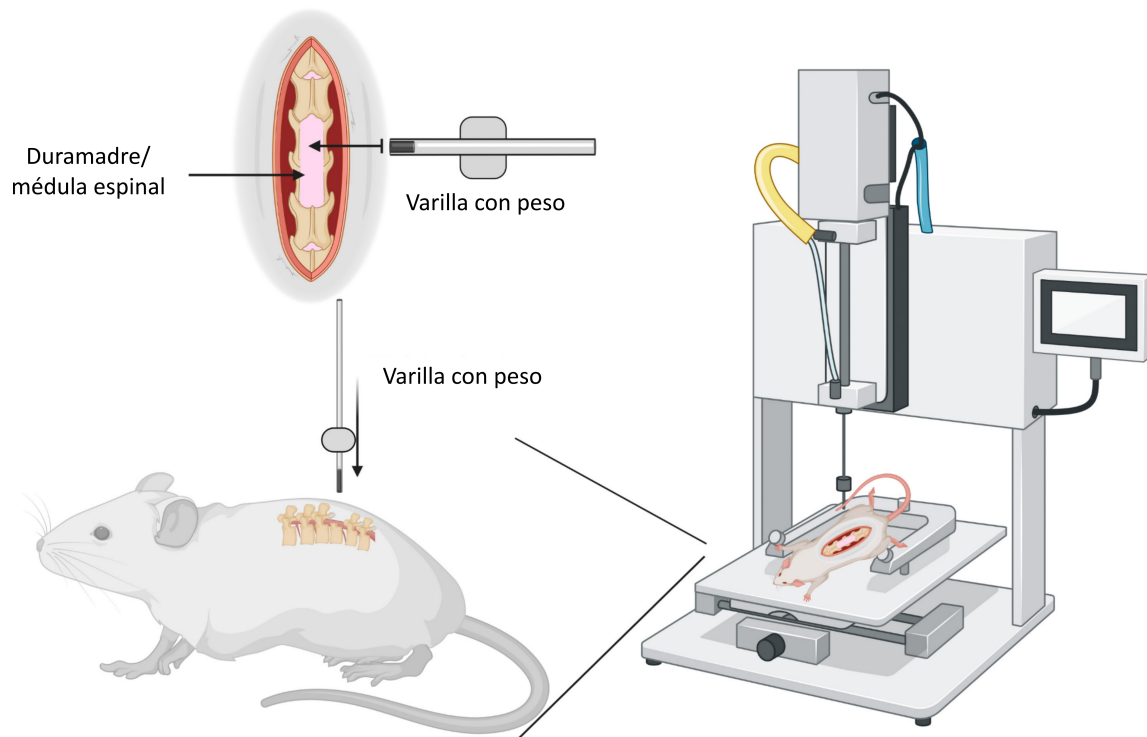


Figura 9. Esquema del dispositivo de contusión mecánica. El peso cae sobre la médula espinal expuesta con una fuerza establecida. Figura adaptada de Ridlen et al. [40].

En las lesiones por transección se corta completamente la médula o se retira completamente parte de ella. Las lesiones por transección son poco frecuentes en humanos, aunque es útil para estudiar estrategias que permitan la regeneración y unión de los axones. A veces también se utiliza una hemisección lateral porque se preserva la integridad y funcionalidad de uno de los lados, lo que permite una mayor supervivencia de los animales, siendo más útil a la hora de estudiar los mecanismos de recuperación funcional.

En cuanto a la duración de las fases temporales, las fases agudas serían similares a las establecidas en humanos, siendo la fase aguda menor a 7 días y la subaguda de 1-2 semanas [41]. En las fases crónicas varían los tiempos ya que estas ratas tienen una esperanza de vida de unos 3 años y una mayor capacidad de

recuperación, alcanzando la fase crónica en aproximadamente un mes, en lugar de en 6 como en humanos [42].

Sin embargo, es importante destacar que existen diferencias que impiden la extrapolación directa de resultados a humanos, como el tamaño corporal, la señalización molecular y variaciones anatómicas, además de los mecanismos intrínsecos de recuperación funcional. Por ello, la investigación en modelos animales más grandes, como primates no humanos, previos a la realización de estudios traslacionales es muy recomendable, a pesar de que no todos los ensayos clínicos actuales se han trasladado desde estudios en primates [41, 43].

1.2. Transcriptómica

El ADN codifica toda la información necesaria para determinar las propiedades y funciones de cada célula. Para ello, la información contenida en el ADN se transcribe, sintetizando moléculas de ARN mensajero (ARNm), que posteriormente serán traducidos para generar proteínas, las cuales son las ejecutoras de los procesos celulares. Al conjunto de ARNm de un tejido o tipo celular, en un momento dado, se le denomina transcriptoma. Disponer de herramientas que permiten detectar y cuantificar simultáneamente todos los transcritos presentes en las células ha supuesto un enorme avance en el conocimiento de muchas enfermedades y de la biología en general [44].

Existen dos principales tecnologías de análisis de expresión génica, los microarrays y el RNA-seq (del inglés, *RNA sequencing*). Los microarrays fueron los primeros en desarrollarse, siendo la tecnología dominante y más popular durante la primera década de este siglo. Consisten en un chip con sondas fijadas de ADN de una sola hebra, complementarias a los transcritos preestablecidos que se van a medir. Estos son marcados con fluorocromos o con radioactividad, por lo que al hibridar con su sonda complementaria pueden detectarse y cuantificarse mediante análisis de imagen. En este caso la cuantificación es relativa, ya que se comparan los niveles de expresión de los genes de dos muestras biológicas diferentes [45, 46].

El RNA-seq empezó a ganar popularidad durante la segunda década hasta llegar a establecerse como técnica predominante para cuantificación de la expresión génica. Se basa en la secuenciación masiva de los transcritos, lo que le confiere ciertas ventajas, como una mayor resolución o la capacidad de identificar nuevas isoformas, ya que no se limita a identificar transcritos preestablecidos. Sin

embargo, hay que destacar que los microarrays siguen siendo una metodología utilizada al requerir un menor coste económico y computacional, por lo que la elección de una u otra tecnología depende de los objetivos concretos del estudio y los gastos que se pueden asumir [45, 46].

Actualmente están empezando a usarse cada vez más RNA-seq de célula única o scRNA-seq (del inglés, *single-cell RNA sequencing*), con el que se puede estudiar el transcriptoma de las células individuales que forman un tejido, lo que permite identificar, cuantificar y caracterizar los diferentes tipos celulares existentes. Esto representa una ventaja con respecto al RNA-seq tradicional, en el que se observa la expresión promedio de cada gen en todo el tejido. Sin embargo, el scRNA-seq también tiene importantes limitaciones, ya que el gasto económico y computacional es mucho más elevado, además de generar datos más ruidosos y menos sensibles, que dificulta su tratamiento y análisis estadístico [47, 48]. Los primeros estudios de LM traumática utilizando scRNA-seq se publicaron a partir de 2020, usando como modelo de estudio el ratón. Hasta 2022 no se publicó el primer estudio de scRNA-seq de LM en rata, lo que ilustra la novedad y el auge de esta nueva tecnología [49].

La tecnología que está emergiendo más recientemente es la transcriptómica espacial, que mediante la combinación de scRNA-seq con otros métodos para estudiar la disposición y expresión génica de los diferentes tipos celulares en un tejido, con resolución a nivel de célula única. Aunque todavía se están poniendo a punto la eficacia de las técnicas y análisis de transcriptómica espacial, promete ser una herramienta muy potente para comprender la compleja heterogeneidad celular y su dinámica espaciotemporal en la LM [49, 50].

1.2.1. Estrategías de análisis transcriptómicos

1.2.1.1. Análisis de expresión diferencial

El análisis de expresión diferencial es la técnica por antonomasia en estudios transcriptómicos. Su objetivo es comparar el nivel de expresión de los genes entre dos o más grupos experimentales, (p. ej. enfermo vs control) para detectar diferencias significativas en genes que se *up/down*-regulan cuando se pasa de un estado a otro.

La obtención de listas de genes diferencialmente expresados es muy útil a la hora de encontrar biomarcadores que permitan, por ejemplo, diagnosticar una enfermedad o predecir cuál será su progresión. De la misma manera, estos genes también pueden ser dianas para el diseño de nuevas estrategias terapéuticas. Con un análisis de expresión diferencial también se genera un perfil o firma transcripcional, el cual es característico de cada sistema biológico. Esto facilita su comparación con otros sistemas de interés, lo que resulta interesante, por ejemplo, para observar la evolución de las diferentes fases de una enfermedad [51–53].

1.2.1.2. Caracterización funcional

Los análisis de caracterización o enriquecimiento funcional son un paso habitual tras realizar un análisis de expresión diferencial. El objetivo es identificar en qué procesos y funciones celulares están implicados los genes desregulados en el estudio, lo que resulta de enorme utilidad a la hora de comprender un sistema biológico y generar hipótesis que guíen nuevas investigaciones.

Estas funciones hacen referencia a grupos de genes relacionados por propiedades biológicas de interés, es decir, son conjuntos de genes que comparten anotaciones extraídas de cualquier base de datos, por ejemplo, las rutas metabólicas donde participan [54].

Existen diferentes métodos de enriquecimiento funcional. A continuación tarán tres de los más habituales y que han sido utilizados en este trabajo: análisis de sobrerrepresentación (ORA, del inglés *over-representation analysis*), análisis de grupos de genes (GSA, del inglés *gene set analysis*) y análisis de redes de interacción proteína-proteína (PPI, del inglés *protein-protein interaction*) (Figura 10).

Los ORA forman parte de la primera generación de métodos de enriquecimiento funcional. Como *input* se parte de una lista de genes de interés, por ejemplo, los genes diferencialmente *up* o *down*-regulados. Después, para cada función, se compara la frecuencia de genes que tienen dicha anotación en la lista de interés con la frecuencia de genes que tienen la anotación en un conjunto de referencia, que suele ser el genoma completo. Aplicando un test estadístico (test exacto de Fisher, hipergeométrico, Chi-cuadrado, etc.), se evalúa si la proporción de genes que realizan una función dada es mayor de lo que cabría esperar por azar, en tal caso, se dice que esa función está sobrerrepresentada o enriquecida. Finalmente, dado que para cada función se hace un test, se realiza un ajuste de p-valores [55].

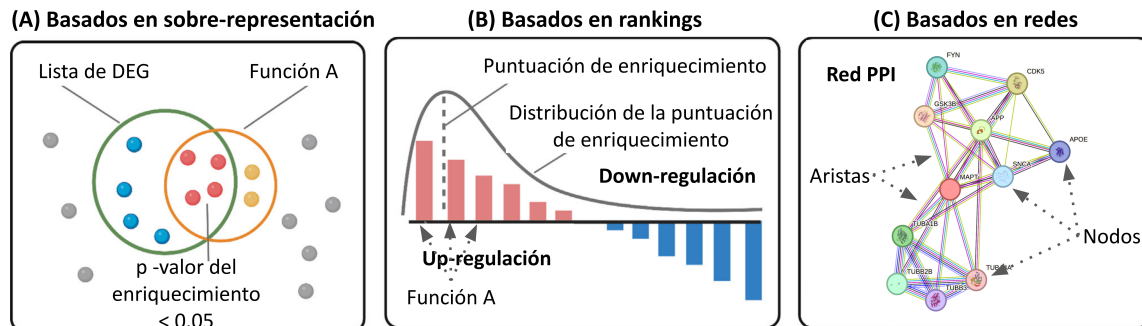


Figura 10. Resumen de tres tipos de métodos de caracterización funcional. (A) Los métodos basados en la sobre-representación examinan si alguna ruta está más presente en una lista de genes de interés de lo que cabría esperar por azar en comparación con un conjunto de referencia (DEG, *differentially expressed genes*). **(B)** Los métodos basados en *rankings* primero se ordenan todos los genes en base a su expresión (p. ej. de mayor a menor logFC) y luego calcula si los genes de una misma función tienden a agruparse en la parte superior o inferior de la lista ordenada. **(C)** En los métodos basados en redes se construye la red en base a las interacciones (físicas o funcionales) de los nodos, lo que permite su posterior análisis topológico y funcional (PPI, *protein-protein interaction*). Figura adaptada de Zhao et al. [56].

Posteriormente, surgió una segunda generación de métodos de enriquecimiento funcional diseñados para solventar las limitaciones asociadas con los métodos ORA. Por un lado, los resultados se pueden ver sesgados ya que el criterio de selección de genes es arbitrario. Por otro lado, también se asume que los genes son independientes y aportan el mismo peso a un proceso biológico [56, 57]. En los métodos GSA se utilizan todos los genes disponibles en el experimento, y además se les da un peso diferente a cada uno de ellos en base de una puntuación biológica, como puede ser el logFC o el p-valor del análisis de expresión diferencial. De esta manera se define un *ranking* de los genes de acuerdo a su expresión diferencial, teniendo en cuenta la dirección del cambio de expresión, es decir, si los genes están *up* o *down*-regulados. Esto se acercaría más a la realidad ya que se esperaría que módulos de genes que trabajan en una misma función celular tuvieran un cambio de expresión coordinado. Al igual que en los ORA, existen diferentes estrategias para determinar si un módulo de genes

tiende a situarse en la parte superior o inferior del *ranking*, por ejemplo, mediante un test de Kolmogorov-Smirnov o usando una regresión logística [57, 58].

Otro de los análisis funcionales más comunes en bioinformática son los análisis de redes PPI. En las redes PPI, los nodos representan las proteínas, mientras que las aristas que los unen son las interacciones entre ellas. STRING es la base de datos por antonomasia para llevar a cabo este tipo de análisis, de hecho, tan solo su versión 11 cuenta con más de 14 000 citas [59]. Para realizar un análisis de redes PPI primero hay que seleccionar una lista de genes de interés. Después, con ellos se construye una red a partir de las interacciones disponibles en una base de datos, que en el caso de STRING pueden ser tanto interacciones físicas como funcionales. Una vez construida, se obtiene un p-valor de enriquecimiento de la red con el que se determina si tiene más conexiones de las que cabría esperar por azar para ese mismo número de nodos, lo que sugeriría la relación funcional entre las proteínas evaluadas. Para identificar las funciones que está llevando a cabo la red, STRING también realiza un ORA a partir de la lista de genes introducida, usando diferentes bases de datos como referencia. Por otro lado, también es interesante llevar a cabo un análisis topológico de la red. STRING permite identificar *hubs*, es decir, grupos de proteínas con un número muy alto de conexiones entre ellas, en comparación con otras de la red. De esta manera, realizando un ORA de las diferentes subredes se pueden identificar los papeles funcionales particulares en los que intervienen cada una de ellas [60–62].

Considerando todo lo anterior, la elección de un método u otro dependerá del contexto y los objetivos a cumplir. Cuando se quiere caracterizar funcionalmente un perfil transcriptómico completo, un GSA ofrecerá resultados más potentes y precisos, que los generados por un ORA. Sin embargo, cuando se pretende caracterizar listas de genes que han sido seleccionadas por otros medios diferentes a un análisis de expresión diferencial u otro en los que no hay un

ranking, aplicar un ORA es una manera óptima de obtener información funcional. Además, si este va acompañado de un análisis de redes PPI se puede incrementar la comprensión de los procesos biológicos que se están llevando a cabo.

1.2.2. Metaanálisis de expresión génica

Desafortunadamente, las listas de genes diferencialmente expresados entre los distintos estudios pueden mostrar una alarmante falta de coincidencia cuando distintos laboratorios investigan el mismo sistema biológico. Este hecho se explica por el pequeño tamaño muestral que se suele utilizar en este tipo de experimentos, las diferencias en los protocolos, instrumentalización y en última instancia, en las estrategias de análisis computacional y métodos estadísticos empleados [63–65].

Dicha falta de robustez y reproducibilidad puede llevar a resultados conflictivos o inconclusos. Por ello, estrategias basadas en revisiones sistemáticas y metaanálisis son fundamentales para resumir la evidencia de manera precisa y fiable. La revisión sistemática identifica y selecciona todos los posibles estudios relacionados por un mismo tema y diseño, mientras que el metaanálisis es la metodología estadística que permite combinar los resultados de los diferentes estudios individuales, incrementando la potencia estadística de esta estrategia integrativa [66].

En el contexto del estudio de la expresión génica, los *datasets* de diferentes estudios pueden combinarse mediante metaanálisis para definir firmas moleculares comunes e identificar biomarcadores con mayor precisión. Existen tres métodos principales para el metaanálisis de expresión génica: metaanálisis basado en la combinación del tamaño del efecto, metaanálisis basado en combinación de p-valores y metaanálisis basado en combinación de *rankings* [67].

A. Metaanálisis basado en la combinación del tamaño del efecto. En esta aproximación, se compara el efecto entre diferentes estudios. El efecto se define como la intensidad de un fenómeno, que en este contexto es la expresión diferencial entre dos grupos experimentales, más concretamente, el logaritmo de magnitud de cambio (logFC). Utilizando un modelo de efectos aleatorios, que asume que el tamaño del efecto varía entre estudios, se calcula el efecto combinado de todos los estudios para cada gen analizado, obteniéndose un valor medio de estos [67].

B. Metaanálisis basado en combinación de p-valores. En esta aproximación, los p-valores obtenidos en el análisis de expresión diferencial de los diferentes estudios se combinan en un único p-valor para cada gen. Esto se puede conseguir mediante diferentes métodos como los de: Wilkinson, Tippett, Fisher, Pearson y Stouffer [67, 68].

En estos métodos todos los estudios tienen la misma influencia, independientemente de su tamaño, aunque pueden ponderarse. Una importante desventaja de este método es que se pierde la dirección del patrón de expresión. Si un gen está significativamente sobreexpresado en algunos estudios y significativamente infraexpresado en otros, se puede obtener un p-valor combinado significativo independientemente de los patrones de expresión en los diferentes estudios. Una posible solución es seleccionar solo los genes que tienen la misma dirección de cambio, aunque puede suponer una pérdida de información [67].

C. Metaanálisis basado en combinación de *rankings*. Estos métodos combinan diferentes estudios basándose en algún estadístico que permita ordenar los genes, como el logFC. Tras el análisis de expresión diferencial los genes se ordenan del más infraexpresado al más sobreexpresado, para después combinar

los *rankings* de los estudios individuales y encontrar consistencias en los extremos del *ranking* mediante la aplicación de métodos no paramétricos [67].

Las ventajas de estos procedimientos son que permiten combinar datos muy heterogéneos y evitan *outliers* de estudios que puedan influenciar mucho los resultados. Sin embargo, son muy sensibles a la diversidad de varianzas entre *datasets*.

2. Justificación y objetivos

2.1. Objetivos generales

A pesar de que se han realizado numerosos estudios transcriptómicos en modelos preclínicos de LM, todavía existe una falta de consenso respecto a la descripción del perfil transcriptómico inducido por la lesión y su correlación funcional. El estudio comparativo entre lesiones de diversa gravedad y su evolución en el tiempo permitirá identificar con mayor profundidad procesos específicos favoreciendo la identificación de dianas terapéuticas más eficientes.

De esta manera, el principal objetivo de esta tesis es obtener perfiles transcripcionales consenso que permitan evaluar de forma robusta y precisa, los cambios de expresión en función del tiempo y la severidad en la LM.

2.2. Objetivos específicos

- Generar perfiles transcripcionales consenso que caractericen la fisiopatología de la LM traumática en modelos de lesión en rata.
- Identificar genes biomarcadores específicos de severidad y fase temporal.
- Caracterizar funcionalmente los perfiles transcripcionales mediante diferentes métodos.
- Validar experimentalmente potenciales genes biomarcadores.

- Explorar la potencial transferibilidad de los genes específicos de severidad identificados en rata para la prognosis de la lesión en muestras de pacientes humanos.
- Elaborar una herramienta web para explorar y compartir los resultados generados.

3. Materiales y métodos

La Figura 11 muestra el flujo de trabajo seguido. Todos los análisis bioinformáticos y estadísticos fueron realizados en R v.4.2.0. Tanto el código desarrollado para reproducir los análisis descritos en este trabajo como las tablas suplementarias están disponibles públicamente en Gitlab (<https://gitlab.com/Biorubens/meta-sci-app>).

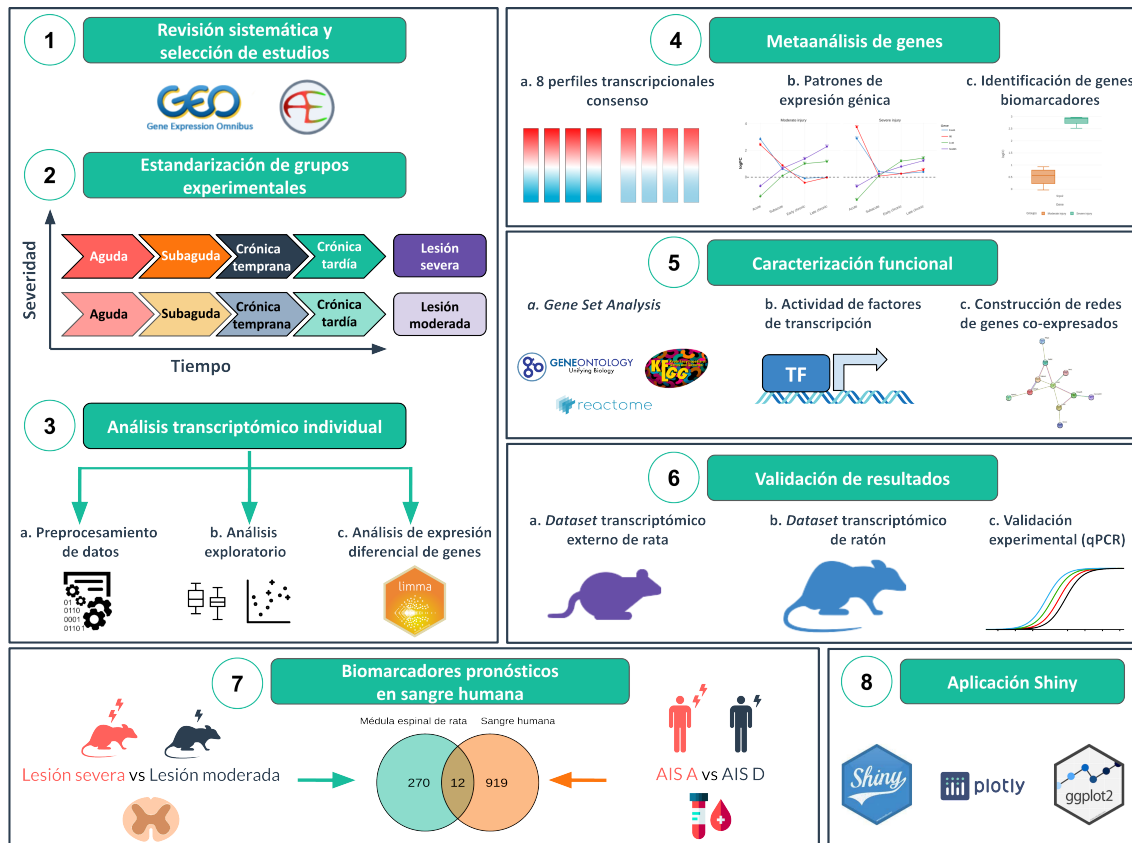


Figura 11. Flujo de trabajo. (1) Se realizó una revisión sistemática y selección de estudios en bases de datos públicas (GEO y ArrayExpress). (2) Los grupos experimentales se establecieron de acuerdo con la severidad y la fase de la lesión. (3) Para cada *dataset* se aplicó un *pipeline* común que incluye el preprocesamiento de datos, análisis exploratorio y análisis de expresión diferencial. (4) Se llevó a cabo un metaanálisis de expresión génica con el objetivo de obtener una firma de consenso de genes para cada comparación. Luego, estas firmas de consenso se utilizaron para (5) identificar biomarcadores específicos de severidad o fase y (6) para caracterizarlas funcionalmente. La caracterización funcional consistió en tres análisis diferentes: *Gene Set Analysis*, estimación de la actividad de factores de transcripción y construcción de redes de genes co-expresados. (7) Se compararon datos de expresión génica tanto de médula espinal de rata como de sangre periférica humana para explorar posibles biomarcadores comunes entre estos sistemas biológicos. (8) Finalmente, se desarrolló una herramienta web para consultar y explorar todos los resultados.

3.1. Revisión sistemática y selección de estudios

La revisión sistemática y selección de estudios se llevó a cabo durante septiembre de 2022. La búsqueda de estudios se realizó en dos bases de datos públicas de datos transcriptómicos, GEO [69] y ArrayExpress [70], siguiendo las directrices de *Preferred Reporting Items for Systematic Reviews and Meta-Analyses* (PRISMA) [71] y usando como palabras clave "spinal cord injury".

Los criterios de inclusión de estudios fueron:

- i Organismo: *Rattus norvegicus*.
- ii Tejido: segmento de médula espinal torácica.
- iii Tecnología de cuantificación de expresión génica: RNA-seq o microarrays de las plataformas Affymetrix o Agilent.

Los criterios de exclusión de estudios fueron:

- i Estudios sin muestras control (*sham* o sin lesión).
- ii Muestras individuales de ratas que habían recibido algún tratamiento.
- iii Muestras individuales rostrales y caudales al epicentro de la lesión.

3.2. Estandarización de grupos experimentales

Dada la diversidad de los estudios, los metadatos se estandarizaron para crear grupos experimentales con una nomenclatura uniforme.

Primero, las muestras se clasificaron en función de la severidad, distinguiéndose entre lesiones moderadas y severas. Se consideraron lesiones severas aquellas

resultantes por impacto en la médula espinal con una fuerza de 200 kdyn o mayor, las provocadas por un impacto de un peso de 10 g a 50 mm de altura o las provocadas por una transección completa de la médula (Tabla S1). Las lesiones que no reunían estos criterios se consideraron moderadas.

Por otro lado, se definieron cuatro fases en base al tiempo que pasa desde que se provoca la lesión hasta que se toman las muestras:

i) Fase aguda (T1): 0-3 días.

ii) Fase subaguda (T2): 4-14 días.

iii) Fase intermedia o crónica temprana (T3): 15-35 días.

iv) Fase crónica tardía (T4): >35 días.

Como grupo control se incluyeron muestras de ratas *sham* (solo se realiza una laminectomía) y ratas *naïve* (sin manipulación).

3.3. Análisis individual de estudios

Para cada uno de los estudios seleccionados seguimos un *pipeline* común que consiste en el preprocesamiento, análisis exploratorio y análisis de expresión diferencial.

3.3.1. Preprocesamiento

Para los estudios de microarrays, se descargaron los datos normalizados con el paquete GEOquery y las sondas fueron anotadas con sus correspondientes

gene symbol usando los paquetes de anotación específicos de cada plataforma. Para los estudios de RNA-seq, se descargaron manualmente las matrices de conteos de GEO y los identificadores se convirtieron a *gene symbol* con el paquete *org.Rn.eg.db*. En todos los casos, se calculó la mediana de la expresión para los identificadores duplicados.

3.3.2. Análisis exploratorio

Se realizó un análisis exploratorio que incluyó diagramas de caja, análisis de componentes principales (PCA) y análisis de *clustering*, para observar la distribución de las muestras en los grupos de interés, evaluar comportamientos anómalos y posibles efectos lote. Las muestras del estudio GSE183591 se tomaron en dos series diferentes, observándose una separación en dos bloques en el PCA. Este efecto de lote se tuvo en consideración posteriormente en el análisis de expresión diferencial. En los diagramas de caja para los *datasets* GSE2599 y GSE464 se identificaron valores negativos por lo que se sumó el valor mínimo de todos los valores y se realizó una transformación logarítmica en base 2 de los datos para modificar su escala.

3.3.3. Análisis de expresión diferencial

Dado que el objetivo es caracterizar la LM en cada una de sus fases y analizar la influencia de la severidad, en el análisis de expresión diferencial cada grupo lesionado se comparó con su correspondiente control dentro de cada estudio. En las plataformas de microarrays se utilizó el *pipeline* estándar de limma, mientras que los *datasets* de RNA-seq se analizaron con el *pipeline* de limma-voom [72], que incluye la eliminación de genes poco expresados y la transformación

logarítmica de la matriz de expresión. Los p-valores se corrigieron mediante el método de Benjamini-Hochberg (BH) [73].

Se realizaron un total de ocho comparaciones. Sus abreviaciones usadas de aquí en adelante por simplicidad son:

1. **M1**: Fase aguda moderada [T1] vs. Control.
2. **M2**: Fase subaguda moderada [T2] vs. Control.
3. **M3**: Fase crónica temprana moderada [T3] vs. Control.
4. **M4**: Fase crónica tardía moderada [T4] vs. Control.
5. **S1**: Fase aguda severa [T1] vs. Control.
6. **S2**: Fase subaguda severa [T2] vs. Control.
7. **S3**: Fase crónica temprana severa [T3] vs. Control.
8. **S4**: Fase crónica tardía severa [T4] vs. Control.

3.4. Metaanálisis de expresión génica

Los resultados del análisis de expresión diferencial se integraron para obtener un perfil transcripcional consenso de cada comparación. Se llevó a cabo un metaanálisis para cada gen presente en al menos dos estudios de cada comparación. Para determinar la medida combinada del efecto se utilizó el modelo de efectos aleatorios propuesto por DerSimonian-Laird [74] e implementado en el paquete metafor [75]. Este método pondera la heterogeneidad de cada estudio e incorpora su variabilidad inherente en la estimación global de la medida del efecto. En consecuencia, los estudios con mayor variabilidad tendrán menor peso en esta

medida. En este caso, se utilizó logFC como medida del efecto, mientras que el error estándar se usó como medida de la variabilidad del logFC. Para cada gen evaluado en el metaanálisis se calculó el logFC, su intervalo de confianza (IC) al 95 % y p-valor que evaluó su significación. Dado que se realizaron múltiples metaanálisis, los p-valores se ajustaron utilizando el método BH, considerándose significativos aquellos menores a 0.05.

3.5. Identificación de genes biomarcadores

En este contexto, se definen genes biomarcadores como aquellos que muestran una mayor desregulación en ciertos grupos de interés con respecto a otros, permitiendo estratificarlos en base a su patrón de expresión.

Para identificar posibles genes biomarcadores primero se construyó una matriz con los valores de logFC de los perfiles transcripcionales consenso. Se seleccionaron aquellos genes evaluados en todos los grupos y significativos en al menos uno de ellos. Para identificar genes biomarcadores de severidad se hizo un test de limma comparando los cuatro perfiles de lesión severa con los cuatro de lesión moderada. Por otro lado, también buscamos genes biomarcadores de fase, tanto comunes a ambas severidades como específicos de lesiones moderadas o severas. La Tabla 1 muestra detalladamente las comparaciones realizadas en cada caso.

Posteriormente, se realizó un análisis de agrupamiento en la matriz de perfiles transcripcionales consenso para evaluar la capacidad de clasificación de los 10 genes con el menor p-valor en cada comparación (top 10 genes). Este análisis incluyó tanto PCA como *clustering* jerárquico. Para el PCA se utilizó el paquete mixOmics [76], mientras que para el *clustering* jerárquico se empleó la función

hclust con los argumentos por defecto, calculando previamente la distancia euclídea.

Tabla 1. Comparaciones aplicadas para la identificación de biomarcadores específicos.

