

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

ESCOLA TÈCNICA SUPERIOR D'ENGINYERIA
AGRONÒMICA I DEL MEDI NATURAL
(ETSIAMN)

GRADO EN BIOTECNOLOGÍA



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA

Caracterización de las diferencias de sexo en las enfermedades
cerebrovasculares mediante el metaanálisis de estudios
transcriptómicos.

AUTOR: YASSIR EL HAMDANI EL AMRI

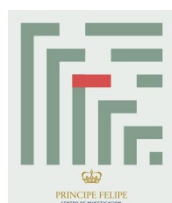
TUTOR ACADÉMICO: PROF D. JAVIER FORMENT MILLET

TUTOR EXTERNO: D. FRANCISCO GARCÍA GARCÍA

COTUTOR EXTERNO: D. SERGIO ROMERA GINER

CURSO ACADÉMICO: 2020 – 2021

VALENCIA, JULIO DE 2021



Unidad de
Bioinformática y
Bioestadística

Centro de Investigación
Príncipe Felipe CIPF



CARACTERIZACIÓN DE LAS DIFERENCIAS DE SEXO EN LAS ENFERMEDADES CEREBROVASCULARES MEDIANTE EL METAANÁLISIS DE ESTUDIOS TRANSCRIPTÓMICOS.

Resumen

Los accidentes cerebrovasculares son la consecuencia de la interrupción del suministro del flujo sanguíneo a una parte del cerebro, desencadenando una hipoxia cerebral y, por consiguiente, la muerte neuronal. Según la naturaleza del accidente, se pueden distinguir dos tipos principales: accidente cerebrovascular isquémico y accidente cerebrovascular hemorrágico. Se trata de una de las enfermedades con mayores índices de mortalidad y morbilidad a nivel mundial.

Esta enfermedad presenta diferencias de prevalencia e incidencia entre mujeres y hombres, siendo la primera causa de muerte para las mujeres a nivel nacional. Esta divergencia en las cifras epidemiológicas es la consecuencia de diferencias en los mecanismos moleculares, cuyo análisis permitiría un mejor diagnóstico y tratamiento de los accidentes. Desafortunadamente, en la mayoría de los estudios biomédicos sobre esta patología, no se tiene en cuenta el sexo como factor clave en la investigación.

La elucidación de los mecanismos moleculares responsables de las diferencias existentes entre sexos permite hallar nuevos biomarcadores que mejoran la precisión y la eficacia en el tratamiento de la enfermedad, el diagnóstico y el pronóstico clínicos.

Con el fin de cumplir este objetivo, se ha llevado a cabo un metaanálisis con datos provenientes de estudios transcriptómicos seleccionados a partir de una revisión bibliográfica sistemática. Con este abordaje se identificaron funciones específicas en mujeres, relacionadas con la integridad del ribosoma y la mitocondria, y la fosforilación oxidativa; en el caso de los hombres, se encontraron funciones específicas vinculadas a la homeostasis de la barrera hematoencefálica, la respuesta inmune y la guía axonal. No obstante, se requiere de un mayor número de estudios para profundizar en estos hallazgos, además de implementar el sexo como factor clave en el diseño experimental de nuevos trabajos, para comprender mejor las discrepancias existentes entre sexos respecto a esta enfermedad.

Palabras clave: Accidentes cerebrovasculares; sexo; transcriptómica; metaanálisis; medicina personalizada.

Autor: Yassir El Hamdani El Amri

Tutor UPV: D. Javier Forment Millet

Tutores CIFP: D. Francisco García García, D. Sergio Romera Giner

Valencia, julio de 2021



CHARACTERIZATION OF SEX DIFFERENCES IN CEREBROVASCULAR ACCIDENTS THROUGH META-ANALYSIS OF TRANSCRIPTOMIC STUDIES.

Abstract

Strokes are the consequence of the interruption of the supply of blood flow to a part of the brain, triggering cerebral hypoxia and, consequently, neuronal death. Depending on the nature of the accident, two main types can be distinguished: ischemic stroke and hemorrhagic stroke. It is one of the diseases with the highest mortality and morbidity rates worldwide.

This disease presents differences in prevalence and incidence between women and men, being the leading cause of death for women nationwide. This divergence in epidemiological figures is the consequence of differences in molecular mechanisms, the analysis of which would allow a better diagnosis and treatment of accidents. Unfortunately, in most biomedical studies on this pathology, sex is not considered as a key factor in the research.

The elucidation of the molecular mechanisms responsible for the existing differences between the sexes allows finding new biomarkers that improve precision and efficacy in the treatment of the disease, the diagnosis and the clinical prognosis.

In order to meet this objective, a meta-analysis has been carried out with data from transcriptomic studies selected from a bibliographic systematic review. With this approach, specific functions were identified in women, related to the integrity of the ribosome and mitochondria, and oxidative phosphorylation; in the case of men, specific functions related to blood-brain barrier homeostasis, immune response and axonal guidance were found. However, a greater number of studies is required to delve into these findings, in addition to implementing sex as a key factor in the experimental design of new studies, to better understand the existing discrepancies between the sexes regarding this disease.

Keywords: Cerebrovascular accidents; sex; transcriptomics; meta-analysis; personalised medicine.

Author: Yassir El Hamdani El Amri

UPV tutor: D. Javier Forment Millet

CIFP tutors: D. Francisco García García, D. Sergio Romera Giner

Valencia, July 2021



ÍNDICE GENERAL

1. INTRODUCCIÓN	1
1.1. Contexto de la enfermedades neurodegenerativas y la implicación del sexo	1
1.2. Accidente cerebrovascular	1
1.2.2. Epidemiología	3
1.2.3. Genética de la enfermedad	6
1.2.4. Factores de riesgo	7
1.2.5. Tratamientos actuales	8
1.3. Datos ómicos y tecnologías de alto rendimiento	8
1.3.1. Transcriptómica	9
1.3.2. Estrategias de análisis transcriptómicos	10
1.3.3. Caracterización funcional	11
1.3.4. Metaanálisis funcional	11
2. OBJETIVOS	13
3. MATERIALES Y MÉTODOS	14
3.1. Revisión sistemática y selección de estudios transcriptómicos	14
3.2. Análisis individualizado de los estudios	15
3.2.1. Procesamiento de los datos	15
3.2.2. Análisis exploratorio	16
3.2.3. Análisis de expresión diferencial	17
3.2.4. Análisis de enriquecimiento funcional	17
3.3. Metaanálisis	18
3.3.1. Determinación de la medida combinada del efecto	18
3.3.2. Representación de los resultados	19
4. RESULTADOS	20
4.1. Revisión sistemática y selección de estudios	20
4.2. Análisis individual de los estudios	21
4.2.1. Análisis exploratorio y procesamiento de los datos	21
4.2.2. Análisis de expresión diferencial	25
4.2.3. Análisis de enriquecimiento funcional	25
4.3. Metaanálisis funcional	26
5. DISCUSIÓN	32

6. CONCLUSIONES	37
7. BIBLIOGRAFÍA.....	38
8. ANEXOS.....	I
8.1. Anexo de figuras.....	I

ÍNDICE DE FIGURAS

- Figura 1.** Ilustraciones de las dos manifestaciones más comunes del accidente cerebrovascular. A la izquierda se representa el ictus isquémico, mientras que a la derecha se encuentra el ictus hemorrágico (*American Heart Association, 2021*). 3
- Figura 2.** Epidemiología de ictus isquémico. Mapa ilustrativo de la distribución de la incidencia del ictus isquémico a partir de los datos de *Global Burden of Disease* de 2017. 4
- Figura 3.** La evolución del número de altas hospitalarias por ictus a nivel nacional en los últimos años. Fuente: Ministerio de Sanidad, Consumo y Bienestar Social. 5
- Figura 5.** Diagrama de flujo que indica los diferentes pasos a seguir para efectuar una revisión sistemática (esquema adaptado de Liberati *et al.*, 2009).....14
- Figura 6.** Ejemplo de gráfico de bosque para el término GO:0005743 (*mitochondrial inner membrane*), obtenido en el metaanálisis de este trabajo.19
- Figura 7.** Diagrama de flujo de la revisión sistemática siguiendo los pasos de la declaración PRISMA (*Preferred Reporting Items for Systematic Reviews and Meta-Analyses*).20
- Figura 8.** Diagrama de cajas de los valores de expresión de las muestras empleadas en el estudio GSE13353. Cada color hace referencia a un grupo experimental distinto.22
- Figura 9.** Análisis de componentes principales (PCA) de las muestras del estudio GSE13353. Cada color hace referencia a un grupo experimental distinto.....22
- Figura 10.** Análisis de agrupamiento de las muestras del estudio GSE13353. Cada color hace referencia a un grupo experimental distinto.23
- Figura 11.** Gráfico de bosque para la función GO:0048821 (*erythrocyte development*). En la parte izquierda se indican los estudios incluidos en el metaanálisis para esta función, mientras que a la derecha la estimación de la medida individual de cada uno con su respectivo intervalo de confianza al 95%. El eje discontinuo vertical indica la posición del LOR cuando su valor es 0. 27
- Figura 12.** Gráfico de bosque para la función GO:0003735 (*structural constituent of ribosome*). En la parte izquierda se indican los estudios incluidos en el metaanálisis para esta función, mientras que a la derecha la estimación de la medida individual de cada uno con su respectivo intervalo de confianza al 95%. El eje discontinuo vertical indica la posición del LOR cuando su valor es 0.28
- Figura 13.** Gráfico de bosque para la función GO:0005178 (*integrin binding*). En la parte izquierda se indican los estudios incluidos en el metaanálisis para esta función, mientras que a la derecha la estimación de la medida individual de cada uno con su respectivo intervalo de confianza al 95%. El eje discontinuo vertical indica la posición del LOR cuando su valor es 0. 28

Figura 14. Gráfico de bosque para la función GO:0062023 (*collagen-containing extracellular matrix*). En la parte izquierda se indican los estudios incluidos en el metaanálisis para esta función, mientras que a la derecha la estimación de la medida individual de cada uno con su respectivo intervalo de confianza al 95%. El eje discontinuo vertical indica la posición del LOR cuando su valor es 0.....30

Figura 15. Gráfico de bosque para la función GO:0031464 (*Cul4A-RING E3 ubiquitin ligase complex*). En la parte izquierda se indican los estudios incluidos en el metaanálisis para esta función, mientras que a la derecha la estimación de la medida individual de cada uno con su respectivo intervalo de confianza al 95%. El eje discontinuo vertical indica la posición del LOR cuando su valor es 0.....30

Figura A.1. Diagrama de cajas de los valores de expresión de las muestras empleadas en el estudio GSE16561. Cada color hace referencia a un grupo experimental distinto.....I

Figura A.2. Análisis de componentes principales (PCA) de las muestras del estudio GSE16561. Cada color hace referencia a un grupo experimental distinto.....I

Figura A.4. Diagrama de cajas de los valores de expresión de las muestras empleadas en el estudio GSE22255. Cada color hace referencia a un grupo experimental distinto.....II

Figura A.5. Análisis de componentes principales (PCA) de las muestras del estudio GSE22255. Cada color hace referencia a un grupo experimental distinto.....II

Figura A.6. Análisis de agrupamiento de las muestras del estudio GSE22255. Cada color hace referencia a un grupo experimental distinto.....II

Figura A.7. Diagrama de cajas de los valores de expresión de las muestras empleadas en el estudio GSE58294. Cada color hace referencia a un grupo experimental distinto..... III

Figura A.8. Análisis de componentes principales (PCA) de las muestras del estudio GSE58294. Cada color hace referencia a un grupo experimental distinto..... III

Figura A.9. Análisis de agrupamiento de las muestras del estudio GSE58294. Cada color hace referencia a un grupo experimental distinto..... III

Figura A.10. Diagrama de cajas de los valores de expresión de las muestras empleadas en el estudio GSE66724. Cada color hace referencia a un grupo experimental distinto..... IV

Figura A.11. Análisis de componentes principales (PCA) de las muestras del estudio GSE66724. Cada color hace referencia a un grupo experimental distinto..... IV

Figura A.12. Análisis de agrupamiento de las muestras del estudio GSE66724. Cada color hace referencia a un grupo experimental distinto..... IV

