



Universidad de Jaén

Escuela de Doctorado

# APLICACIÓN DE LA TEORÍA DE JUEGOS COOPERATIVOS AL ANÁLISIS DE LA EXPRESIÓN GÉNICA EN BIOMEDICINA DE SISTEMAS

Autor: José Alberto Castro Martínez

Director de la tesis: Francisco José Esteban Ruiz  
Departamento: Biología Experimental

Fecha: 09/09/2024

ISBN:  
Licencia CC

RUJJA

Grupo de Biología de Sistemas  
Departamento de Biología Experimental  
Universidad de Jaén, Jaén, España

# Aplicación de la Teoría de Juegos Cooperativos al Análisis de la Expresión Génica en Biomedicina de Sistemas

José Alberto Castro Martínez



**UNIVERSIDAD DE JAÉN**

Jaén, 2024

# Aplicación de la Teoría de Juegos Cooperativos al Análisis de la Expresión Génica en Biomedicina de Sistemas

Tesis Doctoral

Por:

José Alberto Castro Martínez

Tutor y director:

Francisco José Esteban Ruiz,

Grupo de Biología de Sistemas, Departamento de Biología Experimental, Universidad de Jaén

Codirectora:

Leticia Díaz Beltrán,

Unidad de Investigación Clínica, Servicio de Oncología Médica, Hospital Universitario de Jaén

Codirectora:

Eva Vargas Liébanas,

Grupo de Biología de Sistemas, Departamento de Biología Experimental, Universidad de Jaén

## Resumen

Uno de los principales desafíos en la aplicación de técnicas ómicas, como la transcriptómica, reside en las limitaciones inherentes a los métodos estadísticos tradicionales, las cuales se manifiestan a través de parámetros como el intervalo de confianza y los errores estadísticos de tipo I y tipo II. Para aprovechar plenamente el enorme potencial en la detección y clasificación de datos de expresión génica, en particular en el estudio de trastornos complejos biomédicos, proponemos el uso de una técnica basada en la Teoría de Juegos Cooperativos: el ‘Análisis Comparativo del Valor de Shapley’ (CASH, por sus siglas en inglés). Esta técnica, desarrollada originalmente como una respuesta a la complejidad de diversos sistemas naturales, permite la identificación de genes candidatos como posibles biomarcadores sin las restricciones impuestas por métodos más tradicionales. Los fenotipos analizados en este estudio abarcan tanto condiciones neurológicas (Trastorno del Espectro Autista, Esquizofrenia, Trastorno Bipolar y Depresión Mayor) como trastornos relacionados con el endometrio (Endometriosis, Leiomiomatosis Uterina y Cáncer de Endometrio). Como métodos de control, se emplearon el test de Welch y el test empírico de Bayes, el primero una variante de la prueba t de Student que permite comparar grupos con varianzas desiguales y, el segundo, un método que estima distribuciones de probabilidad a partir de datos observados, permitiendo realizar inferencias estadísticas bajo un marco bayesiano, lo cual es particularmente útil cuando se cuenta con información previa o cuando los tamaños de las muestras son pequeños. Los resultados obtenidos señalan que CASH identifica una mayor cantidad de genes candidatos, destacándose como una herramienta exploratoria altamente eficiente cuando otras técnicas ofrecen resultados limitados o se ven restringidas por correcciones estadísticamente conservadoras. Esta característica convierte a CASH en una metodología idónea para el estudio y diagnóstico de trastornos complejos como los abordados en esta investigación. Su simplicidad y eficacia la posicionan como una herramienta clave en el estudio de las bases moleculares de las enfermedades.

## Publicaciones incluidas en esta tesis

- 1) Vargas E.; Castro-Martínez, J.A.; Díaz-Beltrán, L.; Esteban, F.J. La Biología de Sistemas y su aplicación al estudio de enfermedades complejas. *Seminario Médico* **2024**, *64(1)*, 61-71.
- 2) Castro-Martínez, J.A.; Vargas, E.; Díaz-Beltrán, L.; Esteban, F.J. Comparative Analysis of Shapley Values Enhances Transcriptomics Insights across Some Common Uterine Pathologies. *Genes* **2024**, *15*, 723. <https://doi.org/10.3390/genes15060723>
- 3) Castro-Martínez, J.A.; Vargas, E.; Díaz-Beltrán, L.; Esteban, F.J. Enhancing transcriptomics insights into neurological disorders through Comparative Analysis of Shapley Values. Manuscrito preparado para su envío a EMBnet.journal.

# Índice

Lista de abreviaturas .....	7
1. Introducción.....	8
1.1 Sistemas complejos.....	8
1.2 Trastornos neurológicos.....	9
1.2.1 Biomarcadores de trastornos neurológicos .....	12
1.3 Trastornos del endometrio y enfermedades comórbidas .....	16
1.3.1 Biomarcadores de trastornos uterinos .....	17
1.4 Tecnológicas ómicas.....	19
1.5 Teoría de juegos cooperativos .....	21
2. Objetivos.....	23
3. Material y Métodos.....	24
3.1 Datos .....	24
3.2 Métodos .....	26
3.2.1 Preprocesamiento y control de calidad.....	28
3.2.2 Determinación de la expresión diferencial mediante métodos tradicionales .	28
3.2.3 Determinación de la expresión diferencial mediante análisis comparativo del valor de Shapley .....	29
3.2.4 Análisis de enriquecimiento funcional.....	31
4. Resultados y discusión .....	32
<i>Capítulo I: aplicación de la Biología de Sistemas al estudio de enfermedades complejas</i> .....	32
<i>Capítulo II: análisis comparativo del valor de Shapley para el estudio de trastornos</i> <i>endometriales</i> .....	33
<i>Capítulo III: aplicación de la Biología de Sistemas complejos para el estudio de</i> <i>alteraciones neurológicas</i> .....	35

5. Conclusiones.....	39
6. Agradecimientos.....	40
7. Bibliografía.....	42

## Lista de abreviaturas

ASD	Autism Spectrum Disorder
CASh	Comparative Analysis of Shapley value
DAVID	Database for Annotation, Visualization and Integrated Discovery
FC	Fold Change
FDR	False Discovery Rate
GEO	Gene Expression Omnibus
GO	Gene Ontology
GWAS	Genome-Wide Association Studies
NCBI	National Center for Biotechnology Information
NUSE	Normalized Unscaled Standard Error
PCA	Principal Component Analysis
RLE	Relative Log Expression
SNP	Single Nucleotide Polymorphism
TEA	Trastorno del Espectro Autista

# 1. Introducción

## 1.1 Sistemas complejos

La teoría de los sistemas complejos está fundamentada en identificar y caracterizar elementos de diseño comunes observables en diversos sistemas complejos, ya sean naturales, tecnológicos o sociales. La biología de sistemas, que consiste en un enfoque holístico para estudiar elementos biológicos como moléculas o células, ha experimentado un desarrollo acelerado en las últimas décadas (Ma'ayan, 2017). Esta disciplina investiga las interacciones integradas y coordinadas de una o varias redes de genes, de sus variantes, que pueden ser naturales o mutadas, sus proteínas e isoformas, y las moléculas orgánicas e inorgánicas con las que interactúan. Estas interacciones dan lugar a las reacciones bioquímicas (enzimas, sustratos, productos, etc.) que reflejan la función del sistema en cuestión (Clarke, 2017). Esta rama de la Biología necesita la aplicación de la Bioinformática, una disciplina que surgió en la década de 1960 como una herramienta para aplicar métodos computacionales en el análisis de secuencias de proteínas. Posteriormente, los avances simultáneos en métodos de biología molecular y en informática permitieron manipular y secuenciar el ADN con mayor precisión, lo que resultó en el desarrollo de software más potente para tareas de Bioinformática (Gauthier *et al.*, 2019). Gracias a estos avances, se pudieron desarrollar técnicas complejas como las que permitieron la secuenciación del primer genoma humano en 2002, proyecto que duró 13 años y costó aproximadamente 3 millones de dólares. Actualmente, gracias al avance tecnológico, se puede resecuenciar un genoma humano por menos de 100 dólares desde el año 2020 (Tolosa, 2020); no obstante, la función de muchos genes sigue siendo total o parcialmente desconocida (Pavlopoulos *et al.*, 2015), lo que en el contexto del genoma humano constituye uno de los retos más importantes en el camino hacia el desarrollo de la medicina personalizada.

En la actualidad, la biología de sistemas integra la información obtenida a través de la aplicación de tecnologías ómicas. Éstas son herramientas muy valiosas para el análisis integral de cualquier tipo de molécula biológica, que aportan datos que requieren de un tratamiento bioinformático (Karahalil, 2016). Dentro del campo de la Biomedicina, hay investigaciones que las utilizan para la comprensión y tratamiento de alguna situación en concreto, que a menudo no ha podido ser analizada exitosamente con medios más tradicionales. En esta Tesis Doctoral se ha investigado un conjunto de trastornos

complejos, con el propósito de incorporar una nueva metodología que permita una mayor eficacia en la detección de genes candidatos, a partir de datos de expresión génica procedentes de bases de datos públicas. Un trastorno o enfermedad compleja podría definirse como un fenotipo disfuncional cuyo origen, ya sea genético, ambiental o mixto, es desconocido, y la diversidad y heterogeneidad de los síntomas impiden un diagnóstico y tratamiento eficaces. Por la complejidad subyacente, la investigación de estos trastornos sigue en curso, pese a la presencia de técnicas ómicas muy precisas y potentes y, en buena parte de los casos, se desconocen por completo los mecanismos que originan su aparición y desarrollo. Uno de los problemas más acuciantes que poseen estas investigaciones es la carencia de técnicas estadísticas que eviten la asunción de errores estadísticos, ya sea en forma de genes que se obtienen como significativos cuando no lo son (error de tipo I) o en forma de genes que, siendo importantes, no se detectan como tales (error de tipo II). Por ejemplo, si se estima un intervalo de confianza del 95%, el error asociado, al aparecer el problema de las comparaciones múltiples muy incrementado por la gran cantidad de variables que hay en la investigación (muestras, genes, etc.), constituye una dificultad considerable para detectar genes candidatos ya que el error de tipo I afecta al resultado, lo que obliga a los experimentadores a utilizar técnicas de corrección estadística que disminuyen el error de tipo I, pero que incrementan el error de tipo II, lo que no resuelve el problema de modo satisfactorio. Además, cada equipo de investigación hace uso de diferentes técnicas, tales como el análisis de componentes principales (PCA por las siglas en inglés de *Principal Component Analysis*), *Fold Change* (FC), o *False Discovery Rate* (FDR) que, al no aplicarse de un modo estandarizado, hace inviable una comparación directa y fiable de resultados de diferentes laboratorios sobre el mismo trastorno complejo.

Dentro de esta investigación, trataremos como parte de los fenotipos analizados el Trastorno del Espectro Autista, la Esquizofrenia, el Trastorno Bipolar y la Depresión Mayor como trastornos neurológicos, y la Endometriosis, la Leiomiomatosis Uterina y el Cáncer de Endometrio como patologías del endometrio.

## **1.2 Trastornos neurológicos**

El Trastorno del Espectro Autista (TEA, o ASD por sus siglas en inglés, Autism Spectrum Disorder) (Hirota y King, 2023; Lord *et al.*, 2018) es un grupo heterogéneo de

trastornos del neurodesarrollo que abarca desde formas severas de autismo, que pueden incluir discapacidad intelectual, hasta manifestaciones más leves clasificadas como Síndrome de Asperger. Las personas con TEA suelen mostrar dificultades en la interacción social, incluyendo un patrón de intereses restringidos y repetitivos, y presentan déficits en la comunicación y en la atención compartida (Sharma *et al.*, 2018). Según la Organización Mundial de la Salud, uno de cada cien niños sufre este trastorno (Organización Mundial de la Salud, 2023) y su prevalencia está aumentando en los últimos años (Jiang *et al.*, 2022; Kraft *et al.*, 2023). Se ha asociado este trastorno al sistema digestivo (Martin *et al.*, 2018; Srikantha y Mohajeri, 2019), al sistema inmune (Heidari *et al.*, 2021; Meltzer y Van de Water, 2017; Ormstad *et al.*, 2018; Robinson-Agramonte *et al.*, 2022), al sistema circulatorio (Kealy *et al.*, 2020; Mouridsen *et al.*, 2016) y al sistema nervioso (Doroszkiewicz *et al.*, 2021; Matta *et al.*, 2019; Sharon *et al.*, 2016). También se han propuesto factores de riesgo variados que van desde la implicación de la microbiota (Kang *et al.*, 2017) hasta causas genéticas (Qiu *et al.*, 2022), dado que hay una mayor probabilidad de sufrir esta dolencia si un familiar ya está diagnosticado (Medicina Buenos Aires, 2023). Pese a todas estas hipótesis, no ha habido ninguna explicación que contenga todos los síntomas y justifique su origen de manera convincente hasta la fecha (Silverman *et al.*, 2010; Tseng *et al.*, 2022). Debido a que afecta de modo severo a los pacientes y su entorno más cercano, y limita su autonomía, se han realizado experimentos masivos con el propósito de comprender y tratar este trastorno (Dissanayake *et al.*, 2019; Ruzzo *et al.*, 2019), con resultados inconcluyentes.

La esquizofrenia (Andreasen *et al.*, 1994) es una enfermedad neurológica que se manifiesta con síntomas positivos (alucinaciones, falta de habilidades sociales y distorsiones cognitivas) y negativos (apatía general, tanto social como laboral) (Khan *et al.*, 2013). Se ha calculado que hay 24 millones de personas afectadas (1 de cada 300 personas en otra estimación), con una prevalencia en adultos del 0,45% (Saha *et al.*, 2005; McGrath *et al.*, 2008; Organización Mundial de la Salud, 2023). La esquizofrenia disminuye la autonomía de los afectados, causando además un gran sufrimiento en su entorno (Gulayín, 2022; Ribé *et al.*, 2018), lo que se asocia con un aumento del índice de suicidio (Balhara *et al.*, 2012; Carlborg *et al.*, 2010; Sher y Kahn, 2019), además de una mayor vulnerabilidad ante enfermedades cardiovasculares (Barnett *et al.*, 2007; Hagi *et al.*, 2021), metabólicas (Trubetskoy *et al.*, 2022) o infecciosas (Fuglewicz *et al.*, 2017), aumentando el riesgo de una muerte prematura. Su origen sigue siendo una incógnita

(Sekar *et al.*, 2016). Hay una convención acerca de la existencia de factores genéticos que parecen predisponer su aparición (Trubetskoy *et al.*, 2022), pero que no determinan el padecimiento de ésta (Vilain *et al.*, 2013). Asimismo, se habla de otros desencadenantes como el entorno social, el abuso de drogas como el alcohol (Häfner y an der Heiden, 1997; Janoutová *et al.*, 2016; Stilo & Murray, 2019) y la poda neuronal típica de la adolescencia (etapa vital donde se manifiesta más por primera vez) (Germann *et al.*, 2021). Las vainas de mielina que recubren los axones (Valdés-Tovar *et al.*, 2022) y la arquitectura del sistema nervioso central (Bobilev *et al.*, 2020; Cheng *et al.*, 2022; Heckers y Konradi, 2002) son elementos asociados a la esquizofrenia. También se ha utilizado la transcriptómica para tratar de dar respuesta al estudio sobre el origen de esta situación clínica (Gandal *et al.*, 2018). No hay ninguna hipótesis aceptada por la comunidad científica de modo unánime que permita la comprensión integral de la enfermedad (Pino *et al.*, 2014). Pese al esfuerzo de investigación realizado, lo único que se acepta mayoritariamente es que se trata de un trastorno multifactorial (Kahn y Sommer, 2015; Morera-Fumero y Abreu-Gonzalez, 2013). El deterioro de la calidad de vida de los pacientes genera un interés médico y social que concierne a la industria farmacéutica, que busca medicamentos que alivien este sufrimiento minimizando los efectos secundarios asociados a los tratamientos actuales (Krause *et al.*, 2018), los cuales son un obstáculo para los pacientes tratados (Perlick *et al.*, 2010), por lo que se requiere una investigación eficaz para paliar estos problemas de ámbito social y económico (Evans-Lacko *et al.*, 2014).

El trastorno bipolar (Smith *et al.*, 2012; Tondo *et al.*, 2017) es un trastorno neurológico caracterizado por la alternancia de episodios maníacos (euforia, alegría excesiva, entusiasmo desmedido, etc.) con episodios depresivos (anhedonia, tristeza, falta de interés por la vida, etc.) (Fagiolini *et al.*, 2015). En 2019 se estimó que 40 millones de personas se encuentran afectados por el trastorno bipolar (Clemente *et al.*, 2015), lo que causa un sufrimiento elevado y aumenta la tasa de suicidio para estos pacientes (Beyer y Weisler, 2016; Miller y Black, 2020). Pese al desconocimiento de las causas, algunos medicamentos como el litio pueden paliar sus efectos (Katz *et al.*, 2022; Malhi *et al.*, 2017). Se han estudiado las causas genéticas que pudieran estar involucradas (Gandal *et al.*, 2018). Algunos factores ambientales como el alcoholismo y otros tipos de drogadicción han sido relacionados con la causa de esta enfermedad (Aldinger y Schulze, 2017), pero el conocimiento de ésta es aún muy limitado.

La depresión mayor (Filatova *et al.*, 2021) es una enfermedad neurológica con origen desconocido (Gómez *et al.*, 2020), y se diferencia de la depresión común por la severidad de sus síntomas (Kennedy, 2008), caracterizados por anhedonia, tristeza y falta de voluntad para vivir, lo que motiva su investigación (Bauer *et al.*, 2019). La prevalencia mundial es de 350 millones de personas (Gutiérrez-Rojas *et al.*, 2020; Smith *et al.*, 2019), con una alta variabilidad entre regiones (3% en Japón contra 16,3% en EE UU) (Gutiérrez-Rojas *et al.*, 2020), y constituye una gran carga social y económica para el entorno (Greenberg *et al.*, 2021; Keshavarz *et al.*, 2022). Cada año se registran 850.000 suicidios anuales por esta causa (Li *et al.*, 2022; Serra *et al.*, 2022). Se considera que puede haber causas genéticas, lo que motiva estudios transcriptómicos o epigenéticos (Kendall *et al.*, 2021), pero también se baraja con un origen fisiológico, hormonal o factores ambientales tales como el estrés, factores psicológicos y sociales (Suda y Matsuda, 2022). Debido a su origen incierto, se clasifica como trastorno complejo (Harder *et al.*, 2022).