Biomarcador	Comparaciones
Severidad	(S1, S2, S3, S4) vs (M1, M2, M3, M4)
T1	(M1, S1) vs (M2, M3, M4, S2, S3, S4)
T2	(M2, S2) vs (M1, M3, M4, S1, S3, S4)
T3	(M3, S3) vs (M1, M2, M4, S1, S2, S4)
T4	(M4, S4) vs (M1, M2, M3, S1, S3, S4)
M1	M1 vs (M2, M3, M4)
M2	M2 vs (M1, M3, M4)
M3	M3 vs (M1, M2, M4)
M4	M4 vs (M1, M2, M3)
S1	S1 vs (S2, S3, S4)
S2	S2 vs (S1, S3, S4)
S3	S3 vs (S1, S2, S4)
S4	S4 vs (S1, S2, S3)

T1 (fase aguda); **T2** (fase subaguda); **T3** (fase crónica temprana); **T4** (fase crónica tardía); **M1** (fase aguda moderada); **M2** (fase subaguda moderada); **M3** (fase crónica temprana moderada); **M4** (fase crónica tardía moderada); **S1** (fase aguda severa); **S2** (fase subaguda severa); **S3** (fase crónica temprana severa); **S4** (fase crónica tardía severa).

3.6. Caracterización funcional

3.6.1. *Gene Set Analysis*

Para cada firma transcripcional consenso se realizó un análisis de enriquecimiento funcional con el método GSA implementado en el paquete mdgsa [77]. Como *input* el GSA necesita un *ranking* de genes (una lista de genes ordenados según alguna medida objetiva) y una anotación funcional (asociación entre funciones y genes). El *ranking* se obtiene ordenando todos los genes por su p-valor del metaanálisis y el signo del logFC. Como anotaciones se usaron los procesos biológicos de la Gene Ontology y las rutas de KEGG y Reactome [78–80]. Se eliminaron las funciones compuestas por menos de 10 o más de 500 genes. Los p-valores se corrigieron con el método BH.

3.6.2. Estimación de la actividad de factores de transcripción

La identificación de factores de transcripción (TF, del inglés *transcription factor*) diferencialmente activos se llevó a cabo utilizando la función *msviper* del paquete *viper* [81], usando como *input* cada uno de los perfiles transcripcionales consenso. Se utilizaron los regulones (asociación TF-gen diana) de ratón disponibles en el paquete Dorothea [82]. Esta metodología evalúa si las interacciones TF-diana activan o inhiben a un determinado gen (Tabla S2). Se seleccionaron aquellas interacciones con un nivel de confianza de A, B, C o D, excluyéndose los regulones que tuvieran menos de 10 genes diana. Los p-valores se corrigieron con el método BH.

3.6.3. Construcción de redes de genes co-expresados

Con el objetivo de encontrar genes con patrones de expresión similares se utilizó el software *clust* [83] con los siguientes argumentos: `-n 0 -t 100 -cs 4`. Como *input* se utilizaron los perfiles transcripcionales consenso de lesiones moderadas y severas por separado. A continuación, se construyó una red PPI para cada *cluster* de genes co-expresados obtenidos por el software *clust*. Para ello se utilizó el paquete STRINGdb con las interacciones de la versión 11.5 de STRING [84] y una confianza media de 0.400. Después, se aplicó el algoritmo *fastgreedy* implementado en la función *get_clusters* en cada red PPI para obtener redes locales de genes más fuertemente y funcionalmente relacionados entre sí.

Posteriormente, las subredes obtenidas se clasificaron en categorías funcionales de interés manualmente predefinidas, que incluyen: "Cardiovascular system", "Cell cycle, DNA repair/replication", "Cell death", "Cell-cell communication", "Cellular/Organelle organization", "Chemical homeostasis", "Cytoskeleton", "Developmental biology", "Endocrine system", "Extracellular matrix", "Immune system", "Metabolism", "Mitochondrial function", "Nervous system", "Protein metabolism and translation", "Signal transduction", "Transcription/RNA metabolism". Para ello, primero se utilizó la función *get_enrichment* para obtener los términos funcionales sobrerrepresentados en cada subred. A continuación, se seleccionaron los términos pertenecientes a las bases de datos de Gene Ontology, KEGG y Reactome y se asignaron a las categorías generales preestablecidas *ad hoc*. Finalmente, para cada subred se calculó la distribución de frecuencias de cada categoría y se le asignó la categoría con mayor frecuencia como aquella que mejor define la funcionalidad de la red.

3.7. Validación de resultados

Se realizó una validación de los resultados utilizando métodos computacionales y experimentales. La validación computacional empleó datos transcriptómicos de rata y ratón, mientras que la validación experimental se llevó a cabo mediante qPCR en un modelo de lesión medular en ratas. A continuación, se describen los detalles de cada método.

3.7.1. *Dataset* transcriptómico de rata

Se utilizó el *dataset* GSE218088 [85] para la validación computacional de los resultados obtenidos, ya que se publicó después del periodo de búsqueda y selección de estudios y reunía los criterios de inclusión establecidos en el metaanálisis. Se descargó la matriz de expresión normalizada utilizando GEOQuery, se anotaron las sondas con el identificador *gene symbol* y se seleccionaron los genes comunes a los de la matriz transcripcional consenso. Después se realizó un análisis de expresión diferencial, siguiendo el *pipeline* de limma, comparando el grupo lesionado con el grupo control.

Para evaluar si los genes biomarcadores de severidad y fase seleccionados eran capaces de clasificar correctamente un nuevo perfil transcriptómico, se creó una nueva matriz que combina el perfil transcripcional obtenido tras el análisis de expresión diferencial con la matriz de perfiles transcripcionales consenso. Finalmente, se llevó a cabo un análisis de agrupamiento (PCA y *clustering* jerárquico, ver sección 3.5) utilizando tanto el top 10 de genes específicos de severidad, como una combinación de genes biomarcadores de fase (top 10 de T1, T2 y T4).

3.7.2. *Dataset* transcriptómico de ratón

La matriz de expresión normalizada aportada por Li et al. [86] se descargó de la siguiente dirección: <https://doi.org/10.6084/m9.figshare.17702045>. Se seleccionaron las muestras control, así como muestras representativas de las cuatro fases temporales establecidas en nuestro trabajo: 1, 7, 28, y 42 días post-lesión (dpl) (Tabla S3). Para el análisis de expresión diferencial se seleccionaron los genes comunes a los de la matriz transcripcional consenso y se aplicó el mismo *pipeline* de limma utilizado para los estudios individuales, comparando cada grupo lesionado con el grupo control.

Los perfiles transcripcionales obtenidos en cada comparación se unieron en una matriz para poder evaluar los patrones de expresión génica. Para cada gen, se realizó un análisis de correlación de Pearson para comparar los perfiles de expresión en ratón con los obtenidos en el metaanálisis para lesiones moderadas y severas.

Para evaluar si los genes biomarcadores de fase seleccionados en rata eran capaces de clasificar correctamente las fases temporales de ratón, se creó una nueva matriz combinando los perfiles transcripcionales obtenidos tras el análisis de expresión diferencial y la matriz de perfiles transcripcionales consenso. Finalmente, se llevó a cabo un análisis de agrupamiento utilizando un conjunto manualmente curado de potenciales genes específicos de fase (*Fosl1*, *Il6*, *Cfp*, *Sipa1*, *Cp*, y *Dbp*), y que presentan un patrón de expresión similar en los tres sistemas biológicos.

3.7.3. Validación experimental (qPCR)

Se usó un modelo de LM traumática severa en ratas hembras adultas de la cepa Sprague Dawley para la extracción de ARN y validación de resultados in silico mediante qPCR. Los animales se alojaron en la Unidad de Experimentación Animal del Centro de Investigación Príncipe Felipe (Valencia, España) bajo condiciones estándar. Todos los procedimientos experimentales se ajustaron a las directrices establecidas por la Directiva del Consejo de las Comunidades Europeas (86/609/ECC), el Real Decreto Español 53/2013, y el Comité de Cuidado y Uso de Animales del Centro de Investigación Príncipe Felipe (2021/VSC/PEA/0032).

25 ratas fueron pre-medicadas subcutáneamente con morfina (2.5 mg/kg) y anestesiadas con isoflurano al 2 % en un flujo continuo de oxígeno de 1 L/min. Se realizó una laminectomía en las vértebras torácicas 8-9 para exponer la médula espinal, y se indujo una LM severa en el nivel de la vértebra torácica 8 mediante contusión, aplicando una fuerza de 250 kdyn usando el Impactador Infinite Horizon, como se describió previamente [87–89].

El cuidado postquirúrgico incluyó el vaciado manual de la vejiga dos veces al día hasta la recuperación del reflejo vesical y la administración subcutánea de 5 mg/kg de Enrofloxacin (Alsir) durante 7 días, así como 0.1 mg/kg de buprenorfina dos veces al día durante 4 días después de cada intervención. Los animales fueron sacrificados con sobredosis de pentobarbital sódico (100 mg/kg) y perfundidos transcardialmente con una solución salina al 0.9 %, en días representativos de cada fase (T1, T2, T3 y T4): 1 dpl (n = 3), 7 dpl (n = 3), 28 dpl (n = 4), y 56 dpl (n = 4). Se extrajo el tejido de médula del epicentro de la lesión, que fue inmediatamente congelado en nitrógeno líquido y se almacenó a -80°C hasta su uso.

La extracción de ARN se realizó utilizando el método estándar TriZol, seguido de un paso adicional de limpieza usando RNeasy MinElute Cleanup (Qiagen, Alemania) para asegurar la calidad de la muestra ($A_{260}/_{280} \approx 2$ y $A_{260}/_{230} \geq 1.8$). Para la transcripción inversa se usó el kit RNA-to-cDNATM de alta capacidad (Applied Biosystems, Massachusetts, EE.UU.). Los cebadores específicos (Tabla S4) para cada gen de interés se diseñaron usando primer-BLAST (NCBI, Maryland, EE.UU.) y se validaron mediante curvas de eficiencia. La qPCR se realizó por triplicado usando AceQ SYBR qPCR Master Mix (Thermo Fisher) en el sistema de detección Light-Cycler 480 (Roche, Basilea, Suiza). Los datos de Ct se calcularon con el software de cuantificación relativa LightCycler 480 (Roche, Basilea, Suiza). Los niveles de ARNm de GAPDH sirvieron como control interno para la normalización.

3.7.4. Análisis de muestras de sangre humana

La matriz de expresión normalizada del estudio GSE151371 [90] se descargó de GEO y se realizó una transformación logarítmica en base 2 de estos valores. Se seleccionaron las muestras correspondientes a los grupos AIS A, AIS D y controles sanos para los posteriores análisis (Tabla S5). Se llevó a cabo un análisis de expresión diferencial comparando los grupos AIS A vs. AIS D, siguiendo el mismo *pipeline* de limma usado previamente. Finalmente, se seleccionaron los genes significativos ($FDR < 0.1$) para compararlos con la lista de potenciales biomarcadores de severidad en rata. Los genes que intersectan se usaron para el análisis de *clustering* en ambos sistemas biológicos.

3.7.5. Meta-SCI app

Meta-SCI app se ha desarrollado con el paquete shiny [91] y está desplegada en un servidor de shinyapp.io. Los gráficos son generados con los paquetes ggplot2, plotly y heatmaply [92–94]. Todo el procesamiento de datos y análisis se realizan con R. La aplicación está disponible en <https://metasci-cbl.shinyapps.io/metaSCI>.

4. Resultados

4.1. Obtención de perfiles transcripcionales consenso en función de la severidad y la fase de la lesión

La revisión sistemática y selección de estudio se realizó en dos bases de datos públicas para datos transcriptómicos: GEO y ArrayExpress. En el diagrama de flujo PRISMA se presenta la estrategia de selección de estudios (Figura 12). Introduciendo las palabras clave "spinal cord injury" se identificaron 62 estudios únicos en *Rattus norvegicus*. Tras aplicar los criterios de exclusión, se eliminaron 21 estudios porque las muestras no provenían de médula espinal o no fueron analizadas con las plataformas de interés. Además, se descartaron otros 21 estudios al no evaluar lesiones medulares. Tras la evaluación de texto completo, se excluyeron seis estudios más debido a la falta de muestras control, muestras sin tratamiento o información en los metadatos. Finalmente, se incluyeron un total de 14 estudios en el metaanálisis. En la Tabla 2 se detalla la información sobre los estudios seleccionados.

Tras descartar las muestras que habían recibido algún tratamiento o no se ajustaban a los grupos experimentales propuestos se seleccionaron un total de 273 muestras de 14 *datasets* para los análisis individuales. Dada la diversidad entre estudios, se estandarizaron los metadatos para crear grupos experimentales con una nomenclatura uniforme. En la Figura 13A se muestra el número de estudios por grupo, mientras que en la Figura 13B se indica el número de

muestras por estudio, severidad y fase. La Tabla S6 contiene los identificadores y anotaciones de cada muestra.

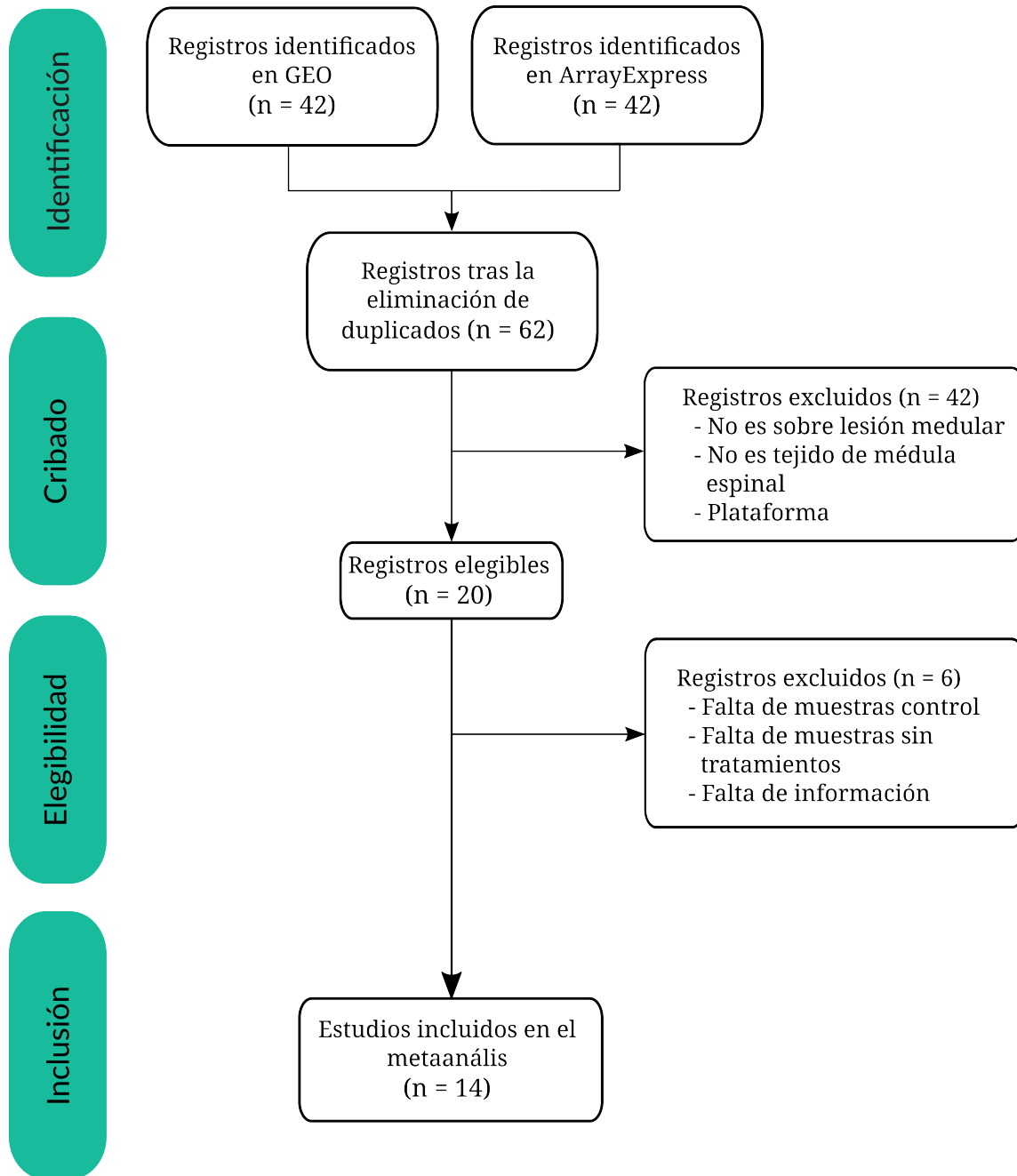


Figura 12. Diagrama de flujo de la revisión sistemática, siguiendo la declaración PRISMA.

Tabla 2. Estudios seleccionados en la revisión sistemática e incluidos en el metaanálisis. Se indica el identificador de GEO, la plataforma y la publicación asociada a cada *dataset*.

GEO ID	Plataforma	Año	Referencia
GSE183591	Agilent-011868 Whole Rat Genome Microarray	2022	[87]
GSE182796	Affymetrix Clariom S Assay, Rat	2021	No publicado
GSE138637	Illumina HiSeq 1500	2020	No publicado
GSE133093	Illumina NextSeq 550	2019	[95]
GSE115067	Illumina NextSeq 550	2018	[96]
GSE102964	Affymetrix Rat Gene 2.0 ST Array	2018	[97]
GSE104317	Affymetrix Rat Genome 230 2.0	2017	[98]
GSE93249	Illumina NextSeq 2000	2017	[99]
GSE52763	Affymetrix Rat Genome 230 2.0	2013	[100]
GSE45006	Affymetrix Rat Genome 230 2.0	2013	[101]
GSE46988	Affymetrix Rat Gene 1.1 ST	2013	[102]
GSE29488	Affymetrix Rat Genome 230 2.0	2012	[103]
GSE2599	Affymetrix Rat Genome U34	2005	[104]
GSE464	Affymetrix Rat Genome U34	2003	[105]

Posteriormente se llevó a cabo el análisis individual de estudios. Los resultados detallados de cada uno, incluyendo análisis exploratorio y análisis de expresión diferencial, pueden consultarse en el módulo "Individual analysis" de la herramienta Meta-SCI app. En la Tabla S7 se indica el número de genes diferencialmente expresados ($FDR < 0.05$) en cada estudio y en cada comparación, clasificados en función de su *up* o *down*-regulación.

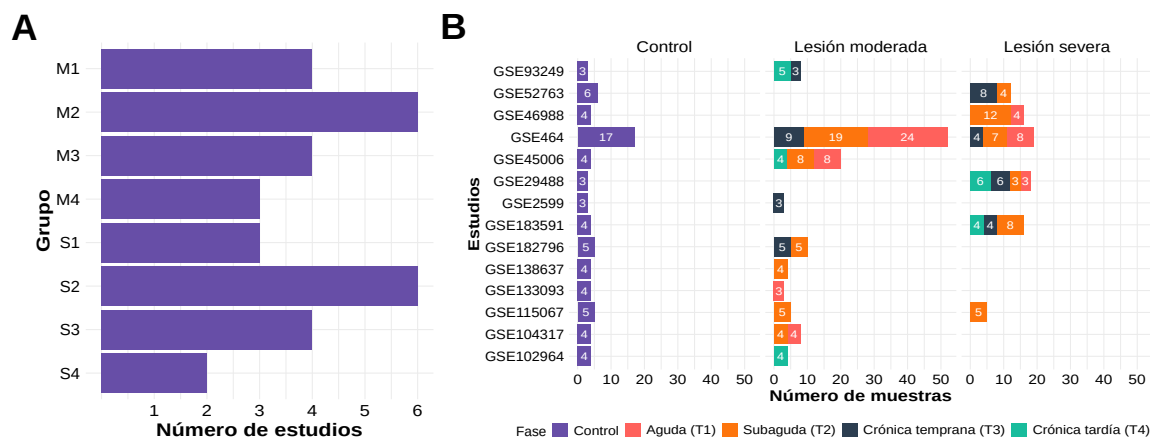


Figura 13. Distribución de estudios y muestras en los grupos experimentales. (A) Número de estudios por grupo. **(B)** Número de muestras por estudio, severidad y fase.

Los resultados del análisis de expresión diferencial se integraron llevando a cabo un metaanálisis para cada comparación, obteniendo así una firma transcripcional consenso de cada grupo experimental. Los diagramas de Venn de la Figura 14A indican el número de genes significativos ($FDR < 0.05$) en cada una de las fases para lesiones moderadas (izquierda) y severas (derecha). En rojo se indican los genes *up*-regulados, mientras que en azul se indican los *down*-regulados. El cuadro negro indica la intersección de genes significativos en las cuatro fases, sugiriendo que existe un núcleo común de genes permanentemente desregulados en todas las fases, tanto para lesiones moderadas como severas. El módulo "Meta-analysis" de la herramienta Meta-SCI app permite consultar detalladamente todos los resultados generados, así como generar diagramas de bosque y de embudo para cada gen.

Para obtener una visión general de las similitudes entre los grupos, realizamos un PCA y un análisis de *clustering* jerárquico de las firmas transcripcionales consenso. En conjunto, observamos una separación entre las lesiones severas y moderadas, con similitudes entre M1 y S1, lo que sugiere que la fase aguda es la fase temporal más distintiva independientemente de la severidad (Figuras 14B y 14C).

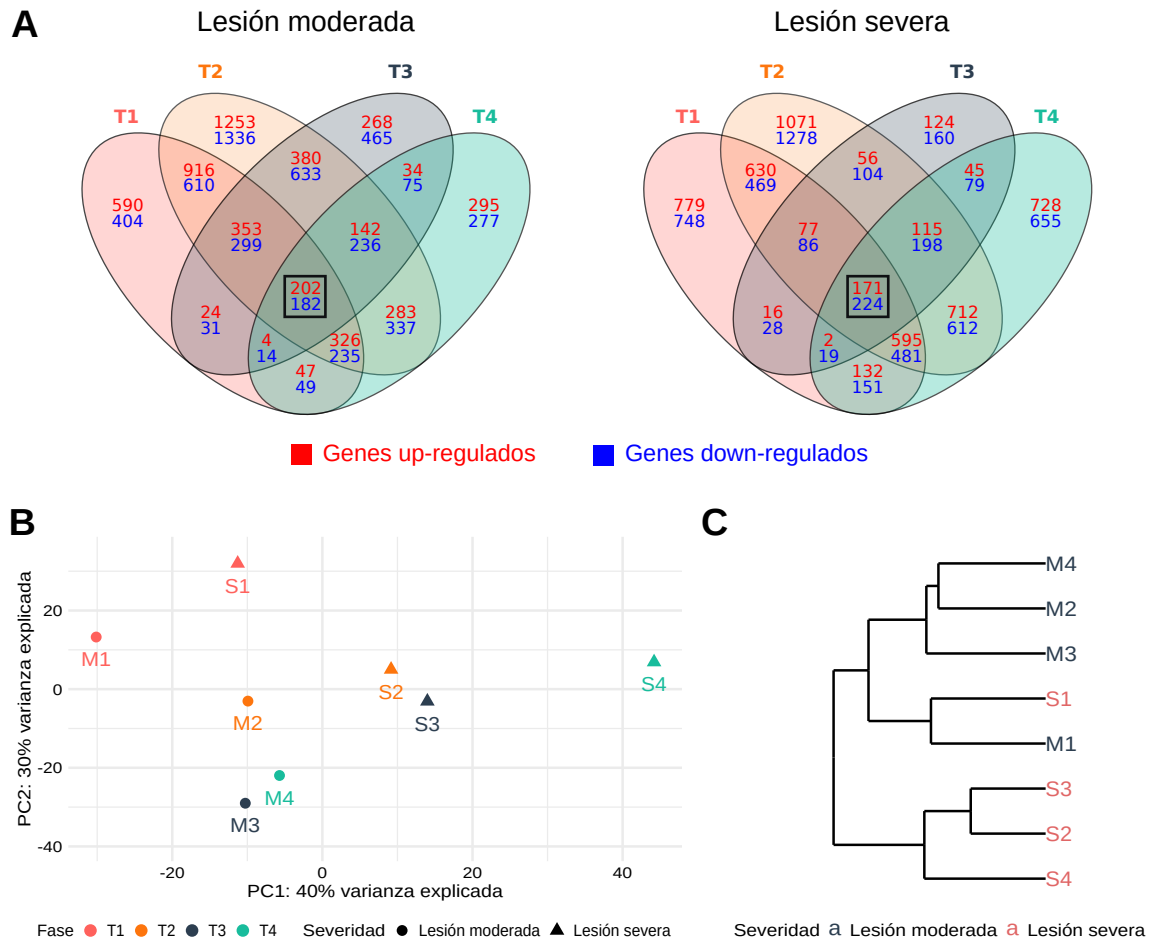


Figura 14. Resumen de resultados del metaanálisis de expresión génica. (A) Diagramas de Venn de genes significativamente alterados en diferentes fases de lesión (T1, aguda; T2, subaguda; T3, crónica temprana; T4, crónica tardía) para lesiones moderadas (izquierda) y severas (derecha). Los genes significativamente *up*-regulados se señalan en rojo, y los genes significativamente *down*-regulados se señalan en azul. El cuadrado negro indica la intersección de los cuatro grupos. **(B)** PCA y **(C)** *clustering* jerárquico de firmas transcripcionales consenso para todos los grupos.

4.2. Los biomarcadores específicos de severidad y fase identificados permiten una clasificación efectiva de los grupos

Para identificar genes biomarcadores específicos de severidad, comparamos los perfiles transcripcionales consenso de los cuatro grupos de lesión severa contra los cuatro grupos de lesión moderada (Tabla 1). Identificamos 282 genes con un $FDR < 0.1$ (Tabla S8). Al realizar un heatmap con *clustering* jerárquico con estos genes se observa una mayor desregulación (tanto en genes *up*-regulados como *down*-regulados) en lesiones severas comparado con lesiones moderadas (Figura 15), resultando en una clara separación del *clustering* jerárquico en dos ramas diferentes en función de la severidad.

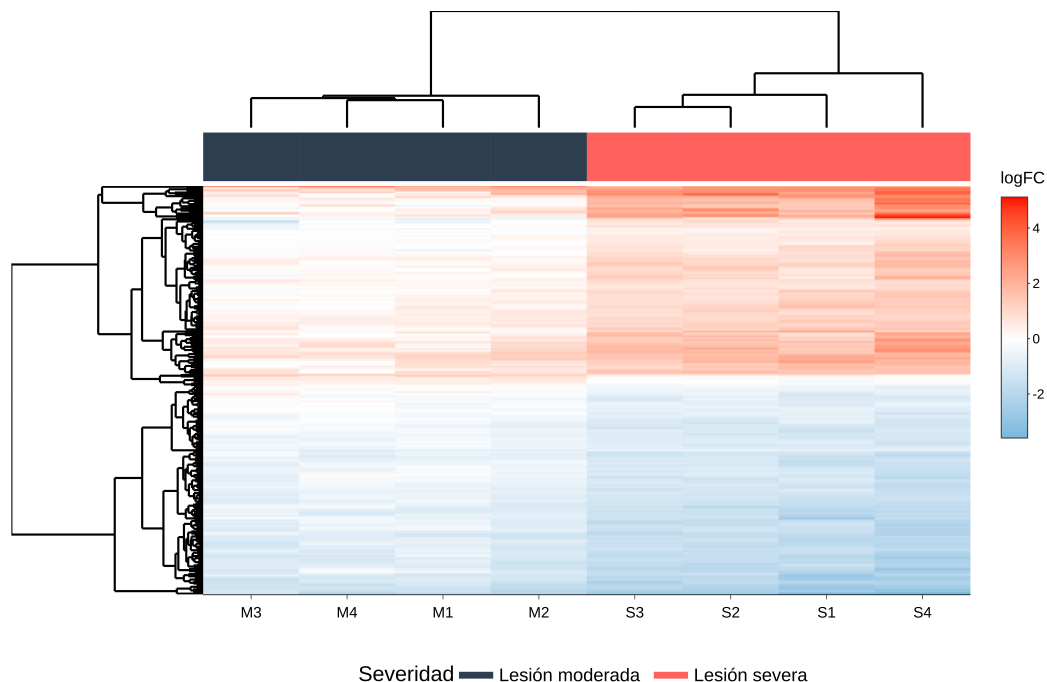


Figura 15. Heatmap con *clustering* jerárquico de biomarcadores específicos de severidad. El heatmap con *clustering* jerárquico muestra una mayor desregulación de estos genes en lesiones severas, tanto en los *up*-regulados como en los *down*-regulados, resultando en una clara separación en dos ramas en función de la severidad.

Las Figuras 16A y 16B ilustran que el top de 10 genes más significativamente alterados (*Srpx2*, *Hoxb8*, *Acap1*, *Snai1*, *Aadat*, *Ppic*, *Lrrc17*, *Map7*, *Actg2* y *H19*) permite una clara separación entre lesiones moderadas y severas, con la componente principal 1 (PC1) explicando un 93 % de la varianza. Los patrones de expresión de una selección de genes específicos de severidad demuestran una desregulación más consistente y sostenida en el tiempo en las lesiones severas, lo cual representa un patrón deseable para la prognosis de la lesión (Figura 16C).

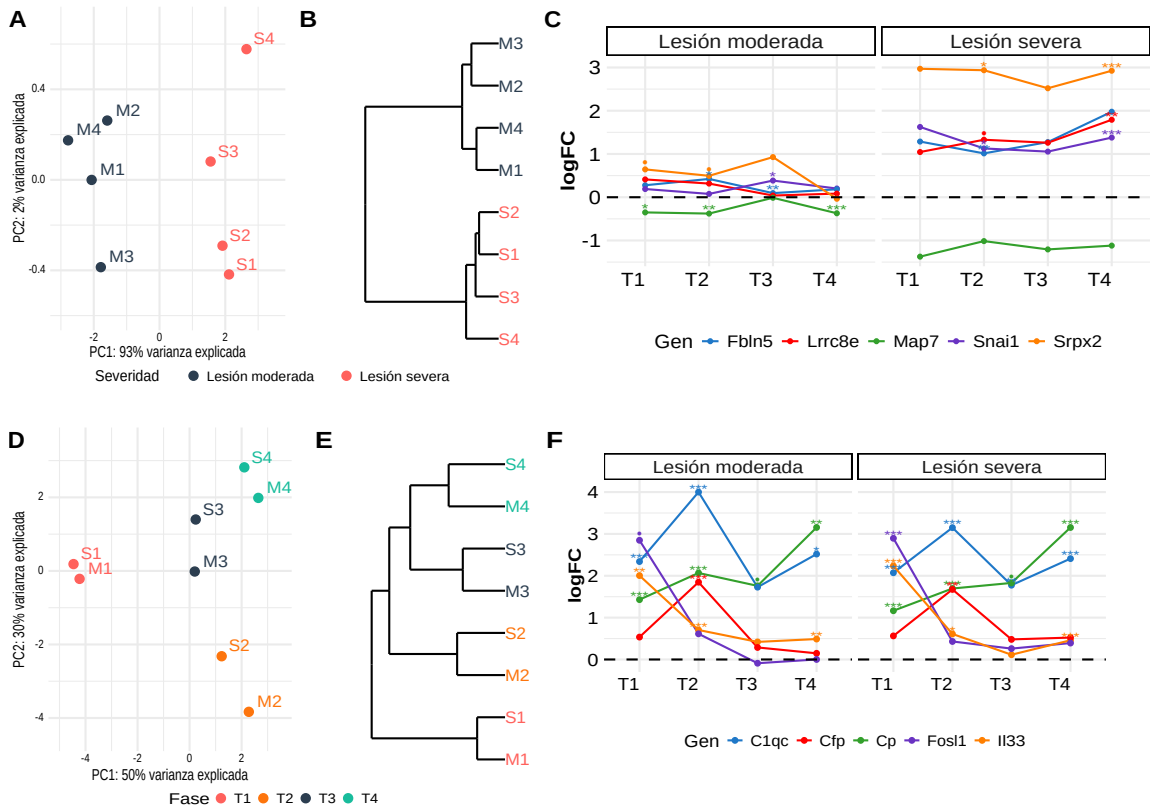


Figura 16. El uso de genes biomarcadores específicos permiten una clasificación efectiva de grupos en base a la severidad o la fase tras la lesión. (A) El PCA y **(B)** el *clustering* jerárquico revelan una clara separación de los grupos en base a la severidad de la lesión usando los top 10 genes específicos de severidad. **(C)** Patrones de expresión génica de una selección de genes específicos de severidad. **(D)** El PCA y **(E)** el *clustering* jerárquico muestran una agrupación por pares de los grupos en base al tiempo. **(F)** Patrones de expresión de una selección de diferentes genes específicos de fase - *Fosl1* e *Il33* específicos de T1; *C1qc* y *Cfp* específicos de T2; y *Cp* específico de T4.