ÍNDICE DE TABLAS

Tabla 1. Número de defunciones por ictus según sexo (filas) y tramo de edad (columnas) en 2017 (INE, 2017).	5
Tabla 2. Desórdenes monogénicos que provocan ictus (Boehme <i>et al.</i> , 2017). En esta tabla se muestran las enfermedades relacionadas con los accidentes cerebrovasculares, su modo de herencia, la proteína implicada, el tipo de accidente cerebrovascular (ACV) que provoca y sus manifestaciones clínicas más comunes.	6
Tabla 3. Variantes genéticas más comunes que provocan ictus (Boehme <i>et al.</i> , 2017). En esta tabla se muestran algunos de los genes, sus respectivas proteínas alteradas y el mecanismo del accidente cerebrovascular (ACV) en el que se manifiesta la alteración.	7
Tabla 4. Factores de riesgo modificables y no modificables. Algunos factores de riesgo son comunes para los dos tipos de accidente cerebrovascular (ACV) principales (p.e. la edad), otros son únicos para cada uno de ellos. En el caso del factor genético, se trata de un factor de riesgo tanto modificable como no modificable en función de la condición estudiada.	7
Tabla 5. Descripción de los estudios seleccionados para el metaanálisis. Se especifica el identificador de cada estudio para la base de datos GEO, la plataforma empleada, el tejido estudiado y las referencias.	21
Tabla 6. Distribución de los individuos incluidos en el estudio GSE13353. Se muestran distribuidos por sexo (hombre o mujer) y estado (accidente cerebrovascular (ACV) o control).	22
Tabla 7. Individuos incluidos en el estudio GSE16561. Se muestran distribuidos por sexo (hombre o mujer) y estado (accidente cerebrovascular (ACV) o control).	23
Tabla 8. Individuos incluidos en el estudio GSE22255. Se muestran distribuidos por sexo (hombre o mujer) y estado (accidente cerebrovascular (ACV) o control).	24
Tabla 9. Individuos incluidos en el estudio. Se muestran distribuidos por sexo (hombre o mujer) y estado (accidente cerebrovascular (ACV) o control).	24
Tabla 10. Individuos incluidos en el estudio GSE66724. Se muestran distribuidos por sexo (hombre o mujer) y estado (accidente cerebrovascular (ACV) o control).	25
Tabla 11. Resumen del resultado del análisis de expresión diferencial. Se llevó a cabo el análisis para la comparación de interés (ME – MC) – (HE – HC) (ME: Mujer Enferma, MC: Mujer Control, HE: Hombre enfermo, HC: Hombre Control). Los genes se encuentran agrupados según su logaritmo de magnitud de cambio (logFC), de forma que están etiquetadas como <i>UP</i> si están sobrerrepresentadas en mujeres (logFC > 0) y como <i>DOWN</i> si están sobrerrepresentadas en hombres (logFC < 0).	25

Tabla 12. Resumen del resultado del análisis de enriquecimiento funcional. Se llevó a cabo el análisis para la comparación de interés (ME – MC) – (HE – HC) (ME: Mujer Enferma, MC: Mujer Control, HE: Hombre enfermo, HC: Hombre Control). PB: proceso biológico, FM: función molecular, CC: componente celular. Las agrupaciones se han efectuado según el valor del logaritmo de *odds ratio*, de manera que aquellas funciones y rutas que se encuentran como *UP* se hallan sobrerrepresentadas en mujeres (LOR > 0); y las que se encuentran como *DOWN* están sobrerrepresentadas en hombres (LOR < 0).26

Tabla 13. Resumen de los procesos biológicos (PB), funciones moleculares (FM), componentes celulares (CC) y rutas KEGG significativos en el metaanálisis. Las agrupaciones se han efectuado según el valor del logaritmo de *odds ratio*, de manera que aquellas funciones y rutas que se encuentran como *UP* se hallan sobrerrepresentadas en mujeres (LOR>0); y las que se encuentran como *DOWN* están sobrerrepresentadas en hombres (LOR<0).....26

Tabla 14. Procesos biológicos (PB) GO significativos en el metaanálisis. Se muestran los identificadores de los términos PB y sus respectivos estimadores de la medida de efecto. Las agrupaciones se han efectuado según el valor del logaritmo de odds ratio, de manera que aquellos procesos biológicos que se encuentran como *UP* se hallan sobrerrepresentados en mujeres (LOR > 0); y los representados como *DOWN* se encuentran sobrerrepresentados en hombres (LOR < 0). LI y LS hacen referencia a los límites inferior y superior del intervalo de confianza del LOR (al 95%). La última columna representa mediante símbolos el sexo en el que se encuentra sobrerrepresentada la función (♂ para hombres, ♀ para mujeres)......27

Tabla 15. Funciones moleculares (FM) GO significativas en el metaanálisis. Se muestran los identificadores de los términos FM y sus respectivos estimadores de la medida de efecto. Las agrupaciones se han efectuado según el valor del logaritmo de *odds ratio*, de manera que aquellas funciones moleculares que se encuentran como *UP* se hallan sobrerrepresentadas en mujeres (LOR > 0); y las representadas como *DOWN* se encuentran sobrerrepresentadas en hombres (LOR < 0). LI y LS hacen referencia a los límites inferior y superior del intervalo de confianza del LOR (al 95%). La última columna representa mediante símbolos el sexo en el que se encuentra sobrerrepresentada la función (♂ para hombres, ♀ para mujeres)......28

Tabla 16. Componentes celulares (CC) GO significativos en el metaanálisis. Se muestran los identificadores de los términos CC y sus respectivos estimadores de la medida de efecto. Las agrupaciones se han efectuado según el valor del logaritmo de *odds ratio*, de manera que aquellas funciones moleculares que se encuentran como *UP* se hallan sobrerrepresentadas en mujeres (LOR > 0); y las representadas como *DOWN* se encuentran sobrerrepresentadas en hombres (LOR < 0). LI y LS hacen referencia a los límites inferior y superior del intervalo de confianza del LOR (al 95%). La última columna representa mediante símbolos el sexo en el que se encuentra sobrerrepresentada la función (♂ para hombres, ♀ para mujeres)......29

Tabla 17. Rutas KEGG significativas en el metaanálisis. Se muestran los identificadores de las rutas KEGG y sus respectivos estimadores de la medida de efecto. Las agrupaciones se han efectuado según el valor del logaritmo de *odds ratio*, de manera que aquellas rutas KEGG que se encuentran como *UP* se hallan sobrerrepresentadas en mujeres ($LOR > 0$); y las representadas como *DOWN* se encuentran sobrerrepresentadas en hombres ($LOR < 0$). LI y LS hacen referencia a los límites inferior y superior del intervalo de confianza del LOR (al 95%).31

ABREVIATURAS

ACV	Accidente cerebrovascular
AIT	Ataque isquémico transitorio
ARN	Ácido ribonucleico
ARNm	Ácido ribonucleico mensajero
BH	Benjamini y Hochberg
CC	Componente celular
CDC	<i>Center for Disease Control</i>
DL	Método DerSimonian y Laird
EP	Enfermedad de Parkinson
EST	<i>Expressed Sequence Tag</i>
FAIR	<i>Findability, Accessibility, Interoperability, and Reusability</i>
FM	Función molecular
GBD	<i>Global Burden Diseases, Injuries and Risk Factors</i>
GEO	<i>Gene Expression Omnibus</i>
GO	<i>Gene Ontology</i>
GSA	<i>Gene Set Analysis</i>
HTS	<i>High-Throughput Sequencing</i> (secuenciación de alto rendimiento)
INE	Instituto Nacional de Estadística
KEGG	<i>Kyoto Encyclopedia of Genes and Genomes</i>
logFC	Log <i>fold change</i> (estimación del cambio de log ₂)
LOR	Logaritmo del <i>odds ratio</i>
MAPK	<i>Mitogen-activated protein kinase</i>
MEC	Matriz extracelular
NA	<i>Not Available</i> (dato no disponible)
NCBI	<i>National Center for Biotechnology Information</i>
OMS	Organización Mundial de la Salud
OXPHOS	<i>Mitochondrial Oxidative Phosphorylation System</i>
PB	Proceso biológico
PBMC	Células mononucleares de sangre periférica
PCA	<i>Principal Component Analysis</i>
PRISMA	<i>Preferred Reporting Items for Systematic reviews and Meta-Analyses</i>
RNA-seq	Secuenciación de ARN
ROS	<i>Reactive Oxygen Species</i>

*Me levanto y me acuesto dando las gracias;
Sé que esto no me lo merezco, apagas mi ansia.
Es difícil estar en tu puesto, mantienes mi calma;
De verdad que te lo agradezco, gracias mama.*

1. INTRODUCCIÓN

1.1. Contexto de la enfermedades neurodegenerativas y la implicación del sexo

Desde que se descubrió la importancia de la incorporación de la variable sexo en los estudios clínicos y experimentales por el *US National Institute of Mental Health* en 2011, se ha podido comprender de mejor forma las diferencias respecto del sexo en diversos aspectos, como la epidemiología o la patofisiología, de numerosos desórdenes neurológicos (*The Lancet Neurology*, 2019). Por ejemplo, se conoce que el estrógeno actúa como neuroprotector en las mujeres, que el alelo APOE4 aumenta el riesgo de manifestar la enfermedad del Alzheimer de forma más pronunciada en mujeres que en hombres, o que las mujeres sufren más accidentes cerebrovasculares y son menos propensas a recuperarse de ellos que los hombres. Asimismo, aun siendo examinada la variable del sexo en este tipo de estudios, las diferencias en la fisiología y las interacciones sociales en hombres y en mujeres (entre otras variables) son raramente estudiadas y sometidas a escrutinio (*Attarian et al.*, 2015; *Tierney et al.*, 2017).

Es muy importante conocer la diferencia entre el sexo y el género a la hora de llevar a cabo estudios clínicos y experimentales, especialmente aquellos relacionados con las enfermedades neurodegenerativas. El sexo hace referencia a las diferencias biológicas y fisiológicas entre hombres y mujeres, donde se incluyen los cromosomas, las hormonas y la anatomía, entre otros. Por otro lado, el género está relacionado con los roles sociales o culturales y las características empleadas para describir masculinidad o feminidad en una sociedad concreta (*Zagni et al.*, 2016; *Tierney et al.*, 2017). Si se emplea la enfermedad de Parkinson como ejemplo, se observa que, de forma generalizada, los hombres presentan mayores índices de prevalencia e incidencia de la enfermedad (diferencia en el sexo) y que están expuestos a un mayor riesgo, puesto que es más común que se encuentren expuestos a agentes tóxicos relacionados con el entorno en el que trabajan (diferencia en el género) (*Ritz et al.*, 2014).

1.2. Accidente cerebrovascular

El accidente cerebrovascular (ACV), también conocido como ictus, derrame o infarto cerebrales, es una enfermedad en la cual se produce una reducción o interrupción del suministro de sangre a una parte del cerebro, por lo que dicha parte no recibe nutrientes ni oxígeno (hipoxia) y, por consiguiente, se produce la muerte celular (*MAYO CLINIC*, 2021).

1.2.1. Etiología

Según el fenómeno que se produzca, se pueden describir dos causas principales que conducen a un ACV. La primera causa es que una arteria que suministra sangre al cerebro quede bloqueada, por lo que se considera un accidente cerebrovascular isquémico. La segunda causa es que se produzca una ruptura o filtración de algún vaso sanguíneo, por lo que se estaría ante un accidente cerebrovascular hemorrágico (Figura 1). Al margen de lo expuesto, existen casos donde el

1. INTRODUCCIÓN

suministro sanguíneo únicamente se interrumpe de forma temporal, considerándose así un ataque isquémico transitorio (AIT), siendo sus síntomas poco duraderos (MAYO CLINIC, 2021).

Accidente cerebrovascular isquémico

Se trata de la manifestación más común del ACV, siendo alrededor del 71% de todos los ictus que se producen a nivel mundial (CamPBell *et al.*, 2019). La mayoría de las isquemias son de origen tromboembólico, siendo las principales causas de origen arterial, así como la aterosclerosis, o de origen cardíaco, como por ejemplo la fibrilación auricular.

Causas arteriales de un ACV isquémico.

Es el tipo de ACV isquémico más frecuente en hombres.

- **Aterosclerosis.** Una embolia en la vasculatura cerebral es bastante común en los ictus isquémicos. Dicho fenómeno puede estar causado por la exposición de la parte lipídica de las placas ateroscleróticas al torrente sanguíneo a partir de la inflamación de la capa fibrosa de dichas placas.
- **Enfermedad cerebrovascular de vaso pequeño.** Esta condición afecta a las arterias y arteriolas más pequeñas del cerebro.

Causas cardíacas de un ACV isquémico.

Es la variante del ACV isquémico más frecuente en mujeres debido a una menor exposición a factores de riesgo, como el tabaquismo o el consumo de alcohol.