### **1.2.1 Biomarcadores de trastornos neurológicos**

La búsqueda de biomarcadores indicadores de la presencia de trastornos neurológicos representa un desafío crucial en el ámbito de la neurociencia moderna. Debido a su naturaleza compleja y heterogénea, la comprensión de las causas que conducen a estos trastornos requiere profundizar a nivel celular y molecular para el desarrollo de tratamientos efectivos. Sin embargo, el acceso al tejido cerebral constituye una limitación para la realización de estudios precisos y detallados. En este contexto, la identificación de biomarcadores no invasivos se torna particularmente útil, pues podría ayudar a superar estas limitaciones y facilitar el diagnóstico de estos trastornos y los consiguientes tratamientos. Hasta la fecha, no se conocen biomarcadores no invasivos que sean capaces de detectar la presencia de los diferentes trastornos neurológicos objeto de estudio de la presente Tesis Doctoral. Sin embargo, la aplicación de tecnologías ómicas ha permitido avanzar en el conocimiento en este sentido.

En el contexto del Trastorno del Espectro Autista, (Ratajczak, 2011) se ha puesto de manifiesto la necesidad de un enfoque holístico para entender el trastorno a través de diversos marcadores indicativos del mismo (sistema gastrointestinal, inmunológico y neurológico, sustancias tóxicas y biomarcadores ubicuos). En este sentido, se ha

propuesto el uso de electroencefalogramas para detectar actividad anormal que fuera predictiva para autismo (Bosl *et al.*, 2011), así como la existencia de una actividad significativa de la *serotonina* de las plaquetas (*5-HT*) y *mTOR* (Veenstra-Vander Weele y Blakely, 2012), sugiriendo una terapia dirigida contra estos genes para mejorar el tratamiento. En 2012, ya se apostaba por la genómica y otras disciplinas biomédicas para obtener biomarcadores (Voineagu, 2012), apuntando como relevante la familia de genes *FOXP*. En 2014, se relacionó el metabolismo del triptófano y otras rutas metabólicas, la influencia de la flora intestinal, la importancia del sistema inmune y la mielinización de las neuronas con el trastorno, proponiendo como biomarcadores *IGF*, la *serotonina* y *anti-BMP* (Wilson, 2014). Además, se trató de anticipar el diagnóstico del TEA con una muestra de sangre en la que se medirían la *serotonina* y la *interleucina 6* (Yang *et al.*, 2015). En 2017, se intentó identificar un patrón específico de actividad encefálica con neuroimagen, sin resultados satisfactorios, por desafíos estadísticos y de complejidad intrínseca (Uddin *et al.*, 2017). También se vincularon las fosfoinositidas y los genes *PTEN* y *mTOR* por su relevancia en el funcionamiento fisiológico del cerebro (Gross, 2017). En 2020, se implicó a las células de la glía y similares en la aparición del autismo, argumentando que el gen *IGF* estaría alterado por una infección bacteriana en el embarazo, afectando a la glía y empeorando la conectividad del tejido neuronal, además de relacionar algunos SNPs con el trastorno (Steinman, 2020). Asimismo, se han asociado las regiones diferenciales de metilación de los espermatozoides de los padres con la probabilidad de que la descendencia sufriera la enfermedad, a través de la realización de estudios de asociación del genoma completo o GWAS (por las siglas en inglés de *Genome-Wide Association Studies*) (Garrido *et al.*, 2021). Más recientemente, en 2022, se defendió la importancia que tenía *CIq TNF-related protein 3 (CTRP3)* dentro del TEA, por sus funciones metabólicas e inmunitarias, aunque con limitaciones de muestra (Alhakhbany *et al.*, 2022).

En cuanto a los biomarcadores de la esquizofrenia, se cree que la melatonina tiene una función muy relevante y que puede ser un biomarcador eficaz, incluso una terapia adecuada. Sin embargo, el desconocimiento sobre esta hormona y sus posibles efectos impide su uso (Morera-Fumero y Abreu-Gonzalez, 2013). Por otra parte, se ha vinculado la arquitectura génica con la esquizofrenia a través de GWAS, afirmando que la complejidad de las enfermedades mentales traspasaba la capacidad de la genética (Sullivan *et al.*, 2012). En 2016, se dividía la sintomatología de la enfermedad en cuatro

aspectos principales (dimensión positiva, negativa, depresiva y cognitiva). Se estimaba que había diferentes biomarcadores para cada dimensión, lo que debería tenerse en cuenta para desarrollar farmacología al respecto (Garcia-Alvarez *et al.*, 2016). En 2017, un estudio explicaba la enorme complejidad en torno a los sistemas y rutas metabólicas del cerebro, desde compuestos inmunitarios (como las citoquinas, que generan inflamación), neuroendocrinos (eje hipotalámico-pituitario-adrenal), metabólicos, neurotróficos, sistemas neuronales como el glutamatérgico, serotoninérgico, colinérgico, adrenérgico, etc., con el objetivo de diagnóstico, prevención y ajuste del tratamiento a través de un biomarcador fiable (Perkovic *et al.*, 2017). En ese mismo año, se relacionaron biomarcadores inflamatorios (*C-Reactive Protein* y *IL-6*) con la esquizofrenia, basándose en un conjunto de pruebas estadísticas, así como la posible influencia de SNPs (Hartwig *et al.*, 2017). En 2018, se estableció en un estudio que *BDNF* (*Brain Derived Neurotrophic Factor*) era un factor central tanto en esquizofrenia como en depresión mayor, sugiriendo aplicar la estimulación magnética transcraneal repetitiva para mejorar el pronóstico (Peng *et al.*, 2018). También existen indicios sobre la limitación de las técnicas ómicas para detectar biomarcadores de esquizofrenia (Davison *et al.*, 2018). En 2000, a través de la epigenética y la transcriptómica, se obtuvieron dos biomarcadores prometedores (*RELN* y *GAD67*), pero no sin reservas debido al limitado conocimiento sobre estos campos. Además, se atribuía una gran importancia al transcriptoma, mostrando una actitud optimista respecto a la capacidad futura de prevención y diagnóstico (Khavari y Cairns, 2020). En 2021, un artículo postulaba la importancia de la regulación del metabolismo del glutamato en el desarrollo de la esquizofrenia. Argumentaba que el déficit de cisteína y serina afectaba a la síntesis de glutatión y a la composición del *sistema X<sup>c-</sup>*, encargado de activar y regular el sistema glutamatérgico, responsable de la motivación. Como posible biomarcador se proponía el ARN mensajero relacionado con la síntesis del *sistema X<sup>c-</sup>* (Hung *et al.*, 2021). En 2022, se abordó la detección de genes candidatos en cerebros *post mortem* de afectados, y se encontraron 15 genes clave que apuntaban a alteraciones de células inmunes con capacidad de infiltrarse, tras la aplicación de múltiples herramientas bioinformáticas (Li *et al.*, 2022). También se ha relacionado la enfermedad con la inflamación y el sistema nervioso, a través de la determinación de los niveles en sangre y otros tejidos de unas sustancias denominadas ‘alarminas’ (*high-mobility group box 1 protein (HMGB1)*, interleucina-1 $\alpha$ , interleucina-33, S100B, proteínas de shock térmico y ácido úrico) (Ma *et al.*, 2022).

Los biomarcadores para la detección del trastorno bipolar son también heterogéneos. En 2014 se vinculó la progranulina, afectada por el tratamiento con litio, un medicamento eficaz, con el trastorno bipolar (Kittel-Schneider *et al.*, 2014). En 2016, tras numerosas aproximaciones experimentales, se relacionó el factor neurotrópico *BDNF* con el trastorno (Sigitova *et al.*, 2017). Además, se estudió el impacto derivado de la ingesta de compuestos dentro de la categoría de ácidos grasos poliinsaturados, como el Omega 3 y el Omega 6, destacando un experimento que demostraba el efecto beneficioso de la ingesta de Omega 3 sobre pacientes de trastorno bipolar (Saunders *et al.*, 2016). En 2017, se propuso la cisteína como indicador para el trastorno bipolar, debido a su importancia en el metabolismo y relevancia en otras enfermedades cerebrales, detectándose niveles diferentes entre afectados y controles, aunque mostrando reservas por el tamaño de la muestra (Salagre *et al.*, 2017). También se ha utilizado el electroencefalograma para investigar la conectividad cerebral de los pacientes (Maggioni *et al.*, 2017). En 2018, se barajó la posibilidad de detectar el trastorno bipolar antes de que se manifieste en jóvenes (Vieta *et al.*, 2018). Igualmente se estableció que los lípidos podían ser importantes en el desarrollo de la enfermedad, destacando la relevancia del fosfatidilinositol (Knowles *et al.*, 2017). En 2019, se estableció una correlación entre la integridad de la barrera hematoencefálica y la gravedad del trastorno (Kamintsky *et al.*, 2019). En 2020, se hacía una comparativa entre trastorno bipolar y esquizofrenia, considerando una naturaleza común entre ambas (Yamada *et al.*, 2020). En 2022 se relacionó el malondialdehído con la enfermedad por sus diferencias de concentración entre pacientes y controles, postulándolo como biomarcador (Capuzzi *et al.*, 2022). Por último, también se consideró que el polipéptido de la luz de los neurofilamentos (*neurofilament light (NfL) polypeptide*), debido a su función dentro del sistema nervioso (crecimiento del radio de los axones, mantenimiento del calibre y transmisión eléctrica), era un buen candidato como biomarcador (Aggio *et al.*, 2022).

En cuanto a la depresión mayor, se ha descrito que los astrocitos y biomarcadores afines están relacionados con esta enfermedad (Rajkowska y Stockmeier, 2013), y hay evidencia de que la utilización de diversas técnicas ómicas y procedimientos como la neuroimagen presentan potencial para detectar y evaluar biomarcadores propios del trastorno, con proyectos como el EMBARC centrados en la estandarización del diagnóstico de la enfermedad (Gadad *et al.*, 2018). Además, se ha tratado de evaluar la influencia de los ácidos grasos poliinsaturados y el metabolismo de las purinas y la

inosina sobre casos de depresión mayor infantil y juvenil, basándose en los niveles diferencialmente distintos entre personas sanas y enfermas (Zhou *et al.*, 2019). En 2019, se publicó un metaanálisis sobre los biomarcadores de depresión mayor, entre los que se encuentran factores gastrointestinales, técnicas de neuroimagen, factores neurotrópicos, neurotransmisores, hormonas, inmunología y estrés oxidativo (Kennis *et al.*, 2020). En 2020 se llevó a cabo un estudio centrado en encontrar diferencias entre posibles biomarcadores diferenciales entre hombres y mujeres (Jentsch *et al.*, 2020). También se ha resaltado la función de los ‘*Toll-like receptors*’ (TLRs), que activan algunos agentes proinflamatorios en el cerebro, y se han sugerido ciertos miRNA como biomarcadores (Figueroa-Hall *et al.*, 2020). La utilización de técnicas bioinformáticas ha permitido obtener 585 genes diferencialmente expresados en depresión mayor al analizar muestras cerebrales de tejidos *post mortem* (Qi y Chen, 2021). Recientemente, se ha vinculado el nivel de *Proteína C Reactiva (CRP)*, la *Interleucina 8*, algunos biomarcadores inmunitarios (principalmente, citoquinas) y la inflamación con la depresión mayor, y se cree que niveles bajos de inflamación tendrían como consecuencia un mejor pronóstico (Gasparini *et al.*, 2022; Harsanyi *et al.*, 2022; Orsolini *et al.*, 2022). Asimismo, se ha asociado el cambio de la flora intestinal con la aparición de la depresión, con una posible influencia de los miRNAs en el proceso (Bai *et al.*, 2022).

### **1.3 Trastornos del endometrio y enfermedades comórbidas**

La endometriosis es un trastorno ginecológico complejo (Burney y Giudice, 2012; Czyzyk *et al.*, 2017) que se caracteriza por la presencia de células ectópicas del endometrio en otras localizaciones corporales, y cuyas principales consecuencias son problemas reproductivos y dolores pélvicos de considerable intensidad (Saunders y Horne, 2021). La prevalencia mundial es de al menos el 10% para mujeres en edad reproductiva (Moradi *et al.*, 2021; Taylor *et al.*, 2021). Se han propuesto numerosos factores, tales como hormonales o inmunológicos, pero debido al gran número de rutas metabólicas afectadas (Agarwal *et al.*, 2019; Saunders y Horne, 2021), los estudios publicados propician una amplia variedad de interpretaciones. La cirugía es el tratamiento más eficaz hasta la fecha (Eltabbakh y Bower, 2008; Leonardi *et al.*, 2020), pero afecta a la fertilidad de las pacientes, y hay numerosos factores que influyen en esta problemática

a tener en cuenta. Debido a esto, la transcriptómica constituye una oportunidad para comprender el trastorno y poder llegar a una solución.

Los leiomiomas uterinos (Koutsilieris, 1992; Lethaby y Vollenhoven, 2015) son tumores de carácter benigno ubicados en el tejido endometrial. Afectan al menos entre el 12 y el 25% de las mujeres en edad reproductiva (De La Cruz y Buchanan, 2017), si no a más debido a su alta frecuencia de carácter asintomático (Giuliani *et al.*, 2020), y se clasifica por tamaño y lugar de implantación en el endometrio (McWilliams y Chennathukuzhi, 2017). Es un fenotipo que altera negativamente la vida de las pacientes y puede llegar a convertirse en tumores malignos (Ip *et al.*, 2010), incluso con metástasis en los pulmones (Jautzke *et al.*, 1996).

El cáncer de endometrio, enfermedad comórbida a la endometriosis (Declas y Lucot, 2019), se caracteriza por un crecimiento tisular descontrolado, y es altamente letal (Giménez-Bachs y Salinas-Sánchez, 2022; Onstad *et al.*, 2016). El cáncer de endometrio afecta a aproximadamente 142.000 mujeres en todo el mundo, y la mortalidad se eleva a 42.000 casos cada año (Amant *et al.*, 2005). La mayor parte de los diagnósticos ocurren después de la menopausia (Wu *et al.*, 2019), afectando en mayor medida a las mujeres con 70 años o más (Amant *et al.*, 2005). A pesar de su mal pronóstico, los síntomas aparecen en la enfermedad temprana, por lo que el diagnóstico se realiza en las primeras etapas (Amant *et al.*, 2005). Los principales factores de riesgo que se consideran son la exposición no regulada del endometrio al estrógeno, menarquia temprana, menopausia tardía, terapia con tamoxifeno, ausencia de hijos concebidos, infertilidad o fallo al ovular y síndrome del ovario poliquístico (Ignatov y Ortmann, 2020). El sangrado después de la menopausia permite su detección temprana. Hay varios métodos terapéuticos, como la histerectomía, que se adecuan al estado de progresión observado en cada mujer.

### **1.3.1 Biomarcadores de trastornos uterinos**

La investigación sobre los trastornos del endometrio es de vital importancia en el campo de la salud reproductiva femenina. Estas condiciones afectan a un gran número de mujeres, con el consiguiente impacto en su calidad de vida y fertilidad. La detección de biomarcadores que permitan diagnosticar de forma temprana estos trastornos representa un avance imprescindible, pues ayuda a comprender mejor las causas que conducen a la

aparición de estas condiciones y hacen posible el desarrollo de tratamientos dirigidos eficaces.

En lo que a la endometriosis respecta, se ha profundizado en el uso de los microARNs como biomarcadores no invasivos, basándose en su relación con ésta y sus propiedades predictivas para el diagnóstico, así como por el hecho de que sus niveles se encuentran afectados como consecuencia de ciertos tratamientos propios de la endometriosis (Agrawal *et al.*, 2018; Bjorkman y Taylor, 2019; Kovács *et al.*, 2021; Raja *et al.*, 2021; Zhuo *et al.*, 2022). Recientemente, un artículo de revisión ponía de manifiesto la existencia de biomarcadores sanguíneos y de orina, pero concluía que no eran suficientes y apelaba a un mayor esfuerzo en la investigación (Anastasiu *et al.*, 2020). En cuanto a los principales procesos biológicos afectados en la endometriosis, un estudio que utilizaba microarrays y análisis de enriquecimiento funcional demostró que la transición epitelio-mesenquimal parecía jugar un rol relevante en el desarrollo de esta condición (Chen *et al.*, 2020). También se ha apuntado a biomarcadores glucídicos y a la composición de las poblaciones bacterianas dentro de la vagina y del intestino como posibles biomarcadores, debido a diferencias significativas detectadas al comparar con personas sanas. Diversos genes, factores inmunes, hormonales e inflamatorios también se han visto relacionados con esta enfermedad (Anastasiu *et al.*, 2020). Entre los genes potencialmente considerados biomarcadores de la endometriosis, destaca *ASPN*, que codifica la proteína asporina y que parece estar relacionado con filtraciones de linfocitos (Wang y Sun, 2022), así como los genes *ARID1a* o los receptores de estrógenos y progesterona (Bartiromo *et al.*, 2022).

La investigación realizada hasta la fecha en el contexto de los leiomiomas uterinos ha permitido proponer diversas moléculas como posibles biomarcadores de la enfermedad. En 2007 se propusieron versican (un proteoglicano), el factor de crecimiento transformante beta 3 (codificado por el gen *TGFB3*), el citocromo P450-26A1 y la dermatopontina, basándose en estudios con técnicas de imagen y cultivos celulares entre situaciones experimental y control, tras lo cual han sido analizados los genes de origen de estas sustancias (Malik y Catherino, 2007). Posteriormente, se trató de analizar en profundidad diversos marcadores como la prolactina, el factor de crecimiento endotelial vascular (codificado por el gen *VEGF*), la grelina, la obestantina, la lactato deshidrogenasa, o el antígeno de cáncer 125, entre otros, pero en todos los casos se consideró que su eficacia era insuficiente para tratar el problema médico (Levy *et al.*,

2013). En 2017, tras un análisis muy exhaustivo y altamente preciso del proteoma, se detectaron 72 proteínas relevantes, y entre ellas se propuso a TRADD (*tumor necrosis factor receptor type I - associated DEATH domain protein*), que se caracteriza por dominar la apoptosis extrínseca y las respuestas inflamatorias, como biomarcador (Liu *et al.*, 2017). También se abordó la complejidad de su detección atendiendo a factores de riesgo como la raza, índice de masa corporal, exposición a disruptores endocrinos, su tipología, las alteraciones citológicas, o las hormonas implicadas, entre otros (McWilliams y Chennathukuzhi, 2017). En 2018, se encuadra la leiomiomatosis uterina dentro de un marco patológico mucho mayor, y se trata sobre los posibles potenciadores génicos como los disruptores endocrinos, con la participación de genes y hormonas implicadas (Chodankar y Critchley, 2019). En 2019, un artículo concluyó con la idea de que los estudios moleculares permitirán una diagnosis diferencial a través de biomarcadores génicos y moleculares, después de usar la bioinformática para procesar la información procedente de inmunohistoquímica y análisis de conglomerados en la detección de posibles biomarcadores (Zhang *et al.*, 2018). Además, recientemente se han considerado factores hormonales, metabólicos, genéticos, cromosómicos, epigenéticos y celulares como posibles biomarcadores eficaces que ayudarán en un futuro a entender la enfermedad (Machado-Lopez *et al.*, 2021).