Con el objetivo de evaluar las relaciones funcionales de estos genes, dividimos la lista de los 282 potenciales genes específicos de severidad en *up*-regulados y *down*-regulados para llevar a cabo un análisis PPI. La red PPI de los 152 genes *down*-regulados muestran enriquecimientos en funciones relacionadas con el desarrollo del sistema nervioso y la regulación de la señalización trans-sináptica (Figura 17). Por otro lado, la red PPI de los 130 genes *up*-regulados está enriquecida en funciones relacionadas con la organización de la matriz extracelular, desarrollo de órganos en animales y morfogénesis de estructuras anatómicas (Figura 18).

Para identificar sistemáticamente genes biomarcadores específicos de fase (independientemente de la severidad) comparamos cada fase contra las otras tres. Además, también hemos buscado potenciales genes específicos de fase para cada una de las severidades (Tabla 1). En la Tabla 3 se indica el top 10 de genes para cada una de las comparaciones. La fase aguda muestra la firma transcripcional más distintiva, con 770 genes con $FDR < 0.1$, mientras que en otras comparaciones no se obtuvieron genes significativos con este umbral. Pese a ello, al usar este top 10 de genes con menor p-valor en cada caso se consigue una clara separación de los grupos de interés (Figura S1-S4 y Tabla S9-S12).

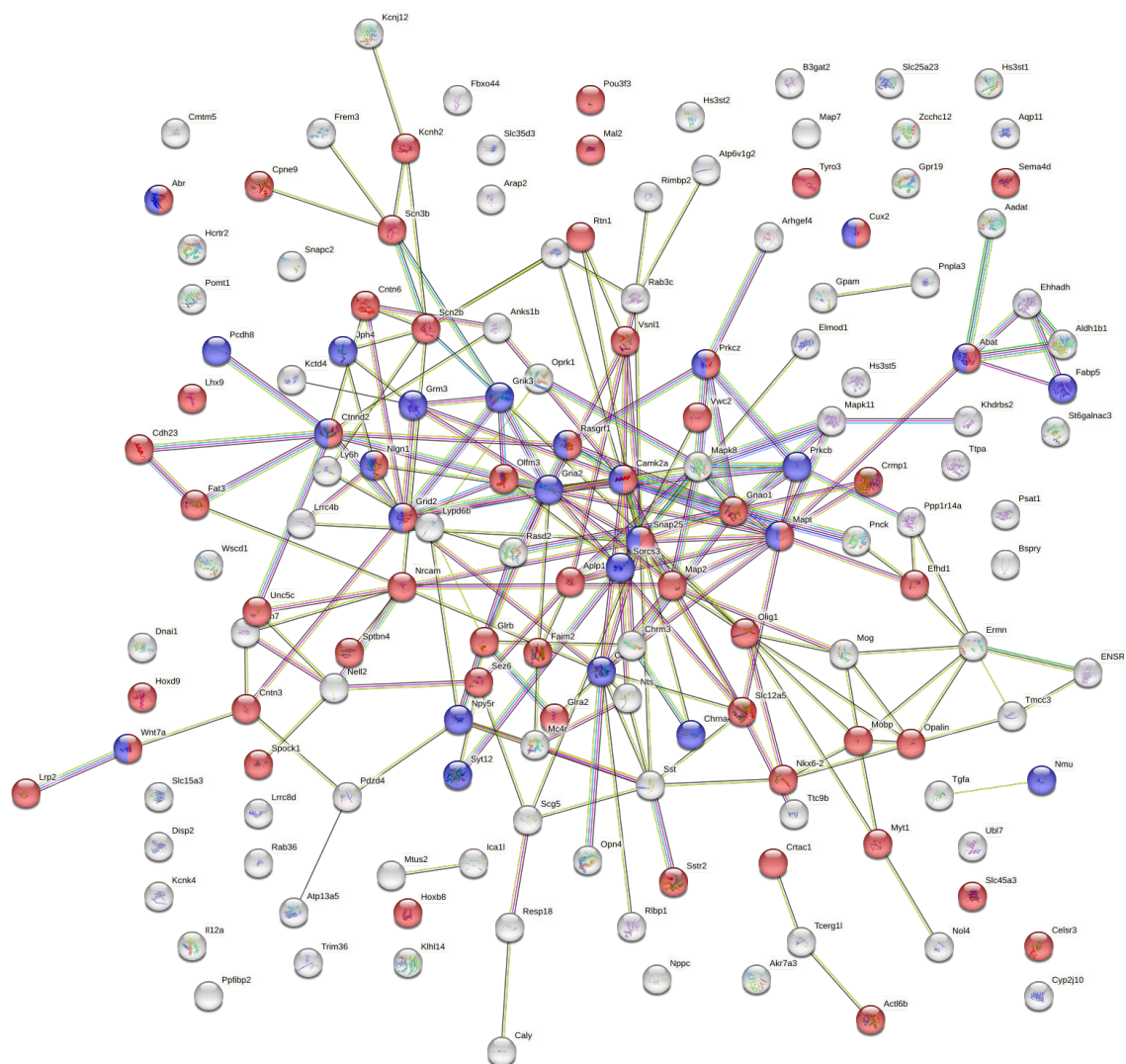


Figura 17. Red PPI para genes marcadores específicos de severidad *down-regulados*. En rojo, proteínas relacionadas con el desarrollo del sistema nervioso (GO:0007399). En azul, proteínas relacionadas con la regulación de la señalización trans-sináptica (GO:0099177). Permalink: <https://version-11-5.string-db.org/cgi/network?networkId=bAOGP0hyvNCg>.

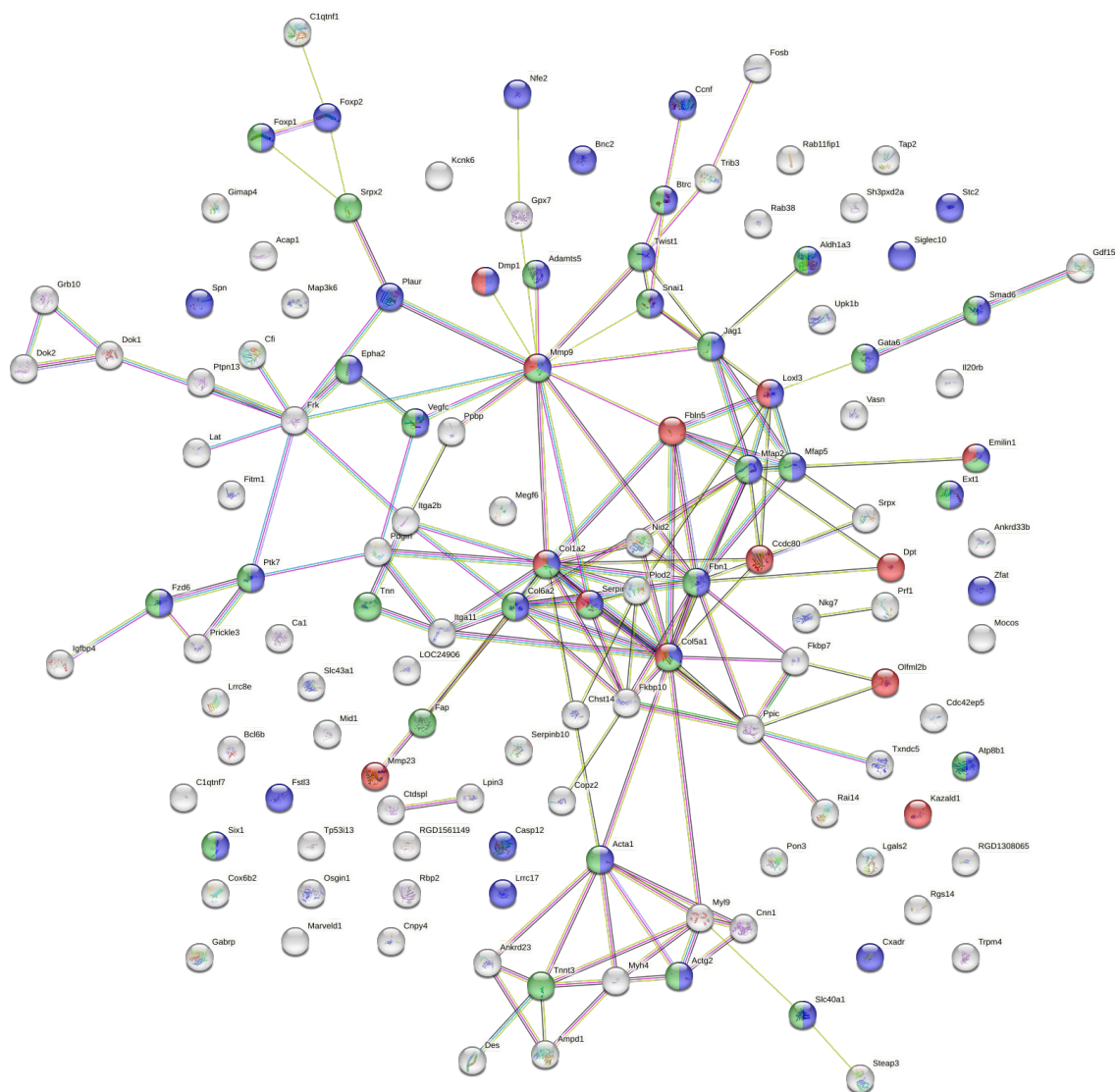


Figura 18. Red PPI para genes marcadores específicos de severidad *up*-regulados. En rojo, proteínas relacionadas con organización de la matriz extracelular (GO:0030198). En azul, proteínas relacionadas con desarrollo de órganos en animales (GO:0048513). En verde, proteínas relacionadas con la morfogénesis de estructuras anatómicas (GO:0009653). Permalink: <https://version-11-5.string-db.org/cgi/network?networkId=bYnq68VI7nAs>.

Tabla 3. Top 10 genes biomarcadores específicos para cada comparación. En negrita se indican los genes *up*-regulados en cada comparación, mientras que el resto están *down*-regulados.

Comparación	Genes
Severidad	Srpx2 , Hoxb8, Acap1 , Snai1 , Aadat, Ppic , Lrrc17 , Map7, Actg2 , H19
T1	Fosl1 , Il33 , Vcam1, Mx2, Mdk, Ldlr , Hmgcr , Alpk3, Dusp5 , Serpina3n
T2	Cfp , Slc46a3 , Scn8a, Ly86 , Casr , C1qc , Slc2a5 , Fgd2 , Aif1 , Slc7a7
T3	Slc28a2, Iqgap1, Cd14, Nt5dc2, Xdh, Msn, Lyve1, Tmem37, Rdh10, Enpp3
T4	Cp , Rab27a , Rab3b , Cybrd1 , Thrsp , Plcb1, Pmp22 , Acot2 , Cyp27a1 , RT1-Db1
M1	Hmgcs1 , Fdps , Cd68, Gpx2 , Alpk3, Crym , Apln , Slc15a2, Pltp, Ldlr
M2	Mmp8 , Cxcl13 , Cd8a , Casr , C1qb , RT1-CE4 , Fap , Cfp , Mmp7 , Slc7a7
M3	Sost , Rbp1, Cd14, Esm1, Hck, Plag1 , Fam102b , Atf3, Slc36a3 , Itgb8
M4	Thrsp , Ccl3, Tmem204 , Calml4 , Cpxm2 , Ckap2, Ttk, Csf3r, G0s2 , Mpz
S1	Fosl1 , Il6 , Kng2 , Tnfrsf12a , Upp1 , Myc , Gprc5a , C3, Exo1 , Chodl
S2	Apoc1 , Cfp , Tnr , Slc28a3 , C1qa , Slc46a3 , Casr , Postn , Slc40a1 , Pnlip
S3	Enpp3, Ccl4, Cd36, Lox, Igsf6, Xdh, Folr2, Msr1, Slc28a2, Rnf4
S4	Mmrn1 , Esm1 , Anxa8 , Crabp2 , Bnc2 , Igfbp6 , Slco2a1 , Ptgis , Tspan8 , Pde5a

Seleccionamos el top 10 de genes de los grupos T1, T2 y T4 para crear una firma fase-específica, que resulta en una separación robusta de los grupos en base al tiempo (Figuras 16D y 16E). En la PCA (Figura 16D), la PC1 separa los grupos M1/S1 del resto, mientras que la PC2 clasifica en el resto de grupos en pares, con una mayor similitud entre M3/S3 y M4/S4, lo que indica que los genes biomarcadores seleccionados permiten la identificación de la fase de la lesión independientemente de la severidad. La Figura 16E, muestra una segregación similar de los grupos, clasificándose dos a dos, con los grupos M1/S1 en una rama independiente en el dendrograma. Aunque hemos identificado que el top de 10 genes específicos de T3 permiten una segregación de M3/S3 del resto de grupos, no los consideramos como biomarcadores apropiados (Figura S3). Esto se debe a que su patrón de expresión característica describe una bajada de expresión entre T2 y T4. Esto podría deberse a la falta de consenso en los resultados debido a

la heterogeneidad de los estudios, más que a una auténtica bajada de expresión durante esta fase. Más probablemente, T3 representa una fase transicional entre T2 y T4, sin que se espere que ningún proceso biológico alcance su punto máximo en T3, como por ejemplo sí cabría esperar con algunas respuestas inmunes en T2.

La red PPI de los 543 genes *up*-regulados específicos de T1 está enriquecida en procesos relacionados con el ciclo celular, respuesta a estrés y procesos metabólicos (Figura 19). Los genes *up*-regulados se distinguen por su aumento de la expresión en la fase aguda que disminuye en las fases posteriores (como lo evidencian los patrones de expresión de *Fos/1* e *Il33* en la Figura 16F, por ejemplo).

Por otro lado, la red PPI de los 227 genes *down*-regulados específicos de T1 está enriquecida en funciones relacionadas con el desarrollo del sistema nervioso y el metabolismo de lípidos (Figura 20). Estos genes, como por ejemplo *Lcat*, *Vcam1* o *Fxyd1*, muestran un patrón de expresión ascendente que se caracteriza por una *down*-regulación en T1 y su posterior *up*-regulación en las siguientes fases.

En combinación, estos dos patrones de expresión divergentes suponen una herramienta valiosa para diferenciar la fase aguda de las demás. De hecho, usando el top 10 de genes específicos de T1, la PCA1 captura el 95% de la varianza, resultando en una clara separación entre los grupos M1/S1 del resto (Figura S1).

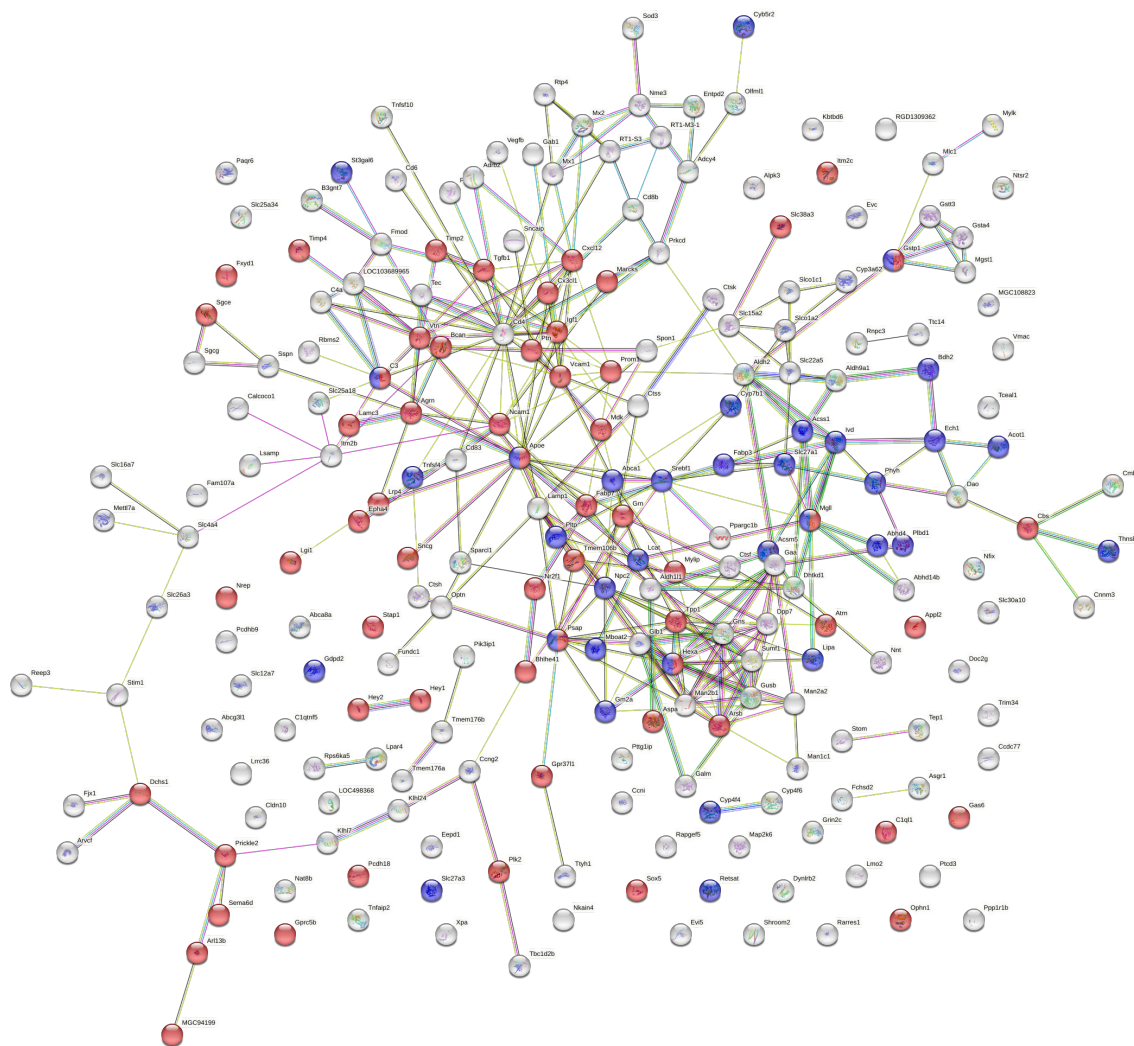


Figura 20. Red PPI para genes marcadores específicos de T1 *down*-regulados. En rojo, proteínas relacionadas con el desarrollo del sistema nervioso (GO:0007399). En azul, proteínas relacionadas con el metabolismo de lípidos (GO:0006629). Permalink: <https://version-11-5.string-db.org/cgi/network?networkId=bSj43bKYgYVn>.

El patrón de expresión que define los genes específicos de la fase subaguda es una *up*-regulación que alcanza su punto máximo en T2 y que después decrece, como por ejemplo *Cfp* y *C1qc* (Figura 16F), genes del complemento asociado con la respuesta inmune secundaria. La red PPI del top 25 genes específicos de T2 (umbral arbitrario), está enriquecida en funciones relacionadas con la respuesta inmune y la activación de macrófagos (Figura 21).

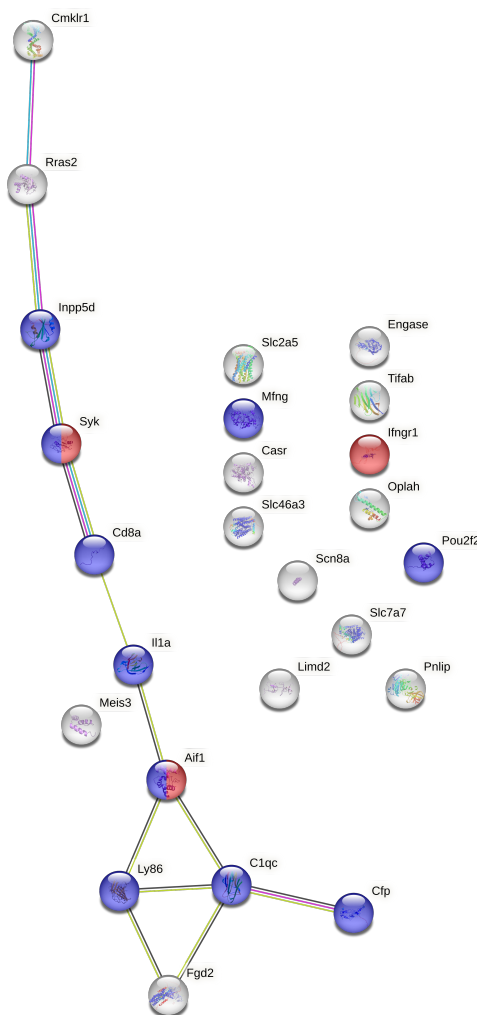


Figura 21. Red PPI para genes marcadores específicos de T2. En rojo, proteínas relacionadas con la activación de macrófagos (GO:0042116). En azul, proteínas relacionadas con la respuesta inmune (GO:0006955). Permalink: <https://version-11-5.string-db.org/cgi/network?networkId=boZP7BxEom6G>.

Los potenciales genes biomarcadores específicos de T4 muestran patrón de expresión ascendente desde T1 que alcanza su máximo en T4, como puede observarse para *Cp* (Figura 16F). La red PPI del top 25 genes específicos de T4 (umbral arbitrario), está enriquecida en funciones relacionadas con el metabolismo de lípidos y respuesta a estímulos (Figura 22).

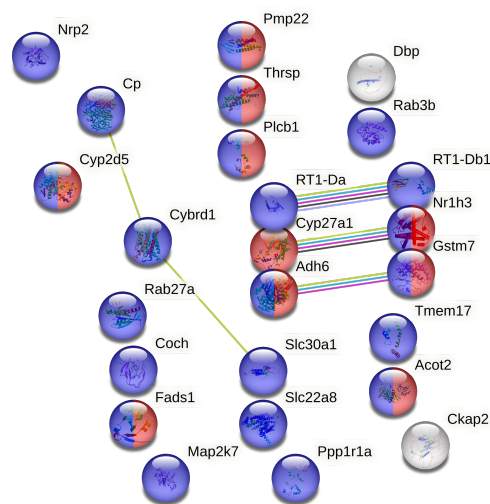


Figura 22. Red PPI para genes marcadores específicos de T4. En rojo, proteínas relacionadas con el metabolismo de lípidos (GO:0042116). En azul, proteínas relacionadas con la respuesta a estímulos (GO:0050896). Permalink: <https://version-11-5.string-db.org/cgi/network?networkId=bleBbz0gyEv6>.

Así, nuestro metaanálisis ha permitido la identificación exitosa de genes biomarcadores que permiten clasificar los perfiles transcriptómicos en base al tiempo que ha transcurrido tras la lesión e independientemente de la severidad.

4.3. Caracterización funcional de las firmas transcripcionales consenso

Para caracterizar los perfiles transcripcionales consenso obtenidos en el metaanálisis, primero llevamos a cabo un análisis de enriquecimiento funcional mediante un GSA. La Figura 23A muestra el número de funciones significativas ($FDR < 0.05$) agrupadas por base de datos, severidad, fase de la lesión y dirección del enriquecimiento. En todos los casos, observamos un mayor número de funciones *up*-reguladas (en rojo) tras la LM en comparación con las funciones *down*-reguladas (en azul), interpretadas como una pérdida de funcionalidad tras la LM. Los diagramas de Venn para lesiones moderadas y graves indicaron un alto número de funciones compartidas entre las cuatro fases temporales tanto para las funciones *up*-regulada como las *down*-reguladas (Figura 23B), lo que sugiere la existencia de mecanismos desregulados inducidos por la LM que persisten en las fases crónicas.

Para obtener una visión general de las categorías biológicas más afectadas tras la LM utilizamos la anotación de Reactome para clasificar las rutas significativamente alteradas. La base de datos Reactome tiene una estructura jerárquica en la que los eventos biológicos en la parte superior incluyen rutas cada vez más específicas [78]. De esta manera, las Figuras 23C y 23D representan la distribución de frecuencias de eventos biológicos en cada comparación de las rutas significativamente *up* y *down*-reguladas, respectivamente.

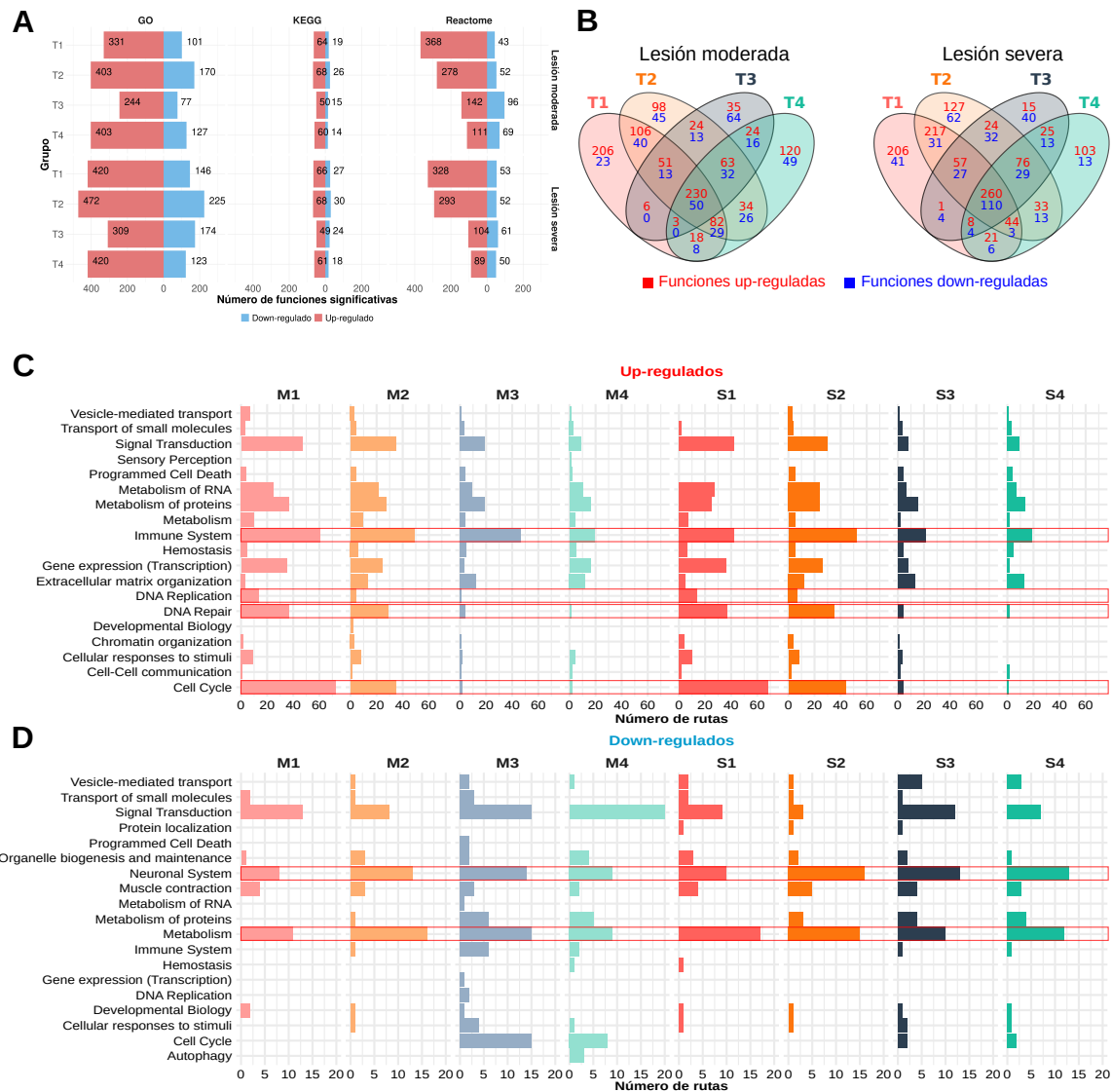


Figura 23. El Gene Set Analysis revela funciones/rutas desreguladas en base al tiempo y la severidad tras la lesión. (A) Número de funciones significativamente alteradas para cada base de datos y grupo experimental. *Up-regulado* (rojo) y *down-regulado* (azul) se refieren a funciones con un valor del logaritmo de odds ratio (LOR) mayor o menor que cero, respectivamente. **(B)** Diagrama de Venn de términos funcionales significativamente alterados. Relación entre funciones comunes y únicas identificadas en diferentes fases de la lesión (T1, aguda; T2, subaguda; T3, crónica temprana; T4, crónica tardía), para lesiones moderadas (izquierda) y severas (derecha). **(C, D)** Distribución de frecuencias de las categorías de Reactome por fase y severidad de la lesión en **(C)** rutas *up-reguladas* y **(D)** *down-reguladas*. Este valor representa la suma de todas las rutas significativamente alteradas dentro de la misma categoría biológica en una comparación. Los cuadros rojos resaltan categorías biológicas de interés.

En cuanto a las rutas *up*-reguladas, el sistema inmune representa una de las categorías más prevalentes en todas las comparaciones, independientemente de la severidad. Estas rutas forman parte del núcleo de rutas desreguladas permanentemente, lo cual puede explicarse por la masiva infiltración de células inflamatorias durante la fase aguda, la activación intrínseca de la microglía en el punto de la lesión y un fallo en la resolución eficiente de la inflamación en la etapa crónica [28].

Por otro lado, se observa una elevada presencia de rutas relacionadas con el ciclo celular, la reparación del ADN o la replicación del ADN en la fase aguda que disminuyen con el tiempo, lo que podría explicarse por la proliferación masiva de células residentes de la médula espinal, como la microglía y la macroglía (astrocitos y células precursoras de oligodendrocitos), así como de células infiltradas, como pericitos, fibroblastos o células de Schwann [105–108].

Estos hallazgos concuerdan con los resultados descritos anteriormente, ya que los genes biomarcadores específicos de T1 muestran enriquecimientos para funciones relacionadas con el ciclo celular. El patrón de expresión de estos genes implica una *up*-regulación sustancial en T1 que disminuye con el tiempo, como se observa con el número de funciones enriquecidas en la Figura 23C.

Las rutas significativamente *down*-reguladas se caracterizan por la pérdida permanente de funciones del sistema nervioso y del metabolismo. En cuanto a las funciones asociadas con el metabolismo, la Figura 24 muestra la *down*-regulación de funciones relacionadas con la síntesis de ácidos grasos, la síntesis de colesterol y el ciclo de Krebs en todos los grupos, como ya ha sido descrito previamente por otros autores [97, 109].