- **Fibrilación auricular.** La fibrilación auricular permite que la sangre se estanque, particularmente en el apéndice auricular izquierdo, lo cual puede provocar una trombosis y posterior embolia a la circulación cerebral isquémica. La prevalencia de este tipo de causas cardíacas del ACV aumenta proporcionalmente respecto de la esperanza de vida y la obesidad, aunque también se deben tener en cuenta otras variables decisivas como la hipertensión crónica, enfermedad valvular, diabetes *mellitus*, hipertiroidismo, consumo enólico o apnea obstructiva del sueño.
- **Endocarditis infectiva.** Una infección bacteriana puede producir una embolia séptica en el cerebro a partir de la acumulación de dichos microorganismos en las vías sanguíneas, obstruyendo así el suministro de sangre al cerebro.

Accidente cerebrovascular hemorrágico

Es otra variante del ACV donde se produce una hemorragia en el cerebro a causa de la ruptura de un vaso sanguíneo. Este tipo de ictus está asociado a una morbilidad severa y mortalidad elevada. Es de gran importancia un diagnóstico y tratamiento tempranos a fin de reducir la expansión de la hemorragia.

1. INTRODUCCIÓN

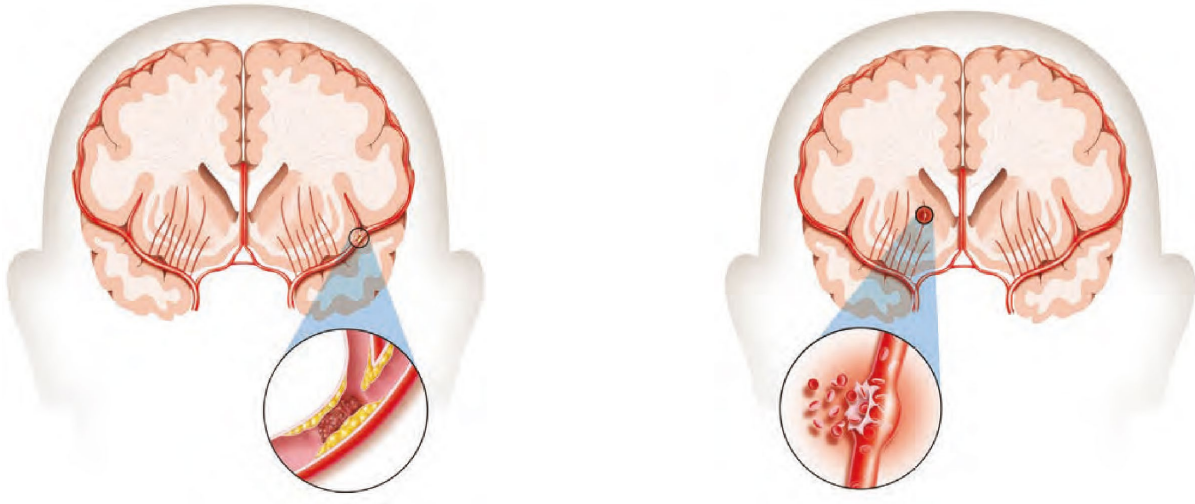


Figura 1. Ilustraciones de las dos manifestaciones más comunes del accidente cerebrovascular. A la izquierda se representa el ictus isquémico, mientras que a la derecha se encuentra el ictus hemorrágico (*American Heart Association, 2021*).

Entre las causas del ACV hemorrágico se encuentra la hipertensión como la más común. Una hipertensión prolongada produce la destrucción de la lámina elástica y la fragmentación de los músculos lisos de las arterias. Los sitios comunes de origen de la hemorragia intracerebral inducida por hipertensión son las pequeñas arterias penetrantes que se originan en las arterias basílicas o en las arterias cerebrales anterior, media o posterior (MAYO CLINIC, 2021).

1.2.2. Epidemiología

El ACV es la tercera causa de muerte más común a nivel mundial y la primera causa de discapacidad permanente en los adultos. La incidencia de dicha enfermedad es de 258 por cada 100.000 habitantes al año. La incidencia fue disminuyendo desde la década de los 70 hasta los últimos años en los países de elevada renta. Sin embargo, la reducción del riesgo vital no ha disminuido de forma paralela a la incidencia, siendo más frecuente en hombres que en mujeres.

Si se centra el foco en el ACV isquémico, se observa que, en 2017, 2,7 millones de personas fallecieron a causa de este ictus. Tanto la incidencia como la prevalencia y las tasas de mortalidad mundiales han descendido desde 1990 a 2013 gracias a la mejora de la prevención secundaria y de la evolución de la detección del accidente en tiempos cortos. No obstante, estos índices sólo se han visto reducidos en países con renta alta, mientras que no se han registrado diferencias significativas en países de rentas media y baja, posiblemente debido a diferencias en la esperanza de vida o los estándares de salud, entre otros factores (Figura 2) (CampBell *et al.*, 2019).

1. INTRODUCCIÓN

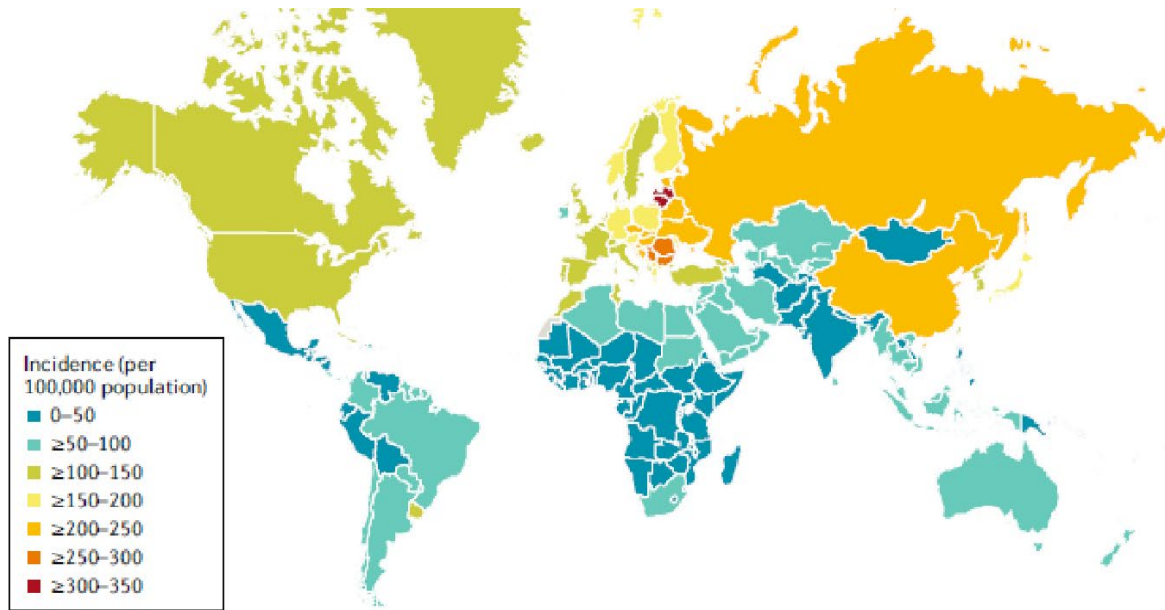


Figura 2. Epidemiología de ictus isquémico. Mapa ilustrativo de la distribución de la incidencia del ictus isquémico a partir de los datos de *Global Burden of Disease* de 2017.

Según los estudios demográficos, las mujeres tienen menor riesgo de sufrir un ACV cuando la comparación se basa en la edad. No obstante, dicho riesgo aumenta significativamente a partir de los 85 años. De hecho, la tasa de letalidad en las mujeres es 1,25 veces mayor comparada con la de los hombres (Carcel *et al.*, 2020).

En España, la prevalencia estimada de ictus en el año 2006 fue del 6,4% en población de 70 años o más, con diferencias significativas en función del sexo (7,3% en hombres y 5,6% en mujeres) y del territorio (8,7% en zonas urbanas y 3,8% en rurales) (The Spanish Epidemiological Study Group on Ageing, 2006). Además, tanto la incidencia como la prevalencia se incrementa en edades avanzadas, particularmente entre las mujeres (Díaz-Guzmán *et al.*, 2008; Díaz-Guzmán *et al.*, 2012), presentando una incidencia de ictus en territorio nacional de 187,4 casos por cada 100.000 habitantes, lo que supone un total de 71.780 nuevos casos considerando la población española de 18 años o más (Díaz-Guzmán *et al.*, 2012) (Figura 3). Se prevé que dicha incidencia se incremente un 35% entre 2015 y 2035 debido, en gran parte, al aumento de la esperanza de vida de la población (The Burden of Stroke in Europe, 2017).

1. INTRODUCCIÓN

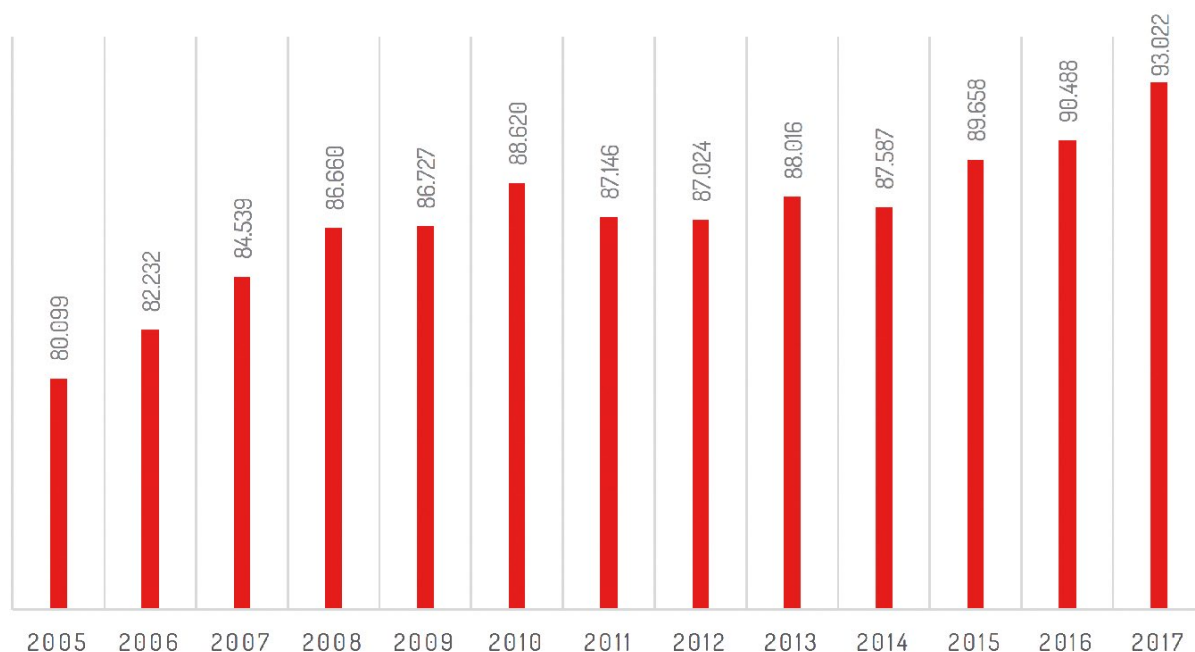


Figura 3. La evolución del número de altas hospitalarias por ictus a nivel nacional en los últimos años. Fuente: Ministerio de Sanidad, Consumo y Bienestar Social.

Asimismo, en 2017, el ictus fue la segunda causa de muerte global, la tercera en hombres y la primera en mujeres, según el Instituto Nacional de Estadística (INE). Teniendo en cuenta el número de altas hospitalarias por ictus producidas en España en 2017, se observa que el 54,6% de las mismas correspondieron a hombres y su número es mayor en los grupos de edad más avanzada. Sin embargo, esta relación se invierte a partir de los 75 años, lo que refleja el perfil epidemiológico del ictus indicado en la literatura: más frecuente entre los varones, excepto en edades más avanzadas, debido, al menos en parte, a la mayor esperanza de vida de las mujeres (Tabla 1) (INE, 2017).

Tabla 1. Número de defunciones por ictus según sexo (filas) y tramo de edad (columnas) en 2017 (INE, 2017).

SEXO Y EDAD	<45	45-54	55-64	65-74	75-84	>85	TOTAL
HOMBRES	113	371	780	1.690	3.815	4.786	11.555
MUJERES	100	207	387	1.004	4.036	9.648	15.382
TOTAL	213	578	1.167	2.694	7.851	14.434	26.937

Las diferencias epidemiológicas entre hombres y mujeres respecto a esta enfermedad se ven mitigadas a una avanzada edad, siendo una de las causas principales la pérdida de la capacidad neuroprotectora de los estrógenos en mujeres posmenopáusicas.