En cuanto al cáncer de endometrio, se han propuesto diferentes biomarcadores como *CA-125*, *HE4*, el microARN *miR-15a-5p*, la proteína de unión fuerte *claudina 9* (*CLDN9*) o la proteína *condroadherina* (Li *et al.*, 2009; Zhou *et al.*, 2021; Endo *et al.*, 2022; Herkiloglu *et al.*, 2022), considerando los resultados obtenidos a partir de diversos proyectos multiómicos (Boroñ *et al.*, 2022) y la cuestión de la especificidad de los biomarcadores en función del tipo de cáncer y los biomarcadores moleculares que influyen sobre el comportamiento clínico y responden a terapias dirigidas (Walsh *et al.*, 2023). Sin embargo, no existen biomarcadores que actualmente se encuentren implementados en la práctica clínica.

## **1.4 Tecnológicas ómicas**

Las tecnologías ómicas (Liu *et al.*, 2022) permiten detectar y analizar moléculas biológicas de un modo sistemático y masivo, con la finalidad de comprender un fenotipo a través de la información obtenida. Algunas de estas populares técnicas son la genómica

(genes) (Del Giacco y Cattaneo, 2012), epigenómica (transformaciones epigenéticas) (Wang y Chang, 2018), transcriptómica (moléculas de ARN) (Jiang *et al.*, 2015), proteómica (proteínas) (Ahram, 2007; Matthiesen y Mutenda, 2007), lipidómica (lípidos) (Han y Gross, 2022) y metabolómica (metabolitos) (Johnson *et al.*, 2016). La principal ventaja de las ómicas es que permiten detectar información a nivel global a partir de una muestra biológica, lo que las ha convertido en una herramienta en auge en los últimos 20 años (Olivier *et al.*, 2019). Dentro de la presente investigación se utilizarán datos procedentes de la aplicación de microarrays, un dispositivo ampliamente utilizado como herramienta para el análisis transcriptómico (Bednár, 2000). Los microarrays consisten en una superficie plana sobre la cual se encuentran fijados fragmentos de un gen concreto e identificado, que al contacto con los ADN copia generados a partir de las muestras biológicas de interés son capaces de hibridarse y emitir una señal detectable mediante fluoróforos. Según la manera de medir la fluorescencia para cuantificar la expresión génica de una muestra, se consideran dos tipos de microarrays: i) los microarrays de un color, en los cuales el material experimental y el control se miden en placas distintas, ofreciendo valores absolutos de luminosidad de cada uno que luego serán comparados, y ii) los microarrays de dos colores, donde material de muestras experimentales y controles son depositados en la misma placa pero con colores distintos, lo que ofrece una comparativa directa entre valores de expresión génica de ambas muestras. Los primeros permiten una comparación más directa entre resultados de diferentes experimentos y los segundos son más intuitivos al detectar diferencias entre fenotipos experimentales y controles, con el inconveniente de que son más difícilmente comparables. La luminosidad medida sobre la placa es transformada en datos brutos de expresión génica según la identidad del gen y la cantidad de material genético hibridado de cada pocillo.

Estos datos serán procesados estadísticamente para obtener información relevante, como la identidad de genes candidatos dentro de cada trastorno complejo. No obstante, gran parte de las técnicas estadísticas existentes hasta la fecha adolecen del problema de comparaciones múltiples (Farcomeni, 2008), ya que al asumir cierto grado de error (a menudo del 5%), los errores derivados del análisis de miles de genes y decenas de pacientes pueden ser significativos en la obtención de información biológica relevante, y si se aplican técnicas correctoras, los resultados pueden llegar a ser demasiado conservadores. La heterogeneidad de la naturaleza de los datos y la metodología utilizada

para su análisis hace que exista una gran disparidad de resultados e interpretaciones de los mismos.

## 1.5 Teoría de juegos cooperativos

La Teoría de Juegos Cooperativos es una rama de las matemáticas que se encarga de determinar la relevancia de la aportación individual de todos los agentes que participan en un sistema, para llegar a un determinado resultado. Dentro de esta metodología, se analizan las diversas posibilidades de combinación de agentes y los resultados obtenidos en conjunto por éstos, lo que permite la medición de la importancia de un agente dentro de un sistema para provocar con su acción un determinado resultado. Lo que la diferencia de la Teoría de Juegos No Cooperativos es que, mientras la última no permite la comunicación ni el acuerdo entre los diversos agentes antes de escoger una estrategia, la Teoría de Juegos Cooperativos permite los pactos previos a la situación de interés para maximizar las ganancias obtenidas por la situación. Esto obliga a utilizar algún criterio para repartir los beneficios entre aportaciones de agentes de diferente valor. Hay diferentes valores utilizados para ello (valor de Shapley, valor de Banzhaf), cuya obtención requiere diversos algoritmos que procesan la información del sistema completo considerado. En el contexto de la presente investigación, los agentes son los genes y el resultado obtenido es el trastorno o enfermedad complejos en cuestión. Esta metodología se utiliza en economía, política, y por supuesto, en biología (Leimar y McNamara, 2022). Dentro de la biología, se ha considerado importante dentro de la evolución de la estructura de sexos 1:1, la comunicación animal y las relaciones de colaboración entre animales (Bshary y Oliveira, 2015; Hinde, 1981; Smith, 1976).

El análisis comparativo del valor de Shapley (por sus siglas en inglés, CASH) es una técnica estadística que combina el algoritmo *Microarray game* (aplicado sobre datos de valores transcriptómicos obtenidos por experimentos de microarrays) con la técnica Bootstrap (que aplica sobre la primera un remuestreo con ausencia aleatoria de ciertos valores, con la finalidad de compensar posibles valores atípicos en la matriz de datos). El algoritmo *Microarray game* asigna un valor de 1 a aquellos genes que se expresan diferencialmente (cuyos valores estén bien por encima de la media más la desviación típica, en caso de genes sobreexpresados, o por debajo de la media menos la desviación típica, en caso de genes subexpresados) y 0 al resto de valores. Una vez aplicado el

algoritmo, tanto en muestras controles como en experimentales, se resta el valor de Shapley obtenido por un gen concreto en muestras controles al obtenido para el mismo gen en muestras experimentales. El resultado es el valor de Shapley definitivo de dicho gen. Las propiedades de esta técnica son cuatro: suficiencia, jugador nulo, simetría y aditividad. El análisis comparativo del valor de Shapley ha sido previamente utilizado para la detección de genes candidatos en el contexto de trastornos complejos (Esteban & Wall, 2011; Moretti *et al.*, 2008).

## 2. Objetivos

El objetivo general de esta Tesis Doctoral es avanzar en el conocimiento acerca de las bases moleculares de enfermedades complejas como las enfermedades neurológicas y las que afectan al tejido uterino a través del uso de aproximaciones *in silico* que permitan identificar potenciales biomarcadores y rutas moleculares implicadas en dichas enfermedades.

Los objetivos específicos de la tesis incluyen:

Capítulo I) contextualizar los principales avances obtenidos gracias a la utilización de las tecnologías ómicas y la biología de sistemas para el estudio de enfermedades complejas, concretamente de trastornos neuropsiquiátricos y desórdenes ginecológicos.

Capítulo II) aplicar técnicas basadas en la Teoría de Juegos Cooperativos para el análisis de datos de expresión génica en trastornos del endometrio y comparar los resultados con los obtenidos mediante la aplicación de técnicas convencionales de análisis de expresión diferencial.

Capítulo III) implementar técnicas basadas en la Teoría de Juegos Cooperativos para el análisis de datos procedentes de microarrays de trastornos neuropsiquiátricos y comparar los resultados con los procedentes del uso de técnicas tradicionales para el análisis de expresión génica diferencial.

## 3. Material y Métodos

En esta sección se recopilan las principales consideraciones y métodos utilizados durante la realización de la presente Tesis Doctoral. Las descripciones detalladas acerca de los protocolos empleados se pueden consultar en los respectivos estudios.

### 3.1 Datos

La selección y obtención precisa de datos biológicos procedentes de experimentos de transcriptómica constituye un pilar fundamental en la investigación en biología molecular y celular. Debido a la naturaleza computacional de los estudios que conforman esta tesis, la mayor parte del material utilizado para su desarrollo son datos biológicos procedentes de diferentes fuentes públicas. Más concretamente, esta investigación ha requerido del uso de conjuntos de datos de expresión génica de acceso público y disponibles gracias a la base de datos del Centro Nacional para Información en Biotecnología o NCBI (por las siglas en inglés de *National Center for Biotechnology Information*), y en particular de su repositorio *Gene Expression Omnibus* (GEO; <https://www.ncbi.nlm.nih.gov/geo/>), donde se recoge información de interés clínico disponible para investigadores de todo el mundo, con el fin de compartir datos clave en la comprensión y el estudio de numerosas enfermedades. Esta plataforma, con origen estadounidense, almacena una ingente cantidad de bases de datos biológicos, además de organizar y parametrizar un conjunto muy amplio de artículos científicos de la rama de Biomedicina. *GEO Datasets* es un subrepositorio de datos genómicos funcionales y públicos, y es de vital importancia en esta investigación porque permite el acceso a la información necesaria en cuanto a la expresión del transcriptoma en los diversos trastornos bajo estudio.

PubMed (<https://pubmed.ncbi.nlm.nih.gov/>) es una parte importante de NCBI que contiene cientos de miles de artículos científicos relativos a la Biomedicina, generalmente dentro de la base de datos de MEDLINE, que está compuesta en gran medida por la Biblioteca Nacional de Medicina de los Estados Unidos. En el primer Capítulo de la presente tesis se realizó una revisión bibliográfica de la literatura disponible, con el objetivo de identificar estudios relevantes que vinculasen la utilización de tecnologías ómicas con avances significativos en la investigación en desórdenes neuropsiquiátricos y

ginecológicos. A lo largo del capítulo se describen los principales hallazgos alcanzados gracias a la aplicación de la biología de sistemas relacionados con la esquizofrenia, el trastorno bipolar, la depresión, el leiomioma uterino, el cáncer de ovario y la endometriosis.

El segundo y tercer Capítulo parten de la búsqueda y selección de conjuntos de datos públicos de interés en los que se ha utilizado la tecnología de microarrays para el análisis del genoma completo como herramienta propia de la transcriptómica. El Capítulo II se centró en conjuntos de datos de expresión génica relacionados con alteraciones ginecológicas a nivel uterino. Por su parte, el Capítulo III giró en torno al análisis de datos transcriptómicos vinculados con alteraciones neuropsiquiátricas. La descripción de la naturaleza de los datos incluidos en los Capítulos II y III se muestra en la Tabla 1.

Tabla 1. Características de los estudios transcriptómicos utilizados para la extracción de datos de los Capítulos II (sombreado en color verde) y III (sombreado en color azul). GEO: *Gene Expression Omnibus*.

<b>Fenotipo</b>	<b>Dataset de GEO</b>	<b>Plataforma</b>	<b>Descripción de las muestras</b>
Cáncer endometrial	GSE36389	Affymetrix Human Genome U133A Array	Cáncer endometrial (n=10) vs. controles (n=6)
	GSE63678	Affymetrix Human Genome U133A 2.0 Array	Carcinoma endometrial (n=4) vs. controles (n=5)
Endometriosis	GSE7846	Affymetrix Human Genome U133 Plus 2.0 Array	Endometriosis (n=4) vs. controles (n=5)
	GSE17504	Affymetrix Human Gene 1.0 ST Array	Endometriosis (n=5) vs. controles (n=6)
Leiomiomatosis uterina	GSE12814	Affymetrix Human Genome U133 Plus 2.0 Array	Leiomioma uterino (n=5) vs. controles (n=9)
	GSE23112	Affymetrix Human Genome U133A Array	Leiomioma uterino (n=3) vs. controles (n=4)
Autismo	GSE6575	Affymetrix Human Genome U133 Plus 2.0 Array	Autismo (n=14) vs. controles (n=11)

	GSE18123	Affymetrix Human Genome U133 Plus 2.0 Array	Autismo (n=13) vs. controles (n=10)
	GSE25507	Affymetrix Human Genome U133 Plus 2.0 Array	Autismo (n=12) vs. controles (n=14)
Esquizofrenia	GSE17612	Affymetrix Human Genome U133 Plus 2.0 Array	Esquizofrenia (n=17) vs. controles (n=13)
	GSE62333	Affymetrix Human Gene 1.1 ST Array	Esquizofrenia (n=11) vs. controles (n=14)
Desorden bipolar	GSE5389	Affymetrix Human Genome U133A Array	Desorden bipolar (n=7) vs. controles (n=10)
	GSE7036	Affymetrix Human Genome U133 Plus 2.0 Array	Desorden bipolar (n=3) vs. controles (n=3)
Miscelánea	GSE12654	Affymetrix Human Genome U95 Version 2 Array	Tejido cerebral (n=24) vs. controles (n=14)
	GSE53987	Affymetrix Human Genome U133 Plus 2.0 Array	Tejido cerebral (n=135) vs. controles (n=51)

### 3.2 Métodos

La aplicación de la tecnología de microarrays para la evaluación de los niveles de expresión génica genera una gran cantidad de datos que deben ser cuidadosamente analizados con el objetivo de obtener resultados con significado biológico. Clásicamente, el análisis de datos procedentes de la aplicación de microarrays se estructura en una serie de bloques bien diferenciados: preprocesamiento-control de calidad, normalización, análisis de expresión diferencial y análisis funcional (Figura 1). Dada la complejidad de los datos biológicos de partida, se utilizan herramientas computacionales que permiten simplificar el análisis. La mayor parte de los análisis realizados en la presente Tesis Doctoral se han ejecutado en R-Studio (versión 2021.09.0), una interfaz computacional muy flexible y potente que utiliza el lenguaje de programación R y que permite la aplicación de numerosos algoritmos, en este caso adaptados al análisis de datos de

expresión génica. Para extender la funcionalidad del software utilizado, se han usado los paquetes de análisis ‘limma’ (Ritchie *et al.*, 2015), ‘[impute](#)’, ‘[multtest](#)’ (versión 2.60.0) y ‘[RankProd](#)’, entre otros.

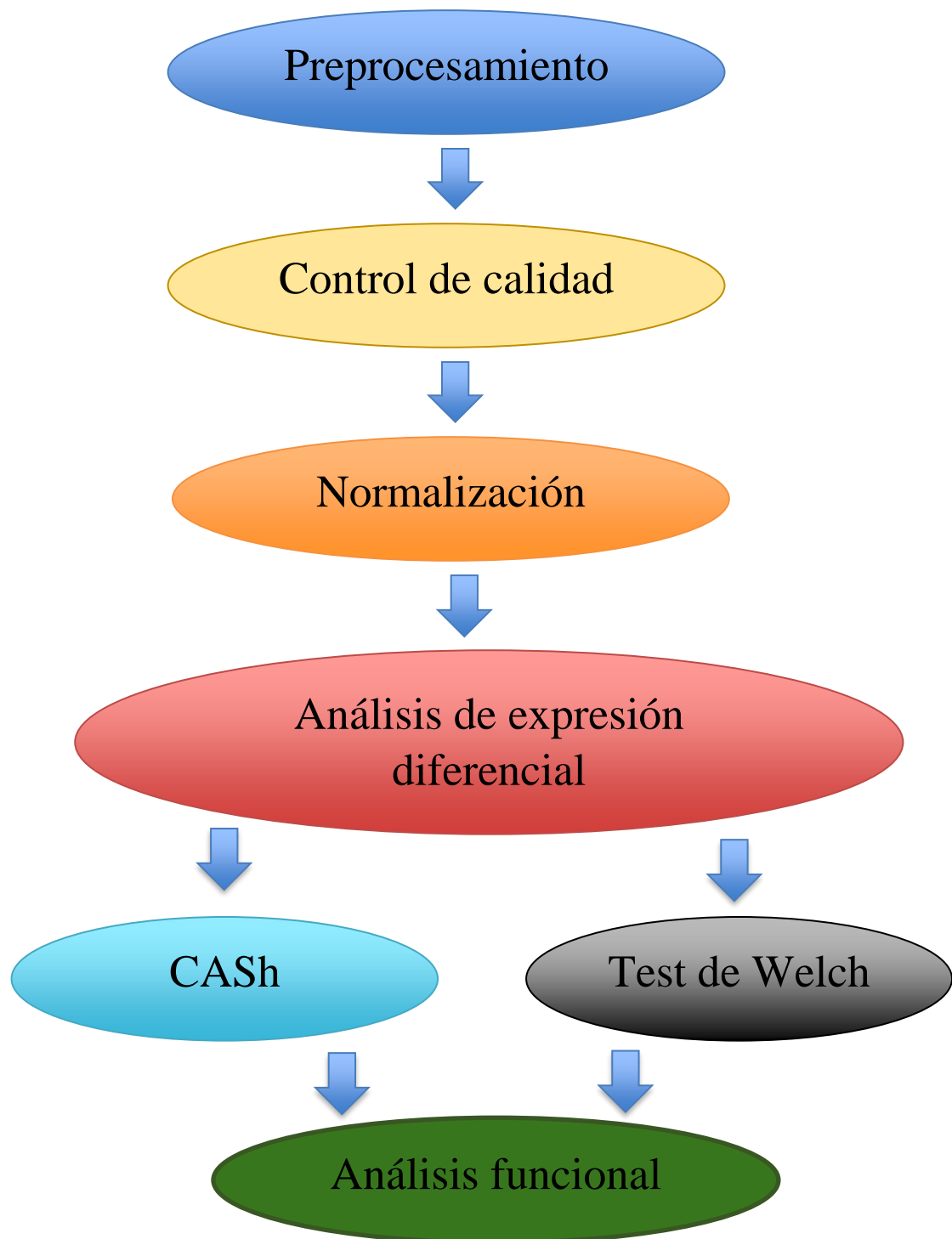


Figura 1. Flujo de trabajo estándar en el análisis de datos de expresión génica. CASH: *Comparative Analysis of Shapley value*.

### 3.2.1 Preprocesamiento y control de calidad

De modo previo al análisis diferencial de los datos contenidos en estos datasets, se ha realizado un preprocesado y un control de calidad de las muestras escogidas para obtener la información más representativa y fidedigna posible, utilizando el paquete ‘[affy](#)’ (versión 1.82.0) y ‘[affyPLM](#)’ (versión 1.80.0), ejecutados ambos en RStudio. El preprocesamiento se ha llevado a cabo mediante una transformación logarítmica, mientras que el control de calidad de las muestras se ha obtenido a través del cálculo de los valores de RLE (*Relative Log Expression*) y NUSE (*Normalized Unscaled Standard Error*), unas métricas de control de calidad utilizadas para evaluar la fiabilidad y consistencia de los datos antes de la realización de análisis posteriores (Abbas-Aghababazadeh, 2018; Tang y Therneau, 2010). La finalidad de la aplicación de estas métricas es la identificación de muestras problemáticas (u *outliers*) antes de proceder con análisis más complejos como el análisis de expresión diferencial.