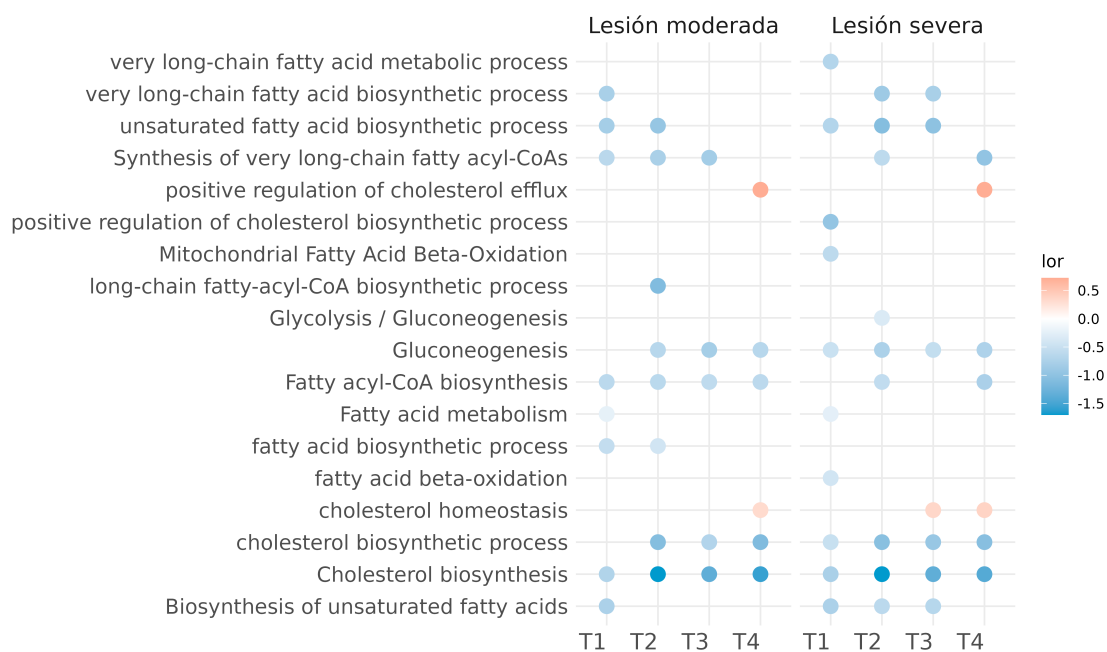


Figura 24. Dotplot de una selección de funciones relacionadas con el metabolismo. En rojo (*up*-regulación) y azul (*down*-regulación) se indican funciones con un valor del logaritmo de odds ratio (LOR) mayor o menor que cero, respectivamente.

La Figura 25 muestra con mayor detalle la *down*-regulación de genes involucrados en la síntesis de colesterol en un heatmap con *clustering* jerárquico, demostrando una *down*-regulación más pronunciada en la fase crónica en comparación con la fase aguda. Nuestros resultados concuerdan con los hallazgos de Spann et al. [97], quienes reportaron una *down*-regulación significativa de la ruta de biosíntesis del colesterol y la *down*-regulación de *Hmgcs1*, *Hmgcr*, *Cyp51*, *Idi1* y *Fdft1*, genes involucrados en la producción de colesterol.

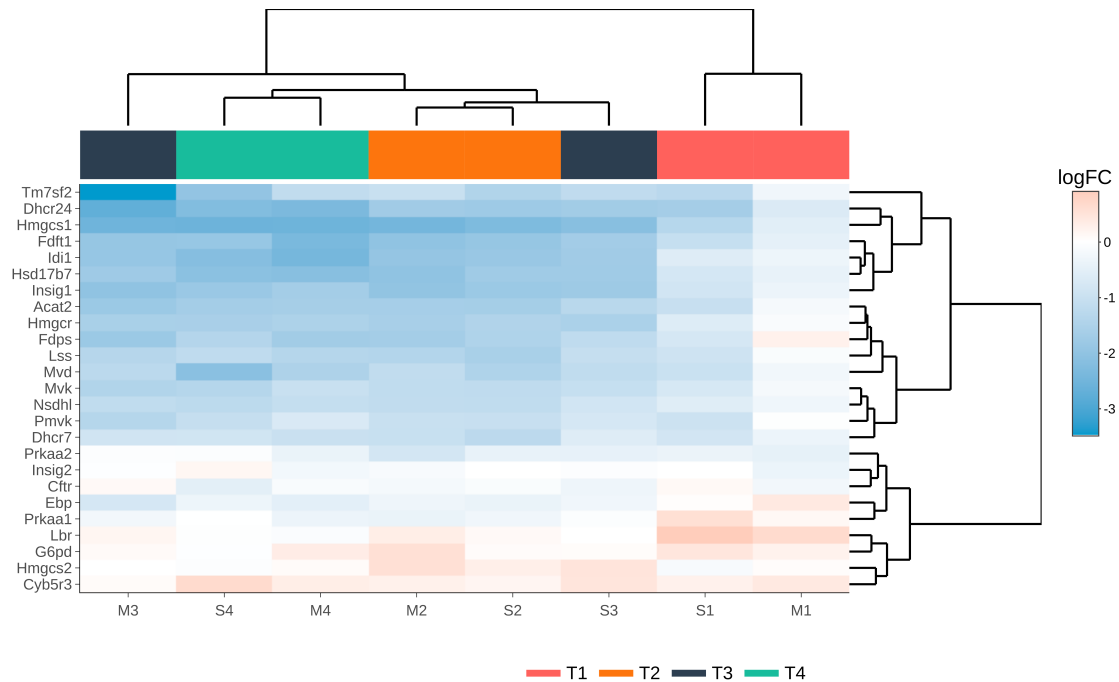


Figura 25. Heatmap con *clustering* jerárquico de los genes del término GO:0006695 (proceso de biosíntesis de colesterol). En rojo (*up*-regulación) y azul (*down*-regulación) se indican genes con un logFC mayor o menor que cero, respectivamente.

La *down*-regulación de funciones/rutas relacionadas con las mitocondrias y la síntesis de ATP también sugiere una disfunción mitocondrial permanente [109] (Figura 26).



Figura 26. Dotplot de una selección de funciones relacionadas con las mitocondrias
En rojo (*up*-regulación) y azul (*down*-regulación) se indican funciones con un valor del logaritmo de odds ratio (LOR) mayor o menor que cero, respectivamente.

El conjunto de eventos descritos afecta negativamente al correcto funcionamiento del sistema nervioso y a cualquier intento de regeneración, como indica por el gran número de funciones *down*-reguladas relacionadas con la regulación de la sinapsis, la regulación de neurotransmisores o el desarrollo del sistema nervioso en todos los grupos (Figuras S5-S7).

El módulo "GSA results" de la herramienta Meta-SCI permite consultar detalladamente los estadísticos de los resultados; la opción "Predefined" de la herramienta "Dotplots" permite explorar gráficos de puntos de funciones específicas relacionadas; y el módulo "Gene2path" permite explorar la relación entre la expresión de los genes y funciones o rutas.

4.4. Inferencia de la actividad de TF en la progresión de la lesión medular

A continuación, estimamos los niveles de actividad de los TF basándonos en los niveles de expresión de sus genes diana. La metodología empleada considera si las interacciones TF-diana activan o inhiben dicho gen. De esta manera, se considera que un TF está diferencialmente activado [Normalized Enrichment Score (NES) > 0], cuando los genes diana de un TF activador estén *up*-regulados o cuando los genes diana de un TF represor estén *down*-regulados. Por el contrario, se considera que un TF está diferencialmente desactivado o con pérdida de función (NES < 0), cuando los genes diana de un TF activador estén *down*-regulados o cuando los genes diana de un TF represor estén *up*-regulados.

La evaluación de la actividad de los TF en los diferentes grupos experimentales reveló una mayor presencia de TF activados frente a TF desactivados en todos los escenarios (Figura 27). Estos resultados subrayan un mayor impacto en la activación de TF durante las etapas tempranas de la LM, que disminuye gradualmente en las fases crónicas. Los diagramas de Venn en la Figura 27B indican la presencia de 19 y 40 TF con patrones de activación diferencial comunes en todas las fases, en lesiones moderadas y severas, respectivamente, lo que sostiene la idea de la existencia de mecanismos desregulados inducidos por la LM que persisten en las fases crónicas, expuesta en la sección anterior.

La Figura 27C (panel izquierdo) muestra los TF consistentemente desregulados compartidos entre lesiones moderadas y severas. TF como Nfkb1, Stat1 y Stat3 se activan durante la progresión de la lesión, contribuyendo a procesos como la modulación de la respuesta inflamatoria [110–112]. Además, también hemos identificado TF diferencialmente activos que previamente no han sido descritos en la LM (por ejemplo, Zfp207, Zfp318 y Zc3h8), y que podrían abrir nuevas vías para el diseño de estrategias terapéuticas.

En el panel derecho de la Figura 27C se muestra el nivel de actividad de una selección de TFs asociados con la severidad o la fase. Sp3, Runx2 y Bhlhe22 están diferencialmente activos en lesiones severas pero no en lesiones moderadas. Por otro lado, TFs como Pin1, Myc, Cebpd (diferencialmente activados), y Hoxb13 y Sox11 (diferencialmente desactivados) sufren una desregulación durante las fases agudas pero vuelven a niveles normales una vez pasada la fase subaguda. Cabe destacar que *Srebf1*, pese a que su nivel de expresión génica exhibe un patrón de *up*-regulación progresiva (Figura 27D), está diferencialmente desactivado en todas las comparaciones excepto en M1. Tanto *Srebf1* como *Srebf2* (también desactivado) representan activadores de genes del metabolismo lipídico [113], lo que podría ayudar a explicar la *down*-regulación previamente descrita de la síntesis de ácidos grasos y colesterol.

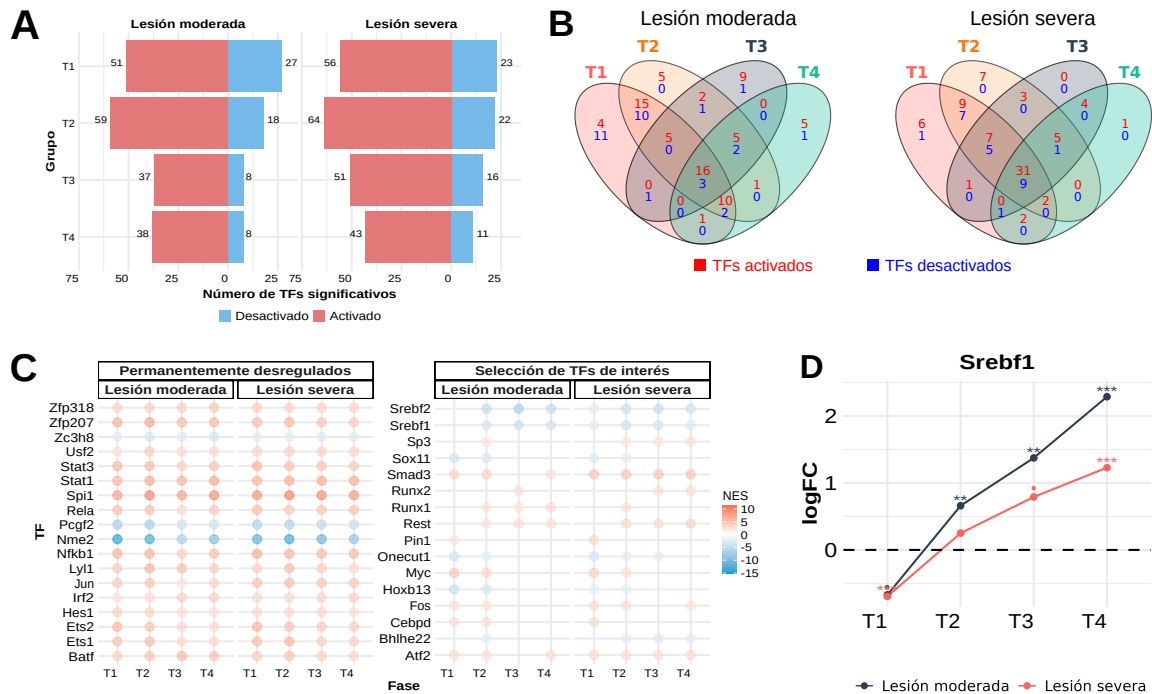


Figura 27. Resumen del análisis de inferencia de la actividad de los factores de transcripción. (A) Número de TF significativamente desregulados para cada grupo experimental. En rojo, TF activados (NES > 0) y en azul TF, desactivados (NES < 0). **(B)** Diagrama de Venn de TF significativamente desregulados. Relación entre TF comunes y únicos identificados en diferentes fases de la lesión (T1, aguda; T2, subaguda; T3, crónica temprana; T4, crónica tardía), para lesiones moderadas (izquierda) y severas (derecha). **(C)** Dotplot de TF significativamente desregulados en todos los grupos (izquierda) y selección de TF de interés (derecha) (Rojo – activación; azul – desactivación). **(D)** Patrones de expresión de *Srebf1*

4.5. Clasificación funcional de las redes de genes co-expresados

Para encontrar genes con patrones de expresión similares, primero dividimos la matriz de firmas transcripcionales consenso en lesiones moderadas y lesiones severas. Analizamos cada conjunto por separado, identificando 19 y 18 *clusters* de genes co-expresados, respectivamente (Figura S8 y S9). Para cada uno de ellos (excepto el *cluster* 0) se construyó una red PPI, las cuales presentan un gran tamaño y un alto nivel de interconectividad, lo que indica una fuerte asociación funcional entre genes con patrones de expresión similares.

Con el objetivo de obtener una mayor comprensión de ellas, se aplicó un algoritmo de *clustering* para encontrar redes locales más estrechamente relacionadas entre sí. De las 37 redes PPI se generaron 261 redes locales, con las que se realizó un análisis de enriquecimiento funcional para clasificarlas en categorías biológicas de interés previamente definidas. De las 261 subredes generadas, clasificamos con éxito 153. Las 108 subredes no clasificadas están formadas tan solo por unas pocas proteínas sin funciones enriquecidas. Los resultados son consistentes con los análisis previos, identificando una mayor representación de las categorías relacionadas con el sistema nervioso, sistema inmune o metabolismo. También, la distribución de frecuencias de las categorías funcionales son similares entre los dos grados de severidad (Figura 28).

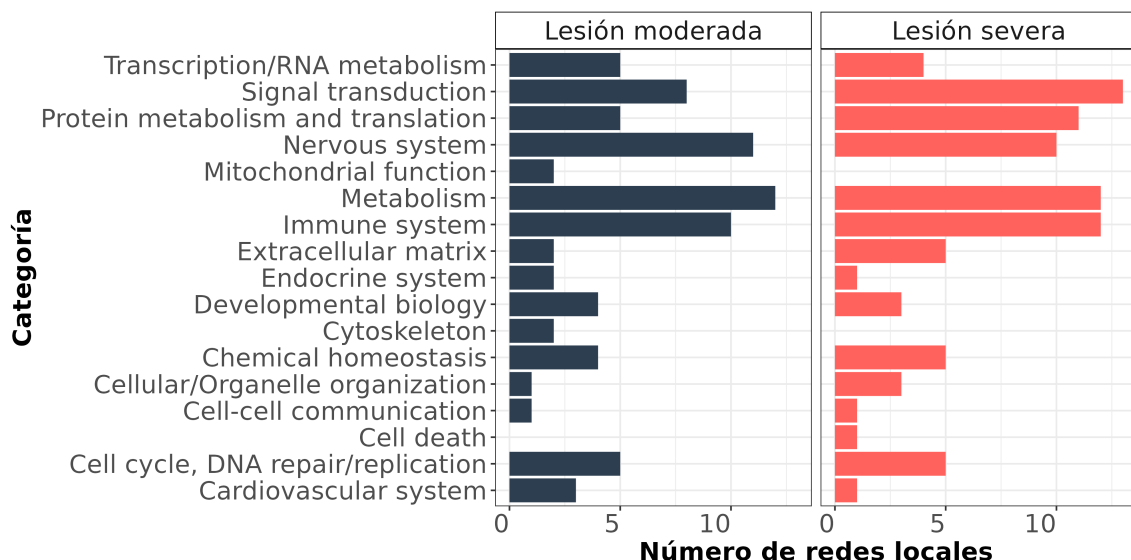


Figura 28. Distribución de frecuencias de las categorías funcionales asociadas a las redes locales formadas por genes co-expresados.

En la Figura 29 se desglosan los patrones de expresión de cada subred local asignados a cada categoría funcional para lesiones moderadas. Se representa el patrón de expresión característico de la red, que se obtiene de calcular la mediana de los valores de logFC de cada uno de los genes que forman la red en un punto temporal.

Puede observarse que todas las subredes asignadas al sistema nervioso están permanentemente *down*-reguladas, indicando dicha pérdida de función de estos procesos, por el contrario, las que pertenecen al sistema inmune permanecen *up*-reguladas.

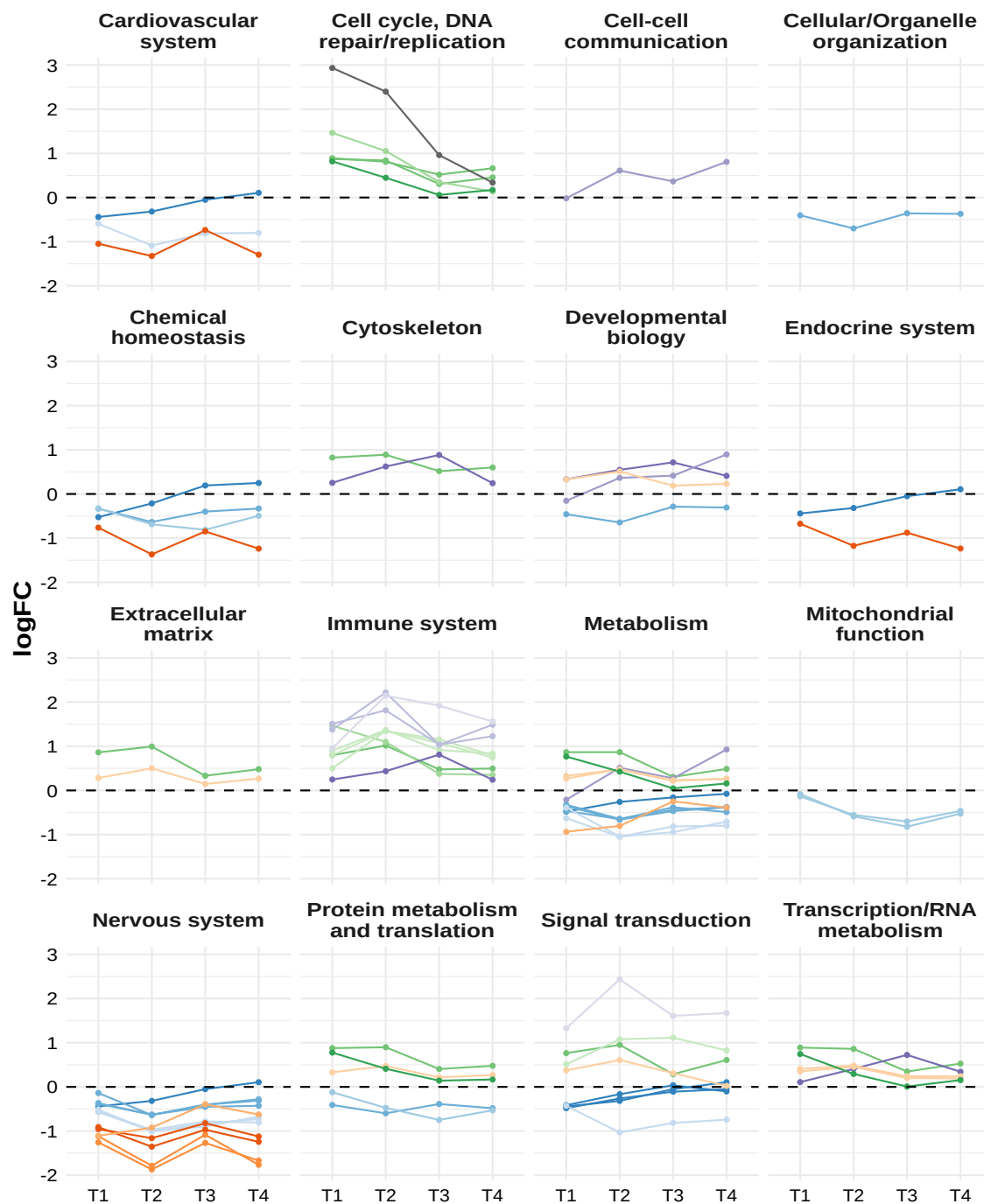


Figura 29. Patrones de expresión de las redes locales de lesión moderada clasificadas por categoría funcional. Se representa el patrón de expresión característico (mediana del $\log FC$ de los genes en un punto temporal) para cada red local obtenida en lesiones moderadas, clasificadas según la categoría funcional asignada. El color indica el *cluster* del que proviene cada red local.

En cuanto a las clasificadas en metabolismo, se aprecia que algunas están *up*-reguladas y otras *down*-reguladas. Esto se debe a que las subredes *up*-reguladas (C8_Mod_subnetwork_4, C8_Mod_subnetwork_5, C9_Mod_subnetwork_1, C10_Mod_subnetwork_8, C14_Mod_subnetwork_1)¹, están implicadas en el metabolismo de proteínas en general, mientras que las redes *down*-reguladas (C1_Mod_subnetwork_2, C2_Mod_subnetwork_3, C2_Mod_subnetwork_8, C2_Mod_subnetwork_9, C4_Mod_subnetwork_4, C4_Mod_subnetwork_7, C7_Mod_subnetwork_3), están implicadas en la función mitocondrial y el metabolismo de lípidos, tal y como se ha visto en los resultados del GSA e inferencia de actividad de TF. Para mayor detalle consultar las secciones "See networks" y "Moderate injury profiles" del módulo "PPI networks" en [Meta-SCI app](#).

En la Figura [S10](#) se representan los patrones de expresión de las redes locales generadas para lesiones severas, clasificadas en base a la categoría funcional, pudiéndose observar un comportamiento similar al descrito en lesiones moderadas.

La Figura [30](#) ilustra de forma más específica dos ejemplos concretos de redes locales. En la Figura [30A](#) se expone la red C11_Mod_subnetwork_11, es decir, la red local número 11 del *cluster* 11 de lesiones moderadas. Todos los genes que conforman esta red tienen un patrón de expresión similar, que se muestra en la parte superior de la misma. Las funciones de esta red están relacionadas con el ciclo celular, y además, las proteínas en verde, rojo y morado forman complejos físicos entre ellas. En la Figura [30B](#) se muestran los patrones de expresión de los genes que forman el complejo verde, que se caracteriza por una sobreexpresión en T1 que decrece progresivamente a lo largo del tiempo, y que, como se discutió

¹La nomenclatura sigue el formato C[número de *cluster*][tipo de lesión (Mod o Sev)][número de red local], donde "C" representa el número de clúster, "Mod" o "Sev" indican el tipo de lesión (moderada o severa), y "subnetwork" especifica el número de subred generada del clúster.

anteriormente, explica la presencia de funciones enriquecidas relacionadas con el ciclo celular en T1 que disminuyen en fases posteriores.

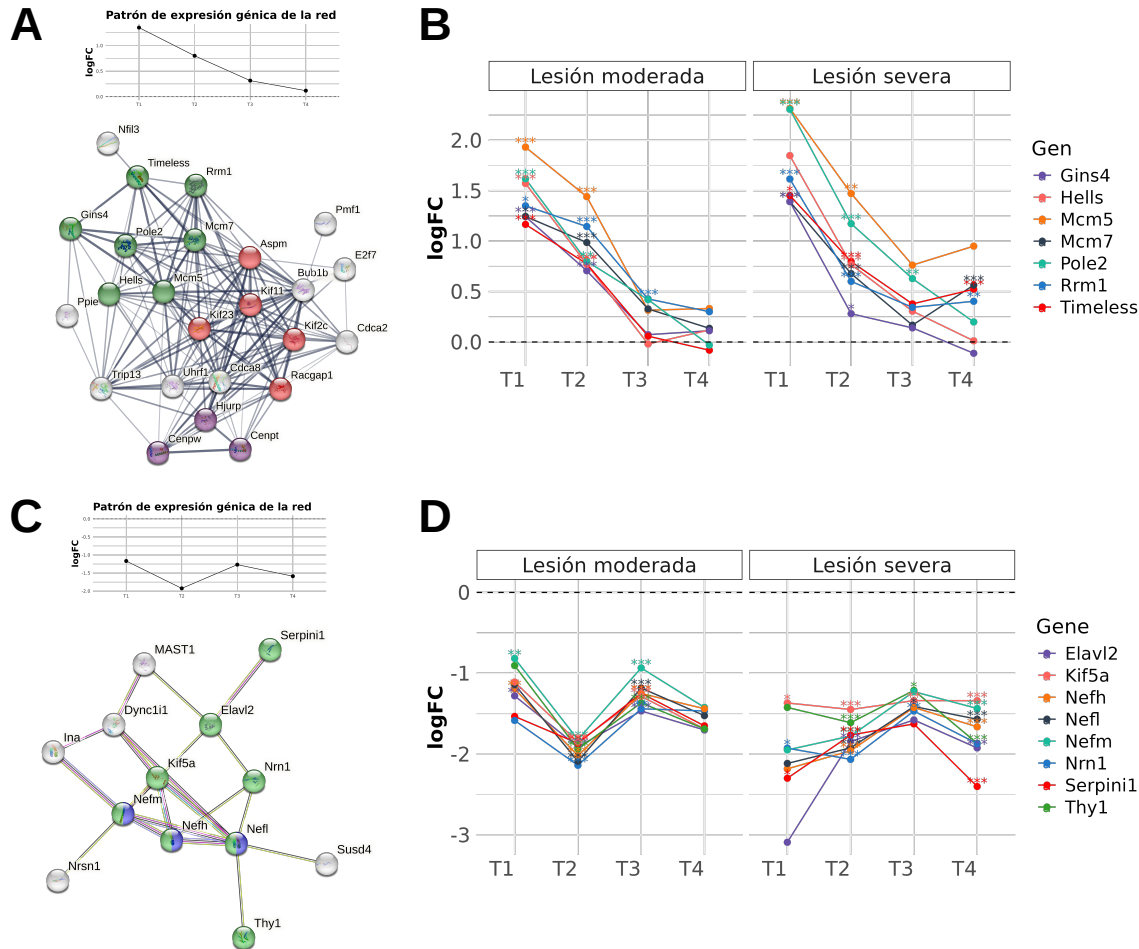


Figura 30. Ejemplos de redes locales de co-expresión. (A) Red PPI C11_Mod_subnetwork_11 (red local número 11 del *cluster* 11 de lesiones moderadas) funcionalmente asociada con el ciclo celular. Todas las proteínas de la red están relacionadas funcionalmente y exhiben el mismo patrón de expresión. Además, las proteínas en verde, rojo y morado forman complejos físicos entre ellas. **(B)** Patrón de expresión génica de las proteínas en verde de la red C11_Mod_subnetwork_11. **(C)** Red PPI C6_Mod_subnetwork_1 (red local número 1 del *cluster* 6 de lesiones moderadas) funcionalmente asociada con el sistema nervioso. En verde se indican las proteínas anotadas con "generación de neuronas" (GO:0048699), y en azul las anotadas con "desarrollo de la médula espinal" (GO:0021510). **(D)** Patrón de expresión génica de las proteínas en verde de la red C6_Mod_subnetwork_1.

En la Figura 30C se muestra la red C6_Mod_subnetwork_1, conformada por genes involucrados en el sistema nervioso. Los genes asociados con "generación de neuronas"(GO:0048699) se muestran en verde, mientras que los involucrados en "desarrollo de la médula espinal"(GO:0021510) se muestran en azul. En este caso, el patrón de expresión de la red se caracteriza por una *down*-regulación persistente (Figura 30D).

El módulo "PPI networks" de la aplicación Meta-SCI permite la exploración detallada de todas las redes generadas con las funcionalidades que ofrece STRING.

4.6. Los genes biomarcadores de severidad y fase propuestos clasifican correctamente nuevos perfiles transcripcionales de rata

Con el objetivo de evaluar si los genes biomarcadores específicos de fase y gravedad seleccionados pueden clasificar correctamente un nuevo perfil transcriptómico escogimos el *dataset* GSE218088 [85] para validación de los resultados. Este *dataset* fue publicado tras el periodo de revisión sistemática y selección de estudios y cumple con los criterios de inclusión establecidos en el metaanálisis. Los autores indican que la lesión se produjo dejando caer un peso de 10 g y un diámetro de 2.5 mm a 8 cm de altura, lo cual se considera una lesión severa [114, 115] y que extrajeron tejido de la médula espinal tres días después de provocar la lesión, lo que corresponde a la fase aguda definida en nuestro trabajo.

Usando el top 10 de genes específicos de severidad (Tabla 3), se observa que el perfil transcripcional del *dataset* GSE218088 se agrupa en una rama independiente del dendrograma junto con los perfiles de lesiones severas (Figura 31A). Además, utilizando biomarcadores específicos de fase para T1, T2 y T4 (Tabla 3), se observa una agrupación del perfil transcriptómico de GSE218088 junto con los grupos M1 y S1 de nuestro metaanálisis (Figura 31B).

De esta manera se demuestra la robustez de los resultados obtenidos y la capacidad de los genes biomarcadores para clasificar correctamente un nuevo perfil transcriptómico.

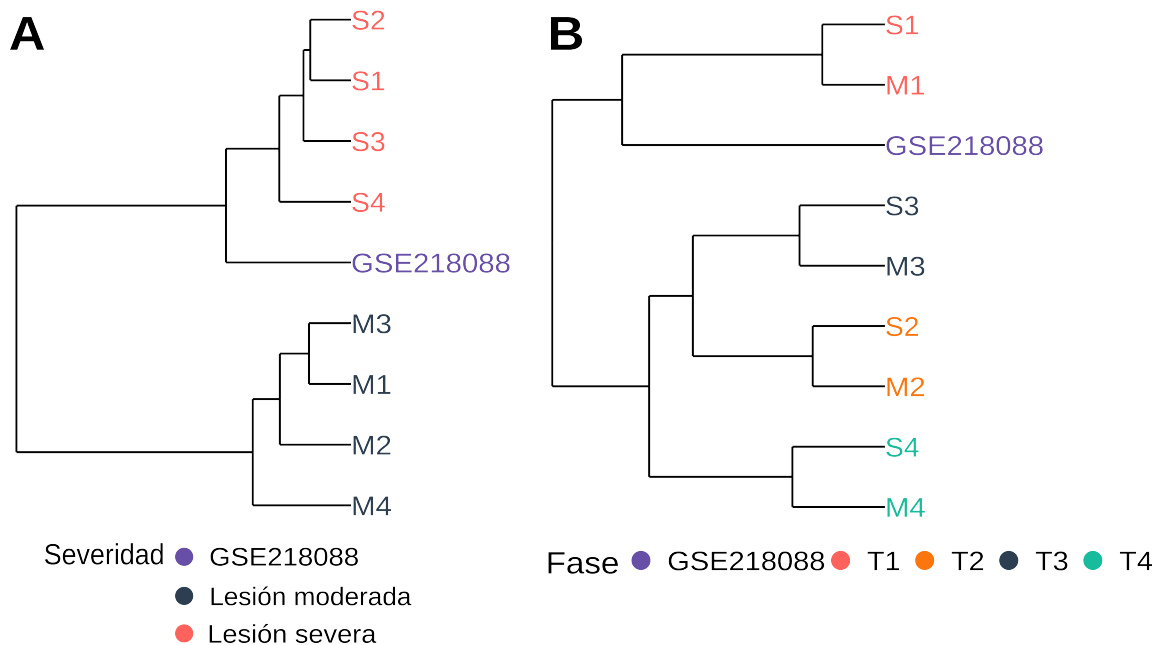


Figura 31. Análisis de *clustering* del *dataset* GSE218088 usando biomarcadores específicos de severidad y fase. (A) *Clustering* jerárquico utilizando genes específicos de severidad con las firmas transcripcionales consenso del metaanálisis y la firma transcriptómica del *dataset* GSE218088. (B) *Clustering* jerárquico utilizando genes específicos de severidad con las firmas transcripcionales consenso del metaanálisis y la firma transcriptómica del *dataset* GSE218088.