Varios estudios afirman que las mujeres sufren discapacidades en mayor medida que los hombres después de sobrevivir a un ACV (Appelros *et al.*, 2009; Bushnell *et al.*, 2018). Las mujeres experimentan mayor limitación de la actividad, peor calidad de vida relativa a la salud, y mayores niveles de depresión post-ictus. Así pues, se llegó a la conclusión de que las mujeres presentan

1. INTRODUCCIÓN

problemas en movilidad, autocuidado, actividades rutinarias, dolor/molestia y ansiedad/depresión (Carcel *et al.*, 2020).

1.2.3. Genética de la enfermedad

El ACV es una enfermedad compleja y multifactorial, causada por la combinación de factores de riesgo vasculares, el entorno y los factores genéticos. En este sentido, los desórdenes monogénicos relacionados con el ictus son poco usuales, estando presentes en apenas el 1% de los accidentes que se producen (Della-Morte *et al.*, 2012).

Dentro de las alteraciones monogénicas, se ha visto que mutaciones en el gen *NOTCH3* dan lugar a arteriopatía cerebral autosómica dominante con infartos subcorticales y leucoencefalopatía (CADASIL), una condición autosómica dominante. Análogamente, la arteriopatía cerebral autosómica recesiva con infartos subcorticales y leucoencefalopatía (CARASIL), una condición autosómica recesiva provocada por la alteración del gen *HTRA1*, se manifiesta en edades mucho más tempranas de lo habitual (entre 10 y 30 años). Por otra parte, la enfermedad de Fabry es una enfermedad ligada al cromosoma X que afecta al almacenamiento lisosómico. Todas estas variaciones genotípicas se traducen en arterosclerosis, cardioembolismo o la enfermedad de vasos pequeños (Tabla 2) (Della-Morte *et al.*, 2012; Dichgans *et al.*, 2019).

Tabla 2. Desórdenes monogénicos que provocan ictus (Boehme *et al.*, 2017). En esta tabla se muestran las enfermedades relacionadas con los accidentes cerebrovasculares, su modo de herencia, la proteína implicada, el tipo de accidente cerebrovascular (ACV) que provoca y sus manifestaciones clínicas más comunes.

Enfermedad	Modo de herencia	Gen/Proteína	Mecanismo de ACV
Arteriopatía cerebral autosómica dominante con infartos subcorticales y leucoencefalopatía	Autosómica dominante	<i>NOTCH3</i> /NOTCH3	Enfermedad de vasos pequeños
Arteriopatía cerebral autosómica recesiva con infartos subcorticales y leucoencefalopatía	Autosómica recesiva	<i>HTRA1</i> /HtrA peptidasa-1 de serina	Enfermedad de vasos pequeños
Angiopatía amiloide cerebral	Autosómica dominante	<i>APP</i> /Proteína precursora de β -amiloide	Enfermedad de vasos pequeños

Por otro lado se encuentran las variantes genéticas más comunes cuando se trata de la manifestación del ACV, la mayoría de los cuales están implicados en los ictus isquémicos. Entre dichas variantes destaca la alteración del gen *FOXF2*, la cual provoca la obturación o la ruptura de los vasos sanguíneos más pequeños del cerebro, entre otros (Tabla 3).

1. INTRODUCCIÓN

Tabla 3. Variantes genéticas más comunes que provocan ictus (Boehme *et al.*, 2017). En esta tabla se muestran algunos de los genes, sus respectivas proteínas alteradas y el mecanismo del accidente cerebrovascular (ACV) en el que se manifiesta la alteración.

Gen	Proteína	Mecanismo de ACV
<i>TSPAN2</i>	Tetraspanina-2	ACV isquémico
<i>FOXF2</i>	Proteína de caja <i>Forkhead</i> F2	ACV isquémico y ACV hemorrágico
<i>ABO</i>	Proteína de grupo sanguíneo	ACV isquémico

1.2.4. Factores de riesgo

Los factores de riesgo se pueden clasificar en dos tipos: modificables y no modificables. Los factores no modificables, también conocidos como marcadores de riesgo, engloban a la edad, el sexo o la genética, entre otros; mientras que los factores modificables son aquellos que permiten la reducción de la gravedad de un ACV si se consigue su mitigación. Algunos de los factores de este tipo más estudiados son la hipertensión, diabetes *mellitus* o el tabaquismo, entre otros (Tabla 4) (Boehme *et al.*, 2017).

Tabla 4. Factores de riesgo modificables y no modificables. Algunos factores de riesgo son comunes para los dos tipos de accidente cerebrovascular (ACV) principales (p.e. la edad), otros son únicos para cada uno de ellos. En el caso del factor genético, se trata de un factor de riesgo tanto modificable como no modificable en función de la condición estudiada.

Tipo de ACV	Factores de riesgo modificables	Factores de riesgo no modificables
ACV isquémico	Edad	Hipertensión
	Sexo	Tabaquismo
	Etnia	Diabetes <i>mellitus</i>
	Genética	
ACV hemorrágico	Edad	Hipertensión
	Sexo	Alcoholismo
	Etnia	Dieta
	Genética	

No obstante, no queda claro si las diferencias de sexo cuando se produce un ictus se limitan a factores biológicos, donde la respuesta al accidente y los riesgos son diferentes entre los sexos (dimorfismo del ictus), o si existe alguna semejanza en los servicios médicos proporcionados.

Teniendo en cuenta las disparidades existentes entre ambos sexos en esta enfermedad, hay algunos factores biológicos y aspectos específicos relacionados con la mujer que son de especial importancia para tener en cuenta como, por ejemplo, el papel de las hormonas sexuales (Ruiz Cantero, 2019; Schrabji *et al.*, 2019). En este sentido, las concentraciones de testosterona son más estables que las de estrógenos a lo largo de la vida de la mujer. Valores bajos de testosterona se han asociado con un mayor riesgo de ictus en el hombre, pero no en las mujeres. Por otro lado, se ha visto que niveles bajos de dehidroepiandrosterona (hormona precursora de estrógenos y testosterona) está vinculado a un mayor riesgo de sufrir un ictus isquémico y con mayor gravedad del ictus (Ruiz Cantero, 2019).

1. INTRODUCCIÓN

Asimismo, el factor de crecimiento similar a la insulina tipo 1 presenta efectos neuroprotectores parecidos a los estrógenos. La caída de los niveles de este factor de crecimiento viene asociada al avance de la edad, y en conjunción con la bajada de los estrógenos en la menopausia, podría explicar la incidencia del ictus y su mayor gravedad en las mujeres mayores (Sohrabji *et al.*, 2019).

1.2.5. Tratamientos actuales

Puesto que existen diferentes tipos de ACV, los tratamientos que se van a llevar a cabo van a ser distintos, aunque un diagnóstico rápido y preciso va a ser crucial para cualquier tipo de ictus, reduciendo así el daño cerebral y permitiendo un tratamiento adecuado y eficaz (CDC, 2021).

Tratamiento del ACV isquémico.

El principal objetivo de este tratamiento se basa en la restauración del flujo del torrente sanguíneo al cerebro. Para ello, se administra Alteplase IV r-tPA, el cual es un activador tisular recombinante de plasminógeno que permite la disolución de los coágulos existentes y la prevención de la formación de nuevos. Sin embargo, el suministro intravenoso de este medicamento se debe de realizar dentro de las 4,5 horas desde que empiezan a manifestarse los síntomas. Este medicamento se puede combinar con intervenciones quirúrgicas, como por ejemplo una angioplastia o una trombectomía. Se ha observado que no hay diferencias significativas entre sexos cuando se lleva a cabo la trombólisis siempre que dicho remedio se lleve a cabo antes de las 3 horas de la manifestación de los síntomas (Cheng y Kim, 2015).

Tratamiento del ACV hemorrágico.

El remedio de urgencia contra este tipo de ictus se centra en controlar el sangrado y reducir la presión en el cerebro causada por el exceso de líquido. A fin de reducir el efecto del accidente, se combinan medicamentos para contrarrestar los efectos de los anticoagulantes o disminuir la presión arterial con cirugía (MAYO CLINIC, 2021).

Las mujeres y los hombres experimentan los accidentes cerebrovasculares de manera diferente. Si bien las diferencias por sexo en la muerte y la discapacidad se atenúan después de tener en cuenta la edad, la función premórbida, la fibrilación auricular, la gravedad del accidente cerebrovascular y las comorbilidades, la realidad es que las mujeres sufren peores consecuencias. Es por ello por lo que es imperativo elucidar las causas de estas diferencias, permitiendo así hallar nuevos biomarcadores que permitan mejorar el diagnóstico, el tratamiento y el pronóstico de los pacientes empleando la vía de la medicina personalizada, cerrando así la brecha entre mujeres y hombres.

1.3. Datos ómicos y tecnologías de alto rendimiento

La información genética se encuentra codificada y almacenada en su totalidad en el genoma, el cual en la mayoría de los organismos se halla representado por el ácido desoxirribonucleico (ADN). Esta molécula está basada en dos hebras antiparalelas formadas por la unión covalente entre nucleótidos, los cuales están constituidos por una molécula de fosfato, una desoxirribosa

1. INTRODUCCIÓN

(monosacárido) y una base nitrogenada, habiendo así diferentes nucleótidos en función de la base (adenina (A), timina (T), citosina (C) o guanina (G)). La información contenida en el ADN puede ser transcrita a otro ácido nucleico, el ácido ribonucleico (ARN). Esta nueva molécula discrepa del ADN en diversos aspectos: el monosacárido del ARN es una ribosa en vez de una desoxirribosa, la timina se sustituye por uracilo y la molécula está compuesta únicamente por una hebra.

El ARN presenta diferentes funciones dependiendo del grupo al que pertenezca. Entre los diferentes tipos de ARN se encuentra el ARN mensajero (ARNm), el cual se emplea como molde para la síntesis de proteínas en el evento de traducción. Los procesos de transcripción y traducción fueron recogidos en el dogma central de la biología molecular postulado por Francis Crick (1970) (Figura 4).

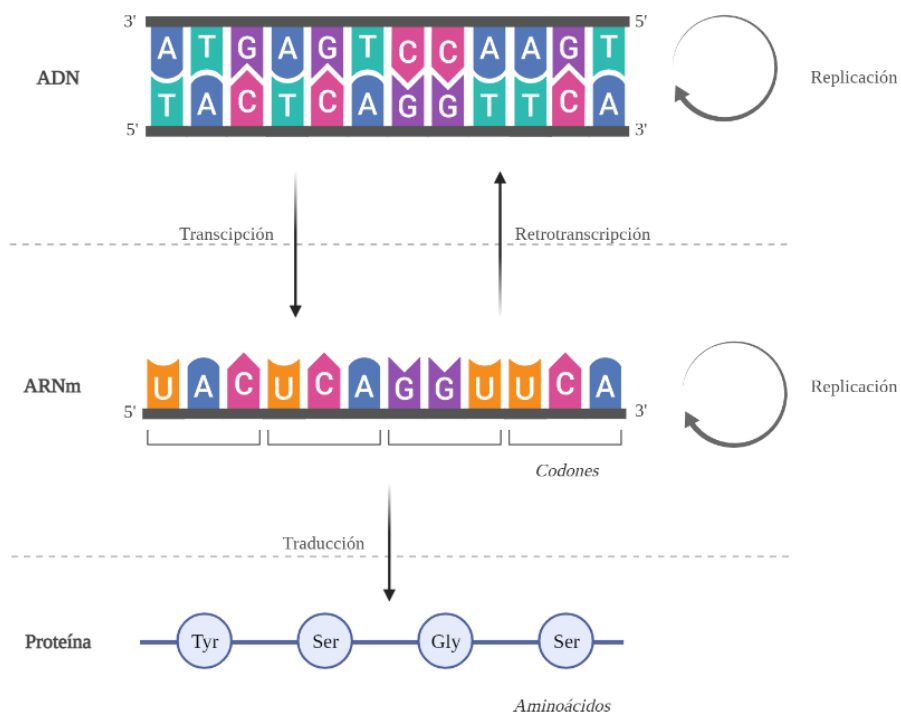


Figura 4. Dogma central de la biología molecular.

1.3.1. Transcriptómica

Consideramos la transcriptómica como el estudio integral de toda la transcripción de ARN codificada por el genoma en condiciones específicas o en una célula específica utilizando tecnologías de alto rendimiento, como los microarrays. También puede considerarse como parte de la genómica funcional que se centra en la expresión génica, especialmente en el ARNm (transcripción). La etiqueta de secuencia expresada (EST), la secuenciación de microarrays, y la secuenciación de ARN (RNA-seq) son herramientas importantes de transcriptómica.

La metodología consiste en el aislamiento, purificación, cuantificación, construcción de bibliotecas de ADN circular (ADNc) y secuenciación de alto rendimiento. Las consideraciones importantes

1. INTRODUCCIÓN

para la selección de herramientas de transcriptómica son la rentabilidad, la sensibilidad, el alto rendimiento y la concentración mínima de ARN inicial.