### 3.2.2 Determinación de la expresión diferencial mediante métodos tradicionales

Para evaluar la expresión diferencial de genes entre grupos de muestras, se emplearon inicialmente métodos estadísticos tradicionales, ampliamente validados en el campo de la bioinformática. El test de Welch (Derrick y White, 2016) fue seleccionado como técnica estadística de control debido a su capacidad para manejar muestras con varianzas desiguales, y su similitud con la prueba t de Student, aunque, a diferencia de esta, no requiere la igualdad de varianzas entre grupos. Sin embargo, es importante destacar que el test de Welch sí asume la normalidad de los datos, por lo que no es ideal para datos génicos, que frecuentemente no siguen esta distribución. El uso de intervalos de confianza en su análisis puede introducir un sesgo que CASH intenta evitar, lo que lo convierte en un método adecuado para compararlo con esta técnica innovadora. Además, se utilizó como método de control el test empírico de Bayes, que estima distribuciones de probabilidad a partir de los datos observados. Este enfoque es útil especialmente cuando se dispone de información previa o cuando el tamaño de la muestra es pequeño, lo que permite realizar inferencias bajo un marco bayesiano. Esto resulta ventajoso en situaciones donde la información a priori puede mejorar la precisión de las estimaciones, especialmente en estudios con muestras limitadas o datos complejos, como los obtenidos

en transcriptómica. En los Capítulos II y III, el valor alfa utilizado para considerar un gen como estadísticamente significativo ha sido de 0.05. Tras este algoritmo, se ha aplicado una técnica de corrección de Benjamini-Hochberg, para la obtención del p-valor corregido tipo *False Discovery Rate* (FDR). Dichos análisis se han ejecutado en RStudio utilizando el paquete '[multtest](#)' (versión 2.60.0) y la función '[ebayes](#)' incluida en el paquete 'limma' (Ritchie *et al.*, 2015).

### **3.2.3 Determinación de la expresión diferencial mediante análisis comparativo del valor de Shapley**

Después de aplicar el test de Welch, y con el objetivo de comparar la utilidad de un método alternativo, se ha aplicado el algoritmo del análisis comparativo del valor de Shapley (Moretti *et al.*, 2008). Para ello, se ha realizado un previo filtrado de genes con una t de Student no pareada a un valor de alfa de 0.05, y otro filtrado con la misma técnica a valor de alfa 0.01, sin ningún tipo de corrección estadística adicional. El motivo inicial para realizar dos análisis con diferentes valores de alfa era comprobar si había diferencias significativas variando el valor de alfa para filtrar el conjunto de los genes disponibles en un experimento de microarrays. Esto mostraría si el tamaño de la muestra obtenida por la criba con la t de Student generaba algún tipo de influencia sobre la calidad de los resultados.

La técnica 'Análisis comparativo del valor de Shapley' se compone de dos algoritmos estadísticos, llamados '*Microarray game*' (Moretti *et al.*, 2007) y 'Bootstrap'. El primero genera una matriz de valores binarios (1s y 0s) en la cual las filas se relacionan con los diferentes genes y las columnas con las muestras utilizadas en el experimento. Mientras que los valores génicos que se destacan por un valor significativamente más alto (sobreexpresados) o más bajo (subexpresados) de la media en el conjunto de los valores génicos dentro de un gen se codifican con el dígito 1, el resto se codifica con el dígito 0. Para el mismo experimento, han de realizarse dos tipos de cálculos, aquellos que obtienen los valores génicos sobreexpresados (mayores que la media más una desviación típica) y los que los obtienen subexpresados (menores que la media menos una desviación típica). Después, sólo se aceptan como genes diferencialmente expresados de forma significativa aquellos cuyo *Fold Change* sea mayor de 1.5, en caso de los genes sobreexpresados, o menores a -1.5, en caso de los genes subexpresados. Una vez que los valores significativos

son cambiados por 1s y el resto por 0s, cada valor 1 que se encuentra en la tabla se divide entre el número de 1s que hay presentes a lo largo de la misma columna (Figura 2). Esto se debe a que se calcula la relevancia de un gen significativo en función del número de genes distintos que también son significativos, dentro de cada muestra control o experimental. Una vez realizado este paso, se suman los valores dentro de cada fila, que representan a los diferentes genes estudiados, de modo que aquellos genes que contengan mayor cantidad de valores génicos significativos tenderán a generar un valor de Shapley mayor que el resto. Esta operación se realiza en las muestras experimentales y las muestras control por separado, ya que se restará el valor obtenido de las muestras controles al obtenido de las muestras experimentales, para asegurar que las diferencias detectadas entre controles y experimentales se deben al fenotipo que se está investigando y no a otros factores. Esta operación se aplica a los genes sobreexpresados y a los genes subexpresados, por separado, obteniendo un valor de Shapley final que indique la relevancia estimada de dicho gen en el fenotipo de interés.

La técnica Bootstrap (Dixon, 2006) es una herramienta de remuestreo en la que se ignoran valores génicos de modo aleatorio para que, en caso de que haya valores atípicos obtenidos de un modo artificial al experimento, estos no alteren el resultado de un modo excesivo para conseguir unos resultados experimentales de genes candidatos más fiables, al ponderar el resultado final basándose en el valor obtenido por todas las iteraciones. En este trabajo, todos los análisis realizados se han completado con 1000 iteraciones.

	E1	E2	E3	E4	C1	C2	C3	C4
A	1	0	1	1	0	0	0	0
B	0	0	1	0	0	1	0	0
C	1	1	1	0	1	0	0	0
D	0	0	0	1	0	0	0	1
E	0	0	0	0	0	0	0	0

Figura 2. Esquema visual del cálculo del valor de Shapley para genes expresados en un trastorno complejo, a partir de los cuales se determinan los genes candidatos para dicho trastorno. Valor de Shapley de la caja X =  $0/1 = 0$  (columna E2). Valor de Shapley de la caja Y =  $1/3 = 0.33$  (columna E3). El valor de Shapley es calculado para todas las casillas. La suma de los valores de Shapley experimentales por cada gen se resta a la suma de los valores Shapley controles del

mismo gen, y el resultado es el valor final de Shapley para dicho gen.  $A = 0.3325$ ;  $B = 0$ ;  $C = 0.3325$ ;  $D = 0$ ;  $E = 0$ . Por lo tanto, A y C son genes candidatos para el fenotipo de interés.

### 3.2.4 Análisis de enriquecimiento funcional

Con el objetivo de proporcionar una interpretación que aporte sentido biológico a los análisis previamente descritos, se realiza el denominado análisis de enriquecimiento funcional una vez obtenidos todos los genes candidatos de sus respectivos experimentos (García-Moreno *et al.*, 2022). Estos análisis vinculan los genes de interés con información publicada sobre éstos, de modo que se puede acceder a información tal como la identidad génica, la localización cromosómica, la función que desempeña dentro de una ruta metabólica concreta, entre otro tipo de descripciones. Durante el desarrollo de la presente tesis, para esta fase del análisis se han utilizado herramientas web como DAVID (*Database for Annotation, Visualization and Integrated Discovery*; <https://david.ncifcrf.gov/>) (Jiao *et al.*, 2012) y g:Profiler (<https://biit.cs.ut.ee/gprofiler/gost>) (Raudvere *et al.*, 2019). En resumen, las herramientas de análisis de enriquecimiento funcional utilizan aproximaciones estadísticas a través de las cuales se identifican grupos de moléculas que estarían sobrerrepresentadas en un conjunto mayor de moléculas, y se trataría de descifrar su posible asociación con fenotipos ya conocidos. Entre las principales fuentes de evidencia, se ha utilizado *Gene Ontology* (GO) para la detección de procesos biológicos, funciones moleculares y componentes celulares significativamente vinculados a los conjuntos de genes diferencialmente expresados obtenidos en cada comparación.

## 4. Resultados y discusión

La investigación en Biomedicina de Sistemas complejos ha despertado un interés creciente a lo largo de los últimos años. Gracias a ello, se ha podido avanzar en el conocimiento acerca de las causas conducentes al desarrollo de trastornos complejos, si bien en la actualidad existe información limitada al respecto. Algunos ejemplos de patologías complejas lo constituyen algunos trastornos neurológicos como el autismo, la esquizofrenia, la depresión mayor o el trastorno bipolar, así como otros desórdenes uterinos como la endometriosis, el cáncer endometrial o los leiomiomas uterinos. Estas patologías, caracterizadas por una compleja arquitectura genética, presentan una alta variabilidad entre diferentes estudios. Estas circunstancias a menudo comprometen el diagnóstico, que en ocasiones ocurre de forma tardía, lo que podría solucionarse a través de la detección de biomarcadores no invasivos y fiables que podrían facilitar la intervención terapéutica temprana. Esta desafiante tarea ha podido ser abordada a lo largo de las últimas décadas gracias a la emergencia de las tecnologías ómicas de alta resolución, que generan una gran cantidad de información, a menudo pobremente explorada. La investigación en este campo ha evidenciado que la biología de sistemas constituye una forma útil de afrontar el reto del descubrimiento de biomarcadores a través de la aplicación de herramientas computacionales capaces de responder preguntas y otorgarles un significado biológico apropiado.

En esta Tesis Doctoral, se han desarrollado tres publicaciones con el objetivo de aportar conocimiento sobre una serie de enfermedades complejas neurológicas y endometriales. En las siguientes secciones, se presentará un resumen de los principales resultados logrados, así como una breve discusión de los principales hallazgos.

### *Capítulo I: aplicación de la Biología de Sistemas al estudio de enfermedades complejas*

En el Capítulo I se presenta una breve contextualización acerca del análisis de datos ómicos y su impacto en la investigación biomédica.

En primer lugar, se describe la necesidad de la investigación de enfermedades complejas a través del uso de herramientas que permitan el abordaje de forma holística de la complejidad inherente a los sistemas biológicos. De igual manera, se mencionan las

principales ventajas de la utilización de las tecnologías ómicas y se hace hincapié en el gran desafío que supone el análisis de los datos derivados de la aplicación de las mismas. Posteriormente, se hace referencia a los principales avances alcanzados gracias a la utilización de estas herramientas en el contexto de algunos trastornos neuropsiquiátricos como la esquizofrenia, el trastorno bipolar, el trastorno depresivo mayor o la ansiedad, así como en desórdenes ginecológicos como los leiomiomas uterinos, la endometriosis o el cáncer de ovario.

En definitiva, el Capítulo I sirve como parte introductoria al presente proyecto de Tesis Doctoral, pues parte de la información descrita es referida en los capítulos posteriores.

## ***Capítulo II: análisis comparativo del valor de Shapley para el estudio de trastornos endometriales***

El Capítulo II se centra en la detección de genes diferencialmente expresados a través del uso de métodos convencionales y el cálculo de valor de Shapley como método alternativo. El objetivo de este trabajo fue comprobar si la técnica no tradicional basada en la Teoría de Juegos Cooperativos es capaz de mejorar la identificación de genes diferencialmente expresados en el contexto de datos públicos de expresión génica centrados en el estudio de tres desórdenes ginecológicos diferentes: el cáncer de endometrio, la leiomiomatosis uterina y la endometriosis. Tras seleccionar los datasets públicos de interés, se descargaron y preprocesaron un total de seis trabajos, dos relativos a cada desorden estudiado. En total, se analizaron 68 muestras correspondientes a datasets de cáncer endometrial (n=27; datasets con códigos GSE36389 y GSE63678), endometriosis (n=20; datasets con códigos GSE7846 y GSE17504) y leiomiomatosis uterina (n=21; datasets con códigos GSE12814 y GSE23112). La utilización de análisis convencionales basados en el test de Welch y el test empírico de Bayes permitieron la detección de un número mínimo de genes diferencialmente expresados en algún dataset, y solamente en algunos casos con valores de significación estadística poco restrictivos (Tabla 2). Sin embargo, en todos los casos, la utilización del método CASH permitió identificar algún gen diferencialmente expresado.

Tabla 2. Número de genes diferencialmente expresados en trastornos endometriales detectados tras el análisis convencional basado en el test de Welch y el test empírico de Bayes (EBayes), y el método alternativo basado en el valor de Shapley (CASH) con p-valores brutos de 0.01 o 0.05, respectivamente. Se indica la corrección de p-valores mediante el uso de *False Discovery Rate* (FDR) cuando corresponda. Sombreado naranja: datasets de cáncer endometrial; azul: endometriosis; verde: leiomiomatosis uterina.

Dataset	Welch	EBayes FDR<0.01	EBayes FDR<0.05	CASH 0.05 FDR<0.05	CASH 0.01	CASH 0.05
GSE36389	0	0	0	0	33	115
GSE63678	0	0	358	33	496	935
GSE7846	0	0	0	140	71	333
GSE17504	0	0	0	17	17	83
GSE12814	0	1	75	0	22	91
GSE23112	0	0	0	16	6	33

La detección de genes diferencialmente expresados gracias a la utilización del método CASH permitió realizar un análisis de enriquecimiento funcional para conocer el contexto biológico en el que se encontraban los conjuntos de transcritos identificados. Así pues, los genes diferencialmente expresados obtenidos tras analizar los datasets de cáncer endometrial (GSE36389 y GSE63678) estaban relacionados con procesos biológicos tales como el desarrollo y la morfogénesis, mientras que los principales componentes celulares y funciones moleculares asociados estaban vinculados con localizaciones extracelulares y unión de diferentes moléculas, respectivamente. En cuanto al análisis funcional de los genes diferencialmente expresados detectados en los datasets de endometriosis (GSE7846 y GSE17504), los principales procesos biológicos asociados se relacionan con el desarrollo, la regulación de distintos procesos celulares, y la morfogénesis. Por su parte, el citoplasma como localización celular y la actividad proteica como función molecular parecen ser procesos en los que se encuentran enriquecidos los genes de interés. Finalmente, los datasets correspondientes a pacientes con leiomiomatosis uterina (GSE12814 y GSE23112) mostraron un particular papel de los genes diferencialmente expresados en la regulación de procesos biológicos, la membrana celular y los procesos de unión a diversas moléculas.

Aunque se trata de resultados preliminares, los resultados obtenidos en nuestro trabajo apuntan hacia eventos con carácter proliferativo como los principales mecanismos

moleculares subyacentes a los genes diferencialmente expresados obtenidos mediante el análisis por CASH. Dichos resultados son esperables, dada la naturaleza y el comportamiento tumoral de los trastornos analizados, y coinciden con los previamente descritos por otros autores (Li *et al.*, 2017; Liu *et al.*, 2021; Takenaka *et al.*, 2022; Zhao *et al.*, 2020).

En lo que respecta a la aplicabilidad en biología reproductiva, ginecología y obstetricia, el uso de estrategias de *machine learning* basadas en el cálculo de los valores de Shapley ha sido recientemente implementado en la detección de biomarcadores de cáncer de ovario, con resultados prometedores (Huang *et al.*, 2023), así como en la investigación sobre el riesgo temporal de preeclampsia en mujeres embarazadas (Eberhard *et al.*, 2024). Por lo tanto, a la luz de todos los resultados descritos, CASH se postula como una herramienta con potencial para el análisis de datos de expresión génica en el contexto de patologías uterinas, si bien es necesario incrementar el número de ensayos y el tamaño de las cohortes analizadas para confirmar la robustez y validez clínica de estos hallazgos.

### ***Capítulo III: aplicación de la Biología de Sistemas complejos para el estudio de alteraciones neurológicas***

Las enfermedades neurológicas tienen un alto impacto en la población mundial y, a pesar de la investigación extensiva, se desconocen muchas de las causas que conducen al origen de las mismas. La llegada de las tecnologías ómicas supuso un importante impulso para la investigación acerca de las bases moleculares de trastornos complejos. Gracias a la aplicación de estas potentes herramientas, se pueden conocer los patrones moleculares asociados con muchas de las condiciones neurológicas más comunes, como el Trastorno del Espectro Autista (TEA), la esquizofrenia, el desorden bipolar o la depresión mayor.

En el presente trabajo, se analizó un conjunto de datos de expresión génica procedentes de la aplicación de la tecnología de microarrays en un total 506 muestras de pacientes diagnosticados con los trastornos neurológicos previamente mencionados. Así pues, se consultaron bases de datos públicas para la obtención de los datos, de forma que se accedió a tres datasets de autismo (GSE6575, GSE18123 y GSE25507), dos datasets

de esquizofrenia (GSE17612 y GSE62333), dos datasets de desorden bipolar (GSE5389 y GSE7036), y dos datasets que combinaban de forma conjunta muestras de esquizofrenia, desorden bipolar y depresión mayor (GSE12654 y GSE53987). Tras la aplicación de métodos de preprocesamiento, control de calidad y normalización, se procedió al análisis de expresión diferencial. Para ello, siguiendo un protocolo experimental similar al descrito para el Capítulo II se aplicaron, por un lado, técnicas de análisis convencional basadas en el test de Welch y el test empírico de Bayes y, por otro, una estrategia alternativa de análisis basada en la Teoría de Juegos Cooperativos a través del análisis mediante el método CASH. Al igual que en el Capítulo II, la aplicación del método CASH tiene efectos muy positivos sobre la detección de genes diferencialmente expresados en los datasets analizados (Tabla 3).

Tabla 3. Número de genes diferencialmente expresados en trastornos neurológicos detectados tras el análisis convencional basado en el test de Welch y el test empírico de Bayes (EBayes), y el método alternativo basado en el valor de Shapley (CASH) con p-valores brutos de 0.01 o 0.05, respectivamente. Se indica la corrección de p-valores mediante el uso de *False Discovery Rate* (FDR) cuando corresponda. Sombreado naranja: datasets de autismo; azul: esquizofrenia; verde: desorden bipolar; amarillo: datasets combinados de esquizofrenia, desorden bipolar y depresión mayor.