4.7. Consistencia en los patrones de expresión temporales entre ratón y rata

Para validar y comparar los patrones de expresión génica obtenidos en el metaanálisis, seleccionamos el *dataset* de ratón publicado por Li et al. [86], ya que posee un diseño experimental similar a los estudios incluidos en nuestro trabajo. Los autores utilizaron un modelo de ratón para estudiar los cambios en la expresión génica tras la LM mediante RNA-seq. Las lesiones se provocaron en la T9 y se tomaron muestras de un segmento de la médula espinal en diferentes días post-lesión, que cubrieron las cuatro fases temporales establecidas en nuestro trabajo. Para generar los patrones de expresión génica seguimos el mismo pipeline aplicado en los estudios individuales, lo que asegura su comparabilidad con los perfiles generados en el metaanálisis. Los resultados del análisis de expresión diferencial de cada comparación y la matriz de perfiles de expresión génica se pueden encontrar en las Tablas S13-S17.

Llevamos a cabo un análisis de correlación para comparar los patrones de expresión entre los modelos de ratón y rata (lesión moderada de rata vs. ratón; lesión severa de rata vs. ratón; lesión moderada de rata vs. lesión severa de rata), que demostró una relación generalmente positiva entre ellos. El gráfico de densidad de los valores de correlación indica una mediana cercana a 0.5 en las tres comparaciones (Figura 32A y Tabla S18). Un total de 892 genes mostraron una fuerte relación positiva ($r > 0.9$) entre la lesión moderada de rata y ratón, otros 892 genes entre lesión severa de rata y ratón y 922 genes entre la lesión moderada y severa en rata. El diagrama de Venn de la Figura 32A muestra el número de genes que intersectan entre estas tres comparaciones, existiendo 198 genes comunes entre ellas.

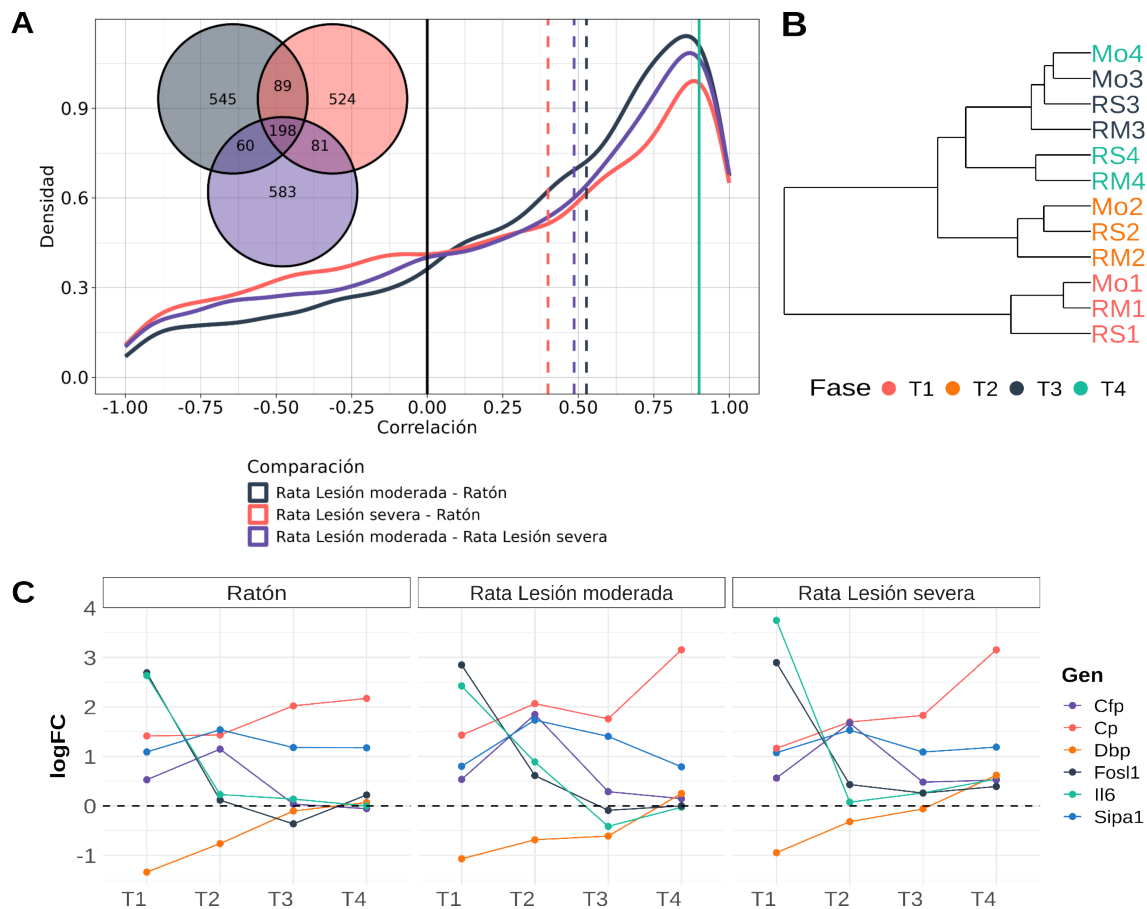


Figura 32. Comparación de los patrones de expresión del metaanálisis en rata con un *dataset* de ratón y análisis de *clustering* usando biomarcadores comunes. **(A)** Gráfico de densidad de los valores de correlación entre los patrones de expresión génica. En negro se representa la lesión moderada de rata vs. ratón; en rojo la lesión severa de rata vs. ratón; y en morado la lesión moderada vs. lesión severa en rata. La línea verde representa un coeficiente de correlación de 0.9. Las líneas discontinuas representan los valores medianos de correlación para cada comparación. El diagrama de Venn muestra la intersección de genes con un coeficiente de correlación mayor de 0.9 en cada comparación. **(B)** *Clustering* jerárquico utilizando genes específicos de fase (*Fos1*, *Il6*, *Cfp*, *Sipa1*, *Cp*, y *Dbp*) en los perfiles transcripcionales consenso y conjunto de datos de ratón (RM, rata lesión moderada; RS, rata lesión severa; Mo, ratón). **(C)** Comparación de patrones de expresión de genes específicos de fase en ratones y lesiones moderadas y severas en ratas.

Después realizamos un análisis de *clustering* (Figura 32B) utilizando un conjunto de genes específicos de fase manualmente curado (*Fosl1*, *Il6*, *Cfp*, *Sipa1*, *Cp*, y *Dbp*) y que muestran patrones de expresión similares en las tres comparaciones (Figura 32C). De manera similar a la Figura 16D y 16E, observamos una correcta estratificación de las fases temporales. El *clustering* jerárquico reveló una clara separación de grupos basados en el tiempo post-lesión, donde T1 se encuentra aislado en una rama independiente, seguido de una segregación de T2 y de las fases crónicas (T3 y T4). En este caso, las firmas de T3 y T4 en el ratón se agrupan dentro de la misma rama que M3 y S3 de la rata, mientras que M4 y S4 permanecen independientes (Figura 32B). Esto podría explicarse debido a la menor diferencia entre los tiempos. Para ratón, las muestras en T3 se toman a 28 dpl, y en T4 a 42 dpl, estando más cerca en el tiempo en comparación con T3 y T4 de los estudios individuales del metaanálisis de rata. T3 en rata cubre de 15 a 35 dpl, y T4 se extiende de 56 a 168 dpl. Intervalos más largos entre fases podrían llevar a diferencias más claras en los perfiles de expresión, ya que la lesión tiene más tiempo para estabilizarse.

Así, este análisis comparativo con un modelo de LM en ratón valida los patrones de expresión obtenidos en el metaanálisis, demostrando una correlación positiva entre las respuestas a la lesión en ratón y rata, apoyando la fiabilidad de los genes biomarcadores identificados.

4.8. Concordancia entre los patrones de expresión obtenidos en el metaanálisis y el análisis qPCR

Nuestro objetivo fue confirmar mediante validación experimental vía qPCR los patrones de expresión identificados en el metaanálisis. Para ello, después de inducir una lesión severa, se extrajo el ARN de la médula espinal de ratas lesionadas en T1, T2, T3 y T4 después de la lesión. Seleccionamos ocho genes para validar sus perfiles de expresión mediante qPCR. Elegimos los TF *Cebpd*, *Hes1*, *Jun* y *Onecut1* debido a su relevancia funcional, así como los genes *Il6*, *Vcam1*, *Vegfa* y *Vim*, que son tanto genes diana de estos TF, como posibles genes biomarcadores específicos de fase. En la Figura 33 se representa una comparativa de los valores de logFC obtenidos en la qPCR, lesión moderada y severa del metaanálisis en rata, y el *dataset* de ratón.

En la Figura 33A se incluyen los patrones de expresión de los cuatro genes diana. En los cuatro sistemas biológicos, *Il6* muestra una gran *up*-regulación en T1 que disminuye en las fases posteriores, aunque con diferencias en la magnitud de expresión entre qPCR y los resultados bioinformáticos. *Vcam1* presentó una *down*-regulación en T1, pero cambia la dirección de la regulación en las fases posteriores, con un aumento gradual de la expresión a lo largo del tiempo, salvo en ratón, donde en T1 hay una ligera *up*-regulación (logFC = 0.36). En *Vegfa* se observa el patrón contrario a *Vcam1*, pasando de estar *up*-regulado en T1 a *down*-regulado en las siguientes fases. Aunque inicialmente *Vim* podría ser un potencial biomarcador específico del grupo S4, debido a su patrón de *up*-regulación ascendente con pico en T4 para lesiones severas, no observamos este patrón en la qPCR aunque sí se confirmó su persistente *up*-regulación.

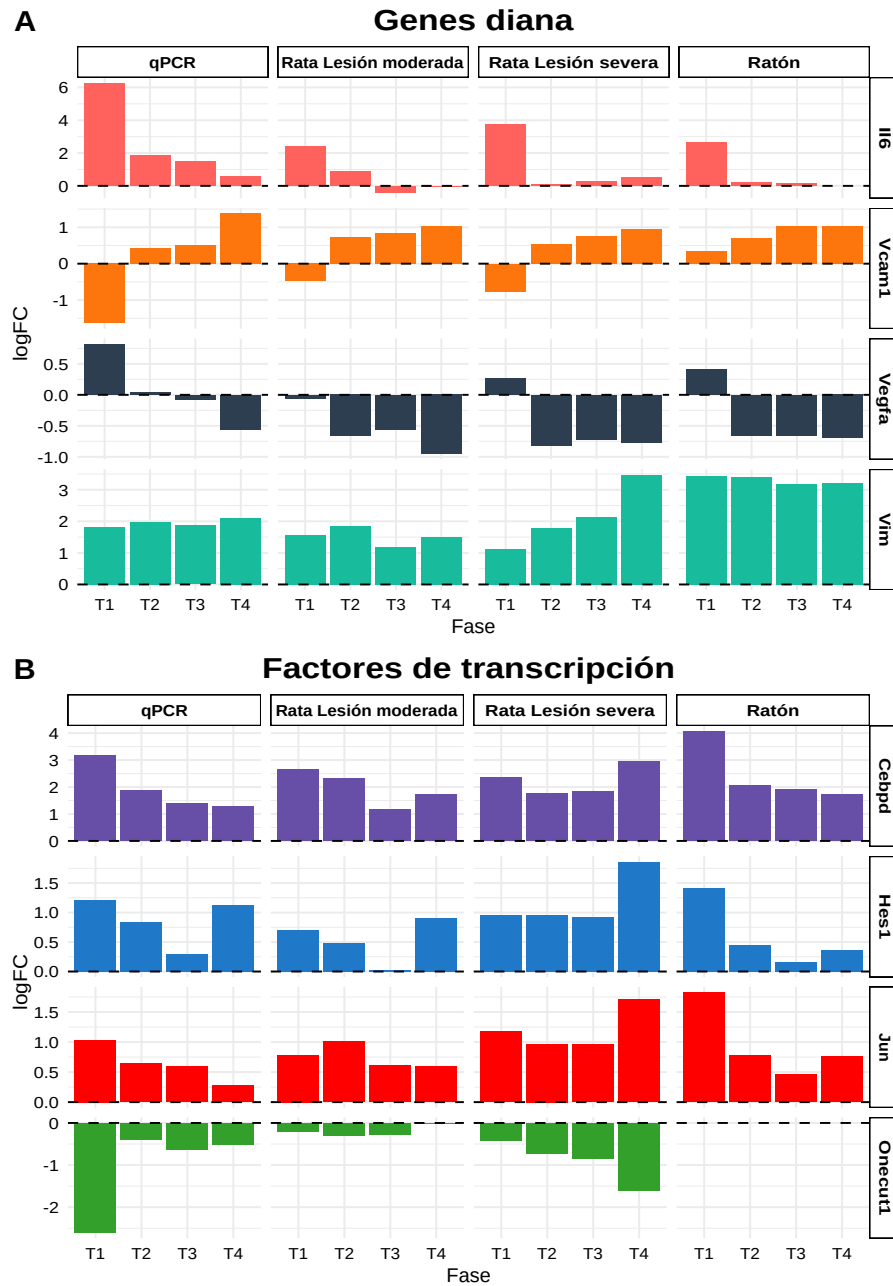


Figura 33. Comparación de los patrones de expresión génica de los factores de transcripción y sus genes diana seleccionados para validación experimental. Los valores de qPCR se obtuvieron calculando el logFC de la división de la expresión en cada tiempo entre la expresión de las muestras no lesionadas. "Rata Lesión moderada" indica el patrón de expresión obtenido en el metaanálisis para lesiones moderadas. "Rata Lesión severa" indica el patrón de expresión obtenido en el metaanálisis para lesiones severas. Ratón indica el patrón de expresión obtenido para el *dataset* de ratón.

Aunque hay una menor consistencia en los patrones de expresión entre los cuatro sistemas para los TF, *Cebpd*, *Hes1* y *Jun*, observamos un acuerdo general dada la persistente *up*-regulación identificada en todos los tiempos (Figura 33B). Por el contrario, *Onecut1* presentó diferentes patrones de cambio pero confirmó su constante *down*-regulación en el tiempo (*Onecut1* no está presente en el *dataset* de ratón).

Para tener una métrica sobre el grado de similitud de los patrones de expresión, realizamos un análisis de correlación, comparando los valores de logFC obtenidos en la qPCR con los derivados del metaanálisis y *dataset* de ratón (Figura 34). Para los genes diana, 10 de 12 comparaciones exhibieron una alta relación positiva, con un coeficiente de correlación de Pearson superior a 0.85. Por el contrario, para los TF tan solo 5 de 11 comparaciones superaron un coeficiente de 0.7.

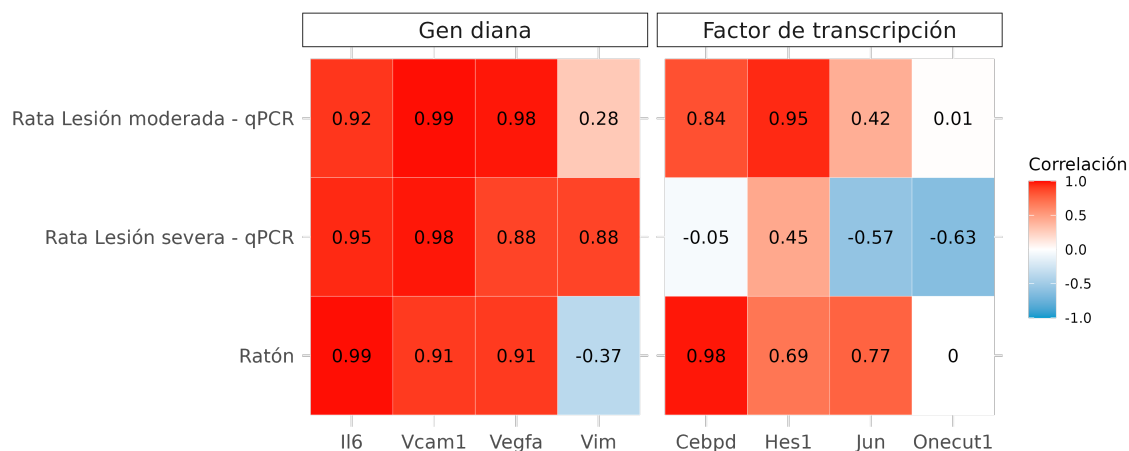


Figura 34. Coeficiente de correlación de Pearson al comparar los valores de logFC de los patrones de expresión obtenidos en la qPCR con los de lesión moderada de rata, lesión severa de rata y el *dataset* de ratón. Los coeficientes de correlación se encuentran disgregados por comparación y gen. A la izquierda los genes diana, a la derecha los factores de transcripción.

En conjunto, aunque reproducir el patrón de expresión exacto resulta complicado, esta validación confirmó los patrones generales de desregulación, o al menos, la dirección de la desregulación en fases específicas tras la lesión.

4.9. Los genes biomarcadores específicos de severidad en rata predicen la prognosis de la lesión de los pacientes a partir de muestras de sangre humana

Analizamos el *dataset* GSE151371 [90] con el objetivo de investigar un posible enfoque traslacional de los genes específicos de severidad identificados en ratas para el pronóstico de la lesión a partir de muestras de sangre de pacientes humanos. En este *dataset* se secuenciaron muestras de sangre humana tras los primeros 10 días de la lesión.

Para lograr un escenario similar al de nuestro metaanálisis basado en ratas, seleccionamos las muestras correspondientes al grado AIS A (pérdida motora y sensorial completa) y AIS D (cierto grado de preservación funcional) [26] para asegurar la similitud en la severidad entre los dos sistemas biológicos. Tras el análisis de expresión diferencial entre los grupos, identificamos 931 genes con $FDR < 0.1$ (Tabla S19). Al comparar esta lista con la de los 282 genes específicos de severidad previamente identificados en modelos de ratas (Figura 35A), encontramos 12 genes comunes (*EXT1*, *FBN1*, *FOSB*, *GNAO1*, *GRB10*, *MMP9*, *NFE2*, *PRF1*, *SLC25A23*, *SNAI1*, *ST6GALNAC3* y *STX11*). En la Figura 35B se representan los valores de logFC obtenidos al comparar AIS A vs. AIS D en muestras de pacientes, así como los obtenidos al comparar los perfiles de lesiones severas con los de lesiones moderadas en rata. En el *dataset* GSE151371 estos genes tienen un patrón de expresión que se caracteriza por una mayor desregulación a medida que aumenta la severidad de la LM (Figura 36).

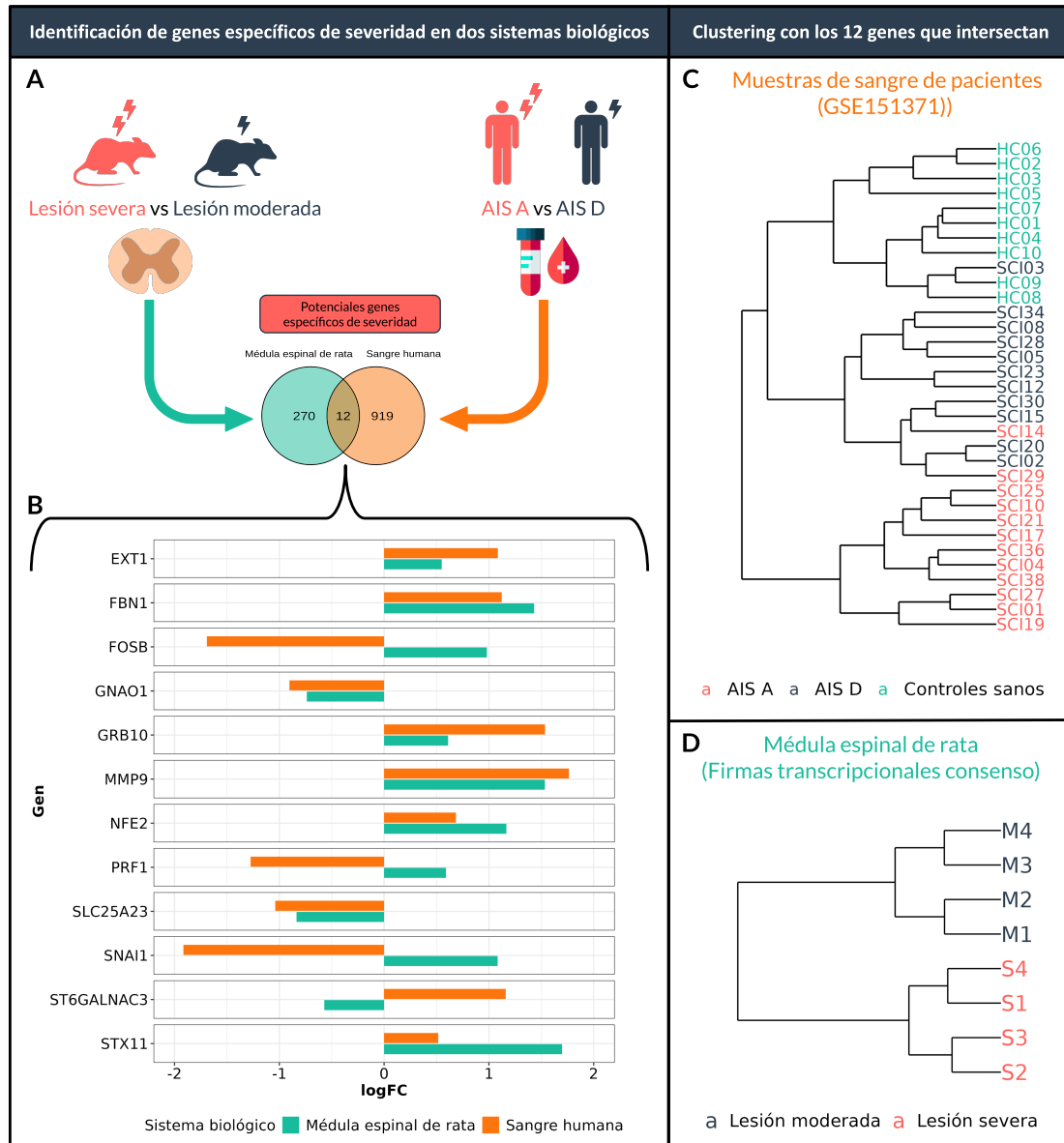


Figura 35. Los genes biomarcadores específicos de severidad estratifican muestras humanas y en modelos de LM en ratas (A) Esquema de la estrategia seguida para identificar genes biomarcadores específicos de severidad comunes a los dos sistemas biológicos. (B) Gráfico de barras que compara el logFC de los 12 genes comunes. Las barras naranjas indican los valores de logFC resultantes de comparar AIS A vs. AIS D en muestras humanas. Las barras verdes indican el logFC después de comparar el logFC de los perfiles transcripcionales consenso en lesiones severas y moderadas obtenidos en rata. (C) Clustering jerárquico utilizando la firma de 12 genes en muestras de sangre humana del conjunto de datos GSE151371. (D) Clustering jerárquico utilizando la firma de 12 genes en las firmas transcripcionales consenso obtenidos para médula espinal de rata.

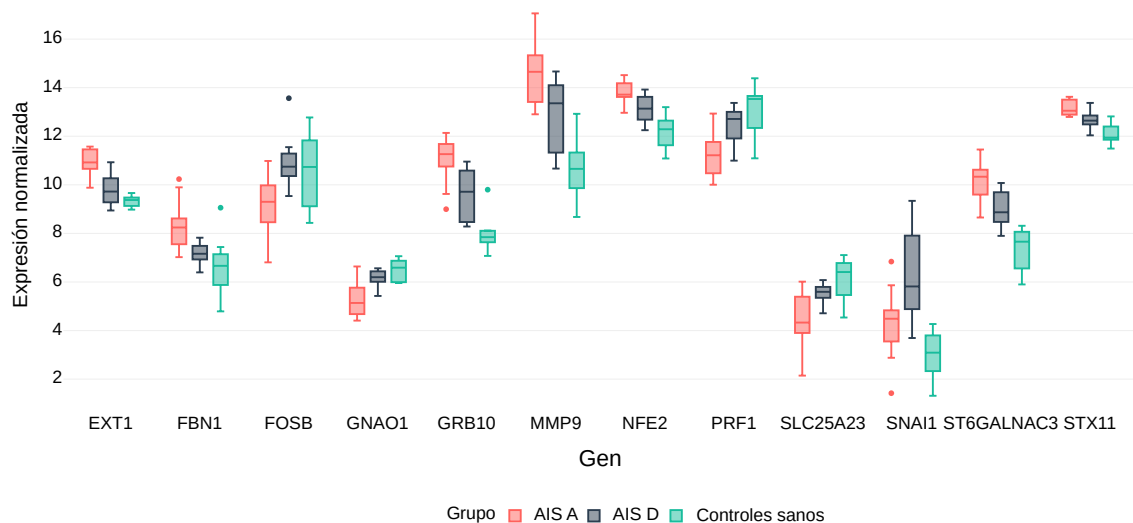


Figura 36. Expresión normalizada de la firma de 12 genes biomarcadores de severidad en el *dataset* GSE151371. Rojo, grupo AIS A; negro, grupo AIS D; verde, controles sanos. Estos genes se caracterizan por una mayor desregulación a medida que aumenta la severidad de la lesión.

Finalmente, analizamos la capacidad de estos 12 genes para estratificar los grupos en función de la severidad mediante análisis de *clustering* en ratas y humanos. En el análisis de muestras de sangre humana se observa una distribución de las muestras en tres *clusters* específicos de cada grupo, donde las muestras de lesión severa (AIS A) se separan en una rama independiente (Figura 35C). Por otro lado, esta firma de 12 genes también permite la estratificación de los perfiles transcripcionales consenso de rata en función de la severidad (Figura 35D).

Estos hallazgos indican que los niveles de expresión de los 12 genes biomarcadores identificados pueden delinear subgrupos específicos de severidad en muestras de sangre humana y de médula espinal de rata.

4.10. Meta-SCI app - una plataforma interactiva e intuitiva para explotar los resultados

Hemos desarrollado Meta-SCI (<https://metasci-cbl.shinyapps.io/metaSCI/>), una aplicación web interactiva que proporciona una interfaz amigable para consultar y visualizar los resultados obtenidos en este trabajo, permitiendo a los usuarios profundizar en la comprensión de los mecanismos subyacentes a la LM.

Este recurso consta de ocho módulos principales (Figura 37). El módulo de "Meta-analysis" permite generar gráficos de bosque y embudo para genes de interés y proporciona acceso a todos los resultados detallados del metaanálisis.

Con la herramienta de "Clustering" se pueden generar PCAs y análisis de *clustering* de los perfiles transcripcionales consenso utilizando genes de interés seleccionados por el usuario o listas predefinidas, con el objetivo de identificar una clasificación de los grupos en base a la severidad o a la fase de la lesión. El análisis de *clustering* se puede realizar en los ocho perfiles consenso o solo en uno de los modelos de severidad. Además, también se puede realizar un *clustering* funcional basado en los resultados del GSA.

En el módulo "Gene patterns" se puede comparar la progresión a lo largo del tiempo de los cambios de expresión entre severidades para genes de interés. Asimismo, también ayuda a identificar patrones de expresión en genes altamente correlacionados entre lesiones moderadas y severas.

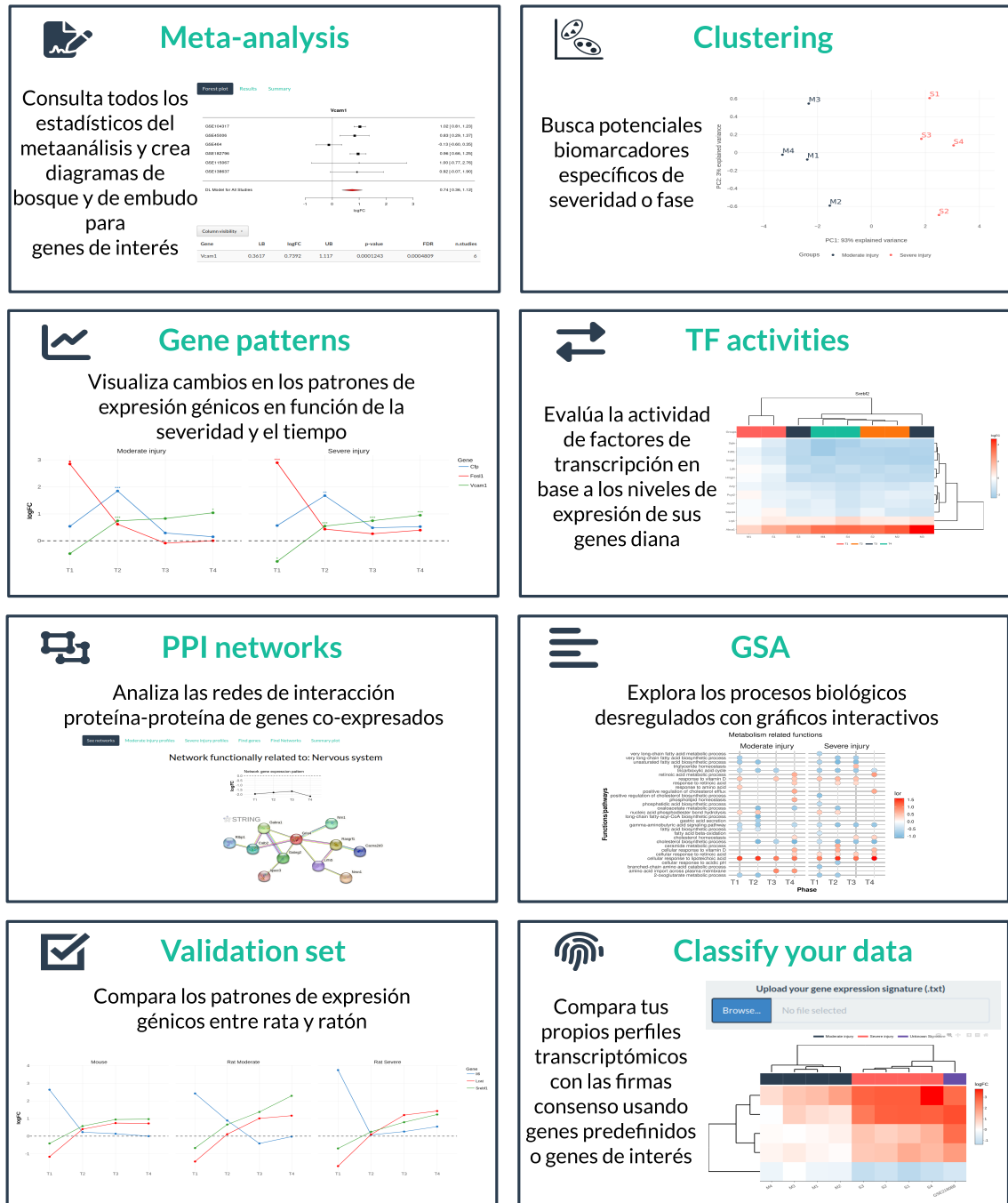


Figura 37. Visión general de los principales módulos de Meta-SCI app. (1) Meta-analysis; (2) Clustering; (3) Gene patterns; (4) TF (transcription factors) activities; (5) PPI (protein-protein interaction) networks; (6) GSA (gene set analysis); (7) Validation set; (8) Classify your data. Meta-SCI app está disponible en <https://metasci-cbl.shinyapps.io/metaSCI/>.

La sección "TF2target" del módulo "TF activities" ofrece la visualización de heatmaps con *clustering* jerárquico basados en la expresión de genes diana de los TF y proporciona las estadísticas detalladas tanto de la inferencia de actividad de los TF, como de sus genes diana. Además, en la sección "Dotplot" se pueden generar gráficos de puntos para comparar los niveles de actividad de cualquier TF evaluado.

En el módulo de "PPI networks", los usuarios pueden explorar todas las redes de co-expresión construidas. Todos los análisis disponibles en la web STRING se pueden utilizar directamente en este módulo para analizar funcionalmente cada red. Los usuarios pueden buscar redes que contengan un gen de interés o bien redes relacionadas con una categoría o término funcional específico.