El RNA-seq es una de las tecnologías avanzadas de alto rendimiento para transcriptómica. El RNA-seq, también conocido como “secuenciación de escopeta de transcriptoma completo”, utiliza herramientas de secuenciación de segunda generación, comúnmente conocida como NGS. Las ventajas del RNA-seq son las siguientes: no depende de la disponibilidad de secuencia del genoma, no tiene límites superiores de cuantificación, muestra una alta reproducibilidad y posee un amplio rango de detección dinámica.

El chip de ADN (también llamado chip de gen, biochip o micromatriz) es una colección de puntos de ADN, ADNc u oligonucleótidos adheridos a un soporte sólido, como vidrio. Esta composición da lugar a un chip de silicio que forma una matriz con el propósito de perfilar la expresión: monitorea los niveles de expresión de miles de genes simultáneamente y es un método basado en la hibridación. A lo largo de los años se han logrado varias innovaciones con respecto a la fabricación de chips, el control de las condiciones de hibridación y el análisis computacional de los datos generados. El chip de alta densidad permite la creación de perfiles de expresión génica a un coste relativamente bajo. La principal limitación de esta tecnología es que la información de la secuencia del genoma es un requisito previo (Yadav *et al.*, 2018).

1.3.2. Estrategias de análisis transcriptómicos

Para realizar el análisis de los datos transcriptómicos es necesario llevar a cabo un procesamiento previo, el cual dependerá de la tecnología empleada para la obtención de dichos datos, así como el objetivo final del estudio (García, 2016).

Como resultado del análisis, se obtiene una matriz de expresión que incluye la cuantificación de las unidades biológicas de interés de todas las muestras (García, 2016). A partir de dicha matriz, se pueden seguir diferentes estrategias de procesamiento:

- **Análisis de expresión diferencial.** Presenta como finalidad la identificación de los elementos biológicos con niveles de expresión significativamente modificados respecto de condiciones establecidas como control. La utilidad de este enfoque se focaliza en la caracterización de patologías, así como el descubrimiento y detección de biomarcadores.
- **Análisis clúster.** También denominado “descubrimiento de clases”, se basa en la agrupación de las muestras de los estudios que presentan un patrón similar en los niveles de expresión transcriptómicos.
- **Predicción de clases.** A partir de un predictor supervisado, se puede clasificar un nuevo individuo atendiendo a las similitudes con las muestras empleadas para el entrenamiento de dicho predictor.

1. INTRODUCCIÓN

1.3.3. Caracterización funcional

Como resultado del análisis de expresión diferencial anteriormente mencionado, se obtiene un listado de diversos elementos biológicos que, junto con diferentes bases de datos que registran anotaciones funcionales de las unidades biológicas, se podrá interpretar y obtener conclusiones sobre el significado de su expresión diferencial. Estas bases de datos empleadas deben de ser revisadas, independientemente de si la información que contienen está completa o no (García, 2016). Algunas de las bases de datos en las que se puede encontrar dicha información son:

- **Gene Ontology (GO)**. Posee información relacionada con los sistemas biológicos en 3 grupos diferentes: Procesos Biológicos (PB), Funciones Moleculares (FM) y Componentes Celulares (CC).
- **Kyoto Encyclopedia of Genes and Genomes (KEGG)**. Presenta información sistemática vinculada a la función de los genes y su implicación en las diferentes rutas biológicas.

Los resultados de los análisis de expresión pueden ser combinados con la información proporcionada por las bases de datos previamente nombrados a través de estrategias como el análisis de sobrerrepresentación o el enriquecimiento funcional (García, 2016). En dichos abordajes se selecciona primeramente un grupo de genes (entre otras unidades biológicas) a partir de un criterio de interés definido, y, a continuación, se comparan las funciones anotadas a dichos genes contra las funciones anotadas a un grupo de genes de referencia. Así pues, se efectúa una evaluación estadística del tamaño del número de genes con la anotación perteneciente a cada función, detectando así sobrerrepresentaciones o enriquecimientos de ésta.

Una aproximación alternativa a las expuestas anteriormente es el análisis de grupos de genes, como el *Gene Set Analysis* (GSA), el cual se basa en la identificación de genes que están sobre o infraexpresados respecto a un gran número de genes (u otras unidades biológicas), es decir, en función su nivel de expresión diferencial (Subramanian *et al.*, 2005).

1.3.4. Metaanálisis funcional

El metaanálisis es la metodología estadística que permite amalgamar resultados de diferentes estudios relacionados con los efectos previamente identificados y estudiados a través de la revisión sistemática (García, 2016).

El metaanálisis persigue cumplir diferentes objetivos, como el de incrementar el poder estadístico o llegar a consenso cuando diferentes estudios con mismo objetivo obtienen resultados diferentes (Russo, 2017).

No sólo se agrupan los datos de estudios independientes cuando se lleva a cabo el metaanálisis, sino que también se tienen en cuenta las diferencias en los tamaños muestrales de cada uno de los estudios individuales y su heterogeneidad (variabilidad). Con estos métodos también se puede comprobar la sensibilidad de los resultados respecto de los protocolos de revisión sistemática

1. INTRODUCCIÓN

empleados, concretamente la selección de estudios y su posterior análisis estadístico (Shorten y Shorten, 2013).

Salvando las diferencias entre las numerosas variantes que existen para realizar un metaanálisis, los pasos que se nombran a continuación son básicos y comunes (Shorten y Shorten, 2013):

1. Planteamiento de la hipótesis.
2. Revisión sistemática de los estudios.
3. Extracción de los datos.
4. Estandarización y ponderación de los estudios.
5. Estimación final del efecto.

En caso del metaanálisis funcional de datos genómicos, es crucial la combinación de la información de los estudios de la misma temática para la detección de clases de funciones biológicas estadísticamente significativas.

2. OBJETIVOS

2. OBJETIVOS

El objetivo general de este trabajo es la identificación y la comprensión de los diferentes mecanismos moleculares diferenciales que hay detrás de los accidentes cerebrovasculares por sexos. Hoy en día se sabe que existen estas diferencias mediante indicadores clínicos y epidemiológicos, pero no se conoce con exactitud a qué se deben. Esta información permitiría ser más precisos en el diagnóstico y en la selección de los tratamientos para cada grupo de pacientes.

Este objetivo se alcanzará mediante el cumplimiento de los siguientes objetivos específicos:

1. Revisión bibliográfica sistemática y selección de estudios de ACV con datos de microarrays de expresión génica o RNA-seq en los repositorios Gene Expression Omnibus (GEO), ArrayExpress, u otras bases de datos.
2. Análisis individual de cada estudio seleccionado, incluyendo un análisis de expresión diferencial, un análisis de enriquecimiento funcional y un análisis de rutas.
3. Metaanálisis funcional y de rutas KEGG que permitirá la detección de biomarcadores específicos por sexo.

3. MATERIALES Y MÉTODOS

3.1. Revisión sistemática y selección de estudios transcriptómicos

La revisión sistemática es un tipo de investigación que proporciona un resumen de estudios sobre una cuestión específica. Utilizando métodos explícitos, se busca evaluar críticamente y sintetizar el amplio conjunto de los estudios hallados en las diferentes bases de datos, siempre y cuando dichos estudios cumplan con los criterios de elegibilidad. Este tipo de revisión es particularmente útil para reunir una serie de estudios realizados por separado, a veces con hallazgos contradictorios, y para resumir sus resultados. Con el fin de llevar a cabo este trabajo, se ha efectuado el flujo de trabajo reflejado en la declaración PRISMA (*Preferred Reporting Items for Systematic Reviews and Meta-Analyses*) (Figura 5) (Liberati *et al.*, 2009).

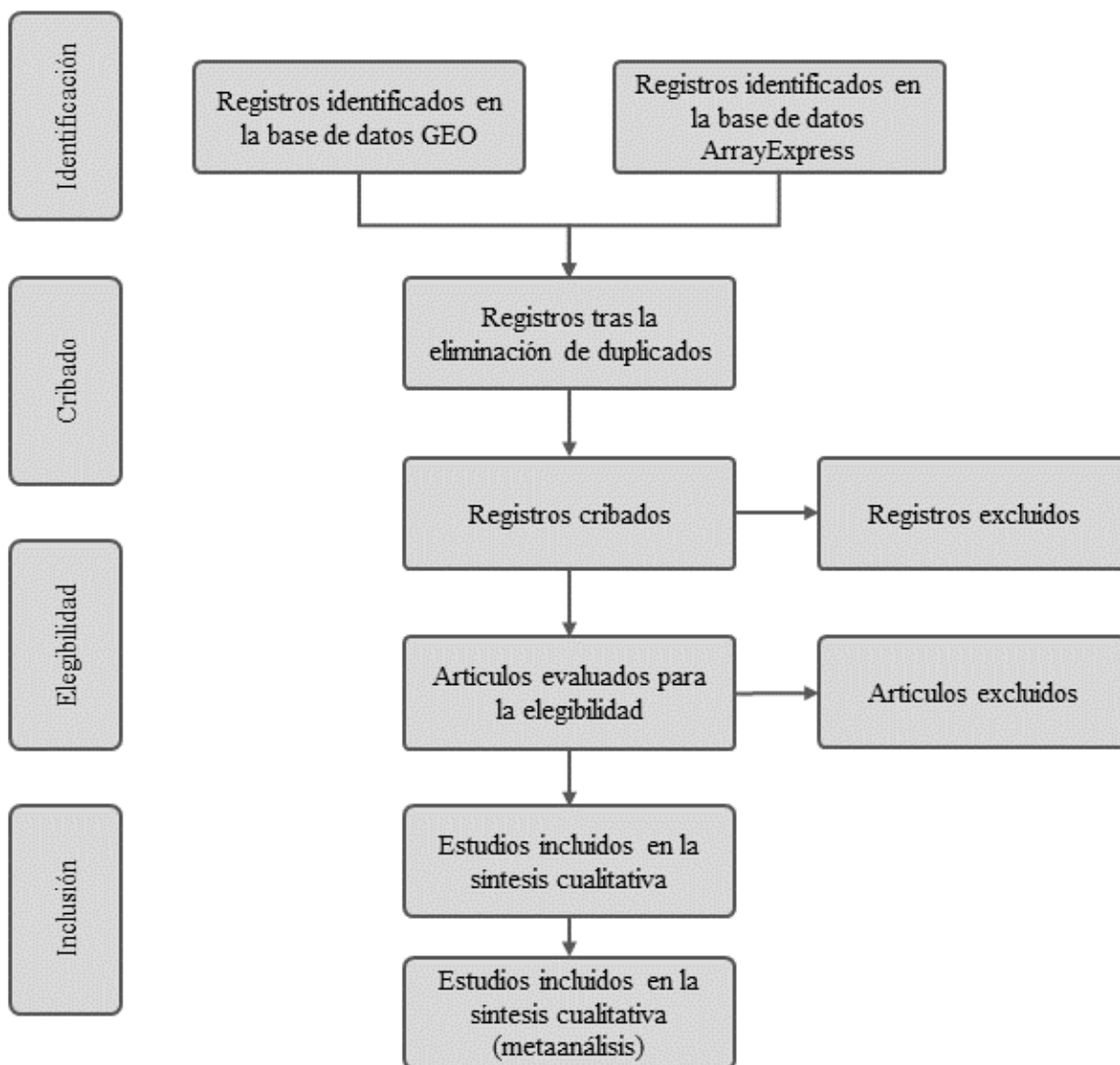


Figura 5. Diagrama de flujo que indica los diferentes pasos a seguir para efectuar una revisión sistemática (esquema adaptado de Liberati *et al.*, 2009).

3. MATERIALES Y MÉTODOS

Con el fin de realizar la selección de los estudios que van a intervenir en este trabajo y su posterior revisión, se efectuó una indagación de los estudios disponibles en las bases de datos públicas sobre los accidentes cerebrovasculares y que recopilaran información sobre la expresión génica y el sexo de las muestras. Dicha exploración se llevó a cabo en las bases de datos GEO (Barrett *et al.*, 2012) y ArrayExpress (Kolesnikov *et al.*, 2015)

Las palabras clave empleadas para la búsqueda fueron “stroke”, “CVA” y “Cerebrovascular accident”. Los estudios hallados a partir de estas palabras clave fueron sometidos a un filtrado posterior en base a dos criterios: que correspondiesen efectivamente a ACV y que presentasen datos de expresión génica. Así pues, se establecieron como tecnologías *expression profiling by high throughput sequencing* y *expression profiling by array*. Se incluyeron aquellos estudios que incluían información sobre el sexo en la descripción del estudio. Se limitaron los estudios únicamente a aquellos que se hubiesen realizado en humanos, y se excluyeron aquellos que se llevaron a cabo empleando líneas celulares. Por otro lado, se excluyeron los estudios que no incluían muestras control o que no presentaban un tamaño muestral igual o superior a 12 muestras, con al menos 3 sujetos por cada grupo experimental.