Dataset	Welch	EBayes FDR<0.01	EBayes FDR<0.05	CASH 0.05 FDR<0.05	CASH 0.01	CASH 0.05
GSE6575	0	0	0	0	204	930
GSE18123	947	205	2973	45	879	1862
GSE25507	0	0	0	0	28	141
GSE17612	0	0	0	0	1	11
GSE62333	5	0	5	0	68	164
GSE5389	1	0	2	0	40	162
GSE7036	0	0	0	0	8	35
GSE12654			consultar en manuscrito			
GSE53987			consultar en manuscrito			

El análisis de enriquecimiento funcional realizado a partir de los genes diferencialmente expresados permitió profundizar más en el sentido biológico de los hallazgos detectados. Así pues, los genes diferencialmente expresados obtenidos tras el

análisis de los datasets relacionados con el TEA (GSE6575, GSE18123 y GSE25507) mostraron el desarrollo de estructuras y de células del músculo cardíaco como principales procesos biológicos asociados. Algunos trabajos previos han identificado numerosos procesos biológicos asociados con cambios de expresión génica en autismo, como la respuesta a estímulos por citoquinas (Ning *et al.*, 2015), la organización cromosómica, el ciclo celular, el metabolismo (Leblond *et al.*, 2021) o la activación de factores de transcripción, entre otros (Anney *et al.*, 2011). Sin embargo, poco se sabe acerca del posible papel que puedan tener los procesos biológicos significativos obtenidos en nuestro análisis.

Por su parte, la regulación de la muerte celular programada y de procesos metabólicos primarios se detectaron como significativamente enriquecidos tras el análisis funcional de los genes diferencialmente expresados en los datasets de esquizofrenia (GSE17612, GSE62333, GSE12654 y GSE53987). La muerte celular programada está asociada con múltiples enfermedades inmunes y se sabe que juega un rol importante en la esquizofrenia, de forma que se podría considerar un indicador diagnóstico de la enfermedad (Feng y Shen, 2023). Este proceso, además, parece que requiere una activación transcripcional y, por lo tanto, un metabolismo activo que, de no llevarse a cabo, podría resultar en un número anormal de neuronas y un desarrollo neuronal patológico (Margolis *et al.*, 1994).

Los datasets relativos al trastorno bipolar (GSE5389, GSE703, GSE12654 y GSE53987) parecían estar relacionados con la regulación de la actividad de la lipoproteína lipasa. Hasta la fecha, algunos trabajos han tratado de determinar la relación existente entre los niveles de ciertos lípidos con la aparición del trastorno bipolar (Bülbül *et al.*, 2014; Xu *et al.*, 2023). Sin embargo, el rol concreto de la lipoproteína lipasa en el desarrollo del trastorno bipolar aún se desconoce.

Finalmente, los genes diferencialmente expresados obtenidos a partir del análisis de los datasets de depresión mayor (GSE12654 y GSE53897) evidenciaron la epibolia (un proceso del desarrollo embrionario) y la cicatrización de heridas como procesos estadísticamente significativos. Algunos procesos biológicos similares previamente relacionados con la depresión mayor a través del análisis bioinformático han sido la hipoxia, la respuesta inmune, la transición epitelio-mesénquima, o las rutas que implican especies reactivas de oxígeno (Liu *et al.*, 2022; Qi y Chen, 2021; Wittenberg *et al.*, 2020).

Los resultados obtenidos en este trabajo nos ayudan a poner de manifiesto el valor que aportan las estrategias bioinformáticas para la realización de estudios libres de hipótesis, análisis del genoma completo o de datos genómicos y transcriptómicos y su integración en enfermedades tan complejas como las neurológicas, en las que resulta complicado acceder a muestras biológicas que permitan observar la dinámica transcripcional del cerebro en este contexto.

## 5. Conclusiones

En esta Tesis Doctoral hemos implementado, por primera vez de manera exhaustiva y aplicada a distintos trastornos biomédicos complejos, el método estadístico denominado 'Análisis Comparativo del Valor de Shapley' (CASH), basado en la Teoría de Juegos Cooperativos. A diferencia de las técnicas convencionales, CASH permite identificar genes candidatos sin depender de los supuestos comunes de error estadístico, utilizando el 'valor de Shapley' como un parámetro clave que añade un enfoque novedoso al análisis del transcriptoma. Para evaluar la eficacia de CASH en la detección de genes diferencialmente expresados en conjuntos de datos de transcriptómica obtenidos mediante microarrays, se emplearon dos técnicas convencionales como métodos de control: el test de Welch y el enfoque de Bayes Empírico.

El 'Análisis Comparativo del Valor de Shapley' (CASH) ha demostrado obtener mejores resultados en la identificación de genes candidatos en comparación con técnicas estadísticas convencionales. En 19 escenarios experimentales, que comparaban a individuos con trastornos complejos frente a una población control, CASH mostró consistentemente la capacidad de identificar un mayor número de genes candidatos bajo condiciones similares, lo que refuerza su potencial en la medicina personalizada. Esta técnica permite una mayor comprensión de los procesos biológicos involucrados en cada caso específico, ofreciendo una valiosa herramienta tanto para la investigación como para el diseño de tratamientos individualizados. CASH se presenta como una alternativa prometedora para el análisis de datos biológicos complejos, extraídos de bases de datos genómicas, lo que puede ser de utilidad tanto para la investigación biomédica como para el diseño de terapias dirigidas en enfermedades poco conocidas. Esta mayor capacidad de detección presenta un potencial significativo para reducir los costos de investigación, además de optimizar la atención sanitaria mediante el perfeccionamiento del diagnóstico y la personalización de los tratamientos, facilitados por la identificación precisa de biomarcadores utilizando el procedimiento presentado en esta Tesis Doctoral.

## 6. Agradecimientos

Quisiera expresar mi más profundo agradecimiento a mi tutor y director de tesis, Francisco José Esteban, porque me dio la oportunidad de introducirme en el interesante y fructífero campo de la Bioinformática, que será uno de los frentes de investigación más complejos y punteros de los siguientes años dentro de la comprensión de los sistemas vivos. Gracias a su apoyo he podido desarrollar habilidades tanto blandas como técnicas que serán imprescindibles no sólo dentro del campo de la investigación biológica, sino de cualquier actividad profesional en general. Eso es algo que me llena no sólo a nivel académico o laboral sino también personal, porque ha fortalecido mi confianza y mi entusiasmo, y espero poner a prueba esas habilidades con nuevos retos que puedan de alguna manera servir a la sociedad, generando nuevo conocimiento y mejorando lo que pueda aportar la Biología al mundo, ya sea en el campo de la Medicina, del Medio Ambiente o lo que haga falta.

Gracias Eva y Leticia, por estar ahí cuando necesitaba algo de dirección. Revisar una interminable lista de biomarcadores obtenida de una interminable lista de artículos de investigación acerca de trastornos complejos no debe ser tarea fácil, agradezco el tiempo y el cariño dedicados al proyecto que es también vuestro, juntos hemos conseguido probar que la Teoría de Juegos Cooperativos es una opción viable y adecuada para enfrentarse con las dificultades típicas del análisis de expresión génica de trastornos complejos tan dispares como el trastorno bipolar o la endometriosis. Estoy seguro de que ese esfuerzo merecerá la pena.

Gracias infinitas a mis padres, por su apoyo a mi proceso de Doctorado, del cual espero que se sientan muy orgullosos. También a mi hermano y toda la familia, que han funcionado como una red de confianza en la cual he podido crecer y superarme a mí mismo, ellos también forman parte de mi éxito.

Sin olvidarse del afecto y ánimo recibido por parte de la Universidad de Jaén, tanto profesores como compañeros de clase y resto de personas que han estado allí cuando lo he necesitado, muchísimas gracias. No conseguiría lo que he conseguido sin su ayuda hospitalaria y su apoyo desinteresado, me han hecho sentir parte de un todo que representa la Universidad de Jaén, espero que ellos también lleguen a alcanzar el triunfo que tanto se merecen.

Y cómo no, mil gracias a los amigos y compañeros que han estado durante mi vida siendo mis pilares, sin ellos estos años hubieran sido muchísimo más duros y quizás no estaría ahora haciendo un Doctorado. Han estado conmigo cuando más lo necesitaba, y eso es algo que yo les agradeceré eternamente. Se merecen lo mejor.

Y, para terminar, muchas gracias a ti, Juan Carlos, por darme el empujón definitivo para lanzarme de cabeza al Doctorado y por acompañarme cuando tenía que saltar los obstáculos que se encuentran dentro de toda gran hazaña. Porque con tu ánimo, sabiduría y consejo, he caminado por la senda del aprendizaje con la seguridad de estar haciendo algo grande y de saber que tenía un fiel aliado para afrontar un proceso tan complejo como es un Doctorado de Bioinformática. Juntos al fin ya hemos llegado a la meta.

## 7. Bibliografia

- Abbas-Aghababazadeh, F., Li, Q., Fridley, B.L. Comparison of normalization approaches for gene expression studies completed with high-throughput sequencing. *PLoS One*. 2018;13(10):e0206312. doi: 10.1371/journal.pone.0206312.
- Agarwal, S. K., Chapron, C., Giudice, L. C., Laufer, M. R., Leyland, N., Missmer, S. A., Singh, S. S., & Taylor, H. S. (2019). Clinical diagnosis of endometriosis: A call to action. *American Journal of Obstetrics and Gynecology*, 220(4), 354.e1-354.e12. doi: 10.1016/j.ajog.2018.12.039
- Aggio, V., Fabbella, L., Finardi, A., Mazza, E. B., Colombo, C., Falini, A., Benedetti, F., & Furlan, R. (2022). Neurofilaments light: Possible biomarker of brain modifications in bipolar disorder. *Journal of Affective Disorders*, 300, 243-248. doi: 10.1016/j.jad.2021.12.122
- Agrawal, S., Tapmeier, T. T., Rahmioglu, N., Kirtley, S., Zondervan, K. T., & Becker, C. M. (2018). The miRNA Mirage: How Close Are We to Finding a Non-Invasive Diagnostic Biomarker in Endometriosis? A Systematic Review. *International Journal of Molecular Sciences*, 19(2), 599. doi: 10.3390/ijms19020599
- Ahram, M. (2007). An introduction into proteomics and its clinical applications. *Saudi Medical Journal*, 28(4), 499-507.
- Aldinger, F., & Schulze, T. G. (2017). Environmental factors, life events, and trauma in the course of bipolar disorder. *Psychiatry and Clinical Neurosciences*, 71(1), 6-17. doi: 10.1111/pcn.12433
- Alhakhbany, M., Al-Ayadhi, L., & El-Ansary, A. (2022). CTRP3 as a novel biomarker in the plasma of Saudi children with autism. *PeerJ*, 10, e12630. doi: 10.7717/peerj.12630
- Amant, F., Moerman, P., Neven, P., Timmerman, D., Van Limbergen, E., & Vergote, I. (2005). Endometrial cancer. *Lancet (London, England)*, 366(9484), 491-505. doi: 10.1016/S0140-6736(05)67063-8
- Anastasiu, C. V., Moga, M. A., Elena Neculau, A., Bălan, A., Scârneciu, I., Dragomir, R. M., Dull, A.-M., & Chicea, L.-M. (2020). Biomarkers for the Noninvasive Diagnosis of Endometriosis: State of the Art and Future Perspectives. *International Journal of Molecular Sciences*, 21(5), 1750. doi: 10.3390/ijms21051750

- Andreasen, N. C., Nopoulos, P., Schultz, S., Miller, D., Gupta, S., Swayze, V., & Flaum, M. (1994). Positive and negative symptoms of schizophrenia: Past, present, and future. *Acta Psychiatrica Scandinavica. Supplementum*, 384, 51-59. doi: 10.1111/j.1600-0447.1994.tb05891.x
- Anney, R.J., Kenny, E.M., O'Dushlaine, C., et al. Gene-ontology enrichment analysis in two independent family-based samples highlights biologically plausible processes for autism spectrum disorders. *Eur J Hum Genet*. 2011;19(10):1082-1089. doi:10.1038/ejhg.2011.75
- Arberas, C. y Ruggieri, V. (2017) Medicina Buenos Aires, Índices de 2010 a 2019 – Suplemento 1, Autismo. Aspectos genéticos y biológicos
- Bai, S., Bai, H., Li, D., Zhong, Q., Xie, J., & Chen, J. (2022). Gut Microbiota-Related Inflammation Factors as a Potential Biomarker for Diagnosing Major Depressive Disorder. *Frontiers in Cellular and Infection Microbiology*, 12, 831186. doi: 10.3389/fcimb.2022.831186
- Balhara, Y. P., & Verma, R. (2012). Schizophrenia and suicide. *East Asian Archives of Psychiatry: Official Journal of the Hong Kong College of Psychiatrists = Dong Ya Jing Shen Ke Xue Zhi: Xianggang Jing Shen Ke Yi Xue Yuan Qi Kan*, 22(3), 126-133.
- Barnett, A. H., Mackin, P., Chaudhry, I., Farooqi, A., Gadsby, R., Heald, A., Hill, J., Millar, H., Peveler, R., Rees, A., Singh, V., Taylor, D., Vora, J., & Jones, P. B. (2007). Minimising metabolic and cardiovascular risk in schizophrenia: Diabetes, obesity and dyslipidaemia. *Journal of Psychopharmacology (Oxford, England)*, 21(4), 357-373. doi: 10.1177/0269881107075509
- Bartiromo, L., Schimberni, M., Villanacci, R., Mangili, G., Ferrari, S., Ottolina, J., Salmeri, N., Dolci, C., Tandoi, I., Candiani, M. A Systematic Review of Atypical Endometriosis-Associated Biomarkers. *Int J Mol Sci*. 2022 Apr 17;23(8):4425. doi: 10.3390/ijms23084425.
- Bauer, M., Rush, A. J., Ricken, R., Pilhatsch, M., & Adli, M. (2019). Algorithms For Treatment of Major Depressive Disorder: Efficacy and Cost-Effectiveness. *Pharmacopsychiatry*, 52(3), 117-125. doi: 10.1055/a-0643-4830
- Bednár, M. (2000). DNA microarray technology and application. *Medical Science Monitor: International Medical Journal of Experimental and Clinical Research*, 6(4), 796-800.

- Beyer, J. L., & Weisler, R. H. (2016). Suicide Behaviors in Bipolar Disorder: A Review and Update for the Clinician. *The Psychiatric Clinics of North America*, 39(1), 111-123. doi: 10.1016/j.psc.2015.09.002
- Bjorkman, S., & Taylor, H. S. (2019). MicroRNAs in endometriosis: Biological function and emerging biomarker candidates. *Biology of Reproduction*, 100(5), 1135-1146. doi: 10.1093/biolre/ioz014
- Bobilev, A. M., Perez, J. M., & Tamminga, C. A. (2020). Molecular alterations in the medial temporal lobe in schizophrenia. *Schizophrenia Research*, 217, 71-85. doi: 10.1016/j.schres.2019.06.001
- Boroń, D., Zmarzły, N., Wierzbik-Strońska, M., Rosińczuk, J., Mieszczanski, P., & Grabarek, B. O. (2022). Recent Multiomics Approaches in Endometrial Cancer. *International Journal of Molecular Sciences*, 23(3), 1237. doi: 10.3390/ijms23031237
- Bosl, W., Tierney, A., Tager-Flusberg, H., & Nelson, C. (2011). EEG complexity as a biomarker for autism spectrum disorder risk. *BMC Medicine*, 9, 18. doi: 10.1186/1741-7015-9-18
- Bshary, R., Oliveira, R. (2015). Cooperation in animals: toward a game theory within the framework of social competence. *Current Opinion in Behavioral Sciences*. Volume 3, June 2015, Pages 31-37. ISSN 2352-1546, doi: 10.1016/j.cobeha.2015.01.008.
- Bülbül, F., Eryiğit, A.G., Erbağci, A.B., Selek, S., Savaş, H. Alterations of Lipid-Lipoprotein and Leptin in Bipolar Disorder Associated with Clinic Process. *Noro Psikiyatr Ars*. 2014;51(1):52-56. doi:10.4274/npa.y6668
- Burney, R. O., & Giudice, L. C. (2012). Pathogenesis and pathophysiology of endometriosis. *Fertility and Sterility*, 98(3), 511-519. doi: 10.1016/j.fertnstert.2012.06.029
- Capuzzi, E., Ossola, P., Caldiroli, A., Auxilia, A. M., & Buoli, M. (2022). Malondialdehyde as a candidate biomarker for bipolar disorder: A meta-analysis. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 113, 110469. doi: 10.1016/j.pnpbp.2021.110469
- Carlborg, A., Winnerbäck, K., Jönsson, E. G., Jokinen, J., & Nordström, P. (2010). Suicide in schizophrenia. *Expert Review of Neurotherapeutics*, 10(7), 1153-1164. doi: 10.1586/ern.10.82

- Charles, S.. (2014). Aberrant Tryptophan Metabolism: The Unifying Biochemical Basis for Autism Spectrum Disorders?. *Biomarkers in medicine*, 8, 313-5. doi: 10.2217/bmm.14.11.
- Chen, M., Zhou, Y., Xu, H., Hill, C., Ewing, R. M., He, D., Zhang, X., & Wang, Y. (2020). Bioinformatic analysis reveals the importance of epithelial-mesenchymal transition in the development of endometriosis. *Scientific Reports*, 10, 8442. doi: 10.1038/s41598-020-65606-9
- Cheng, W., van der Meer, D., Parker, N., Hindley, G., O'Connell, K. S., Wang, Y., Shadrin, A. A., Alnæs, D., Bahrami, S., Lin, A., Karadag, N., Holen, B., Fernandez-Cabello, S., Fan, C.-C., Dale, A. M., Djurovic, S., Westlye, L. T., Frei, O., Smeland, O. B., & Andreassen, O. A. (2022). Shared genetic architecture between schizophrenia and subcortical brain volumes implicates early neurodevelopmental processes and brain development in childhood. *Molecular Psychiatry*, 27(12), 5167-5176. doi: 10.1038/s41380-022-01751-z
- Chodankar, R., & Critchley, H. O. D. (2019). Biomarkers in abnormal uterine bleeding†. *Biology of Reproduction*, 101(6), 1155-1166. doi: 10.1093/biolre/i0y231
- Clarke R. (2017). Introduction: Cancer Gene Networks. *Methods Mol Biol*. 2017;1513:1-9. doi: 10.1007/978-1-4939-6539-7\_1.
- Clemente, A. S., Diniz, B. S., Nicolato, R., Kapczinski, F. P., Soares, J. C., Firmo, J. O., & Castro-Costa, É. (2015). Bipolar disorder prevalence: A systematic review and meta-analysis of the literature. *Revista Brasileira De Psiquiatria (Sao Paulo, Brazil: 1999)*, 37(2), 155-161. doi: 10.1590/1516-4446-2012-1693
- Czyzyk, A., Podfigurna, A., Szeliga, A., & Meczekalski, B. (2017). Update on endometriosis pathogenesis. *Minerva Ginecologica*, 69(5), 447-461. doi: 10.23736/S0026-4784.17.04048-5
- Davison, J., O'Gorman, A., Brennan, L., & Cotter, D. R. (2018). A systematic review of metabolite biomarkers of schizophrenia. *Schizophrenia Research*, 195, 32-50. doi: 10.1016/j.schres.2017.09.021
- De La Cruz, M. S. D., & Buchanan, E. M. (2017). Uterine Fibroids: Diagnosis and Treatment. *American Family Physician*, 95(2), 100-107.
- Declas, E., & Lucot, J. P. (2019). [Extra uterine leiomyomatosis: Review of the literature]. *Gynecologie, Obstetrique, Fertilité & Senologie*, 47(7-8), 582-590. doi: 10.1016/j.gofs.2019.06.010