El módulo "GSA" comprende tres secciones diferentes. Con "GSA results", se puede acceder a todas las estadísticas detalladas del GSA. Utilizando la herramienta "Dotplot", los usuarios pueden generar gráficos de puntos para funciones/rutas de interés, buscarlas por palabras clave o visualizar gráficos predefinidos de categorías biológicas relevantes. Todos los gráficos son dinámicamente reactivos a los filtros aplicados. Similar a la sección "TF2target", la sección "gene2path" muestra heatmaps con *clustering* jerárquico con el nivel de expresión de los genes que pertenecen a un término funcional. Las tablas de la parte inferior permiten consultar tanto las estadísticas del GSA del término funcional que se está visualizando, así como la de todos los genes individuales que pertenecen al mismo.

El módulo "Validation set" integra las herramientas mencionadas anteriormente, "Gene patterns" y "Clustering". En este caso, el módulo facilita la comparación de patrones de expresión de rata con el *dataset* de ratón, permitiendo visualizar los genes altamente correlacionados entre ambas especies. Por otro lado,

realiza análisis de *clustering* de los perfiles transcriptómicos de ratones y ratas, permitiendo la identificación de biomarcadores comunes para los dos modelos.

Finalmente, la herramienta "Classify your data" permite a los usuarios analizar perfiles de expresión que introduzcan como *input* para evaluar su similitud con los perfiles transcripcionales consenso generados en metaanálisis, ya sea usando genes de interés seleccionados por el usuario o bien listas de genes predefinidas.

Adicionalmente se incluye un módulo de ayuda, "Help", en el que se puede encontrar información detallada de las diferentes secciones de Meta-SCI app, y el módulo "Others", donde se puede consultar información general sobre el diseño de nuestro metaanálisis, consultar los resultados individuales de los estudios y descargar tanto metadatos como las matrices de expresión utilizadas.

Para facilitar su uso y agilizar su difusión, también se ha creado un canal de Youtube con tutoriales donde se explica el uso e interpretación de algunas de las principales secciones, disponible en <https://www.youtube.com/@metasci>.

Así, presentamos Meta-SCI app como una herramienta innovadora para la comunidad de investigadores en LM. Ofrecemos una batería de herramientas para consultar los diferentes análisis generados en este trabajo, donde todos los gráficos y datos presentados son dinámicos e interactivos, lo que permite a los usuarios explorar los resultados con mayor detalle y facilidad, pudiendo descargarlos para llevar a cabo estudios adicionales de los mismos.

5. Discusión

La LM representa un complejo desafío que requiere de nuevas aproximaciones para una mejor comprensión de los mecanismos subyacentes a ella. En este trabajo presentamos un metaanálisis transcriptómico que reúne el mayor número de estudios y muestras en *datasets* de rata, hasta donde llega nuestro conocimiento. Además, hemos integrado los diferentes modelos experimentales en base a la severidad y al tiempo tras la lesión, permitiendo evaluar estos factores en la progresión de la LM tanto a nivel de expresión génica como a nivel funcional, creando por primera vez una referencia transcriptómica de libre acceso para el estudio de la LM.

Una de las fortalezas de nuestro trabajo radica en la generación de firmas transcripcionales consenso de los grupos establecidos. La aplicación de un procesamiento uniforme de los *datasets* individuales y su integración mediante técnicas de metaanálisis, aumenta la robustez estadística al determinar la magnitud de los cambios de expresión. Una posible limitación surge de la inclusión de *datasets* derivados de diversas tecnologías y plataformas (4 RNA-seq, 10 microarrays). Sin embargo, con este reanálisis estandarizado y la integración tras el análisis de expresión diferencial conseguimos mitigar la heterogeneidad inherente de cada estudio y los sesgos derivados de tecnologías y *pipelines* bioinformáticos diferentes [116]. Esto supone un logro significativo teniendo en cuenta que estudios con diseños experimentales idénticos pueden reportar cambios de expresión opuestos para un mismo gen [63–65]. La validación de resultados supone otro punto fuerte de nuestro trabajo. Por un lado, con la validación experimental mediante qPCR de ocho genes (cuatro TF y cuatro genes diana), confirmamos la dirección de la desregulación en todos los casos y el

patrón de expresión de los cuatro genes diana. Por otro lado, con la validación bioinformática identificamos genes con el mismo patrón de expresión temporal en un *dataset* externo de ratón. También demostramos la validez de los genes biomarcadores de severidad y fase propuestos dada su capacidad de clasificar correctamente una nueva firma transcripcional de rata. Nuestro trabajo facilita la comparación de los niveles de expresión génica, funciones/rutas y actividad de TF a lo largo del tiempo y en base a la severidad de la lesión. Además, posibilitamos la identificación directa de los niveles de expresión génica asociados a funciones/rutas o TF específicos, siendo accesible de manera interactiva mediante Meta-SCI app.

Si bien una de las fortalezas de nuestro trabajo reside en la recopilación de 14 estudios y 273 muestras, la representación desigual de los grupos experimentales supone la principal limitación. Existe un mayor número de estudios que contienen muestras de fases agudas (T1 y T2) con respecto a los que incluyen muestras de las fases crónicas (T3 y T4). Por ejemplo, los grupos M2 y S2 incluyen muestras de seis estudios diferentes, mientras que el grupo S4 contiene muestras de solo dos estudios. Hasta donde sabemos, previamente el trabajo de Squair et al. [96] integró el mayor número de *datasets* transcriptómicos de LM en ratas hasta 2018. Sin embargo, nuestra investigación representa un avance significativo, ya que incorpora un mayor número de *datasets* (14 vs. 3) y emplea métodos de integración más sofisticados mediante el uso de metaanálisis, en lugar de una intersección de resultados. La limitación de la disponibilidad de muestras revela dos problemáticas. Primero, a nivel biológico, se pone de manifiesto la necesidad de más estudios que investiguen las fases crónicas, con el objetivo de ganar una mayor comprensión de sus mecanismos que permitan el desarrollo de terapias que supongan una recuperación funcional en los pacientes que ya tienen una lesión estable. Segundo, a nivel de comunidad científica, se subraya la necesidad

de publicar datos siguiendo los principios FAIR (del inglés, *findability, accessibility, interoperability, and reusability*) [117], ya que ciertos autores no hacen públicos los *datasets*, no realizan descripciones adecuadas de las muestras o se niegan a compartir cualquier información, lo que dificulta el avance de la ciencia.

En este trabajo se han generado ocho perfiles transcripcionales consenso que caracterizan cada uno de los grupos establecidos. Estos perfiles transcriptómicos sirven en primer lugar para observar el patrón de expresión a través del tiempo y la severidad en genes concretos. Posteriormente, se utilizaron con un doble propósito: por un lado la identificación de genes biomarcadores de ambas variables; y por otro, su caracterización funcional mediante diferentes métodos (GSA, inferencia de actividad de TF, construcción de redes de co-expresión). La solidez de estos perfiles de expresión se respaldan por diferentes fuentes. Anteriormente, Tica et al. [118] integraron un *dataset* transcriptómico y uno proteómico para identificar genes/proteínas persistentemente desregulados a los 7 días y 8 semanas tras la lesión. Encontraron 40 genes/proteínas *up*-regulados y 48 *down*-regulados, los cuales coinciden consistentemente con nuestros resultados. Cabe destacar que *Vim* es uno de los genes que se identificó como permanentemente *up*-regulado. En nuestro trabajo también hemos identificado una *up*-regulación persistente de *Vim* en ambas severidades del metaanálisis, en el *dataset* de ratón y en la validación por qPCR.

También validamos los perfiles de expresión del metaanálisis comparándolos con los perfiles de expresión de un *dataset* de ratón con un diseño experimental semejante a los estudios incluidos en nuestro trabajo [86]. Observamos una correlación general positiva entre los patrones de expresión génica de ratones y ratas, tanto para lesiones moderadas como severas. Al comparar estos tres sistemas, 2080 genes únicos presentaron coeficientes de correlación superiores a 0.9 en al menos una de las tres comparaciones, con 198 genes comunes entre

todos los sistemas biológicos, lo que refuerza la validez de nuestros resultados. Si bien la replicación exacta de los patrones de expresión resulta complicada, la validación experimental vía qPCR confirmó la dirección de la desregulación (*up*-regulación o *down*-regulación) en los diferentes puntos temporales. Así, estas firmas transcripcionales consenso proporcionan información sobre cómo evolucionan los patrones de expresión, ofreciendo información valiosa para el desarrollo de estrategias terapéuticas adecuadas para cada fase de la LM.

Hemos identificado con éxito genes biomarcadores que distinguen claramente entre lesiones moderadas y severas, los cuales se caracterizan por tener una desregulación significativa y persistente en las lesiones severas en comparación con las lesiones moderadas. De estos, los genes *down*-regulados desempeñan diferentes funciones relacionadas con el desarrollo del sistema nervioso y las sinapsis. Estos hallazgos reflejan una mayor pérdida de funcionalidad asociada a lesiones severas debido a un daño tisular más extenso y una ruptura axonal completa. Por otro lado, los genes *up*-regulados están asociados con la remodelación tisular y la organización de la matriz extracelular. Notablemente, 45 de estos genes se traducen a proteínas que se transportan a la región extracelular. En lesiones severas, estos genes podrían activarse para regular mecanismos involucrados en la cicatrización de heridas y fibrosis [119]. Usando el top 10 genes específicos de severidad conseguimos clasificar correctamente un *dataset* externo de modelo de lesión severa en ratas, demostrando la validez de los mismos. Dichos genes presentan un perfil de expresión adecuado para ser utilizados como posibles biomarcadores pronóstico, ya que en lesiones moderadas no hay cambios significativos de expresión mientras que en lesiones severas presentan una desregulación notable y persistente en el tiempo. Así, medir la expresión de estos genes en etapas tempranas de la lesión podría ayudar a predecir su progresión.

Entre los potenciales biomarcadores de severidad se encuentran metaloproteinasas de matriz (MMP), *Mmp9* y *Mmp23*, que tienen una *up*-regulación significativamente mayor en lesiones severas en comparación con las moderadas. En estudios previos se han propuesto las MMP, y concretamente *Mmp9*, como indicadores predictivos de un peor resultado neurológico en pacientes humanos, así como en modelos de rata, ratón y perro [120–124]. Esta consistencia entre especies subraya su fiabilidad para una potencial aplicación en investigación clínica como biomarcadores pronóstico. Una posible explicación del aumento de expresión de MMP es que tras la lesión hay una rotura de la matriz extracelular que contribuye a la destrucción del tejido nervioso. Tanto las células residentes como las células infiltradas del sistema inmune secretan MMP que contribuyen a una mayor destrucción de la matriz extracelular. De esta manera, una lesión más severa y que afecta a un área más extensa, conllevaría a una mayor producción de MMP [119, 125]. Por otro lado, la proteína asociada a microtúbulos 2 (Map2) también ha sido previamente propuesta como un biomarcador de severidad [120, 121, 126]. Nuestros resultados indican una mayor *down*-regulación tanto de *Map2* como de *Map7* en lesiones severas, por lo que podrían ser biomarcadores candidatos interesantes para futuros estudios.

También hemos identificado genes biomarcadores específicos de fase que permitieron la correcta estratificación de los perfiles transcriptómicos en base al tiempo tras la lesión, independientemente de la severidad. En la fase aguda (T1) es donde identificamos un mayor número de potenciales genes biomarcadores, lo que la hace la fase más característica y fácilmente distinguible, como se pudo observar tras la distribución de los grupos obtenidos en el análisis de *clustering* (Figuras 14B-C y 17D-E). El patrón de expresión característico de estos genes es una gran desregulación en T1 que vuelve a niveles basales en las siguientes fases. Esto puede explicarse por la cascada de eventos que se desencadenan

tras el trauma inicial, incluyendo la activación de mecanismos de respuesta al estrés e inflamación para contener la lesión [28, 32]. Además, observamos una gran activación de genes relacionados con el ciclo celular, impulsada por la proliferación/activación de células mitóticas (como microglía y astrocitos), así como el inicio de procesos apoptóticos en células dañadas [105, 106]. De hecho, hay estudios que indican que la inhibición de la expresión de genes del ciclo celular puede mejorar la recuperación motora y cognitiva [105, 106, 127, 128].

Los genes específicos de la fase subaguda (T2), que incluyen genes relacionados con el complemento (por ejemplo, *Cfp*, *C1qa*, *C1qb* y *C1qc*), exhiben un patrón distintivo en el que se alcanza el punto máximo de expresión en T2 y después decrece. Este patrón concuerda con el comportamiento de células como macrófagos y microglía que expresan y secretan proteínas del complemento [129, 130]. Los monocitos migran al lugar de la lesión y se diferencian en macrófagos aproximadamente a los 3 días post-lesión, culminando esta respuesta con un pico alrededor de los 7 días post-lesión. De igual manera, la microglía residente alcanza su pico de activación una semana después de la lesión [131, 132]. La progresiva *up*-regulación a lo largo del tiempo que alcanza su pico en T4 caracteriza a los genes específicos de esta fase, permitiendo que se distingan eficazmente de las otras. El uso de una combinación de genes específicos de fase ha permitido una clasificación precisa tanto de los perfiles transcriptómicos obtenidos en el metaanálisis como el del *dataset* de validación GSE218088. Además, aunque algunos de estos genes son específicos de especie, hay genes comunes (como *Fosl1*, *Il6*, *Cfp*, *Sipa1*, *Cp* y *Dbp*) que también pueden usarse para la clasificación temporal de perfiles transcriptómicos de ratón. La validación experimental mediante qPCR de *Il6*, *Vcam1* y *Vegfa* confirmó los patrones de expresión observados en el metaanálisis y en el *dataset* de ratón. Así, nuestro trabajo ha permitido la identificación de biomarcadores específicos de fase que

podrían ayudar a mejorar la estratificación de pacientes. Además, conocer cómo varía la expresión de los genes a lo largo de la lesión es un factor importante para el diseño de estrategias terapéuticas, ya que permite precisar el momento óptimo de aplicación [133].

En nuestra investigación hemos proporcionado un marco funcional para evaluar el grado de desregulación de procesos biológicos de interés. Los resultados obtenidos concuerdan con investigaciones previas sobre LM, caracterizadas por una mantenida *up*-regulación de funciones relacionadas con el sistema inmune y una *down*-regulación de procesos asociados con la función y desarrollo del sistema nervioso. Este fenómeno está influenciado por otros procesos que sostienen un entorno desfavorable para la recuperación funcional, tales como desequilibrio iónico, lesión vascular, isquemia, estrés por radicales libres, muerte celular [28, 31, 101], disfunción mitocondrial [109], alteraciones metabólicas (incluida la *down*-regulación del metabolismo de lípidos y colesterol) [97] o la remodelación de la matriz extracelular [125]. Considerando la complejidad de este contexto y teniendo en cuenta la naturaleza interconectada de estos sistemas, cualquier estrategia para desarrollar terapias efectivas debería incluir una combinación de múltiples dianas.

A nivel funcional, observamos diferencias entre las fases, siendo particularmente notable la disminución en la importancia de las rutas relacionadas con el ciclo celular, la reparación y replicación del ADN a medida que progresa la lesión. Sin embargo, en contraste con los resultados a nivel de gen, no hemos identificado diferencias sustanciales entre severidades. Asimismo, pese a que las redes de co-expresión obtenidas comprenden diferentes genes dependiendo de la severidad, son similares en cuanto al número y la naturaleza de categorías funcionales enriquecidas. Este análisis corrobora la relación funcional entre genes con patrones de expresión similares [134, 135] y coincide con el estudio de

De Biase et al. [136], donde no se observaron alteraciones consistentes en los procesos biológicos asociados a la severidad de la lesión, a pesar de las aparentes variaciones que ocurren a nivel de expresión génica. Este fenómeno puede explicarse por el diseño de las técnicas de enriquecimiento funcional, en las que se captura la acción coordinada de grupos de genes. En consecuencia, aunque genes específicos pueden mostrar una desregulación más pronunciada en lesiones severas, el cambio consistente en la dirección de la expresión apunta hacia el mismo resultado.

En nuestro trabajo también se ha inferido la actividad de TF a partir de los niveles de expresión de sus genes diana, ofreciendo una nueva perspectiva para el desarrollo de intervenciones terapéuticas más eficientes. De manera similar a los resultados obtenidos en el GSA, hemos identificado cambios temporales, habiendo un mayor número de TF diferencialmente activos en las fases iniciales, que disminuyen en las fases crónicas, existiendo un núcleo común de TF desregulados permanentemente. Aunque no observamos discrepancias generales en función de la severidad, sí identificamos tres TF diferencialmente activos en lesiones severas pero no en moderadas: Bhlhe22, Sp3 y Runx2. Bhlhe22 pertenece a la familia bHLH (del inglés, *basic-helix-loop-helix*), conocida por su papel regulador en el desarrollo del sistema nervioso. Específicamente, Bhlhe22 actúa reprimiendo la expresión de sus genes diana para controlar la diferenciación neuronal [137, 138]. La *up*-regulación que muestran sus genes diana lesiones severas evidencia una disminución de funcionalidad de Bhlhe22, que se correlaciona con la pérdida de la identidad neuronal. Esto podría explicar el mayor impacto en los procesos relacionados con el funcionamiento y el desarrollo del sistema nervioso en lesiones severas. Tanto Sp3 como Runx2 son activadores transcripcionales que modulan la expresión de genes que intervienen en procesos celulares generales como el ciclo celular, apoptosis, metabolismo, así como en el

sistema inmune y vascular [139–143]. Teniendo en cuenta que el papel que juegan Sp3 y Runx2 en la LM sigue siendo poco comprendido, y además es dependiente de cada tipo celular en específico, hace que estos TF sean objetivos interesantes para futuros estudios.

Hemos validado experimentalmente cuatro factores de transcripción mediante qPCR: *Hes1*, *Onecut1*, *Cebpd* y *Jun*. Nuestros resultados indican la persistente *up*-regulación de *Hes1* tanto en el metaanálisis como por qPCR, además de su activación diferencial en todos los grupos. *Hes1* es un represor transcripcional involucrado en la neurogénesis [144]. Estudios previos han demostrado la sobreexpresión de *Hes1* tras una LM y que su inhibición mediante microRNA-381 tiene un efecto protector en la LM [145, 146]. *Onecut1* es un activador transcripcional que regula la producción, diversificación, distribución y mantenimiento de varias poblaciones neuronales [147–149]. Aunque por qPCR no pudimos validar el patrón de expresión obtenido en el metaanálisis, posiblemente debido a la variabilidad inter-experimental, sí pudimos confirmar su *down*-regulación tras la LM. El análisis de inferencia de actividad indicó una pérdida de función de *Onecut1*, especialmente en las primeras fases de la lesión. Estos resultados sugieren que la sobreexpresión o el incremento de la actividad de *Onecut1* podrían ser beneficiosos para el desarrollo y la función neuronal. Nuestros resultados bioinformáticos y experimentales también indicaron una *up*-regulación de *Cebpd*, así como su activación diferencial en fases agudas. *Cebpd* es un activador transcripcional que regula genes involucrados en respuestas inmunitarias [150]. Estudios previos en ratones han demostrado la sobreexpresión de *Cebpd* tras una LM y cómo su inhibición mejora la recuperación funcional [150, 151]. En nuestros resultados también hemos identificado la *up*-regulación y activación diferencial en todos los grupos experimentales de *Jun*, un activador transcripcional altamente inducido tras el daño neuronal [152, 153].

Zhang et al. [153] propusieron que la inhibición de Jun mediante miR-152 promovía la recuperación funcional al reducirse la respuesta inflamatoria.

Encontramos de gran interés la identificación de una firma de 12 genes capaces de clasificar la severidad de la lesión a partir de perfiles transcriptómicos de sangre de pacientes y de médula espinal de ratas. El principal desafío radica en la imposibilidad de obtener datos transcriptómicos de médula espinal en pacientes humanos, sumado a la ausencia de *datasets* que permitan comparar lesiones moderadas y severas en muestras de sangre de rata, lo que impide la comparación directa entre tejidos. Por estas razones, resulta valiosa la identificación de biomarcadores comunes en médula y sangre entre ambas especies. Estos resultados sugerirían que algunos procesos biológicos podrían estar conservados y ser detectables en sangre periférica, incluso si la lesión se localiza en la médula espinal. Cabe destacar que de estos 12 genes, *MMP9* es el que más ha sido mencionado en la literatura. Previamente *MMP9* se ha propuesto como un biomarcador pronóstico de severidad en diferentes especies, incluyendo humanos, ratas, ratones y perros [120–124]. Aunque identificar un biomarcador tanto en sangre como en tejido es complicado, *MMP9* también cuenta con este respaldo. En el trabajo de Qi et al. [154] identificaron una firma de genes, entre las que se encontraba *MMP9*, que permitía identificar pacientes con glioblastoma a partir de perfiles transcriptómicos de sangre y tejido tumoral. En el contexto de la LM, una mayor expresión de *MMP9* en sangre podría deberse a una mayor activación de diferentes poblaciones del sistema inmune, que secretan esta proteína para degradar la barrera hematoencefálica e incrementar la permeabilidad vascular para poder extravasarse al lugar de lesión [101, 155]. Por otro lado, en la médula espinal tanto las células residentes como infiltradas, junto a células endoteliales secretan *MMP9* para favorecer la permeabilidad vascular y remodelar la matriz extracelular [28, 119, 123, 125].

Los pacientes con LM son propensos a desarrollar complicaciones crónicas secundarias como disfunciones respiratorias, complicaciones cardiovasculares, síndromes dolorosos, osteoporosis e inmunodeficiencia, entre otras. Así pues, los biomarcadores para el diagnóstico y pronóstico precoz que permitan la prevención de complicaciones secundarias tras una LM son fundamentales para la calidad de vida de los pacientes. El valor de la detección de biomarcadores en sangre periférica humana radica en su aplicación práctica. La extracción de sangre periférica es un procedimiento simple y no invasivo, lo que facilita su aplicación en entornos clínicos, como hospitales. Esta accesibilidad significa que, si se validan, estos biomarcadores podrían servir como indicadores de pronóstico, proporcionando una forma sencilla y económica de controlar la progresión o la gravedad de la LM en pacientes.

Actualmente el sistema ASIA, acompañado de evidencia anatómica de la extensión del tejido afectado y preservado mediante resonancia magnética (MRI), representa el estándar para predecir la gravedad de la lesión de la médula espinal. Sin embargo, este método presenta limitaciones. El sistema ASIA se basa en evaluaciones del estado neurológico a través de exámenes sensoriales y motores. En la fase aguda puede ser poco preciso y difícil de obtener en pacientes que no responden o están inconscientes. El sistema ASIA tampoco considera la heterogeneidad de los pacientes, por ejemplo, individuos más jóvenes tienen mayor potencial de tener una recuperación funcional [90, 156]. Por ello existe la necesidad de encontrar nuevos métodos operativos y no invasivos para el pronóstico de la severidad, como una firma transcripcional de sangre periférica. Aunque existe evidencia de genes biomarcadores con capacidad pronóstica comunes entre especies, nuestro trabajo subraya la necesidad de más estudios, y sobre todo, tamaños muestrales más grandes, para evaluar su efectividad en pacientes o para identificar nuevos genes específicos de humanos que puedan

llegar a ser una herramienta práctica en la clínica.

Actualmente los trabajos bioinformáticos generan un volumen de resultados cada vez mayor. Esto se debe a que cada vez es más habitual que los estudios incluyan el análisis de múltiples *datasets*, ya sea con motivo de integración, como trabajos multiómicos o metaanálisis, o como *datasets* de validación. No solo puede haber diferentes *datasets*, sino que también suele haber diversos grupos experimentales, comparaciones y tipos de análisis bioinformáticos. Además, en los cada vez más habituales experimentos de célula única, todo lo anterior se aplica a cada tipo celular identificado. Esta situación hace que los estudios actuales generen un volumen de resultados muy elevado, de manera que publicarlos mediante la manera tradicional, haciendo uso del material suplementario con tablas en Excel y documentos PDF sea ineficiente y poco práctico. Por ello, es necesario que se apliquen y establezcan nuevos recursos y formatos que permitan compartir y facilitar la interpretación de los resultados a la comunidad científica.

Con ese objetivo, hemos desarrollado Meta-SCI app. Esta plataforma va más allá de ser un mero repositorio de resultados; Meta-SCI app sirve como una herramienta de análisis para la comunidad investigadora. La aplicación permite a los investigadores acceder a los resultados detallados de los diferentes análisis, generar gráficos y explorar las relaciones entre genes y funciones que coincidan con sus intereses de investigación, todo ello de manera dinámica e interactiva. Además, Meta-SCI app incluye dos modelos de lesión, permitiendo a los investigadores elegir el que mejor se adapte a sus necesidades. Así, Meta-SCI app se convierte en un recurso valioso para la generación y validación de hipótesis y constituye una base de datos de referencia transcripcional para el estudio de la LM.

6. Conclusiones

1. Hemos recopilado un total de 14 estudios y 273 muestras para llevar a cabo el primer metaanálisis transcriptómico en ratas con LM, incorporando la perspectiva de severidad y tiempo post-lesión.
2. Se han obtenido ocho firmas transcripcionales consenso que permiten caracterizar la LM en función de la severidad y su progresión a lo largo de la lesión.
3. Los genes biomarcadores específicos de severidad identificados permiten diferenciar la lesión severa de la moderada. El patrón de expresión que presentan los hace adecuados para predecir la prognosis de la lesión. Los genes *up*-regulados están involucrados en la remodelación de la matriz extracelular y desarrollo de tejido, mientras que los genes *down*-regulados tienen funciones relacionadas con el sistema nervioso.
4. Los genes biomarcadores específicos de fase identificados estratifican correctamente los perfiles transcriptómicos, independientemente de la severidad, debido a los patrones de expresión característicos de cada punto temporal. La fase aguda es la más distintiva de las cuatro. Los genes *up*-regulados de T1 están involucrados en ciclo celular, respuesta a estrés y diferentes procesos metabólicos, mientras que los *down*-regulados intervienen en el desarrollo del sistema nervioso y el metabolismo de lípidos. Para T2 los genes específicos están asociados con la respuesta inmune y activación de macrófagos, encontrándose entre ellos genes del complemento. Las funciones asociadas a los genes específicos de T4 están relacionadas con el metabolismo de lípidos y respuesta a estrés.

5. Los resultados de los diferentes abordajes de caracterización funcional de los perfiles transcriptómicos consenso son consistentes entre sí. No se observan diferencias notables a nivel de severidad pero sí a nivel temporal, con un mayor número de procesos desregulados en la fase aguda que decrece en las siguientes fases. En la fase aguda se identifica una *up*-regulación de funciones relacionadas con el ciclo celular, la reparación del ADN o la replicación del ADN que disminuyen en las fases posteriores. Por otro lado, hay un conjunto de funciones permanentemente desreguladas, donde destacan la constante activación del sistema inmune, incluso en fases crónicas, así como la pérdida de las funciones relacionadas con el sistema nervioso, la función mitocondrial o el metabolismo de lípidos.
6. Se ha demostrado la capacidad de los genes biomarcadores específicos de severidad y fase seleccionados de clasificar correctamente un nuevo perfil transcriptómico de rata independiente del metaanálisis.
7. Los patrones de expresión obtenidos en el metaanálisis muestran una correlación positiva general con los patrones de expresión de un *dataset* externo de ratón, identificándose 1497 genes con un coeficiente de correlación superior a 0.9 entre rata y ratón.
8. Se ha demostrado mediante validación experimental que el sentido de la desregulación obtenido mediante la qPCR y el metaanálisis coincide de forma consistente. Los patrones de expresión de la qPCR y los del metaanálisis muestran una mayor correlación para los genes diana (10 de 12 comparaciones tienen un coeficiente de correlación superior a 0.85) que para los TFs (5 de 11 comparaciones tienen un coeficiente de correlación superior a 0.7).
9. Hemos identificado una firma de 12 genes que permite estratificar los grupos en función de la severidad tanto en muestras humanas de sangre como los

perfiles consenso de médula rata, lo que podría ayudar en la prognosis de la lesión de los pacientes.

10. El desarrollo de Meta-SCI app permite compartir los resultados obtenidos y facilitar su interpretación a la comunidad científica, en un entorno dinámico e interactivo, pudiendo establecerse como una referencia transcriptómica para el estudio de la LM.

Bibliografía

- [1] María Angustias Torres Alaminos. Aspectos epidemiológicos de la lesión medular en el hospital nacional de parapléjicos. *Ene*, 12(2), 2018.
- [2] Jarred M Griffin and Frank Bradke. Therapeutic repair for spinal cord injury: combinatory approaches to address a multifaceted problem. *EMBO molecular medicine*, 12(3):e11505, 2020.
- [3] Xu-Chang Hu, Yu-Bao Lu, Yong-Na Yang, Xue-Wen Kang, Yong-Gang Wang, Bing Ma, and Shuai Xing. Progress in clinical trials of cell transplantation for the treatment of spinal cord injury: how many questions remain unanswered? *Neural regeneration research*, 16(3):405–413, 2021.
- [4] Beatriz F Ribeiro, Bruna C da Cruz, Bárbara M de Sousa, Patrícia D Correia, Nuno David, Camila Rocha, Ramiro D Almeida, Maria Ribeiro da Cunha, António A Marques Baptista, and Sandra I Vieira. Cell therapies for spinal cord injury: a review of the clinical trials and cell-type therapeutic potential. *Brain*, 146(7):2672–2693, 2023.
- [5] Yiran Li, Ting Dong, Zhiwei Li, Shilei Ni, Fang Zhou, Olawale A Alimi, Shaojuan Chen, Bin Duan, Mitchell Kuss, and Shaohua Wu. Review of advances in electrospinning-based strategies for spinal cord regeneration. *Materials Today Chemistry*, 24:100944, 2022.
- [6] Christopher S Ahuja, Jefferson R Wilson, Satoshi Nori, Mark RN Kotter, Claudia Druschel, Armin Curt, and Michael G Fehlings. Traumatic spinal cord injury. *Nature reviews Disease primers*, 3:17018, 2017.

- [7] Anoushka Singh, Lindsay Tetreault, Suhkvinder Kalsi-Ryan, Aria Nouri, and Michael G Fehlings. Global prevalence and incidence of traumatic spinal cord injury. *Clinical epidemiology*, pages 309–331, 2014.
- [8] YDHZH Kang, Han Ding, Hengxing Zhou, Zhijian Wei, Lu Liu, Dayu Pan, and Shiqing Feng. Epidemiology of worldwide spinal cord injury: a literature review. *J Neurorestorator*, 6(1):3, 2018.
- [9] A Huete, E Díaz, E Ortega, and A Esclarín. Análisis sobre la lesión medular en españa. *Federación Nacional de ASPAYM. Toledo: Intersocial*, 2012.
- [10] Avital Yohann, Laura N Purcell, and Anthony Charles. Traumatic spinal cord injury and outcomes in low-resource settings. In *Diagnosis and Treatment of Spinal Cord Injury*, pages 3–14. Elsevier, 2022.
- [11] Luz Helena Lugo, Fabio Salinas, and Héctor Iván García. Out-patient rehabilitation programme for spinal cord injured patients: Evaluation of the results on motor fim score. *Disability and Rehabilitation*, 29(11-12):873–881, 2007.
- [12] BB Lee, Raymond A Cripps, Michael Fitzharris, and PC Wing. The global map for traumatic spinal cord injury epidemiology: update 2011, global incidence rate. *Spinal cord*, 52(2):110–116, 2014.
- [13] N Derakhshanrad, MS Yekaninejad, F Vosoughi, F Sadeghi Fazel, and H Saberi. Epidemiological study of traumatic spinal cord injuries: experience from a specialized spine center in iran. *Spinal cord*, 54(10):901–907, 2016.
- [14] Diana M Molinares, David R Gater, Scott Daniel, and Nicole L Pontee. Nontraumatic spinal cord injury: epidemiology, etiology and management. *Journal of personalized medicine*, 12(11):1872, 2022.