Una vez realizada la selección final de los estudios, se descargaron los respectivos datos de expresión normalizados.

Para aquellos estudios que, a priori, recogían la información de sexo, y que, sin embargo, en los datos descargados no se hacía referencia explícita a esta variable, se ejecutó una imputación del sexo de las muestras empleando la función *imputeSex* del paquete de R *MetaIntegrator* (Haynes *et al.*, 2016). Dicha librería permite la asignación del sexo de la muestra a partir de la comprobación de la expresión de una serie de genes hallados en los cromosomas sexuales los cuales se encuentran diferencialmente expresados en hombres (*KDM5D*, *RPS4Y1*, *EIF1AY*) o en mujeres (*XIST*, *RPS4X* o *CD40LG*).

3.2. Análisis individualizado de los estudios

3.2.1. Procesamiento de los datos

A fin de que la estrategia del metaanálisis sea aplicable y fructífera, es indispensable llevar a cabo un análisis individualizado de los estudios, pudiendo combinarlos posteriormente si cumplen los requisitos pertinentes. El procesamiento de los datos se efectuó completamente con el lenguaje de programación R haciendo uso del entorno de desarrollo integrado RStudio (R Core Team, 2019). Todo el código de programación escrito y empleado en este trabajo se encuentra disponible en el repositorio GitLab (<https://gitlab.com/ubb-cipf/tfgs-stroke/tfg-yassir>).

Homogeneización de las covariables de las matrices de expresión

Puesto que los estudios seleccionados no emplearon una metodología estandarizada debido a la naturaleza de las plataformas y tecnologías empleadas, se efectuó una normalización de la denominación de los datos de los estudios escogidos para el trabajo. Para ello, se nombró

3. MATERIALES Y MÉTODOS

“Control” a aquellas muestras que pertenecían a pacientes sanos y “Stroke” a las muestras que hacían referencia a pacientes que habían sufrido un ACV. Asimismo, se clasificaron las muestras en función del sexo, habiendo así dos opciones: “Male” o “Female”.

Por consiguiente, se obtuvieron 4 grupos experimentales en cada estudio: “Stroke female”, “Control female”, “Stroke male”, “Control male”.

Homogeneización de la anotación de las sondas

La homogeneización de la anotación de las sondas es un apartado crucial en el procesamiento de los datos de expresión procedentes de microarrays, ya que los estudios pueden proceder de plataformas diferentes y, por ende, presentar nomenclatura distinta. Los identificadores de las sondas de los microarrays se transformaron a identificadores *Entrez*, obteniendo genes con identificadores únicos haciendo uso de la base de datos de *National Center for Biotechnology Information* (NCBI, Estados Unidos). Puesto que para un mismo identificador de *Entrez* se encontraban varias sondas, se eligió el valor con la primera ocurrencia de la sonda. Así, sólo se obtenía una entrada *Entrez* por sonda.

3.2.2. Análisis exploratorio

Una vez efectuada la selección y estandarización de los estudios, se procedió a evaluar la distribución de los perfiles de expresión, identificando así, si se encontraban muestras que presentaban comportamientos anómalos.

Se partió de datos previamente normalizados, evitando la incorporación de nuevos sesgos y mayor variabilidad.

Cuando se utilizan datos provenientes de repositorios públicos, cabe la posibilidad de que haya ciertos valores no disponibles representados como NA (*Not Available*). Debido a esto y, para tener en cuenta este tipo de complicaciones, se lleva a cabo un análisis exploratorio. En este trabajo, se ha optado por la imputación de los datos NA teniendo en cuenta los valores vecinos a partir de la función *knn.impute* del paquete *bnstruct* (Franzin *et al.*, 2016). Seguidamente, se efectuó una transformación logarítmica en base 2 a fin de homogeneizar la escala de los datos. Este paso es crucial, puesto que los datos de los diferentes estudios se han obtenido a partir de diversas plataformas y el uso directo de estos datos no aportaría resultados fiables.

Una vez llevada a cabo la imputación de los NAs, se representaron los datos de expresión de cada una de las muestras en un diagrama de cajas a fin de comprobar que las medianas eran semejantes en todas las muestras y que los datos estaban normalizados. Seguidamente, se realizó un análisis de componentes principales (PCA) con la finalidad de observar comportamientos similares en los niveles de expresión de las muestras en función del estado de los pacientes y del sexo, pudiéndose así generar agrupaciones notables. De la misma forma, se llevó a cabo un análisis de agrupamiento (*clustering*) empleando el índice de correlación con el mismo objetivo. Para la obtención del PCA se hizo uso de las funciones *pcaGenes* y *plotPCA*, mientras que para el *clustering* se empleó la

3. MATERIALES Y MÉTODOS

función *plotTreeClust*, todas disponibles en el repositorio de *metaFunR* (<https://gitlab.com/ubb-cipf/metafunr>).

3.2.3. Análisis de expresión diferencial

Una vez realizado el análisis exploratorio, se evaluaron las diferencias estadísticamente significativas relacionadas con la expresión génica, en los distintos grupos experimentales anteriormente mencionados, pudiendo identificar de esta forma genes sobreexpresados o infraexpresados. Se compararon las diferencias en mujeres debidas a la patología con las diferencias en hombres debidas a la enfermedad a través del siguiente contraste:

$$(Stroke\ female - Control\ female) - (Stroke\ male - Control\ male)$$

Ecuación 1

Dicho contraste fue valorado mediante los modelos lineales introducidos en el paquete *limma* disponible a través de *Bioconductor* (Ritchie *et al.*, 2015). Así pues, se calcularon los coeficientes estimados y el error estándar a partir de una matriz de contrastes y el modelo lineal, así como los estadísticos de expresión diferencial. Por último, se corrigió el valor de p a partir del método de Benjamini y Hochberg (BH) (Benjamini y Hochberg 1995).

Las funciones y el flujo de trabajo empleados se pueden encontrar en el repositorio de *metaFunR* en la función *diffExp*.

3.2.4. Análisis de enriquecimiento funcional

El objetivo de este análisis es la obtención de grupos de genes enriquecidos funcionalmente indicando la función biológica, el proceso molecular y/o el componente molecular que se halla representado por dichos grupos génicos a partir de GSA.

Obtenidos los datos de expresión diferencial de los genes, se llevó a cabo el análisis de enriquecimiento funcional mediante GSA con la librería *mdgsa*. Las funciones y el flujo de trabajo se encuentran en la función *doGsea* en el repositorio de *metaFunR*.

A fin de realizar un enriquecimiento funcional con los términos de GO, se efectuó anteriormente una correspondencia entre los identificadores *Entrez* y sus respectivos códigos GO a través de la librería *GO.db* (Carlson, 2019). Para finalizar, se filtraron las anotaciones muy específicas o demasiado genéricas (anotaciones con un número de genes asociados menor a 10 o superior a 500, respectivamente) y se consideraron significativos únicamente aquellos términos que poseían un valor ajustado de p inferior a 0,05.

3. MATERIALES Y MÉTODOS

3.3. Metaanálisis

Generalmente, la determinación de las funciones génicas suele carecer de significancia estadística si se basan en estudios individuales e independientes, puesto que diversos factores, así como la tecnología utilizada para la obtención de datos, el tamaño muestral empleado o el abordaje experimental pueden afectar a los resultados finales. Es por ello por lo que se combinan los datos generados a partir de varios estudios para determinar funciones enriquecidas.

Una vez llevado a cabo el paso del enriquecimiento funcional, se dispone de una medida que informa del nivel de sobrerrepresentación de cada función, a través de los *odds ratios* y sus respectivos errores estándar. Ambas medidas se recogerán en dos matrices diferentes para llevar a cabo el metaanálisis.

3.3.1. Determinación de la medida combinada del efecto

A partir de las matrices de funciones enriquecidas para cada uno de los términos GO y para las rutas metabólicas, se determina la medida combinada del efecto, así como su respectiva variabilidad. En función de dichos parámetros, se puede decidir qué tipo de modelo de efectos se quiere aplicar a los estudios:

- Efectos fijos. Este método asume que los datos presentan diferencias generadas únicamente por azar.
- Efectos aleatorios. Este método es aplicado cuando se asume que los estudios tienen errores de muestreo diferentes, lo cual es bastante común cuando se trabaja con estudios que generan datos biológicos.

Una vez efectuado el metaanálisis, se extraen una serie de estimadores de la medida del efecto combinada para las funciones y rutas estudiadas, así como los indicadores de heterogeneidad de los mismos:

- QE es el estadístico de contraste, mientras que el QEp es el valor de p. Ambos se obtienen a partir del método de DerSimonian y Laird (DL) (DerSimonian y Laird, 1986), empleados para la detección de la heterogeneidad presente entre los estudios. En este método, la hipótesis nula hace referencia a la no heterogeneidad entre dichos estudios.
- LOR es el valor de la estimación del efecto combinado del conjunto de los estudios. El signo de dicho valor hace referencia a la sobrerrepresentación de una función determinada en uno de los grupos de comparación previamente establecida. La magnitud del valor es cuantitativa y es proporcional a la sobrerrepresentación de dicha función.
- Intervalo de confianza del LOR (al 95%), formado a partir del valor de la variabilidad obtenida.
- El valor de p indica el nivel de significancia del efecto combinado. Dicho indicador debe ser corregido ya que no tiene en cuenta que se están analizando múltiples funciones (multiplicidad del análisis). Una vez corregido, se obtiene el valor de p ajustado. En el caso de este trabajo se ajustó mediante el método BH.

3. MATERIALES Y MÉTODOS

3.3.2. Representación de los resultados

Los resultados del metaanálisis se resumen por medio de tablas y mediante gráficos de bosque para cada una de las funciones obtenidas. Estos gráficos permiten evaluar los resultados globales y el peso en el resultado de cada función de cada estudio individual (García, 2016). En la parte izquierda se indican los estudios incluidos en el metaanálisis, mientras que a la derecha la estimación de la medida individual de cada uno con su respectivo intervalo de confianza al 95%. En el centro del gráfico se muestra la medida de efecto (cuadrado negro) y su intervalo de confianza, siendo proporcional el tamaño del cuadrado a la precisión de las estimaciones. En la parte inferior central se visualiza el resultado del metaanálisis, de forma que su posición respecto a la línea de efecto nulo informa sobre la significación estadística del resultado, y su anchura nos informa de su intervalo de confianza. Finalmente, en la parte inferior izquierda encontramos el modelo de análisis empleado (DL) y a la derecha la estimación de la medida resultante del metaanálisis y su intervalo de confianza al 95% (Figura 6).

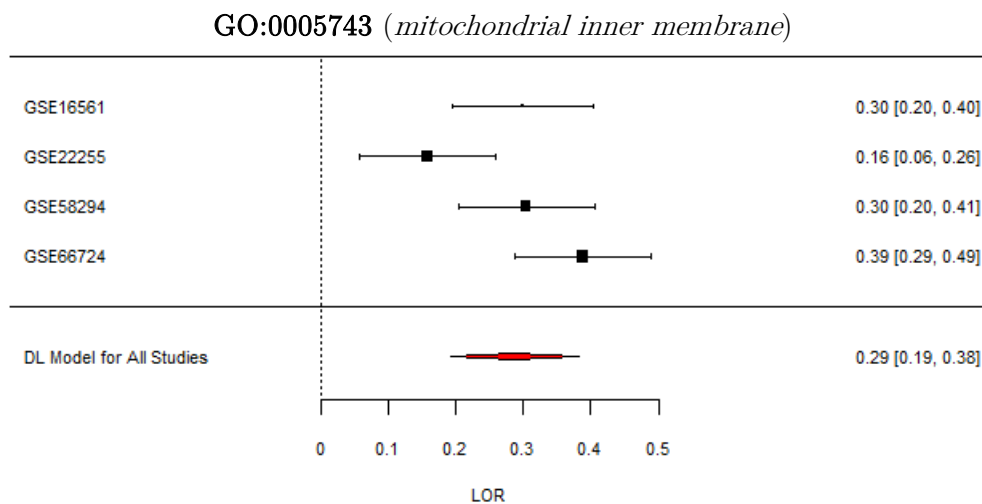


Figura 6. Ejemplo de gráfico de bosque para el término GO:0005743 (*mitochondrial inner membrane*), obtenido en el metaanálisis de este trabajo.

4. RESULTADOS

A continuación, se exponen los resultados obtenidos en este trabajo, siguiendo el mismo orden que en el apartado de Materiales y Métodos.