- Del Giacco, L., & Cattaneo, C. (2012). Introduction to genomics. *Methods in Molecular Biology (Clifton, N.J.)*, 823, 79-88. doi: 10.1007/978-1-60327-216-2\_6
- Derrick, B., & White, P. (2016). Why Welch's test is Type I error robust. *The Quantitative Methods for Psychology*, 12(1), 30-38. doi: 10.20982/tqmp.12.1.p030
- Dissanayake, C., Searles, J., Barbaro, J., Sadka, N., & Lawson, L. P. (2019). Cognitive and behavioral differences in toddlers with autism spectrum disorder from multiplex and simplex families. *Autism Research: Official Journal of the International Society for Autism Research*, 12(4), 682-693. doi: 10.1002/aur.2074
- Dixon, P. M. (2006). Bootstrap Resampling. En *Encyclopedia of Environmetrics*. John Wiley & Sons, Ltd. doi: 10.1002/9780470057339.vab028
- Doroszkiewicz, J., Groblewska, M., & Mroczko, B. (2021). The Role of Gut Microbiota and Gut–Brain Interplay in Selected Diseases of the Central Nervous System. *International Journal of Molecular Sciences*, 22(18), 10028. doi: 10.3390/ijms221810028
- Eltabbakh, G. H., & Bower, N. A. (2008). Laparoscopic surgery in endometriosis. *Minerva Ginecologica*, 60(4), 323-330.
- Endo, Y., Sugimoto, K., Kobayashi, M., Kobayashi, Y., Kojima, M., Furukawa, S., Soeda, S., Watanabe, T., Higashi, A. Y., Higashi, T., Hashimoto, Y., Fujimori, K., & Chiba, H. (2022). Claudin-9 is a novel prognostic biomarker for endometrial cancer. *International Journal of Oncology*, 61(5), 135. doi: 10.3892/ijo.2022.5425
- Eberhard, B.W., Gray, K.J., Bates, D.W., Kovacheva, V.P. Deep survival analysis for interpretable time-varying prediction of preeclampsia risk. *J Biomed Inform.* 2024;156:104688. doi:10.1016/j.jbi.2024.104688
- Esteban, F. J., & Wall, D. P. (2011). *Gene expression analysis by game theory and microarray technology with an application in Autism Spectrum Disorder*.
- Moretti, S., van Leeuwen, D., Gmuender, H., Bonassi, S., van Delft, J., Kleinjans, J., Patrone, F., & Merlo, D. F. (2008). Combining Shapley value and statistics to the analysis of gene expression data in children exposed to air pollution. *BMC Bioinformatics*, 9(1), 361. doi: 10.1186/1471-2105-9-361
- Evans-Lacko, S., Courtin, E., Fiorillo, A., Knapp, M., Luciano, M., Park, A.-L., Brunn, M., Byford, S., Chevreur, K., Forsman, A. K., Gulacsi, L., Haro, J. M., Kennelly, B., Knappe, S., Lai, T., Lasalvia, A., Miret, M., O'Sullivan, C., Obradors-Tarragó,

- C., Thornicroft, G. (2014). The state of the art in European research on reducing social exclusion and stigma related to mental health: A systematic mapping of the literature. *European Psychiatry: The Journal of the Association of European Psychiatrists*, 29(6), 381-389. doi: 10.1016/j.eurpsy.2014.02.007
- Fagiolini, A., Coluccia, A., Maina, G., Forgione, R. N., Goracci, A., Cuomo, A., & Young, A. H. (2015). Diagnosis, Epidemiology and Management of Mixed States in Bipolar Disorder. *CNS Drugs*, 29(9), 725-740. doi: 10.1007/s40263-015-0275-6
- Farcomeni, A. (2008). A review of modern multiple hypothesis testing, with particular attention to the false discovery proportion. *Statistical Methods in Medical Research*, 17(4), 347-388. doi: 10.1177/0962280206079046
- Feng, Y., Shen, J. Machine learning-based predictive models and drug prediction for schizophrenia in multiple programmed cell death patterns. *Front Mol Neurosci*. 2023;16:1123708. doi:10.3389/fnmol.2023.1123708
- Figuroa-Hall, L. K., Paulus, M. P., & Savitz, J. (2020). Toll-Like Receptor Signaling in Depression. *Psychoneuroendocrinology*, 121, 104843. doi: 10.1016/j.psyneuen.2020.104843
- Filatova, E. V., Shadrina, M. I., & Slominsky, P. A. (2021). Major Depression: One Brain, One Disease, One Set of Intertwined Processes. *Cells*, 10(6), 1283. doi: 10.3390/cells10061283
- Fuglewicz, A. J., Piotrowski, P., & Stodolak, A. (2017). Relationship between toxoplasmosis and schizophrenia: A review. *Advances in Clinical and Experimental Medicine: Official Organ Wroclaw Medical University*, 26(6), 1031-1036. doi: 10.17219/acem/61435
- Gadad, B. S., Jha, M. K., Czysz, A., Furman, J. L., Mayes, T. L., Emslie, M. P., & Trivedi, M. H. (2018). Peripheral Biomarkers of Major Depression and Antidepressant Treatment Response: Current Knowledge and Future Outlooks. *Journal of affective disorders*, 233, 3-14. doi: 10.1016/j.jad.2017.07.001
- Gandal, M. J., Zhang, P., Hadjimichael, E., Walker, R. L., Chen, C., Liu, S., Won, H., van Bakel, H., Varghese, M., Wang, Y., Shieh, A. W., Haney, J., Parhami, S., Belmont, J., Kim, M., Moran Losada, P., Khan, Z., Mleczko, J., Xia, Y., ... Geschwind, D. H. (2018). Transcriptome-wide isoform-level dysregulation in

- ASD, schizophrenia, and bipolar disorder. *Science (New York, N.Y.)*, 362(6420), eaat8127. doi: 10.1126/science.aat8127
- Garcia-Alvarez, L., Garcia-Portilla, M. P., Gonzalez-Blanco, L., Saiz Martinez, P. A., de la Fuente-Tomas, L., Menendez-Miranda, I., Iglesias, C., & Bobes, J. (2016). Biomarcadores sanguíneos diferenciales de las dimensiones psicopatológicas de la esquizofrenia. *Revista de Psiquiatría y Salud Mental*, 9(4), 219-227. doi: 10.1016/j.rpsm.2016.04.003
- Garcia-Moreno, A., López-Domínguez, R., Villatoro-García, J. A., Ramirez-Mena, A., Aparicio-Puerta, E., Hackenberg, M., Pascual-Montano, A., & Carmona-Saez, P. (2022). Functional Enrichment Analysis of Regulatory Elements. *Biomedicines*, 10(3), Article 3. doi: 10.3390/biomedicines10030590
- Garrido, N., Cruz, F., Egea, R. R., Simon, C., Sadler-Riggelman, I., Beck, D., Nilsson, E., Ben Maamar, M., & Skinner, M. K. (2021). Sperm DNA methylation epimutation biomarker for paternal offspring autism susceptibility. *Clinical Epigenetics*, 13, 6. doi: 10.1186/s13148-020-00995-2
- Gasparini, A., Callegari, C., Lucca, G., Bellini, A., Caselli, I., & Ielmini, M. (2022). Inflammatory Biomarker and Response to Antidepressant in Major Depressive Disorder: A Systematic Review and Meta-Analysis. *Psychopharmacology Bulletin*, 52(1), 36-52.
- Gauthier, J., Vincent, A. T., Charette, S. J., & Derome, N. (2019). A brief history of bioinformatics. *Briefings in Bioinformatics*, 20(6), 1981-1996. doi: 10.1093/bib/bby063
- Germann, M., Brederoo, S. G., & Sommer, I. E. C. (2021). Abnormal synaptic pruning during adolescence underlying the development of psychotic disorders. *Current Opinion in Psychiatry*, 34(3), 222-227. doi: 10.1097/YCO.0000000000000696
- Giménez-Bachs, J. M., & Salinas-Sánchez, A. S. (2022). [Biomarkers in renal cancer.]. *Archivos Espanoles De Urologia*, 75(2), 118-125.
- Giuliani, E., As-Sanie, S., & Marsh, E. E. (2020). Epidemiology and management of uterine fibroids. *International Journal of Gynaecology and Obstetrics: The Official Organ of the International Federation of Gynaecology and Obstetrics*, 149(1), 3-9. doi: 10.1002/ijgo.13102
- Gómez Maquet, Y., Ángel, J. D., Cañizares, C., Lattig, M. C., Agudelo, D. M., Arenas, Á., & Ferro, E. (2020). The role of stressful life events appraisal in major

- depressive disorder. *Revista Colombiana De Psiquiatria (English Ed.)*, 49(2), 68-75. doi: 10.1016/j.rcp.2018.07.004
- Greenberg, P. E., Fournier, A.-A., Sisitsky, T., Simes, M., Berman, R., Koenigsberg, S. H., & Kessler, R. C. (2021). The Economic Burden of Adults with Major Depressive Disorder in the United States (2010 and 2018). *Pharmacoeconomics*, 39(6), 653-665. doi: 10.1007/s40273-021-01019-4
- Gross, C. (2017). Defective Phosphoinositide Metabolism in Autism. *Journal of neuroscience research*, 95(5), 1161-1173. doi: 10.1002/jnr.23797
- Gulayín, M. E. (2022). [Burden in family caregivers of people with schizophrenia: A literature review]. *Vertex (Buenos Aires, Argentina)*, XXXIII(155), 50-65. doi: 10.53680/vertex.v33i155.135
- Gutiérrez-Rojas, L., Porrás-Segovia, A., Dunne, H., Andrade-González, N., & Cervilla, J. A. (2020). Prevalence and correlates of major depressive disorder: A systematic review. *Revista Brasileira De Psiquiatria (Sao Paulo, Brazil: 1999)*, 42(6), 657-672. doi: 10.1590/1516-4446-2020-0650
- Häfner, H., & an der Heiden, W. (1997). Epidemiology of schizophrenia. *Canadian Journal of Psychiatry. Revue Canadienne De Psychiatrie*, 42(2), 139-151. doi: 10.1177/070674379704200204
- Hagi, K., Nosaka, T., Dickinson, D., Lindenmayer, J. P., Lee, J., Friedman, J., Boyer, L., Han, M., Abdul-Rashid, N. A., & Correll, C. U. (2021). Association Between Cardiovascular Risk Factors and Cognitive Impairment in People With Schizophrenia: A Systematic Review and Meta-analysis. *JAMA Psychiatry*, 78(5), 510-518. doi: 10.1001/jamapsychiatry.2021.0015
- Han, X., & Gross, R. W. (2022). The foundations and development of lipidomics. *Journal of Lipid Research*, 63(2), 100164. doi: 10.1016/j.jlr.2021.100164
- Harder, A., Nguyen, T.-D., Pasman, J. A., Mosing, M. A., Hägg, S., & Lu, Y. (2022). Genetics of age-at-onset in major depression. *Translational Psychiatry*, 12(1), 124. doi: 10.1038/s41398-022-01888-z
- Harsanyi, S., Kupcova, I., Danisovic, L., & Klein, M. (2022). Selected Biomarkers of Depression: What Are the Effects of Cytokines and Inflammation? *International Journal of Molecular Sciences*, 24(1), 578. doi: 10.3390/ijms24010578
- Hartwig, F. P., Borges, M. C., Horta, B. L., Bowden, J., & Davey Smith, G. (2017). Inflammatory Biomarkers and Risk of Schizophrenia. *JAMA Psychiatry*, 74(12), 1226-1233. doi: 10.1001/jamapsychiatry.2017.3191

- Heckers, S., & Konradi, C. (2002). Hippocampal neurons in schizophrenia. *Journal of Neural Transmission (Vienna, Austria: 1996)*, 109(5-6), 891-905. doi: 10.1007/s007020200073
- Heidari, A., Rostam-Abadi, Y., & Rezaei, N. (2021). The immune system and autism spectrum disorder: Association and therapeutic challenges. *Acta Neurobiologiae Experimentalis*, 81(3), 249-263. doi: 10.21307/ane-2021-023
- Herkiloglu, D., Gokce, S., Gumus, E., Kaygusuz, E. I., & Cevik, O. (2022). Chondoadherin as a biomarker in patients with endometrial cancer. *Annali Italiani Di Chirurgia*, 92, 325-330.
- Hinde, R.A. (1981). Animal signals: Ethological and games-theory approaches are not incompatible. *Animal Behaviour*, Volume 29, Issue 2, May 1981, Pages 535-542. ISSN 0003-3472, doi: 10.1016/S0003-3472(81)80116-9.
- Hirota, T., & King, B. H. (2023). Autism Spectrum Disorder: A Review. *JAMA*, 329(2), 157-168. doi: 10.1001/jama.2022.23661
- Huang, W., Suominen, H., Liu, T., Rice, G., Salomon, C., Barnard, A.S. Explainable discovery of disease biomarkers: The case of ovarian cancer to illustrate the best practice in machine learning and Shapley analysis. *J Biomed Inform.* 2023;141:104365. doi: 10.1016/j.jbi.2023.104365
- Hung, C.-C., Lin, C.-H., & Lane, H.-Y. (2021). Cystine/Glutamate Antiporter in Schizophrenia: From Molecular Mechanism to Novel Biomarker and Treatment. *International Journal of Molecular Sciences*, 22(18), 9718. doi: 10.3390/ijms22189718
- Ignatov, A., & Ortmann, O. (2020). Endocrine Risk Factors of Endometrial Cancer: Polycystic Ovary Syndrome, Oral Contraceptives, Infertility, Tamoxifen. *Cancers*, 12(7), 1766. doi: 10.3390/cancers12071766
- Ip, P. P. C., Tse, K. Y., & Tam, K. F. (2010). Uterine smooth muscle tumors other than the ordinary leiomyomas and leiomyosarcomas: A review of selected variants with emphasis on recent advances and unusual morphology that may cause concern for malignancy. *Advances in Anatomic Pathology*, 17(2), 91-112. doi: 10.1097/PAP.0b013e3181cfb901
- Janoutová, J., Janácková, P., Serý, O., Zeman, T., Ambroz, P., Kovalová, M., Varechová, K., Hosák, L., Jirík, V., & Janout, V. (2016). Epidemiology and risk factors of schizophrenia. *Neuro Endocrinology Letters*, 37(1), 1-8.

- Jautzke, G., Müller-Ruchholtz, E., & Thalmann, U. (1996). Immunohistological detection of estrogen and progesterone receptors in multiple and well differentiated leiomyomatous lung tumors in women with uterine leiomyomas (so-called benign metastasizing leiomyomas). A report on 5 cases. *Pathology, Research and Practice*, 192(3), 215-223. doi: 10.1016/S0344-0338(96)80224-X
- Jentsch, M. C., Burger, H., Meddens, M. B. M., Beijers, L., van den Heuvel, E. R., Meddens, M. J. M., & Schoevers, R. A. (2020). Gender Differences in Developing Biomarker-Based Major Depressive Disorder Diagnostics. *International Journal of Molecular Sciences*, 21(9), 3039. doi: 10.3390/ijms21093039
- Jiang, C.-C., Lin, L.-S., Long, S., Ke, X.-Y., Fukunaga, K., Lu, Y.-M., & Han, F. (2022). Signalling pathways in autism spectrum disorder: Mechanisms and therapeutic implications. *Signal Transduction and Targeted Therapy*, 7(1), 229. doi: 10.1038/s41392-022-01081-0
- Jiang, Z., Zhou, X., Li, R., Michal, J. J., Zhang, S., Dodson, M. V., Zhang, Z., & Harland, R. M. (2015). Whole transcriptome analysis with sequencing: Methods, challenges and potential solutions. *Cellular and Molecular Life Sciences: CMLS*, 72(18), 3425-3439. doi: 10.1007/s00018-015-1934-y
- Jiao, X., Sherman, B.T., Huang, D.W., Stephens, R., Baseler, M.W., Clifford Lane, H., Lempicki, R.A., David, W.S.: a stateful web service to facilitate gene/protein list analysis, *Bioinformatics*, Volume 28, Issue 13, July 2012, Pages 1805–1806, doi: 10.1093/bioinformatics/bts251
- Johnson, C. H., Ivanisevic, J., & Siuzdak, G. (2016). Metabolomics: Beyond biomarkers and towards mechanisms. *Nature Reviews. Molecular Cell Biology*, 17(7), 451-459. doi: 10.1038/nrm.2016.25
- Kahn, R. S., & Sommer, I. E. (2015). The neurobiology and treatment of first-episode schizophrenia. *Molecular Psychiatry*, 20(1), 84-97. doi: 10.1038/mp.2014.66
- Kamintsky, L., Cairns, K. A., Veksler, R., Bowen, C., Beyea, S. D., Friedman, A., & Calkin, C. (2019). Blood-brain barrier imaging as a potential biomarker for bipolar disorder progression. *NeuroImage: Clinical*, 26, 102049. doi: 10.1016/j.nicl.2019.102049
- Kang, D.-W., Adams, J. B., Gregory, A. C., Borody, T., Chittick, L., Fasano, A., Khoruts, A., Geis, E., Maldonado, J., McDonough-Means, S., Pollard, E. L., Roux, S., Sadowsky, M. J., Lipson, K. S., Sullivan, M. B., Caporaso, J. G., & Krajmalnik-