- [15] Lukas Grassner, Julia Marschallinger, Martin W Dünser, Helmut F Novak, Alexander Zerbs, Ludwig Aigner, Eugen Trinkla, and Johann Sellner. Nontraumatic spinal cord injury at the neurological intensive care unit: spectrum, causes of admission and predictors of mortality. *Therapeutic advances in neurological disorders*, 9(2):85–94, 2016.
- [16] SNS Cosar, OU Yemisci, P Oztop, Nuri Çetin, B Sarifakioglu, SA Yalbuzzdag, K Ustaomer, and M Karatas. Demographic characteristics after traumatic and non-traumatic spinal cord injury: a retrospective comparison study. *Spinal cord*, 48(12):862, 2010.
- [17] Katherine Pérez, Ana M Novoa, Elena Santamariña-Rubio, Yislenz Narvaez, Vita Arrufat, Carme Borrell, Elena Cabeza, Eva Cirera, Josep Ferrando, Anna García-Altés, et al. Incidence trends of traumatic spinal cord injury and traumatic brain injury in spain, 2000–2009. *Accident Analysis & Prevention*, 46:37–44, 2012.
- [18] Merce Avellanet and Miguel Angel Gonzalez-Viejo. People with spinal cord injury in spain. *American Journal of Physical Medicine & Rehabilitation*, 96(2):S112–S115, 2017.
- [19] Miguel Angel González-Viejo, Merce Avellanet, Anna Boada-Pladellorens, Lluïsa Montesinos-Magraner, María Luisa Jaúregui-Abrisqueta, Enrique Bárbara-Bataller, Bosco Méndez-Ferrer, Judith Sánchez-Raya, Nora Cívicos, José Luis Méndez-Suarez, et al. International spinal cord injury community survey: Socioeconomic and healthcare satisfaction in spain. *Global Spine Journal*, 14(8):2381–2388, 2024.
- [20] Michael DeVivo, Yuying Chen, Stephen Mennemeyer, and Anne Deutsch. Costs of care following spinal cord injury. *Topics in spinal cord injury rehabilitation*, 16(4):1–9, 2011.

- [21] Ian Peate and Muralitharan Nair. *Fundamentals of anatomy and physiology: For nursing and healthcare students (2nd ed.)*. John Wiley & Sons, 2016.
- [22] Kenneth Saladin. *Human Anatomy (5th ed.)*. McGraw-Hill Education, 2017.
- [23] Kenneth V Kardong et al. *Vertebrates: comparative anatomy, function, evolution*. McGraw-Hill New York, 2006.
- [24] M. V. Pankiv, Ye. V. Paltov, and Z. Z. Masna. Structural components of the normal thoracic and lumbar vertebrae in rats. *Journal of Education, Health and Sport*, 11(1):163–169, Jan. 2021.
- [25] Gerard J Tortora and Bryan Derrickson. *Principios de anatomía y fisiología (13th ed.)*. Médica Panamericana, 2013.
- [26] Timothy T Roberts, Garrett R Leonard, and Daniel J Cepela. Classifications in brief: American spinal injury association (asia) impairment scale, 2017.
- [27] Kristine O'Phelan and Indira De Jesus. Management of the spinal cord injury in the neurocritical care unit, 2019.
- [28] Arsalan Alizadeh, Scott Matthew Dyck, and Soheila Karimi-Abdolrezaee. Traumatic spinal cord injury: an overview of pathophysiology, models and acute injury mechanisms. *Frontiers in neurology*, 10:282, 2019.
- [29] Miguel A Ortega, Oscar Fraile-Martinez, Cielo García-Montero, Sergio Haro, Miguel Ángel Álvarez-Mon, Diego De Leon-Oliva, Ana M Gomez-Lahoz, Jorge Monserrat, Mar Atienza-Pérez, David Díaz, et al. A comprehensive look at the psychoneuroimmunoendocrinology of spinal cord injury and its progression: mechanisms and clinical opportunities. *Military Medical Research*, 10(1):26, 2023.

- [30] Huber Padilla-Zambrano, Yancarlos Ramos-Villegas, Hernando Raphael Alvis-Miranda, Andrei F Joaquin, and Luis Rafael Moscote-Salazar. Fisiopatología del trauma raquimedular. *Revista Mexicana de Neurociencia*, 18(5):45–53, 2017.
- [31] George I Mataliotakis and Athanasios I Tsirikos. Spinal cord trauma: pathophysiology, classification of spinal cord injury syndromes, treatment principles and controversies. *Orthopaedics and Trauma*, 30(5):440–449, 2016.
- [32] Nuno A Silva, Nuno Sousa, Rui L Reis, and Antonio J Salgado. From basics to clinical: a comprehensive review on spinal cord injury. *Progress in neurobiology*, 114:25–57, 2014.
- [33] Mark A Anderson, Joshua E Burda, Yilong Ren, Yan Ao, Timothy M O’Shea, Riki Kawaguchi, Giovanni Coppola, Baljit S Khakh, Timothy J Deming, and Michael V Sofroniew. Astrocyte scar formation aids central nervous system axon regeneration. *Nature*, 532(7598):195–200, 2016.
- [34] Ina B Wanner, Mark A Anderson, Bingbing Song, Jaclynn Levine, Ana Fernandez, Zachary Gray-Thompson, Yan Ao, and Michael V Sofroniew. Glial scar borders are formed by newly proliferated, elongated astrocytes that interact to corral inflammatory and fibrotic cells via stat3-dependent mechanisms after spinal cord injury. *Journal of Neuroscience*, 33(31):12870–12886, 2013.
- [35] James W Rowland, Gregory WJ Hawryluk, Brian Kwon, and Michael G Fehlings. Current status of acute spinal cord injury pathophysiology and emerging therapies: promise on the horizon. *Neurosurgical focus*, 25(5):E2, 2008.

- [36] Arsalan Alizadeh and Soheila Karimi-Abdolrezaee. Microenvironmental regulation of oligodendrocyte replacement and remyelination in spinal cord injury. *The Journal of physiology*, 594(13):3539–3552, 2016.
- [37] Jacob Kjell and Lars Olson. Rat models of spinal cord injury: from pathology to potential therapies. *Disease models & mechanisms*, 9(10):1125–1137, 2016.
- [38] M Sharif-Alhoseini, M Khormali, M Rezaei, M Safdarian, A Hajighadery, MM Khalatbari, S Meknatkhah, M Rezvan, M Chalangari, P Derakhshan, et al. Animal models of spinal cord injury: a systematic review. *Spinal cord*, 55(8):714–721, 2017.
- [39] AN Minakov, AS Chernov, DS Asutin, NA Konovalov, and GB Telegin. Experimental models of spinal cord injury in laboratory rats. *Acta naturae*, 10(3):4, 2018.
- [40] Reggie Ridlen, Kristine McGrath, and Catherine A Gorrie. Animal models of compression spinal cord injury. *Journal of Neuroscience Research*, 100(12):2201–2212, 2022.
- [41] Norihiko Nakano, Yoshiyasu Nakai, Tae-Beom Seo, Tamami Homma, Yoshihiro Yamada, Masayoshi Ohta, Yoshihisa Suzuki, Toshio Nakatani, Masanori Fukushima, Miki Hayashibe, et al. Effects of bone marrow stromal cell transplantation through csf on the subacute and chronic spinal cord injury in rats. *PLoS One*, 8(9):e73494, 2013.
- [42] KR Pritchett and BF Corning. Biology and medicine of rats. *Laboratory Animal Medicine and Management. International Veterinary Information Service. Ithaca NY*, 2004.

- [43] John D Houle and Alan Tessler. Repair of chronic spinal cord injury. *Experimental neurology*, 182(2):247–260, 2003.
- [44] ZhiCheng Dong and Yan Chen. Transcriptomics: advances and approaches. *Science China Life Sciences*, 56:960–967, 2013.
- [45] Daniel Castillo, Juan Manuel Gálvez, Luis Javier Herrera, Belén San Román, Fernando Rojas, and Ignacio Rojas. Integration of rna-seq data with heterogeneous microarray data for breast cancer profiling. *BMC bioinformatics*, 18(1):506, 2017.
- [46] Rohan Lowe, Neil Shirley, Mark Bleackley, Stephen Dolan, and Thomas Shafee. Transcriptomics technologies. *PLoS computational biology*, 13(5): e1005457, 2017.
- [47] David Lähnemann, Johannes Köster, Ewa Szczurek, Davis J McCarthy, Stephanie C Hicks, Mark D Robinson, Catalina A Vallejos, Kieran R Campbell, Niko Beerenwinkel, Ahmed Mahfouz, et al. Eleven grand challenges in single-cell data science. *Genome biology*, 21:1–35, 2020.
- [48] Peter V Kharchenko. The triumphs and limitations of computational methods for scrna-seq. *Nature methods*, 18(7):723–732, 2021.
- [49] Yuqi Cao, Shunxing Zhu, Bin Yu, and Chun Yao. Single-cell rna sequencing for traumatic spinal cord injury. *The FASEB Journal*, 36(12):e22656, 2022.
- [50] Leilei Gong, Yun Gu, Xiaoxiao Han, Chengcheng Luan, Chang Liu, Xinghui Wang, Yufeng Sun, Mengru Zheng, Mengya Fang, Shuhai Yang, et al. Spatiotemporal dynamics of the molecular expression pattern and intercellular interactions in the glial scar response to spinal cord injury. *Neuroscience bulletin*, 39(2):213–244, 2023.

- [51] Sonia Verma, Abhay Kumar, Rajiv Narang, Akshya K Bisoi, and Dipendra K Mitra. Signature transcriptome analysis of stage specific atherosclerotic plaques of patients. *BMC Medical Genomics*, 15(1):99, 2022.
- [52] Irene Pérez-Díez, Zoraida Andreu, Marta R Hidalgo, Carla Perpiñá-Clérigues, Lucía Fantín, Antonio Fernandez-Serra, María de la Iglesia-Vaya, José A Lopez-Guerrero, and Francisco García-García. A comprehensive transcriptional signature in pancreatic ductal adenocarcinoma reveals new insights into the immune and desmoplastic microenvironments. *Cancers*, 15(11):2887, 2023.
- [53] Julie G Burel and Bjoern Peters. Discovering transcriptional signatures of disease for diagnosis versus mechanism. *Nature reviews Immunology*, 18(5):289–290, 2018.
- [54] Joaquín Dopazo. Functional profiling methods in cancer. In *Cancer Gene Profiling*, pages 363–374. Springer, 2009.
- [55] Joaquin Dopazo. Functional interpretation of microarray experiments. *Omics: a journal of integrative biology*, 10(3):398–410, 2006.
- [56] Kangmei Zhao and Seung Yon Rhee. Interpreting omics data with pathway enrichment analysis. *Trends in Genetics*, 39(4):308–319, 2023.
- [57] Farhad Maleki, Katie Ovens, Daniel J Hogan, and Anthony J Kusalik. Gene set analysis: challenges, opportunities, and future research. *Frontiers in genetics*, 11:654, 2020.
- [58] Joaquin Dopazo. Formulating and testing hypotheses in functional genomics. *Artificial intelligence in medicine*, 45(2-3):97–107, 2009.
- [59] Damian Szklarczyk, Annika L Gable, David Lyon, Alexander Junge, Stefan Wyder, Jaime Huerta-Cepas, Milan Simonovic, Nadezhda T Doncheva,

- John H Morris, Peer Bork, et al. String v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic acids research*, 47(D1): D607–D613, 2019.
- [60] Samadhi S Wimalagunasekara, Janith WJK Weeraman, Shamala Tirimanne, and Pasan C Fernando. Protein-protein interaction (ppi) network analysis reveals important hub proteins and sub-network modules for root development in rice (*oryza sativa*). *Journal of Genetic Engineering and Biotechnology*, 21(1):69, 2023.
- [61] James E Tomkins and Claudia Manzoni. Advances in protein-protein interaction network analysis for parkinson’s disease. *Neurobiology of Disease*, 155:105395, 2021.
- [62] Jianguo Xia, Maia J Benner, and Robert EW Hancock. Networkanalyst-integrative approaches for protein–protein interaction network analysis and visual exploration. *Nucleic acids research*, 42(W1): W167–W174, 2014.
- [63] Aravind Subramanian, Pablo Tamayo, Vamsi K Mootha, Sayan Mukherjee, Benjamin L Ebert, Michael A Gillette, Amanda Paulovich, Scott L Pomeroy, Todd R Golub, Eric S Lander, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences*, 102(43): 15545–15550, 2005.
- [64] Alison K Merikangas, Matthew Shelly, Alexys Knighton, Nicholas Kotler, Nicole Tanenbaum, and Laura Almasy. What genes are differentially expressed in individuals with schizophrenia? a systematic review. *Molecular psychiatry*, 27(3):1373–1383, 2022.

- [65] Toxicogenomics Research Consortium Weis BK weis@niehs.nih.gov 1 b. Standardizing global gene expression analysis between laboratories and across platforms. *Nature methods*, 2(5):351–356, 2005.
- [66] Nour Shaheen, Ahmed Shaheen, Alaa Ramadan, Mahmoud Tarek Hefnawy, Abdelraouf Ramadan, Ismail A Ibrahim, Maged Elsayed Hassanein, Mohamed E Ashour, and Oliver Flouty. Appraising systematic reviews: a comprehensive guide to ensuring validity and reliability. *Frontiers in research metrics and analytics*, 8:1268045, 2023.
- [67] Daniel Toro-Domínguez, Juan Antonio Villatoro-García, Jordi Martorell-Marugán, Yolanda Román-Montoya, Marta E Alarcón-Riquelme, and Pedro Carmona-Saéz. A survey of gene expression meta-analysis: methods and applications. *Briefings in Bioinformatics*, 22(2):1694–1705, 2021.
- [68] Nicholas A Heard and Patrick Rubin-Delanchy. Choosing between methods of combining-values. *Biometrika*, 105(1):239–246, 2018.
- [69] Tanya Barrett, Stephen E Wilhite, Pierre Ledoux, Carlos Evangelista, Irene F Kim, Maxim Tomashevsky, Kimberly A Marshall, Katherine H Phillippy, Patti M Sherman, Michelle Holko, et al. Ncbi geo: archive for functional genomics data sets—update. *Nucleic acids research*, 41(D1):D991–D995, 2012.
- [70] Nikolay Kolesnikov, Emma Hastings, Maria Keays, Olga Melnichuk, Y Amy Tang, Eleanor Williams, Mirosław Dylag, Natalja Kurbatova, Marco Brandizi, Tony Burdett, et al. Arrayexpress update—simplifying data submissions. *Nucleic acids research*, 43(D1):D1113–D1116, 2014.
- [71] Alessandro Liberati, Douglas G Altman, Jennifer Tetzlaff, Cynthia Mulrow, Peter C Gøtzsche, John PA Ioannidis, Mike Clarke, Philip J Devereaux,

- Jos Kleijnen, and David Moher. The prisma statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *PLoS medicine*, 6(7):e1000100, 2009.
- [72] Matthew E Ritchie, Belinda Phipson, DI Wu, Yifang Hu, Charity W Law, Wei Shi, and Gordon K Smyth. limma powers differential expression analyses for rna-sequencing and microarray studies. *Nucleic acids research*, 43(7):e47–e47, 2015.
- [73] Yoav Benjamini and Yosef Hochberg. 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, 1995.
- [74] Rebecca DerSimonian and Nan Laird. Meta-analysis in clinical trials. *Controlled clinical trials*, 7(3):177–188, 1986.
- [75] Wolfgang Viechtbauer. Conducting meta-analyses in r with the metafor package. *Journal of statistical software*, 36(3):1–48, 2010.
- [76] Florian Rohart, Benoît Gautier, Amrit Singh, and Kim-Anh Lê Cao. mixomics: An r package for ‘omics feature selection and multiple data integration. *PLoS computational biology*, 13(11):e1005752, 2017.
- [77] David Montaner and Joaquin Dopazo. Multidimensional gene set analysis of genomic data. *PloS one*, 5(4):e10348, 2010.
- [78] Imre Vastrik, Peter D'Eustachio, Esther Schmidt, Geeta Joshi-Tope, Gopal Gopinath, David Croft, Bernard de Bono, Marc Gillespie, Bijay Jassal, Suzanna Lewis, et al. Reactome: a knowledge base of biologic pathways and processes. *Genome biology*, 8(3):R39, 2007.

- [79] Michael Ashburner, Catherine A Ball, Judith A Blake, David Botstein, Heather Butler, J Michael Cherry, Allan P Davis, Kara Dolinski, Selina S Dwight, Janan T Eppig, et al. Gene ontology: tool for the unification of biology. *Nature genetics*, 25(1):25, 2000.
- [80] Minoru Kanehisa, Susumu Goto, Shuichi Kawashima, Yasushi Okuno, and Masahiro Hattori. The kegg resource for deciphering the genome. *Nucleic acids research*, 32(suppl_1):D277–D280, 2004.
- [81] Mariano J Alvarez, Federico Giorgi, and Andrea Califano. Using viper, a package for virtual inference of protein-activity by enriched regulon analysis. *Bioconductor*, pages 1–14, 2014.
- [82] Luz Garcia-Alonso, Christian H Holland, Mahmoud M Ibrahim, Denes Turei, and Julio Saez-Rodriguez. Benchmark and integration of resources for the estimation of human transcription factor activities. *Genome research*, 29(8): 1363–1375, 2019.
- [83] Basel Abu-Jamous and Steven Kelly. Clust: automatic extraction of optimal co-expressed gene clusters from gene expression data. *Genome biology*, 19(1):172, 2018.
- [84] Damian Szklarczyk, Annika L Gable, Katerina C Nastou, David Lyon, Rebecca Kirsch, Sampo Pyysalo, Nadezhda T Doncheva, Marc Legeay, Tao Fang, Peer Bork, et al. The string database in 2021: customizable protein–protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic acids research*, 49(D1):D605–D612, 2021.
- [85] Jian Cao, Tao Tang, Jianye Tan, Qi Chen, Jinghong Yuan, Tao Li, and Xigao Cheng. Expression profiles of long noncoding rnas and messenger rnas in

a rat model of spinal cord injury. *Computational and Mathematical Methods in Medicine*, 2023(1):6033020, 2023.

- [86] Chen Li, Zhourui Wu, Liqiang Zhou, Jingliang Shao, Xiao Hu, Wei Xu, Yilong Ren, Xingfei Zhu, Weihong Ge, Kunshan Zhang, et al. Temporal and spatial cellular and molecular pathological alterations with single-cell resolution in the adult spinal cord after injury. *Signal transduction and targeted therapy*, 7(1):65, 2022.
- [87] Beatriz Martínez-Rojas, Esther Giraldo, Rubén Grillo-Risco, Marta R Hidalgo, Eric López-Mocholi, Ana Alastrue-Agudo, Francisco García-García, and Victoria Moreno-Manzano. Npc transplantation rescues sci-driven camp/epac2 alterations, leading to neuroprotection and microglial modulation. *Cellular and Molecular Life Sciences*, 79(8):455, 2022.
- [88] Ana Alastrue-Agudo, Francisco Javier Rodriguez-Jimenez, Eric López Mocholi, Francesca De Giorgio, Slaven Erceg, and Victoria Moreno-Manzano. Fm19g11 and ependymal progenitor/stem cell combinatory treatment enhances neuronal preservation and oligodendrogenesis after severe spinal cord injury. *International Journal of Molecular Sciences*, 19(1):200, 2018.
- [89] Victoria Moreno-Manzano, Francisco Javier Rodríguez-Jiménez, Mireia García-Roselló, Sergio Laínez, Slaven Erceg, Maria Teresa Calvo, Mohammad Ronaghi, Maria Lloret, Rosa Planells-Cases, Jose María Sánchez-Puelles, et al. Activated spinal cord ependymal stem cells rescue neurological function. *Stem cells*, 27(3):733–743, 2009.
- [90] Nikos Kyritsis, Abel Torres-Espín, Patrick G Schupp, J Russell Huie, Austin Chou, Xuan Duong-Fernandez, Leigh H Thomas, Rachel E Tsolinas,

Debra D Hemmerle, Lisa U Pascual, et al. Diagnostic blood rna profiles for human acute spinal cord injury. *Journal of Experimental Medicine*, 218(3): e20201795, 2021.

- [91] Winston Chang, Joe Cheng, JJ Allaire, Carson Sievert, Barret Schloerke, Yihui Xie, Jeff Allen, Jonathan McPherson, Alan Dipert, and Barbara Borges. *shiny: Web Application Framework for R*, 2024. R package version 1.9.1.9000, <https://github.com/rstudio/shiny>.
- [92] Hadley Wickham. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York, 2016. ISBN 978-3-319-24277-4.
- [93] Carson Sievert. *Interactive Web-Based Data Visualization with R, plotly, and shiny*. Chapman and Hall/CRC, 2020. ISBN 9781138331457.
- [94] Galili, Tal, O’Callaghan, Alan, Sidi, Jonathan, Sievert, and Carson. heatmaply: an r package for creating interactive cluster heatmaps for online publishing. *Bioinformatics*, 2017.
- [95] Walter J Jermakowicz, Melissa M Carballosa-Gautam, Alberto A Vitores, and Ian D Hentall. Brainstem-evoked transcription of defensive genes after spinal cord injury. *Frontiers in Cellular Neuroscience*, 13:510, 2019.
- [96] Jordan W Squair, Seth Tigchelaar, Kyung-Mee Moon, Jie Liu, Wolfram Tetzlaff, Brian K Kwon, Andrei V Krassioukov, Christopher R West, Leonard J Foster, and Michael A Skinnider. Integrated systems analysis reveals conserved gene networks underlying response to spinal cord injury. *Elife*, 7:e39188, 2018.
- [97] Redin A Spann, William J Lawson, Raymond J Grill, Michael R Garrett, and Bernadette E Grayson. Chronic spinal cord changes in a high-fat diet-fed

male rat model of thoracic spinal contusion. *Physiological genomics*, 49(9): 519–529, 2017.

- [98] Gerardo Avila-Martin, Manuel Mata-Roig, Iriana Galan-Arriero, Julian S Taylor, Xavier Busquets, and Pablo V Escriba. Treatment with albumin-hydroxyoleic acid complex restores sensorimotor function in rats with spinal cord injury: Efficacy and gene expression regulation. *PloS one*, 12(12):e0189151, 2017.
- [99] Raquel Cuevas-Diaz Duran, Han Yan, Yiyan Zheng, Xingfan Huang, Raymond Grill, Dong H Kim, Qilin Cao, and Jia Qian Wu. The systematic analysis of coding and long non-coding rnas in the sub-chronic and chronic stages of spinal cord injury. *Scientific reports*, 7:41008, 2017.
- [100] Hae Young Shin, Hyosil Kim, Min Jung Kwon, Dong Hoon Hwang, KiYoung Lee, and Byung Gon Kim. Molecular and cellular changes in the lumbar spinal cord following thoracic injury: regulation by treadmill locomotor training. *PLoS One*, 9(2):e88215, 2014.
- [101] Mahmood Chamankhah, Eftekhari Eftekharpour, Soheila Karimi-Abdolrezaee, Paul C Boutros, Serban San-Marina, and Michael G Fehlings. Genome-wide gene expression profiling of stress response in a spinal cord clip compression injury model. *BMC genomics*, 14(1):583, 2013.
- [102] Abel Torres-Espín, Joaquim Hernández, and Xavier Navarro. Gene expression changes in the injured spinal cord following transplantation of mesenchymal stem cells or olfactory ensheathing cells. *PLoS One*, 8(10): e76141, 2013.
- [103] Hongmei Duan, Weihong Ge, Aifeng Zhang, Yue Xi, Zhihua Chen, Dandan Luo, Yin Cheng, Kevin S Fan, Steve Horvath, Michael V Sofroniew, et al.

Transcriptome analyses reveal molecular mechanisms underlying functional recovery after spinal cord injury. *Proceedings of the National Academy of Sciences*, 112(43):13360–13365, 2015.

- [104] James B Aimone, J Leigh Leasure, Victoria M Perreau, Michaela Thallmair, Christopher Reeve Paralysis Foundation Research Consortium, et al. Spatial and temporal gene expression profiling of the contused rat spinal cord. *Experimental neurology*, 189(2):204–221, 2004.
- [105] Simone Di Giovanni, Susan M Knoblach, Cinzia Brandoli, Sadia A Aden, Eric P Hoffman, and Alan I Faden. Gene profiling in spinal cord injury shows role of cell cycle in neuronal death. *Annals of Neurology: Official Journal of the American Neurological Association and the Child Neurology Society*, 53(4):454–468, 2003.
- [106] Junfang Wu, Bogdan A Stoica, and Alan I Faden. Cell cycle activation and spinal cord injury. *Neurotherapeutics*, 8(2):221–228, 2011.
- [107] Yulong Zhou, Zhouguang Wang, Jiawei Li, Xiaokun Li, and Jian Xiao. Fibroblast growth factors in the management of spinal cord injury. *Journal of cellular and molecular medicine*, 22(1):25–37, 2018.
- [108] Haruo Kanno, Damien D Pearce, Hiroshi Ozawa, Eiji Itoi, and Mary Bartlett Bunge. Schwann cell transplantation for spinal cord injury repair: its significant therapeutic potential and prospectus. *Reviews in the Neurosciences*, 26(2):121–128, 2015.
- [109] Alexander G Rabchevsky, Felicia M Michael, and Samir P Patel. Mitochondria focused neurotherapeutics for spinal cord injury. *Experimental neurology*, 330:113332, 2020.

- [110] Kristyna Karova, John V Wainwright, Lucia Machova-Urdzikova, Rishikaysh V Pisal, Meic Schmidt, Pavla Jendelova, and Meena Jhanwar-Uniyal. Transplantation of neural precursors generated from spinal progenitor cells reduces inflammation in spinal cord injury via $\text{nf-}\kappa\text{b}$ pathway inhibition. *Journal of neuroinflammation*, 16:1–11, 2019.
- [111] Tianyi Wang, Wenqi Yuan, Yong Liu, Yanjun Zhang, Zhijie Wang, Xianhu Zhou, Guangzhi Ning, Liang Zhang, Liwei Yao, Shiqing Feng, et al. The role of the jak-stat pathway in neural stem cells, neural progenitor cells and reactive astrocytes after spinal cord injury. *Biomedical Reports*, 3(2): 141–146, 2015.
- [112] Irawan Satriotomo, Kellie K Bowen, and Raghu Vemuganti. Jak2 and stat3 activation contributes to neuronal damage following transient focal cerebral ischemia. *Journal of neurochemistry*, 98(5):1353–1368, 2006.
- [113] Maria Teresa Bengoechea-Alonso and Johan Ericsson. The phosphorylation-dependent regulation of nuclear srebp1 during mitosis links lipid metabolism and cell growth. *Cell cycle*, 15(20):2753–2765, 2016.
- [114] Renuka Verma, Jasleen Kaur Viridi, Nirmal Singh, and Amteshwar Singh Jaggi. Animals models of spinal cord contusion injury. *The Korean journal of pain*, 32(1):12–21, 2019.
- [115] Batbayar Khuyagbaatar, Kyungsoo Kim, and Yoon Hyuk Kim. Conversion equation between the drop height in the new york university impactor and the impact force in the infinite horizon impactor in the contusion spinal cord injury model. *Journal of neurotrauma*, 32(24):1987–1993, 2015.
- [116] Ricardo O Ramirez Flores, Jan D Lanzer, Christian H Holland, Florian Leuschner, Patrick Most, Jobst-Hendrik Schultz, Rebecca T Levinson, and Julio Saez-Rodriguez. Consensus transcriptional landscape of human

end-stage heart failure. *Journal of the American Heart Association*, 10(7): e019667, 2021.

- [117] Mark D Wilkinson, Michel Dumontier, IJsbrand Jan Aalbersberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg, Jan-Willem Boiten, Luiz Bonino da Silva Santos, Philip E Bourne, et al. The fair guiding principles for scientific data management and stewardship. *Scientific data*, 3, 2016.
- [118] Jure Tica, Elizabeth J Bradbury, and Athanasios Didangelos. Combined transcriptomics, proteomics and bioinformatics identify drug targets in spinal cord injury. *International journal of molecular sciences*, 19(5):1461, 2018.
- [119] Agnes E Haggerty, Megan M Marlow, and Martin Oudega. Extracellular matrix components as therapeutics for spinal cord injury. *Neuroscience letters*, 652:50–55, 2017.
- [120] Brian K Kwon, Steve Casha, R John Hurlbert, and V Wee Yong. Inflammatory and structural biomarkers in acute traumatic spinal cord injury. *Clinical chemistry and laboratory medicine*, 49(3):425–433, 2011.
- [121] Ahmed A Albayar, Abigail Roche, Przemyslaw Swiatkowski, Sarah Antar, Nouran Ouda, Eman Emara, Douglas H Smith, Ali K Ozturk, and Basem I Awad. Biomarkers in spinal cord injury: Prognostic insights and future potentials. *Frontiers in neurology*, 10, 2019.
- [122] Romulo de Castro Jr, Cynthia L Burns, David J McAdoo, and Anne M Romanic. Metalloproteinase increases in the injured rat spinal cord. *Neuroreport*, 11(16):3551–3554, 2000.
- [123] Linda J Noble, Frances Donovan, Takuji Igarashi, Staci Goussev, and Zena Werb. Matrix metalloproteinases limit functional recovery after spinal cord

- injury by modulation of early vascular events. *Journal of Neuroscience*, 22 (17):7526–7535, 2002.
- [124] S Nagano, SH Kim, S Tokunaga, K Arai, M Fujiki, and K Misumi. Matrix metalloprotease-9 activity in the cerebrospinal fluid and spinal injury severity in dogs with intervertebral disc herniation. *Research in veterinary science*, 91(3):482–485, 2011.
- [125] Andrew D Gaudet and Phillip G Popovich. Extracellular matrix regulation of inflammation in the healthy and injured spinal cord. *Experimental neurology*, 258:24–34, 2014.
- [126] Shu-Xin Zhang, Mark Underwood, Alex Landfield, Feng-fa Huang, Sam Gibson, and James W Geddes. Cytoskeletal disruption following contusion injury to the rat spinal cord. *Journal of Neuropathology & Experimental Neurology*, 59(4):287–296, 2000.
- [127] Shruti V Kabadi, Bogdan A Stoica, Kimberly R Byrnes, Marie Hanscom, David J Loane, and Alan I Faden. Selective cdk inhibitor limits neuroinflammation and progressive neurodegeneration after brain trauma. *Journal of Cerebral Blood Flow & Metabolism*, 32(1):137–149, 2012.
- [128] Jacob W Skovira, Junfang Wu, Jessica J Matyas, Alok Kumar, Marie Hanscom, Shruti V Kabadi, Raymond Fang, and Alan I Faden. Cell cycle inhibition reduces inflammatory responses, neuronal loss, and cognitive deficits induced by hypobaric exposure following traumatic brain injury. *Journal of neuroinflammation*, 13:1–16, 2016.
- [129] Nicolas S Merle, Remi Noe, Lise Halbwachs-Mecarelli, Veronique Fremeaux-Bacchi, and Lubka T Roumenina. Complement system part ii: role in immunity. *Frontiers in immunology*, 6:257, 2015.