4.1. Revisión sistemática y selección de estudios

Cuando se realizó la búsqueda de estudios relacionados con ACV (sin tener en cuenta aquellos que aparecieron por duplicado), aplicando los criterios de selección (tecnología empleada, especie, número de muestras mínimo y tipo de estudio), se hallaron 103 estudios candidatos para ser empleados en el metaanálisis.

Una vez revisados de forma individualizada esos 103 estudios, 5 fueron los elegidos para desarrollar el metaanálisis (Tabla 5), siendo excluidos los 98 estudios restantes por diversos motivos (Figura 7).

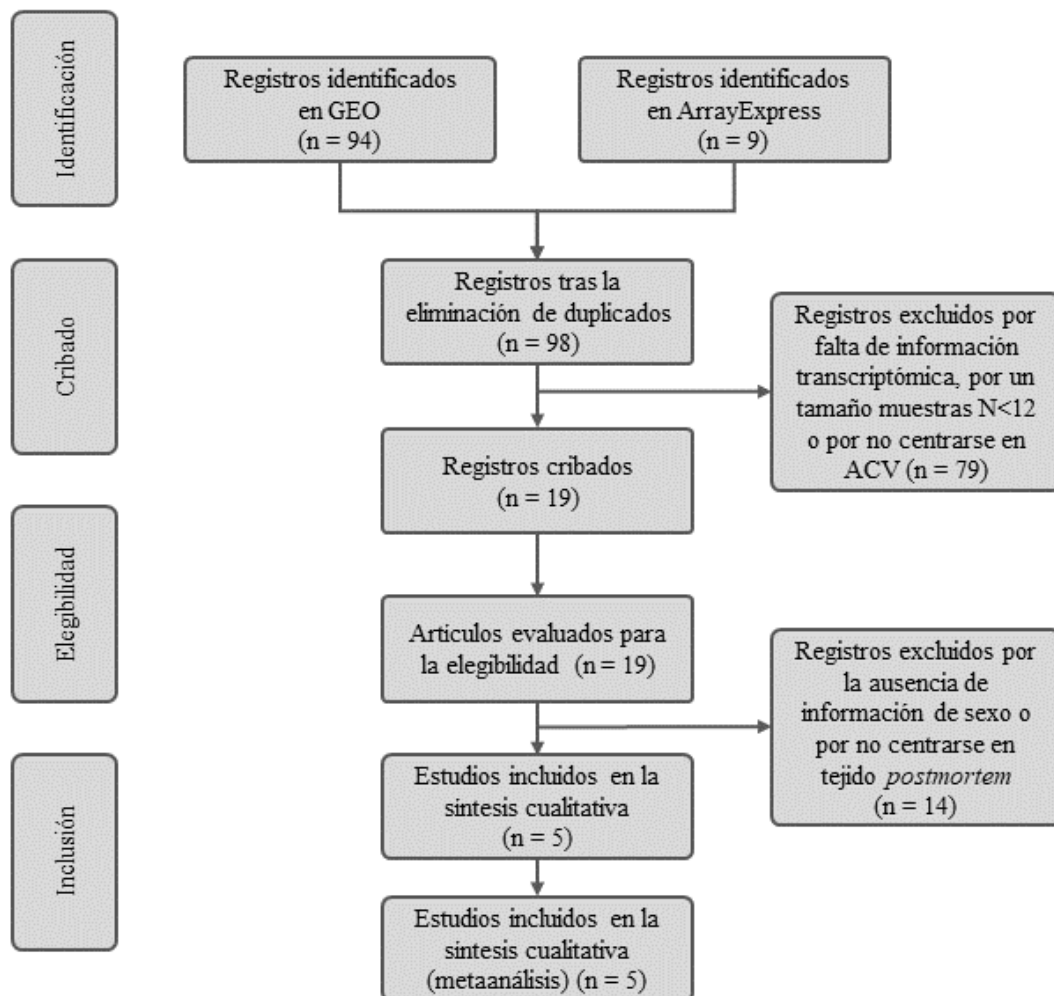


Figura 7. Diagrama de flujo de la revisión sistemática siguiendo los pasos de la declaración PRISMA (*Preferred Reporting Items for Systematic Reviews and Meta-Analyses*).

4. RESULTADOS

Tabla 5. Descripción de los estudios seleccionados para el metaanálisis. Se especifica el identificador de cada estudio para la base de datos GEO, la plataforma empleada, el tejido estudiado y las referencias.

ESTUDIO	PLATAFORMA	TEJIDO	CONTRIBUIDORES
GSE13353	Affymetrix Human Genome U133 Plus 2.0 Array	Pared del aneurisma intracraneal sacular	Kurki <i>et al.</i> , 2011
GSE22255	Affymetrix Human Genome U133 Plus 2.0 Array	Células mononucleares de sangre periférica (PBMC)	Krug <i>et al.</i> , 2012
GSE16561	Illumina HumanRef-8 v3.0 expression beadchip	Células mononucleares de sangre periférica (PBMC)	Barr <i>et al.</i> , 2010
GSE58294	Affymetrix Human Genome U133 Plus 2.0 Array	Células mononucleares de sangre periférica (PBMC)	Stamova <i>et al.</i> , 2014
GSE66724	Affymetrix Human Genome U133 Plus 2.0 Array	Células mononucleares de sangre periférica (PBMC)	Allende <i>et al.</i> , 2016

4.2. Análisis individual de los estudios

Una vez efectuada la descarga de los datos de los estudios escogidos y la estandarización de la nomenclatura de las variables que se van a tener en cuenta en este trabajo, se describió el número de muestras que contenía cada estudio en función del estado de la enfermedad (paciente o control) y del sexo (mujer u hombre). Seguidamente, se llevó a cabo la evaluación de los perfiles de expresión diferencial génica y se realizó un análisis de enriquecimiento funcional para los procesos biológicos (PB), las funciones moleculares (FM) y las componentes celulares (CC) individualmente, así como el enriquecimiento de las rutas metabólicas. Acto seguido, se integró esta información mediante el metaanálisis para cada uno de los grupos funcionales indicados.

4.2.1. Análisis exploratorio y procesamiento de los datos

GSE13353

Este estudio (Kurki *et al.*, 2011) tuvo como objetivo la comparación de la expresión génica entre individuos que habían sufrido una ruptura del aneurisma intracraneal de aquellos que no presentaban ruptura de ésta. A continuación, se muestra la distribución del número de muestras en los cuatro grupos experimentales, según el sexo y el estado (Tabla 6). A este estudio se le imputó el sexo ya que no presentaba dicha información en los datos descargados. El número de muestras para cada grupo experimental concuerda con lo descrito en el artículo correspondiente. Los datos tuvieron que ser ajustados a escala logarítmica en base 2. La Figura 8 muestra este ajuste, representando los datos en un diagrama de cajas. Así pues, para visualizar el agrupamiento de las muestras, se llevó a cabo un PCA (Figura 9) y un análisis de grupos (Figura 10).

4. RESULTADOS

Tabla 6. Distribución de los individuos incluidos en el estudio GSE13353. Se muestran distribuidos por sexo (hombre o mujer) y estado (accidente cerebrovascular (ACV) o control).

	MUJER	HOMBRE	TOTAL
ACV	8	3	11
CONTROL	5	3	8
TOTAL	13	6	19

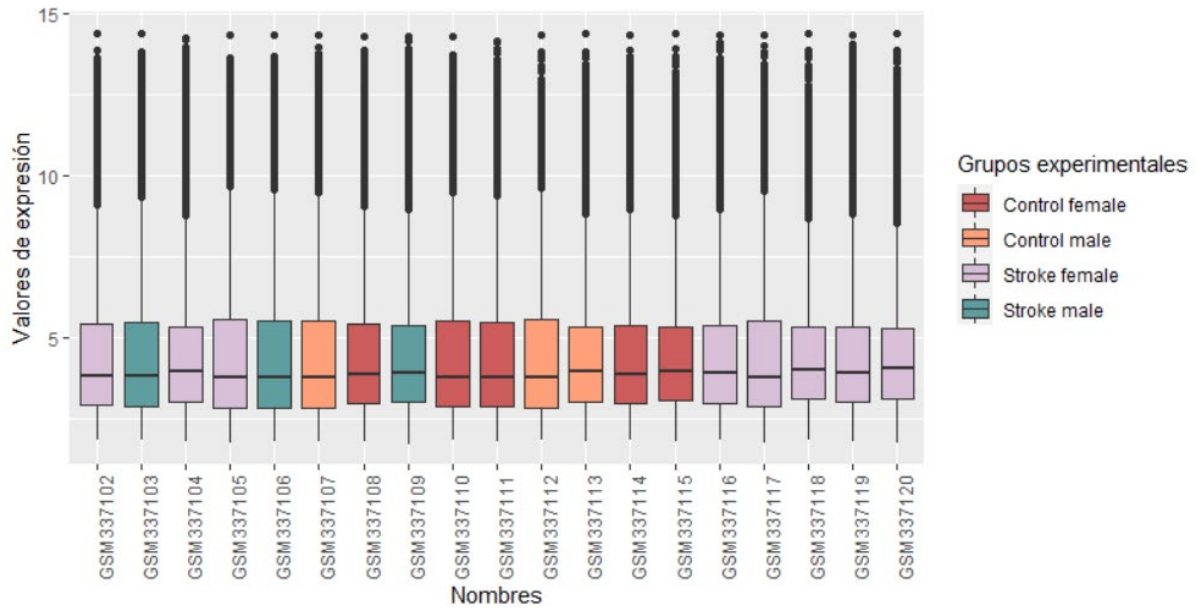


Figura 8. Diagrama de cajas de los valores de expresión de las muestras empleadas en el estudio GSE13353. Cada color hace referencia a un grupo experimental distinto.

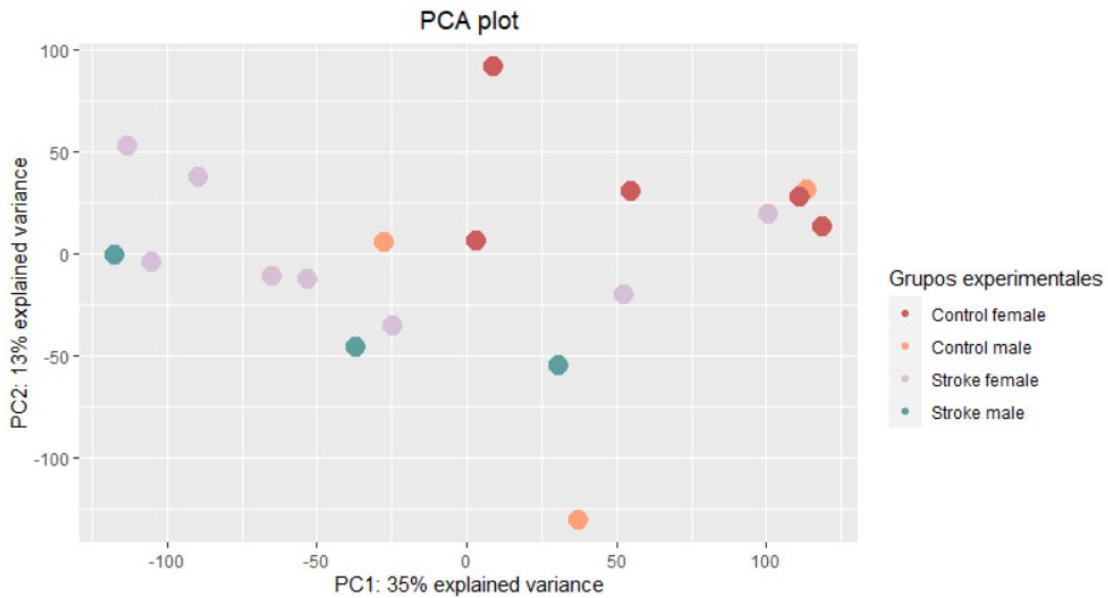


Figura 9. Análisis de componentes principales (PCA) de las muestras del estudio GSE13353. Cada color hace referencia a un grupo experimental distinto.

4. RESULTADOS



Figura 10. Análisis de agrupamiento de las muestras del estudio GSE13353. Cada color hace referencia a un grupo experimental distinto.

GSE16561

El propósito de este proyecto (Barr *et al.* 2010) fue elucidar los niveles de expresión génica en la sangre periférica de pacientes con ACV isquémico agudo, a fin de identificar un panel de genes para el diagnóstico de dicha patología. Se recogieron muestras de sangre periférica de pacientes con ictus que tenían una edad superior a los 18 años, con ictus isquémico diagnosticado por resonancia magnética y controles de individuos que no tenían ictus neurológicamente sanos. A continuación se expone el número de muestras utilizadas en este estudio (Tabla 7). Los datos tuvieron que ser ajustados a escala logarítmica en base 2. El diagrama de cajas, el PCA y el agrupamiento de las muestras se recogen en el Anexo (Figura A1, Figura A2 y Figura A3, respectivamente).