- Brown, R. (2017). Microbiota Transfer Therapy alters gut ecosystem and improves gastrointestinal and autism symptoms: An open-label study. *Microbiome*, 5(1), 10. doi: 10.1186/s40168-016-0225-7
- Karahalil, B. (2016). Overview of Systems Biology and Omics Technologies. *Current Medicinal Chemistry*, 23(37), 4221-4230. doi: 10.2174/0929867323666160926150617
- Katz, I. R., Rogers, M. P., Lew, R., Thwin, S. S., Doros, G., Ahearn, E., Ostacher, M. J., DeLisi, L. E., Smith, E. G., Ringer, R. J., Ferguson, R., Hoffman, B., Kaufman, J. S., Paik, J. M., Conrad, C. H., Holmberg, E. F., Boney, T. Y., Huang, G. D., Liang, M. H., & Li+ plus Investigators. (2022). Lithium Treatment in the Prevention of Repeat Suicide-Related Outcomes in Veterans With Major Depression or Bipolar Disorder: A Randomized Clinical Trial. *JAMA Psychiatry*, 79(1), 24-32. doi: 10.1001/jamapsychiatry.2021.3170
- Kealy, J., Greene, C., & Campbell, M. (2020). Blood-brain barrier regulation in psychiatric disorders. *Neuroscience Letters*, 726, 133664. doi: 10.1016/j.neulet.2018.06.033
- Kendall, K. M., Van Assche, E., Andlauer, T. F. M., Choi, K. W., Luykx, J. J., Schulte, E. C., & Lu, Y. (2021). The genetic basis of major depression. *Psychological Medicine*, 51(13), 2217-2230. doi: 10.1017/S0033291721000441
- Kennedy, S. H. (2008). Core symptoms of major depressive disorder: Relevance to diagnosis and treatment. *Dialogues in Clinical Neuroscience*, 10(3), 271-277. doi: 10.31887/DCNS.2008.10.3/shkennedy
- Kennis, M., Gerritsen, L., van Dalen, M., Williams, A., Cuijpers, P., & Bockting, C. (2020). Prospective biomarkers of major depressive disorder: A systematic review and meta-analysis. *Molecular Psychiatry*, 25(2), 321-338. doi: 10.1038/s41380-019-0585-z
- Keshavarz, K., Hedayati, A., Rezaei, M., Goudarzi, Z., Moghimi, E., Rezaee, M., & Lotfi, F. (2022). Economic burden of major depressive disorder: A case study in Southern Iran. *BMC Psychiatry*, 22(1), 577. doi: 10.1186/s12888-022-04220-7
- Khan, Z. U., Martin-Montañez, E., & Muly, E. C. (2013). Schizophrenia: Causes and treatments. *Current Pharmaceutical Design*, 19(36), 6451-6461. doi: 10.2174/1381612811319360006

- Khavari, B., & Cairns, M. J. (2020). Epigenomic Dysregulation in Schizophrenia: In Search of Disease Etiology and Biomarkers. *Cells*, 9(8), 1837. doi: 10.3390/cells9081837
- Kittel-Schneider, S., Weigl, J., Volkert, J., Geßner, A., Schmidt, B., Hempel, S., Kiel, T., Olmes, D. G., Bartl, J., Weber, H., Kopf, J., & Reif, A. (2014). Further evidence for plasma progranulin as a biomarker in bipolar disorder. *Journal of Affective Disorders*, 157, 87-91. doi: 10.1016/j.jad.2014.01.006
- Knowles, E. E., Meikle, P. J., Huynh, K., Göring, H. H., Olvera, R. L., Mathias, S. R., Duggirala, R., Almasy, L., Blangero, J., Curran, J. E., & Glahn, D. C. (2017). Serum phosphatidylinositol as a biomarker for bipolar disorder liability. *Bipolar Disorders*, 19(2), 107. doi: 10.1111/bdi.12468
- Koutsilieris, M. (1992). Pathophysiology of uterine leiomyomas. *Biochemistry and Cell Biology = Biochimie Et Biologie Cellulaire*, 70(5), 273-278. doi: 10.1139/o92-043
- Kovács, Z., Glover, L., Reidy, F., MacSharry, J., & Saldova, R. (2021). Novel diagnostic options for endometriosis – Based on the glycome and microbiome. *Journal of Advanced Research*, 33, 167-181. doi: 10.1016/j.jare.2021.01.015
- Kraft, C., Chamanadjian, C., & Aylward, B. S. (2023). Autism Spectrum Disorder: The New Asthma? *Clinical Pediatrics*, 62(7), 673-677. doi: 10.1177/00099228221144146
- Krause, M., Zhu, Y., Huhn, M., Schneider-Thoma, J., Bighelli, I., Chaimani, A., & Leucht, S. (2018). Efficacy, acceptability, and tolerability of antipsychotics in children and adolescents with schizophrenia: A network meta-analysis. *European Neuropsychopharmacology: The Journal of the European College of Neuropsychopharmacology*, 28(6), 659-674. doi: 10.1016/j.euroneuro.2018.03.008
- Leblond, C.S., Le, T.L., Malesys, S., et al. Operative list of genes associated with autism and neurodevelopmental disorders based on database review. *Mol Cell Neurosci*. 2021;113:103623. doi:10.1016/j.mcn.2021.103623
- Leimar, O., & McNamara, J. M. (s. f.). Game theory in biology: 50 years and onwards. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 378(1876), 20210509. doi: 10.1098/rstb.2021.0509

- Leonardi, M., Gibbons, T., Armour, M., Wang, R., Glanville, E., Hodgson, R., Cave, A. E., Ong, J., Tong, Y. Y. F., Jacobson, T. Z., Mol, B. W., Johnson, N. P., & Condous, G. (2020). When to Do Surgery and When Not to Do Surgery for Endometriosis: A Systematic Review and Meta-analysis. *Journal of Minimally Invasive Gynecology*, 27(2), 390-407.e3. doi: 10.1016/j.jmig.2019.10.014
- Lethaby, A., & Vollenhoven, B. (2015). Fibroids (uterine myomatosis, leiomyomas). *BMJ Clinical Evidence*, 2015, 0814.
- Levy, G., Hill, M. J., Plowden, T. C., Catherino, W. H., & Armstrong, A. Y. (2013). Biomarkers in uterine leiomyoma. *Fertility and sterility*, 99(4), 1146-1152. doi: 10.1016/j.fertnstert.2012.10.048
- Li, J., Dowdy, S., Tipton, T., Podratz, K., Lu, W.-G., Xie, X., & Jiang, S.-W. (2009). HE4 as a biomarker for ovarian and endometrial cancer management. *Expert Review of Molecular Diagnostics*, 9(6), 555-566. doi: 10.1586/erm.09.39
- Li, X., Mu, F., Liu, D., Zhu, J., Yue, S., Liu, M., Liu, Y., & Wang, J. (2022). Predictors of suicidal ideation, suicide attempt and suicide death among people with major depressive disorder: A systematic review and meta-analysis of cohort studies. *Journal of Affective Disorders*, 302, 332-351. doi: 10.1016/j.jad.2022.01.103
- Li, Z., Li, X., Jin, M., Liu, Y., He, Y., Jia, N., Cui, X., Liu, Y., Hu, G., & Yu, Q. (2022). Identification of potential biomarkers and their correlation with immune infiltration cells in schizophrenia using combinative bioinformatics strategy. *Psychiatry Research*, 314, 114658. doi: 10.1016/j.psychres.2022.114658
- Liu, J., Li, W., Wang, L., Li, J., Li, E., & Luo, Y. (2022). [Multi-omics technology and its applications to life sciences: A review]. *Sheng Wu Gong Cheng Xue Bao = Chinese Journal of Biotechnology*, 38(10), 3581-3593. doi: 10.13345/j.cjb.220724
- Liu, J., Zuo, Y., Qu, G.M., et al. CypB promotes cell proliferation and metastasis in endometrial carcinoma. *BMC Cancer*. 2021;21(1):747. doi:10.1186/s12885-021-08374-7
- Li, M., Xin, X.Y., Jin, Z.S., Hua, T., Wang, H.B., Wang, H.B. Transcriptomic analysis of stromal cells from patients with endometrial carcinoma. *Int J Clin Exp Pathol*. 2017;10(9):9853-9857.

- Liu, S., Lu, T., Zhao, Q, et al. A machine learning model for predicting patients with major depressive disorder: A study based on transcriptomic data. *Front Neurosci.* 2022;16:949609. doi:10.3389/fnins.2022.949609
- Liu, Y., Lu, D., Sheng, J., Luo, L., & Zhang, W. (2017). Identification of TRADD as a potential biomarker in human uterine leiomyoma through iTRAQ based proteomic profiling. *Molecular and Cellular Probes*, 36, 15-20. doi: 10.1016/j.mcp.2017.07.001
- Lord, C., Elsabbagh, M., Baird, G., & Veenstra-Vanderweele, J. (2018). Autism spectrum disorder. *Lancet (London, England)*, 392(10146), 508-520. doi: 10.1016/S0140-6736(18)31129-2
- Ma, H., Cheng, N., & Zhang, C. (2022). Schizophrenia and Alarmins. *Medicina*, 58(6), 694. doi: 10.3390/medicina58060694
- Ma'ayan, A. (2017). Complex systems biology. *Journal of The Royal Society Interface*, 14(134), 20170391. doi: 10.1098/rsif.2017.0391
- Machado-Lopez, A., Simón, C., & Mas, A. (2021). Molecular and Cellular Insights into the Development of Uterine Fibroids. *International Journal of Molecular Sciences*, 22(16), 8483. doi: 10.3390/ijms22168483
- Maggioni, E., Bianchi, A. M., Altamura, A. C., Soares, J. C., & Brambilla, P. (2017). The putative role of neuronal network synchronization as a potential biomarker for bipolar disorder: A review of EEG studies: Special Section on “Translational and Neuroscience Studies in Affective Disorders” Section Editor, Maria Nobile MD, PhD. *Journal of Affective Disorders*, 212, 167-170. doi: 10.1016/j.jad.2016.12.045
- Malhi, G. S., Gessler, D., & Outhred, T. (2017). The use of lithium for the treatment of bipolar disorder: Recommendations from clinical practice guidelines. *Journal of Affective Disorders*, 217, 266-280. doi: 10.1016/j.jad.2017.03.052
- Malik, M., & Catherino, W. H. (2007). Novel method to characterize primary cultures of leiomyoma and myometrium with the use of confirmatory biomarker gene arrays. *Fertility and Sterility*, 87(5), 1166-1172. doi: 10.1016/j.fertnstert.2006.08.111
- Margolis, R.L., Chuang, D.M., Post, R.M. Programmed cell death: implications for neuropsychiatric disorders. *Biol Psychiatry*. 1994;35(12):946-956. doi: 10.1016/0006-3223(94)91241-6

- Martin, C. R., Osadchiy, V., Kalani, A., & Mayer, E. A. (2018). The Brain-Gut-Microbiome Axis. *Cellular and Molecular Gastroenterology and Hepatology*, 6(2), 133-148. doi: 10.1016/j.jcmgh.2018.04.003
- Matta, S. M., Hill-Yardin, E. L., & Crack, P. J. (2019). The influence of neuroinflammation in Autism Spectrum Disorder. *Brain, Behavior, and Immunity*, 79, 75-90. doi: 10.1016/j.bbi.2019.04.037
- Matthiesen, R., & Mutenda, K. E. (2007). Introduction to proteomics. *Methods in Molecular Biology (Clifton, N.J.)*, 367, 1-35. doi: 10.1385/1-59745-275-0:1
- McGrath, J., Saha, S., Chant, D., & Welham, J. (2008). Schizophrenia: A concise overview of incidence, prevalence, and mortality. *Epidemiologic Reviews*, 30, 67-76. doi: 10.1093/epirev/mxn001
- McWilliams, M. M., & Chennathukuzhi, V. M. (2017). Recent Advances in Uterine Fibroid Etiology. *Seminars in reproductive medicine*, 35(2), 181-189. doi: 10.1055/s-0037-1599090
- Meltzer, A., & Van de Water, J. (2017). The Role of the Immune System in Autism Spectrum Disorder. *Neuropsychopharmacology: Official Publication of the American College of Neuropsychopharmacology*, 42(1), 284-298. doi: 10.1038/npp.2016.158
- Miller, J. N., & Black, D. W. (2020). Bipolar Disorder and Suicide: A Review. *Current Psychiatry Reports*, 22(2), 6. doi: 10.1007/s11920-020-1130-0
- Moradi, Y., Shams-Beyranvand, M., Khateri, S., Gharahjeh, S., Tehrani, S., Varse, F., Tiyuri, A., & Najmi, Z. (2021). A systematic review on the prevalence of endometriosis in women. *The Indian Journal of Medical Research*, 154(3), 446-454. doi: 10.4103/ijmr.IJMR\_817\_18
- Morera-Fumero, A. L., & Abreu-Gonzalez, P. (2013). Role of Melatonin in Schizophrenia. *International Journal of Molecular Sciences*, 14(5), 9037-9050. doi: 10.3390/ijms14059037
- Moretti, S., Patrone, F., & Bonassi, S. (2007). The class of microarray games and the relevance index for genes. *TOP: An Official Journal of the Spanish Society of Statistics and Operations Research*. 15. 256-280. doi: 10.1007/s11750-007-0021-4.
- Mouridsen, S. E., Rich, B., & Isager, T. (2016). Diseases of the circulatory system among adult people diagnosed with infantile autism as children: A longitudinal case

control study. *Research in Developmental Disabilities*, 57, 193-200. doi: 10.1016/j.ridd.2016.07.002

Nikolac Perkovic, M., Nedic Erjavec, G., Svob Strac, D., Uzun, S., Kozumplik, O., & Pivac, N. (2017). Theranostic Biomarkers for Schizophrenia. *International Journal of Molecular Sciences*, 18(4), 733. doi: 10.3390/ijms18040733

Ning, L.F., Yu, Y.Q., Guo Ji, E.T., et al. Meta-analysis of differentially expressed genes in autism based on gene expression data. *Genet Mol Res*. 2015;14(1):2146-2155. Published 2015 Mar 27. doi: 10.4238/2015.March.27.1

Olivier, M., Asmis, R., Hawkins, G. A., Howard, T. D., & Cox, L. A. (2019). The Need for Multi-Omics Biomarker Signatures in Precision Medicine. *International Journal of Molecular Sciences*, 20(19), 4781. doi: 10.3390/ijms20194781

Onstad, M. A., Schmandt, R. E., & Lu, K. H. (2016). Addressing the Role of Obesity in Endometrial Cancer Risk, Prevention, and Treatment. *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology*, 34(35), 4225-4230. doi: 10.1200/JCO.2016.69.4638

Organización Mundial de la Salud, 2023

Ormstad, H., Bryn, V., Saugstad, O. D., Skjeldal, O., & Maes, M. (2018). Role of the Immune System in Autism Spectrum Disorders (ASD). *CNS & Neurological Disorders Drug Targets*, 17(7), 489-495. doi: 10.2174/1871527317666180706123229

Orsolini, L., Pompili, S., Tempia Valenta, S., Salvi, V., & Volpe, U. (2022). C-Reactive Protein as a Biomarker for Major Depressive Disorder? *International Journal of Molecular Sciences*, 23(3), 1616. doi: 10.3390/ijms23031616

Pavlopoulos, G. A., Malliarakis, D., Papanikolaou, N., Theodosiou, T., Enright, A. J., & Iliopoulos, I. (2015). Visualizing genome and systems biology: Technologies, tools, implementation techniques and trends, past, present and future. *GigaScience*, 4, 38. doi: 10.1186/s13742-015-0077-2

Peng, S., Li, W., Lv, L., Zhang, Z., & Zhan, X. (2018). BDNF as a Biomarker in Diagnosis and Evaluation of Treatment for Schizophrenia and Depression. *Discovery Medicine*, 26(143), 127-136.

- Perlick, D. A., Rosenheck, R. A., Kaczynski, R., Swartz, M. S., Canive, J. M., & Lieberman, J. A. (2010). Impact of antipsychotic medication on family burden in schizophrenia: Longitudinal results of CATIE trial. *Schizophrenia Research*, *116*(2-3), 118-125. doi: 10.1016/j.schres.2009.09.026
- Pino, O., Guilera, G., Gómez-Benito, J., Najas-García, A., Rufián, S., & Rojo, E. (2014). Neurodevelopment or neurodegeneration: Review of theories of schizophrenia. *Actas Espanolas De Psiquiatria*, *42*(4), 185-195.
- Qi, D., & Chen, K. (2021). Bioinformatics Analysis of Potential Biomarkers and Pathway Identification for Major Depressive Disorder. *Computational and Mathematical Methods in Medicine*, *2021*, 3036741. doi: 10.1155/2021/3036741
- Qiu, S., Qiu, Y., Li, Y., & Cong, X. (2022). Genetics of autism spectrum disorder: An umbrella review of systematic reviews and meta-analyses. *Translational Psychiatry*, *12*(1), 249. doi: 10.1038/s41398-022-02009-6
- Raja, M. H. R., Farooqui, N., Zuberi, N., Ashraf, M., Azhar, A., Baig, R., Badar, B., & Rehman, R. (2021). Endometriosis, infertility and MicroRNA's: A review. *Journal of Gynecology Obstetrics and Human Reproduction*, *50*(9), 102157. doi: 10.1016/j.jogoh.2021.102157
- Rajkowska, G., & Stockmeier, C. A. (2013). ASTROCYTE PATHOLOGY IN MAJOR DEPRESSIVE DISORDER: INSIGHTS FROM HUMAN POSTMORTEM BRAIN TISSUE. *Current drug targets*, *14*(11), 1225-1236.
- Ratajczak, H. V. (2011). Theoretical aspects of autism: Biomarkers—a review. *Journal of Immunotoxicology*, *8*(1), 80-94. doi: 10.3109/1547691X.2010.538749
- Raudvere, U., Kolberg, L., Kuzmin, I., Arak, T., Adler, P., Peterson, H., & Vilo, J. (2019). g:Profiler: A web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Research*, *47*(W1), W191-W198. doi: 10.1093/nar/gkz369
- Ribé, J. M., Salamero, M., Pérez-Testor, C., Mercadal, J., Aguilera, C., & Cleris, M. (2018). Quality of life in family caregivers of schizophrenia patients in Spain: Caregiver characteristics, caregiving burden, family functioning, and social and professional support. *International Journal of Psychiatry in Clinical Practice*, *22*(1), 25-33. doi: 10.1080/13651501.2017.1360500
- Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W., Smyth, G.K. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 2015 Apr 20;43(7):e47. doi: 10.1093/nar/gkv007.