- [130] R Lubbers, MF Van Essen, C Van Kooten, and LA Trouw. Production of complement components by cells of the immune system. *Clinical & Experimental Immunology*, 188(2):183–194, 2017.
- [131] Lindsay M Milich, Christine B Ryan, and Jae K Lee. The origin, fate, and contribution of macrophages to spinal cord injury pathology. *Acta neuropathologica*, 137:785–797, 2019.
- [132] Timothy J Kopper and John C Gensel. Myelin as an inflammatory mediator: Myelin interactions with complement, macrophages, and microglia in spinal cord injury. *Journal of neuroscience research*, 96(6):969–977, 2018.
- [133] Seyoung Mun, Kyudong Han, and Jung Keun Hyun. The time sequence of gene expression changes after spinal cord injury. *Cells*, 11(14):2236, 2022.
- [134] Rosario Michael Piro, Ugo Ala, Ivan Molineris, Elena Grassi, Chiara Bracco, Gian Paolo Perego, Paolo Provero, and Ferdinando Di Cunto. An atlas of tissue-specific conserved coexpression for functional annotation and disease gene prediction. *European Journal of Human Genetics*, 19(11):1173–1180, 2011.
- [135] Emma Pierson, GTEx Consortium, Daphne Koller, Alexis Battle, and Sara Mostafavi. Sharing and specificity of co-expression networks across 35 human tissues. *PLoS computational biology*, 11(5):e1004220, 2015.
- [136] Andrea De Biase, Susan M Knoblach, Simone Di Giovanni, Chenguang Fan, Annamaria Molon, Eric P Hoffman, and Alan I Faden. Gene expression profiling of experimental traumatic spinal cord injury as a function of distance from impact site and injury severity. *Physiological genomics*, 22(3):368–381, 2005.

- [137] Daniel J Dennis, Sisu Han, and Carol Schuurmans. bhlh transcription factors in neural development, disease, and reprogramming. *Brain research*, 1705: 48–65, 2019.
- [138] Sarah E Ross, Alejandra E McCord, Cynthia Jung, Denize Atan, Stephanie I Mok, Martin Hemberg, Tae-Kyung Kim, John Salogiannis, Linda Hu, Sonia Cohen, et al. Bhlhb5 and prdm8 form a repressor complex involved in neuronal circuit assembly. *Neuron*, 73(2):292–303, 2012.
- [139] Zhenjun Lou, Veronica M Maher, and J Justin McCormick. Identification of the promoter of human transcription factor sp3 and evidence of the role of factors sp1 and sp3 in the expression of sp3 protein. *Gene*, 351:51–59, 2005.
- [140] Marjolein Meinders, Divine I Kulu, Harmen JG van de Werken, Mark Hoogenboezem, Hans Janssen, Rutger WW Brouwer, Wilfred FJ van Ijcken, Erik-Jan Rijkers, Jeroen AA Demmers, Imme Krüger, et al. Sp1/sp3 transcription factors regulate hallmarks of megakaryocyte maturation and platelet formation and function. *Blood, The Journal of the American Society of Hematology*, 125(12):1957–1967, 2015.
- [141] Ding Wan, Jin Feng, Peng Wang, Zhenxing Yang, and Tao Sun. Hypoxia-and inflammation-related transcription factor sp3 may be involved in platelet activation and inflammation in intracranial hemorrhage. *Frontiers in Neurology*, 13:886329, 2022.
- [142] Yabing Chen, Xinyang Zhao, and Hui Wu. Transcriptional programming in arteriosclerotic disease: a multifaceted function of the runx2 (runt-related transcription factor 2). *Arteriosclerosis, thrombosis, and vascular biology*, 41(1):20–34, 2021.

- [143] Jae-Hwan Jeong, Jung-Sook Jin, Hyun-Nam Kim, Sang-Min Kang, Julie C Liu, Christopher J Lengner, Florian Otto, Stefan Mundlos, Janet L Stein, Andre J Van Wijnen, et al. Expression of runx2 transcription factor in non-skeletal tissues, sperm and brain. *Journal of cellular physiology*, 217 (2):511–517, 2008.
- [144] Risa Sueda, Itaru Imayoshi, Yukiko Harima, and Ryoichiro Kageyama. High hes1 expression and resultant ascl1 suppression regulate quiescent vs. active neural stem cells in the adult mouse brain. *Genes & development*, 33 (9-10):511–523, 2019.
- [145] Xin Wang, Qiaoxia Wang, Huiling Tian, Wei Lv, Liangyu Song, Zhigang Li, Haijiang Yao, and Suhua Shi. Electroacupuncture in promoting neural repair after spinal cord injury: Inhibiting the notch signaling pathway and regulating downstream proteins expression. *The Anatomical Record*, 304 (11):2494–2505, 2021.
- [146] Wendong Ruan, Guangzhi Ning, Shiqing Feng, Shijie Gao, and Yan Hao. Microrna-381/hes1 is a potential therapeutic target for spinal cord injury. *International Journal of Molecular Medicine*, 42(2):1008–1017, 2018.
- [147] Karolina U Kabayiza, Gauhar Masgutova, Audrey Harris, Vincent Rucchin, Benvenuto Jacob, and Frédéric Clotman. The onecut transcription factors regulate differentiation and distribution of dorsal interneurons during spinal cord development. *Frontiers in Molecular Neuroscience*, 10:157, 2017.
- [148] Agnès Roy, Cédric Francius, David L Rousso, Eve Seuntjens, Joke Debruyne, Georg Luxenhofer, Andrea B Huber, Danny Huylebroeck, Bennett G Novitch, and Frédéric Clotman. Onecut transcription factors act upstream of isl1 to regulate spinal motoneuron diversification. *Development*, 139(17):3109–3119, 2012.

- [149] Mathilde Toch, Audrey Harris, Olivier Schakman, Elena Kondratskaya, Jean-Luc Boulland, Nicolas Dauguet, Stéphanie Debrulle, Charlotte Baudouin, Maria Hidalgo-Figueroa, Xiuqian Mu, et al. Onecut-dependent nkx6. 2 transcription factor expression is required for proper formation and activity of spinal locomotor circuits. *Scientific reports*, 10(1):996, 2020.
- [150] Robert Chevreau, Hussein Ghazale, Chantal Ripoll, Chaima Chalfouh, Quentin Delarue, Anne Laure Hemonnot-Girard, Daria Mamaeva, Helene Hirbec, Bernard Rothhut, Shalaka Wahane, et al. Rna profiling of mouse ependymal cells after spinal cord injury identifies the oncostatin pathway as a potential key regulator of spinal cord stem cell fate. *Cells*, 10(12):3332, 2021.
- [151] Shao-Ming Wang, Jung-Yu C Hsu, Chiung-Yuan Ko, Nai-En Chiu, Wai-Ming Kan, Ming-Derg Lai, and Ju-Ming Wang. Astrocytic ccaat/enhancer-binding protein delta contributes to glial scar formation and impairs functional recovery after spinal cord injury. *Molecular Neurobiology*, 53:5912–5927, 2016.
- [152] Gennadij Raivich, Marion Bohatschek, Clive Da Costa, Osuke Iwata, Matthias Galiano, Maria Hristova, Abdolrahman S Nateri, Milan Makwana, Lluís Riera-Sans, David P Wolfer, et al. The ap-1 transcription factor c-jun is required for efficient axonal regeneration. *Neuron*, 43(1):57–67, 2004.
- [153] T Zhang, G Gao, and F Chang. mir-152 promotes spinal cord injury recovery via c-jun amino terminal kinase pathway. *European Review for Medical & Pharmacological Sciences*, 23(1), 2019.
- [154] Dan Qi, Yiqun Geng, Jacob Cardenas, Jinghua Gu, S Stephen Yi, Jason H Huang, Ekokobe Fonkem, and Erxi Wu. Transcriptomic analyses of patient

peripheral blood with hemoglobin depletion reveal glioblastoma biomarkers. *NPJ Genomic Medicine*, 8(1):2, 2023.

- [155] Smriti M Agrawal, Lorraine Lau, and V Wee Yong. Mmps in the central nervous system: where the good guys go bad. In *Seminars in cell & developmental biology*, volume 19, pages 42–51. Elsevier, 2008.
- [156] Pascal Mputu Mputu, Marie Beauséjour, Andréane Richard-Denis, and Jean-Marc Mac-Thiong. Early predictors of neurological outcomes after traumatic spinal cord injury: a systematic review and proposal of a conceptual framework. *American Journal of Physical Medicine & Rehabilitation*, 100(7):700–711, 2021.

7. Material suplementario

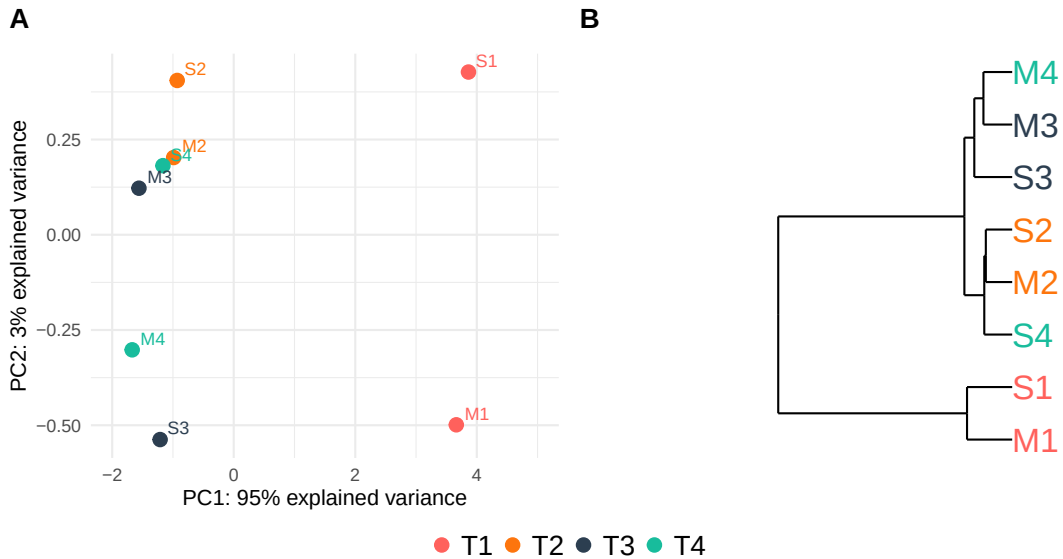


Figura S1. Análisis de *clustering* de las firmas transcripcionales consenso con el top 10 genes específicos de T1. (A) PCA y (B) *clustering* jerárquico usando los genes: *Fos11*, *Il33*, *Vcam1*, *Mx2*, *Mdk*, *Ldlr*, *Hmgcr*, *Alpk3*, *Dusp5*, *Serpina3n*.

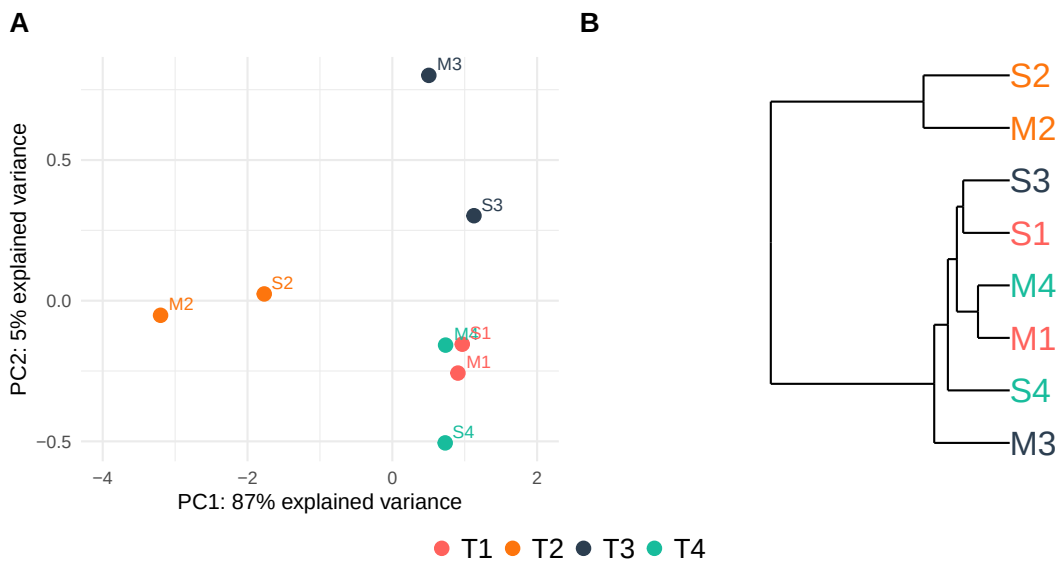


Figura S2. Análisis de *clustering* de las firmas transcripcionales consenso con el top 10 genes específicos de T1. (A) PCA y (B) *clustering* jerárquico usando los genes: *Cfp*, *Slc46a3*, *Scn8a*, *Ly86*, *Casr*, *C1qc*, *Slc2a5*, *Fgd2*, *Aif1*, *Slc7a7*.

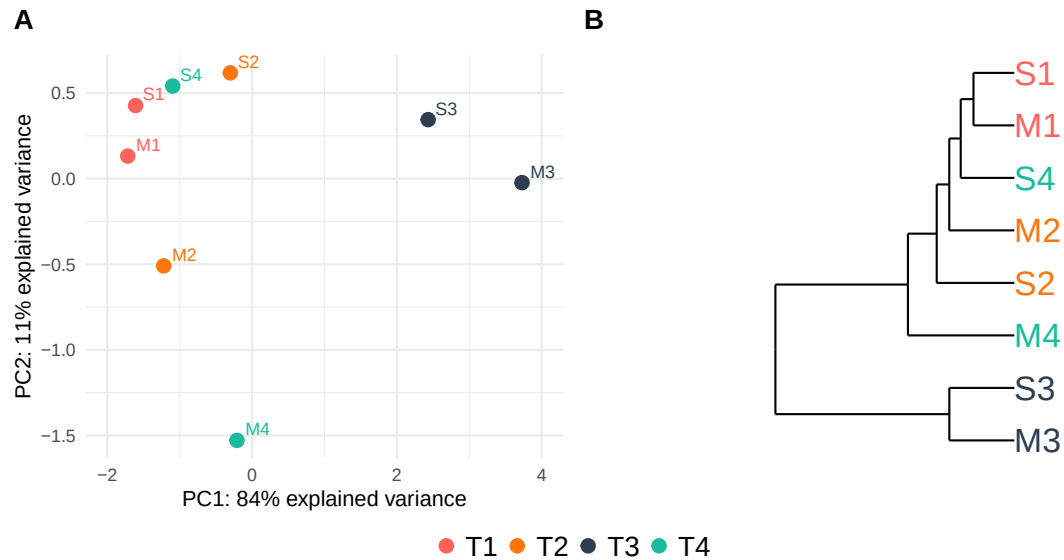


Figura S3. Análisis de *clustering* de las firmas transcripcionales consenso con el top 10 genes específicos de T1. (A) PCA y (B) *clustering* jerárquico usando los genes: *Slc28a2*, *Iqgap1*, *Cd14*, *Nt5dc2*, *Xdh*, *Msn*, *Lyve1*, *Tmem37*, *Rdh10*, *Enpp3*.

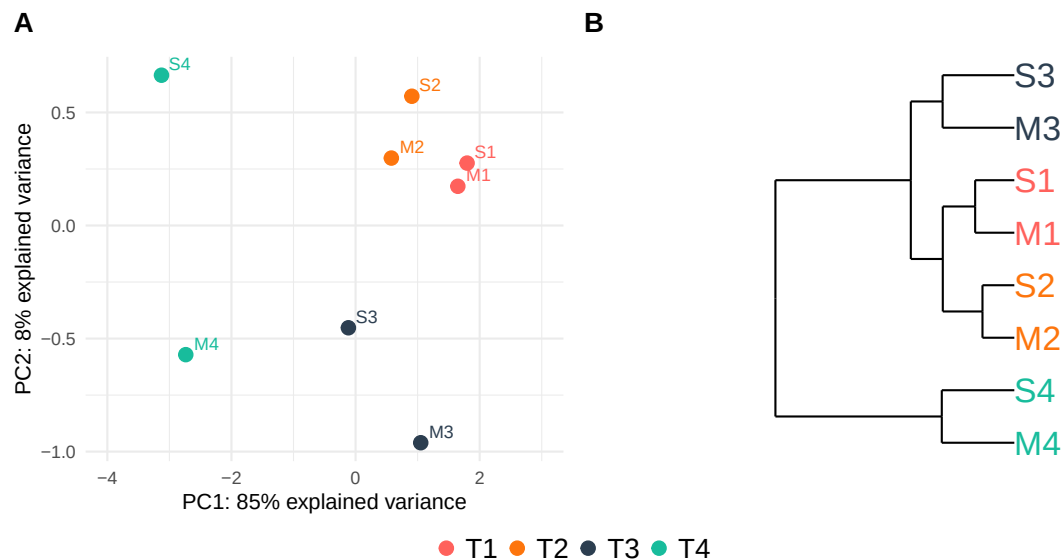


Figura S4. Análisis de *clustering* de las firmas transcripcionales consenso con el top 10 genes específicos de T1. (A) PCA y (B) *clustering* jerárquico usando los genes: *Cp*, *Rab27a*, *Rab3b*, *Cybrd1*, *Thrsp*, *Plcb1*, *Pmp22*, *Acot2*, *Cyp27a1*, *RT1-Db1*.



Figura S5. Dotplot de una selección de funciones relacionadas con neurotransmisores En rojo (*up*-regulación) y azul (*down*-regulación) se indican funciones con un valor del logaritmo de odds ratio (LOR) mayor o menor que cero, respectivamente.

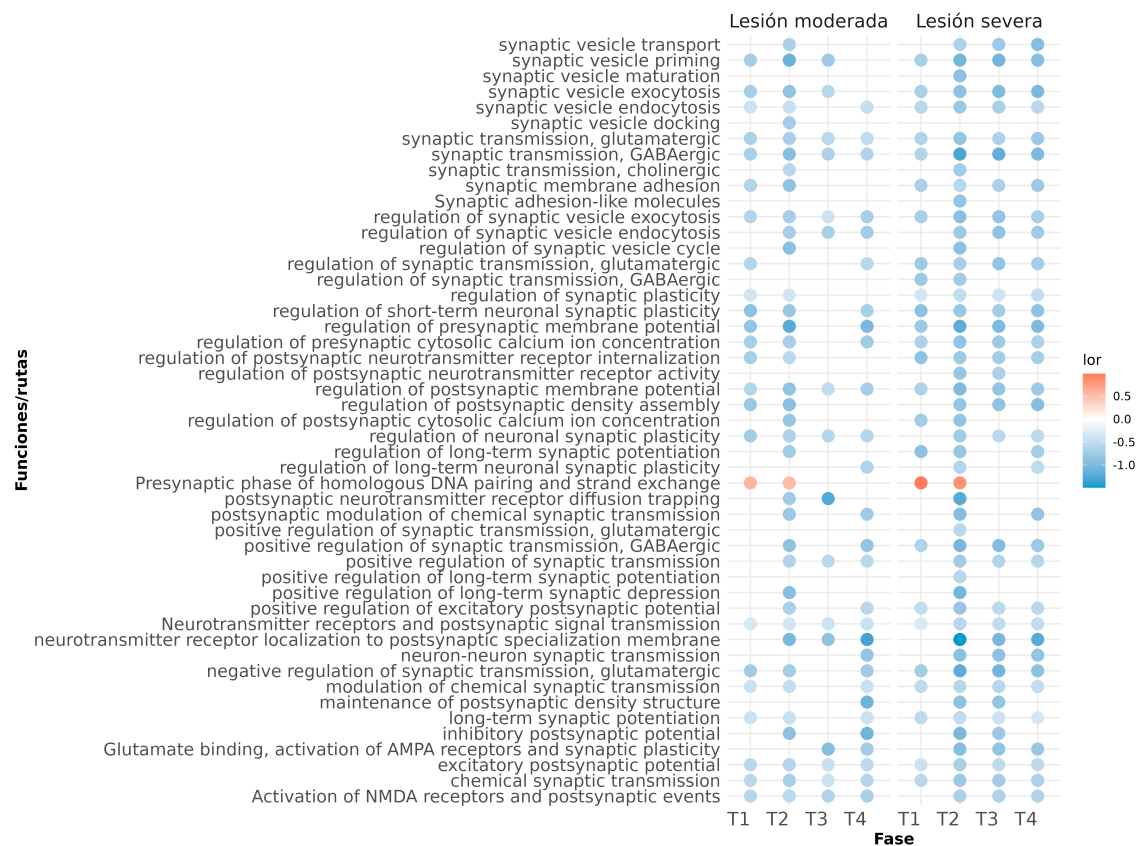


Figura S6. Dotplot de una selección de funciones relacionadas con la sinapsis
En rojo (*up*-regulación) y azul (*down*-regulación) se indican funciones con un valor del logaritmo de odds ratio (LOR) mayor o menor que cero, respectivamente.

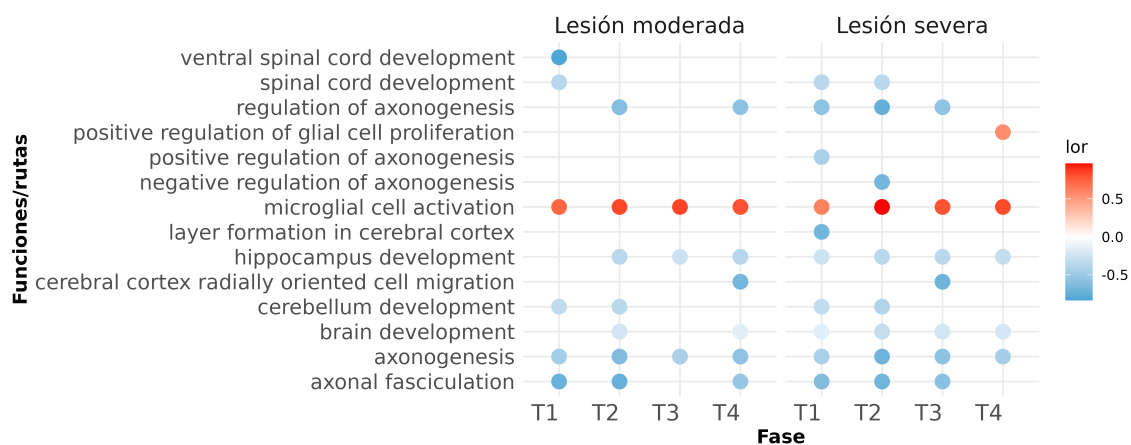


Figura S7. Dotplot de una selección de funciones relacionadas con el desarrollo del sistema nervioso En rojo (*up*-regulación) y azul (*down*-regulación) se indican funciones con un valor del logaritmo de odds ratio (LOR) mayor o menor que cero, respectivamente.

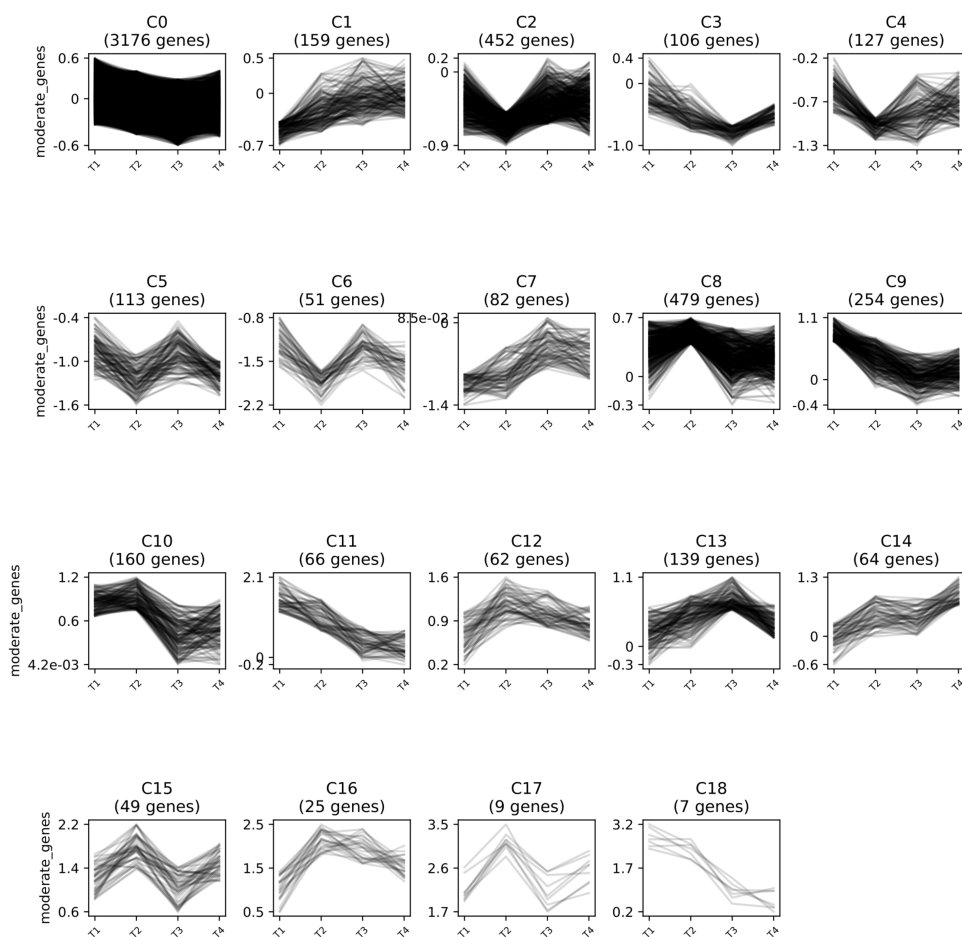


Figura S8. Clusters de genes co-expresados para lesión moderada.

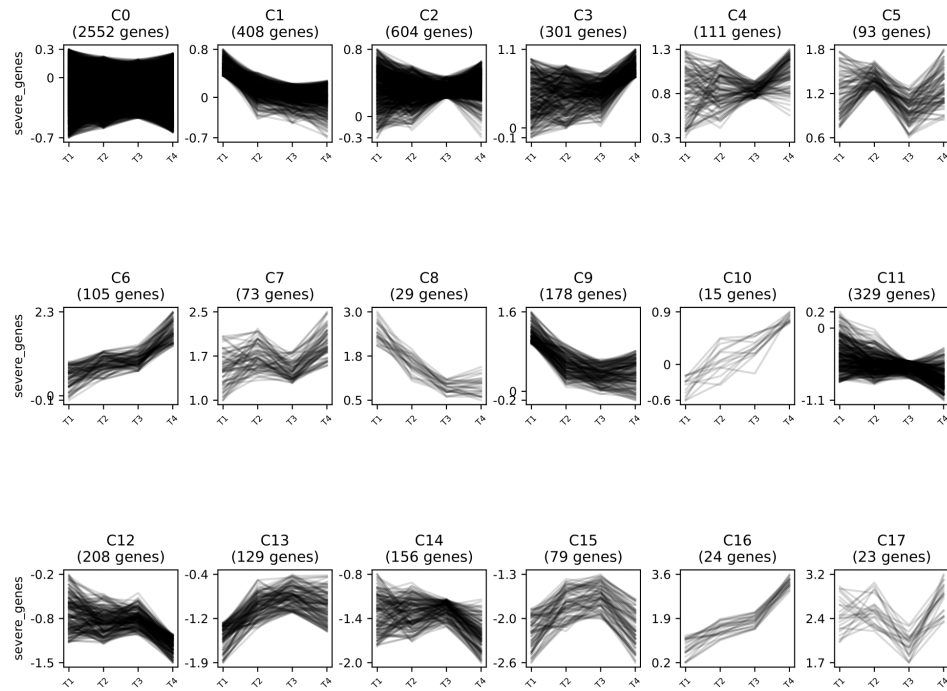


Figura S9. Clusters de genes co-expresados para lesión severa.

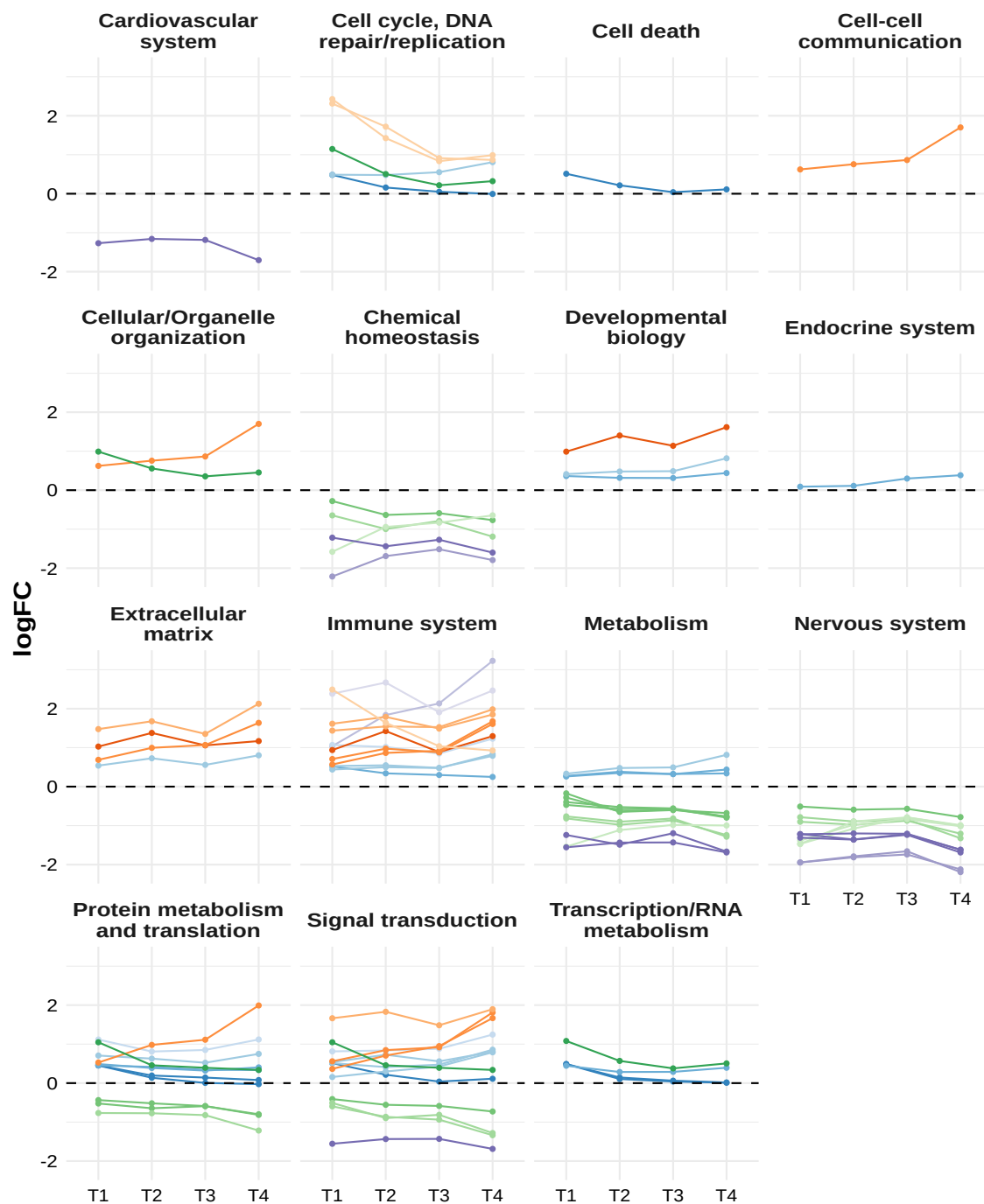


Figura S10. Patrones de expresión de las redes locales de lesión severa clasificadas por categoría funcional. Se representa el patrón de expresión característico (mediana del logFC de los genes en un punto temporal) para cada red local obtenida en lesiones severas, clasificadas según la categoría funcional asignada. El color indica el *cluster* del que proviene cada red local.