- Robinson-Agramonte, M. de L. A., Noris García, E., Fraga Guerra, J., Vega Hurtado, Y., Antonucci, N., Semprún-Hernández, N., Schultz, S., & Siniscalco, D. (2022). Immune Dysregulation in Autism Spectrum Disorder: What Do We Know about It? *International Journal of Molecular Sciences*, *23*(6), 3033. doi: 10.3390/ijms23063033
- Ruzzo, E. K., Pérez-Cano, L., Jung, J.-Y., Wang, L.-K., Kashef-Haghighi, D., Hartl, C., Singh, C., Xu, J., Hoekstra, J. N., Leventhal, O., Leppä, V. M., Gandal, M. J., Paskov, K., Stockham, N., Polioudakis, D., Lowe, J. K., Prober, D. A., Geschwind, D. H., & Wall, D. P. (2019). Inherited and De Novo Genetic Risk for Autism Impacts Shared Networks. *Cell*, *178*(4), 850-866.e26. doi: 10.1016/j.cell.2019.07.015
- Saha, S., Chant, D., Welham, J., & McGrath, J. (2005). A systematic review of the prevalence of schizophrenia. *PLoS Medicine*, *2*(5), e141. doi: 10.1371/journal.pmed.0020141
- Salagre, E., Vizuete, A. F., Leite, M., Brownstein, D. J., McGuinness, A., Jacka, F., Dodd, S., Stubbs, B., Köhler, C. A., Vieta, E., Carvalho, A. F., Berk, M., & Fernandes, B. S. (2017). Homocysteine as a peripheral biomarker in bipolar disorder: A meta-analysis. *European Psychiatry*, *43*, 81-91. doi: 10.1016/j.eurpsy.2017.02.482
- Saunders, E. F. H., Ramsden, C. E., Sherazy, M. S., Gelenberg, A. J., Davis, J. M., & Rapoport, S. I. (2016). Omega-3 and Omega-6 Polyunsaturated Fatty Acids in Bipolar Disorder. *The Journal of clinical psychiatry*, *77*(10), e1301-e1308. doi: 10.4088/JCP.15r09925
- Saunders, P. T. K., & Horne, A. W. (2021). Endometriosis: Etiology, pathobiology, and therapeutic prospects. *Cell*, *184*(11), 2807-2824. doi: 10.1016/j.cell.2021.04.041
- Sekar, A., Bialas, A. R., de Rivera, H., Davis, A., Hammond, T. R., Kamitaki, N., Tooley, K., Presumey, J., Baum, M., Van Doren, V., Genovese, G., Rose, S. A., Handsaker, R. E., Schizophrenia Working Group of the Psychiatric Genomics Consortium, Daly, M. J., Carroll, M. C., Stevens, B., & McCarroll, S. A. (2016). Schizophrenia risk from complex variation of complement component 4. *Nature*, *530*(7589), 177-183. doi: 10.1038/nature16549
- Serra, G., De Crescenzo, F., Maisto, F., Galante, J. R., Iannoni, M. E., Trasolini, M., Maglio, G., Tondo, L., Baldessarini, R. J., & Vicari, S. (2022). Suicidal behavior in juvenile bipolar disorder and major depressive disorder patients: Systematic

- review and meta-analysis. *Journal of Affective Disorders*, 311, 572-581. doi: 10.1016/j.jad.2022.05.063
- Sharma, S. R., Gonda, X., & Tarazi, F. I. (2018). Autism Spectrum Disorder: Classification, diagnosis and therapy. *Pharmacology & Therapeutics*, 190, 91-104. doi: 10.1016/j.pharmthera.2018.05.007
- Sharon, G., Sampson, T. R., Geschwind, D. H., & Mazmanian, S. K. (2016). The Central Nervous System and the Gut Microbiome. *Cell*, 167(4), 915-932. doi: 10.1016/j.cell.2016.10.027
- Sher, L., & Kahn, R. S. (2019). Suicide in Schizophrenia: An Educational Overview. *Medicina (Kaunas, Lithuania)*, 55(7), 361. doi: 10.3390/medicina55070361
- Sigitova, E., Fišar, Z., Hroudová, J., Cikánková, T., & Raboch, J. (2017). Biological hypotheses and biomarkers of bipolar disorder. *Psychiatry and Clinical Neurosciences*, 71(2), 77-103. doi: 10.1111/pcn.12476
- Silverman, J. L., Yang, M., Lord, C., & Crawley, J. N. (2010). Behavioural phenotyping assays for mouse models of autism. *Nature Reviews. Neuroscience*, 11(7), 490-502. doi: /10.1038/nrn2851
- Smith, D. J., Whitham, E. A., & Ghaemi, S. N. (2012). Bipolar disorder. *Handbook of Clinical Neurology*, 106, 251-263. doi: 10.1016/B978-0-444-52002-9.00015-2
- Smith, J.M. (1976). *American Scientist*, Vol. 64, No. 1 (January-February 1976), pp. 41-45 (5 pages).
- Smith, R. C., Osborn, G. G., Dwamena, F. C., D’Mello, D., Freilich, L., & Laird-Fick, H. S. (2019). Major Depression and Related Disorders. En *Essentials of Psychiatry in Primary Care: Behavioral Health in the Medical Setting*. McGraw-Hill Education.
- Srikantha, P., & Mohajeri, M. H. (2019). The Possible Role of the Microbiota-Gut-Brain-Axis in Autism Spectrum Disorder. *International Journal of Molecular Sciences*, 20(9), 2115. doi: 10.3390/ijms20092115
- Steinman, G. (2020). The putative etiology and prevention of autism. *Progress in Molecular Biology and Translational Science*, 173, 1-34. doi: 10.1016/bs.pmbts.2020.04.013
- Stilo, S. A., & Murray, R. M. (2019). Non-Genetic Factors in Schizophrenia. *Current Psychiatry Reports*, 21(10), 100. doi: 10.1007/s11920-019-1091-3

- Suda, K., & Matsuda, K. (2022). How Microbes Affect Depression: Underlying Mechanisms via the Gut–Brain Axis and the Modulating Role of Probiotics. *International Journal of Molecular Sciences*, 23(3), 1172. doi: 10.3390/ijms23031172
- Sullivan, P. F., Daly, M. J., & O'Donovan, M. (2012). Genetic Architectures of Psychiatric Disorders: The Emerging Picture and Its Implications. *Nature reviews. Genetics*, 13(8), 537-551. doi: 10.1038/nrg3240
- Takenaka, K., Curry-Hyde, A., Olzomer, E.M., Farrell, R., Byrne, F.L., Janitz, M. Investigation of Transcriptome Patterns in Endometrial Cancers from Obese and Lean Women. *Int J Mol Sci*. 2022;23(19):11471. doi:10.3390/ijms231911471
- Tang, H., Therneau, T.M. Statistical metrics for quality assessment of high-density tiling array data. *Biometrics*. 2010 Jun;66(2):630-5. doi: 10.1111/j.1541-0420.2009.01298.x
- Taylor, H. S., Kotlyar, A. M., & Flores, V. A. (2021). Endometriosis is a chronic systemic disease: Clinical challenges and novel innovations. *Lancet (London, England)*, 397(10276), 839-852. doi: 10.1016/S0140-6736(21)00389-5
- Tolosa, A. (2020). Genotipia ([https://genotipia.com/genetica\\_medica\\_news/secuenciacion-de-genomas-humanos/](https://genotipia.com/genetica_medica_news/secuenciacion-de-genomas-humanos/))
- Tondo, L., Vázquez, G. H., & Baldessarini, R. J. (2017). Depression and Mania in Bipolar Disorder. *Current Neuropharmacology*, 15(3), 353-358. doi: 10.2174/1570159X14666160606210811
- Trubetskoy, V., Pardiñas, A. F., Qi, T., Panagiotaropoulou, G., Awasthi, S., Bigdeli, T. B., Bryois, J., Chen, C.-Y., Dennison, C. A., Hall, L. S., Lam, M., Watanabe, K., Frei, O., Ge, T., Harwood, J. C., Koopmans, F., Magnusson, S., Richards, A. L., Sidorenko, J., ... Schizophrenia Working Group of the Psychiatric Genomics Consortium. (2022). Mapping genomic loci implicates genes and synaptic biology in schizophrenia. *Nature*, 604(7906), 502-508. doi: 10.1038/s41586-022-04434-5
- Tseng, C.-E. J., McDougle, C. J., Hooker, J. M., & Zürcher, N. R. (2022). Epigenetics of Autism Spectrum Disorder: Histone Deacetylases. *Biological Psychiatry*, 91(11), 922-933. doi: 10.1016/j.biopsych.2021.11.021
- Uddin, L. Q., Dajani, D. R., Voorhies, W., Bednarz, H., & Kana, R. K. (2017). Progress and roadblocks in the search for brain-based biomarkers of autism and attention-

- deficit/hyperactivity disorder. *Translational Psychiatry*, 7(8), e1218. doi: 10.1038/tp.2017.164
- Valdés-Tovar, M., Rodríguez-Ramírez, A. M., Rodríguez-Cárdenas, L., Sotelo-Ramírez, C. E., Camarena, B., Sanabrais-Jiménez, M. A., Solís-Chagoyán, H., Argueta, J., & López-Riquelme, G. O. (2022). Insights into myelin dysfunction in schizophrenia and bipolar disorder. *World Journal of Psychiatry*, 12(2), 264-285. doi: 10.5498/wjp.v12.i2.264
- Veenstra-VanderWeele, J., & Blakely, R. D. (2012). Networking in Autism: Leveraging Genetic, Biomarker and Model System Findings in the Search for New Treatments. *Neuropsychopharmacology*, 37(1), 196-212. doi: 10.1038/npp.2011.185
- Vieta, E., Salagre, E., Grande, I., Carvalho, A. F., Fernandes, B. S., Berk, M., Birmaher, B., Tohen, M., & Suppes, T. (2018). Early Intervention in Bipolar Disorder. *American Journal of Psychiatry*, 175(5), 411-426. doi: 10.1176/appi.ajp.2017.17090972
- Vilain, J., Galliot, A.-M., Durand-Roger, J., Leboyer, M., Llorca, P.-M., Schürhoff, F., & Szöke, A. (2013). [Environmental risk factors for schizophrenia: A review]. *L'Encephale*, 39(1), 19-28. doi: 10.1016/j.encep.2011.12.007
- Voineagu, I. (2012). Autism: From Genetics to Biomarkers. *Disease markers*, 33(5), 223-224. doi: 10.3233/DMA-2012-0931
- Walsh, C. S., Hacker, K. E., Secord, A. A., DeLair, D. F., McCourt, C., & Urban, R. (2023). Molecular testing for endometrial cancer: An SGO clinical practice statement. *Gynecologic Oncology*, 168, 48-55. doi: 10.1016/j.ygyno.2022.10.024
- Wang, K. C., & Chang, H. Y. (2018). Epigenomics: Technologies and Applications. *Circulation Research*, 122(9), 1191-1199. doi: 10.1161/CIRCRESAHA.118.310998
- Wang, L., & Sun, J. (2022). ASPN Is a Potential Biomarker and Associated with Immune Infiltration in Endometriosis. *Genes*, 13(8). doi: 10.3390/genes13081352
- Wilson, H. (2014). Potential biomarkers of autism. *Biomarkers in Medicine*, Vol. 8, No. 3. Theme: Potential biomarkers of Autism – Foreword
- Wittenberg, G.M., Greene, J., Vértes, P.E., Drevets, W.C., Bullmore, E.T. (2020). Major Depressive Disorder Is Associated With Differential Expression of Innate Immune and Neutrophil-Related Gene Networks in Peripheral Blood: A

- Quantitative Review of Whole-Genome Transcriptional Data From Case-Control Studies [published correction appears in *Biol Psychiatry*. 2020 Oct 15;88(8):668. doi: 10.1016/j.biopsych.2020.07.007]. *Biol Psychiatry*. 88(8):625-637. doi: 10.1016/j.biopsych.2020.05.006
- Wu, Y., Sun, W., Liu, H., & Zhang, D. (2019). Age at Menopause and Risk of Developing Endometrial Cancer: A Meta-Analysis. *BioMed Research International*, 2019, 8584130. doi: 10.1155/2019/8584130
- Xu, H., Li, R., Wang, L., et al. (2023). Non-enzymatic antioxidants, macro-minerals and monocyte/high-density lipoprotein cholesterol ratio among patients with bipolar disorder. *J Affect Disord*. 322:76-83. doi: 10.1016/j.jad.2022.11.017
- Yamada, Y., Matsumoto, M., Iijima, K., & Sumiyoshi, T. (2020). Specificity and Continuity of Schizophrenia and Bipolar Disorder: Relation to Biomarkers. *Current Pharmaceutical Design*, 26(2), 191-200. doi: 10.2174/1381612825666191216153508
- Yang, C.-J., Liu, C.-L., Sang, B., Zhu, X.-M., & Du, Y.-J. (2015). The combined role of serotonin and interleukin-6 as biomarker for autism. *Neuroscience*, 284, 290-296. doi: 10.1016/j.neuroscience.2014.10.011
- Zhang, Q., Kanis, M. J., Ubago, J., Liu, D., Scholtens, D. M., Strohl, A. E., Lurain, J. R., Shahabi, S., Kong, B., & Wei, J.-J. (2018). The selected biomarker analysis in five types of uterine smooth muscle tumors. *Human pathology*, 76, 17-27. doi: 10.1016/j.humpath.2017.12.005
- Zhou, L., Wang, W., Wang, F., Yang, S., Hu, J., Lu, B., Pan, Z., Ma, Y., Zheng, M., Zhou, L., Lei, S., Song, P., Liu, P., Lu, W., & Lu, Y. (2021). Plasma-derived exosomal miR-15a-5p as a promising diagnostic biomarker for early detection of endometrial carcinoma. *Molecular Cancer*, 20, 57. doi: 10.1186/s12943-021-01352-4
- Zhou, X., Liu, L., Lan, X., Cohen, D., Zhang, Y., Ravindran, A. V., Yuan, S., Zheng, P., Coghill, D., Yang, L., Hetrick, S. E., Jiang, X., Benoliel, J.-J., Cipriani, A., & Xie, P. (2019). Polyunsaturated fatty acids metabolism, purine metabolism and inosine as potential independent diagnostic biomarkers for major depressive disorder in children and adolescents. *Molecular Psychiatry*, 24(10), 1478-1488. doi: 10.1038/s41380-018-0047-z

Zhuo, Z., Wang, C., & Yu, H. (2022). Plasma microRNAs can be a potential diagnostic biomarker for endometriosis. *Ginekologia Polska*, 93(6), Article 6. doi: 10.5603/GP.a2021.0127

**CAPÍTULO I:**  
**BIOLOGÍA DE SISTEMAS Y SU APLICACIÓN**  
**EN EL ESTUDIO DE ENFERMEDADES**  
**COMPLEJAS**

## 1. Referencia

Vargas E, Castro-Martínez JA, Díaz-Beltrán L, Esteban FJ. La Biología de Sistemas y su aplicación al estudio de enfermedades complejas. *Seminario Médico* 2024, 64(1): 61-71.

## 2. Breve resumen

En la era contemporánea de la investigación biomédica, el imparable avance de las tecnologías ómicas y la biología de sistemas ha desencadenado una transformación fundamental en nuestra comprensión de las enfermedades complejas. Las tecnologías ómicas, que abarcan disciplinas como la genómica, transcriptómica, epigenómica, proteómica, metabolómica y microbiómica, han proporcionado un acceso sin precedentes a la intrincada red molecular que subyace a fenómenos fisiológicos y patológicos. En este contexto, el presente trabajo se propone explorar y describir el impacto que estas herramientas han supuesto en el avance sobre el conocimiento de enfermedades complejas. Desde la capacidad de analizar y detectar múltiples biomarcadores simultáneamente hasta la integración holística de datos mediante la biología de sistemas, estas tecnologías emergentes han allanado el camino para revelar los entresijos de enfermedades hasta ahora consideradas enigmáticas, ofreciendo perspectivas prometedoras para el desarrollo de enfoques diagnósticos y terapéuticos más precisos y eficaces.

3. **Fecha de publicación:** 2024

4. **Estado:** publicado

5. **Tipo de publicación:** artículo

6. **Categoría:** revista nacional

7. **Ubicación:** Instituto de Estudios Giennenses

**CAPÍTULO II:**

**ANÁLISIS COMPARATIVO DEL VALOR DE  
SHAPLEY PARA EL ESTUDIO DE PATOLOGÍAS  
UTERINAS**

## 1. Referencia

Castro-Martínez JA, Vargas E, Díaz-Beltrán L, Esteban FJ. Comparative Analysis of Shapley value enhances transcriptomics insights into some common uterine pathologies. *Genes* 2024; 15: 723. doi: <https://doi.org/10.3390/genes15060723>

## 2. Breve resumen

Las patologías uterinas representan un reto para la salud de la mujer a escala global. A pesar de la investigación en el campo, la causa y origen de algunas de las patologías que afectan al útero no se encuentran bien definidas aún. Este estudio presenta un análisis exhaustivo de datos transcriptómicos procedentes de diferentes datasets relativos a patologías uterinas tan relevantes como la endometriosis, el cáncer de endometrio o los leiomiomas uterinos. A través del uso del Análisis Comparativo del valor de Shapley (CASH, por sus siglas en inglés), hemos demostrado la eficacia de este método en la mejora de los resultados obtenidos respecto a los análisis de expresión diferencial clásicos. Los resultados obtenidos arrojan luz sobre los patrones moleculares subyacentes a estos desórdenes ginecológicos, y apuntan a CASH como una herramienta valiosa para la mejora de la precisión de los análisis transcriptómicos en contextos biológicos complejos. Esta investigación contribuye a una mejor comprensión de los patrones de expresión génica y potenciales biomarcadores asociados con estas patologías, ofreciendo implicaciones para futuras estrategias diagnósticas y terapéuticas.

3. **Fecha de publicación:** 2024

4. **Estado:** publicado

5. **Tipo de publicación:** artículo

6. **Categoría:** revista internacional

7. **Ubicación:** Genes (Basilea, Suiza)

**CAPÍTULO III:**  
**PROFUNDIZANDO EN LA TRANSCRIPTÓMICA**  
**DE DESÓRDENES NEUROLÓGICOS A TRAVÉS**  
**DEL USO DEL VALOR DE SHAPLEY**

## 1. Referencia

Castro-Martínez JA, Vargas E, Díaz-Beltrán L, Esteban FJ. Enhancing transcriptomics insights into neurological disorders through Comparative Analysis of Shapley values. *EMBnet.journal*, manuscrito preparado para envío.

## 2. Breve resumen

Las patologías neurológicas afectan a millones de personas en todo el mundo, a pesar de que las causas y orígenes de muchos trastornos comunes permanecen ocultos. En este estudio se realiza un análisis en profundidad de datos transcriptómicos en una variedad de condiciones neuropsiquiátricas, entre las que se incluye el trastorno del espectro autista, la esquizofrenia, el trastorno bipolar y el trastorno depresivo mayor. Este trabajo demuestra la superioridad del método basado en Teoría de Juegos Cooperativos denominado Análisis Comparativo del valor de Shapley (CASH) en la detección de genes diferencialmente expresados cuando se compara con metodologías tradicionales basadas en el uso del test de Welch y el test empírico de Bayes. Nuestros resultados revelan firmas moleculares críticas asociadas con los trastornos neurológicos objeto de estudio, y señalan a CASH como una herramienta poderosa para mejorar la precisión de los análisis transcriptómicos en escenarios biológicos complejos. Esta investigación permite avanzar en la comprensión de patrones de expresión génica asociados a determinadas patologías e identifica posibles biomarcadores, con importantes implicaciones para el desarrollo de estrategias diagnósticas y terapéuticas más precisas.

## 3. Fecha de publicación: -

4. **Estado:** manuscrito preparado para su envío

5. **Tipo de publicación:** artículo

6. **Categoría:** revista internacional

7. **Ubicación:** EMBnet.journal (Uppsala, Suecia)