Tabla 7. Individuos incluidos en el estudio GSE16561. Se muestran distribuidos por sexo (hombre o mujer) y estado (accidente cerebrovascular (ACV) o control).

	MUJER	HOMBRE	TOTAL
ACV	22	17	39
CONTROL	14	10	24
TOTAL	36	27	63

GSE22255

En este estudio publicado en 2011, Krug *et al.* tenían como objetivo entender la etiología del ictus isquémico. Para ello, realizaron un perfil de expresión genómica entre pacientes que habían sufrido este tipo de accidente e individuos controles. A continuación se muestra el número de individuos que participaron en el estudio (Tabla 8). Los datos tuvieron que ser ajustados a escala

4. RESULTADOS

logarítmica en base 2. El diagrama de cajas, el PCA y el agrupamiento de las muestras se encuentran en el Anexo (Figura A4, Figura A5 y Figura A6, respectivamente).

Tabla 8. Individuos incluidos en el estudio GSE22255. Se muestran distribuidos por sexo (hombre o mujer) y estado (accidente cerebrovascular (ACV) o control).

	MUJER	HOMBRE	TOTAL
ACV	10	10	20
CONTROL	10	10	20
TOTAL	20	20	40

GSE58294

En este estudio (Stamova *et al.*, 2014), se extrajo sangre de pacientes que habían sufrido un ACV isquémico (concretamente el ACV cardioembólico) y de controles con la finalidad de obtener perfiles de expresión diferenciales. La tabla muestra la distribución de individuos por sexo y por estado (Tabla 9). A este estudio se le imputó el sexo, ya que no presentaba dicha información en los datos descargados. El número de muestras para cada grupo experimental concuerda con lo descrito en el artículo correspondiente. Los datos tuvieron que ser ajustados a escala logarítmica en base 2. El diagrama de cajas, el PCA y el agrupamiento de las muestras se encuentra en el Anexo (Figura A7, Figura A8 y Figura A9, respectivamente).

Tabla 9. Individuos incluidos en el estudio. Se muestran distribuidos por sexo (hombre o mujer) y estado (accidente cerebrovascular (ACV) o control).

	MUJER	HOMBRE	TOTAL
ACV	22	17	39
CONTROL	14	10	24
TOTAL	36	27	63

GSE66724

Se conoce que, para prevenir un ACV causado por fibrilación arterial, se debe administrar al paciente principios activos con carácter anticoagulante. Sin embargo, la elevada frecuencia de hemorragia producida por estos medicamentos es un factor a tener en cuenta. Para ello, en este proyecto (Allende *et al.*, 2016) se estudian las células mononucleares de sangre periférica (PBMC) a través de la identificación de patrones de expresión diferencial entre individuos que han padecido el accidente, y un grupo de controles. La tabla expone la distribución de las muestras en función de las variables sexo y enfermedad (Tabla 10). A este estudio se le imputó el sexo ya que no presentaba dicha información en los datos descargados. El número de muestras para cada grupo experimental concuerda con lo descrito en el artículo correspondiente. Asimismo, los datos tuvieron que ser ajustados a escala logarítmica en base 2. El diagrama de cajas, el PCA y el agrupamiento de las muestras se encuentra en el Anexo (Figura A10, Figura A11 y Figura A12, respectivamente).

4. RESULTADOS

Tabla 10. Individuos incluidos en el estudio GSE66724. Se muestran distribuidos por sexo (hombre o mujer) y estado (accidente cerebrovascular (ACV) o control).

	MUJER	HOMBRE	TOTAL
ACV	3	5	8
CONTROL	3	5	8
TOTAL	6	10	16

4.2.2. Análisis de expresión diferencial

En la siguiente tabla se muestra el número de genes que se han identificado diferencialmente expresados para cada estudio, considerando que el valor ajustado de p por BH debe ser inferior a 0,05 para la comparación de interés (Tabla 11). Aquellos genes que se encuentran agrupados en *Up* están sobrerrepresentados en mujeres porque presentan un valor positivo del logaritmo de la magnitud de cambio (\logFC); mientras que los genes que se han clasificado como *Down* se encuentran sobrerrepresentados en hombres, puesto que tienen asociado un valor negativo de \logFC .

Tabla 11. Resumen del resultado del análisis de expresión diferencial. Se llevó a cabo el análisis para la comparación de interés (ME – MC) – (HE – HC) (ME: Mujer Enferma, MC: Mujer Control, HE: Hombre enfermo, HC: Hombre Control). Los genes se encuentran agrupados según su logaritmo de magnitud de cambio (\logFC), de forma que están etiquetadas como *UP* si están sobrerrepresentadas en mujeres ($\logFC > 0$) y como *DOWN* si están sobrerrepresentadas en hombres ($\logFC < 0$).

ESTUDIO	Nº GENES TOTALES	Nº GENES SIGNIFICATIVOS	<i>UP</i>	<i>DOWN</i>
GSE13353	20947	0	0	0
GSE22255	20947	3984	2017	1967
GSE58294	20947	4	4	0
GSE16561	17791	0	0	0
GSE66724	20947	0	0	0

4.2.3. Análisis de enriquecimiento funcional

Al igual en el análisis de expresión diferencial, se ha hecho uso de la Ecuación 1 para la obtención de una tabla donde se indica el número de funciones anotadas en funciones significativas de la GO. Dichas funciones presentan un valor ajustado de p por BH menor a 0,05. Asimismo, estas funciones biológicas están agrupadas según el valor de LOR. De este modo, aquellas funciones que poseen un valor de LOR positivo se encuentran sobrerrepresentadas en mujeres, mientras que las funciones que presentan un valor de LOR negativo se hallan sobrerrepresentadas en hombres.

4. RESULTADOS

Tabla 12. Resumen del resultado del análisis de enriquecimiento funcional. Se llevó a cabo el análisis para la comparación de interés (ME – MC) – (HE – HC) (ME: Mujer Enferma, MC: Mujer Control, HE: Hombre enfermo, HC: Hombre Control). PB: proceso biológico, FM: función molecular, CC: componente celular. Las agrupaciones se han efectuado según el valor del logaritmo de *odds ratio*, de manera que aquellas funciones y rutas que se encuentran como *UP* se hallan sobrerrepresentadas en mujeres (LOR > 0); y las que se encuentran como *DOWN* están sobrerrepresentadas en hombres (LOR < 0).

	Nº PB GO significativos			Nº FM GO significativos			Nº CC GO significativos			Nº rutas KEGG significativas		
	<i>Up</i>	<i>Down</i>	Total	<i>Up</i>	<i>Down</i>	Total	<i>Up</i>	<i>Down</i>	Total	<i>Up</i>	<i>Down</i>	Total
GSE13353	52	130	182	17	38	55	47	56	103	6	45	51
GSE16561	15	10	25	3	7	10	8	8	16	4	2	6
GSE22255	14	6	20	0	3	3	33	5	38	35	14	49
GSE58294	49	29	78	9	2	11	38	12	50	7	4	11
GSE66724	14	33	47	1	9	10	8	37	45	4	11	15

4.3. Metaanálisis funcional

Tras realizar los análisis de enriquecimiento funcional en cada estudio, se llevó a cabo un metaanálisis funcional para las categorías proceso biológico (PB), función molecular (FM) y componente celular (CC), así como las rutas KEGG, con el objetivo de integrar la información del conjunto de estudios seleccionados. Los resultados significativos se muestran en la siguiente tabla (valor ajustado de p por BH < 0,05) (Tabla 13).

Tabla 13. Resumen de los procesos biológicos (PB), funciones moleculares (FM), componentes celulares (CC) y rutas KEGG significativos en el metaanálisis. Las agrupaciones se han efectuado según el valor del logaritmo de *odds ratio*, de manera que aquellas funciones y rutas que se encuentran como *UP* se hallan sobrerrepresentadas en mujeres (LOR>0); y las que se encuentran como *DOWN* están sobrerrepresentadas en hombres (LOR<0).

	FUNCIONES PB SIGNIFICATIVAS	FUNCIONES FM SIGNIFICATIVAS	FUNCIONES CC SIGNIFICATIVAS	Nº RUTAS KEGG SIGNIFICATIVAS
<i>UP</i>	0	1	5	2
<i>DOWN</i>	5	3	9	6
TOTAL	5	4	14	8

Términos GO para procesos biológicos

Se hallaron 5 procesos biológicos significativamente sobrerrepresentados en total, siendo todos en hombres. En la siguiente tabla se exponen dichos términos GO encontrados con el metaanálisis, así como los respectivos estimadores de la medida del efecto (Tabla 14). Asimismo, tanto el LOR y el intervalo de confianza para cada estudio individual, y el global del metaanálisis, para el término GO:0048821 puede observarse en la Figura 11.

4. RESULTADOS

Tabla 14. Procesos biológicos (PB) GO significativos en el metaanálisis. Se muestran los identificadores de los términos PB y sus respectivos estimadores de la medida de efecto. Las agrupaciones se han efectuado según el valor del logaritmo de odds ratio, de manera que aquellos procesos biológicos que se encuentran como *UP* se hallan sobrerrepresentados en mujeres ($LOR > 0$); y los representados como *DOWN* se encuentran sobrerrepresentados en hombres ($LOR < 0$). LI y LS hacen referencia a los límites inferior y superior del intervalo de confianza del LOR (al 95%). La última columna representa mediante símbolos el sexo en el que se encuentra sobrerrepresentada la función (♂ para hombres, ♀ para mujeres).

ID	DESCRIPCIÓN	LI 95%	LOR	LS 95%	SEXO
GO:0048821	<i>erythrocyte development</i>	-1,149	-0,829	-0,508	♂
GO:0030036	<i>actin cytoskeleton organization</i>	-0,486	-0,378	-0,27	♂
GO:0002576	<i>platelet degranulation</i>	-0,47	-0,345	-0,219	♂
GO:0019221	<i>cytokine-mediated signalling pathway</i>	-0,302	-0,218	-0,135	♂
GO:0006954	<i>inflammatory response</i>	-0,264	-0,192	-0,119	♂

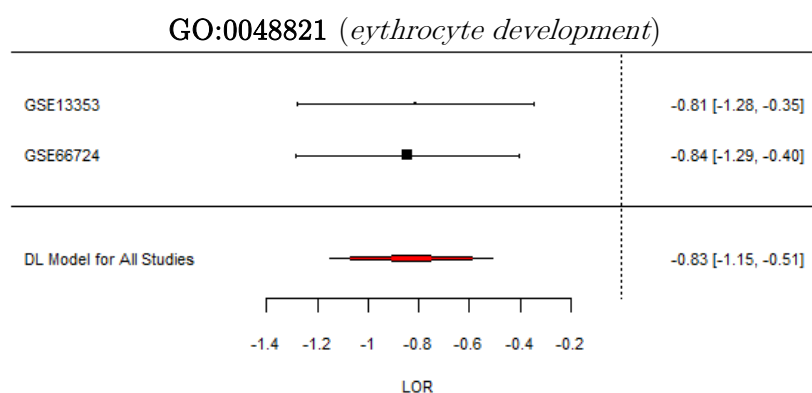


Figura 11. Gráfico de bosque para la función GO:0048821 (*erythrocyte development*). En la parte izquierda se indican los estudios incluidos en el metaanálisis para esta función, mientras que a la derecha la estimación de la medida individual de cada uno con su respectivo intervalo de confianza al 95%. El eje discontinuo vertical indica la posición del LOR cuando su valor es 0.

Con los resultados obtenidos respecto a los procesos biológicos, se puede afirmar que las funciones sobrerrepresentadas en los hombres están relacionadas con el desarrollo eritrocitario, la organización citoesquelética de actina, la degranulación plaquetaria, la vía de señalización mediada por citoquinas o la respuesta inflamatoria.

Términos GO para funciones moleculares

Se identificaron 4 funciones moleculares significativas: 3 sobrerrepresentadas en hombres y 1 en mujeres. En la siguiente tabla se detallan para estas 4 funciones los respectivos estimadores de la medida del efecto (Tabla 15). Así pues, tanto el LOR y el intervalo de confianza para cada estudio individual, y el global del metaanálisis, para los términos GO:0048821 y GO:0005178 pueden observarse en la Figura 12 y 13, respectivamente.

4. RESULTADOS

Tabla 15. Funciones moleculares (FM) GO significativas en el metaanálisis. Se muestran los identificadores de los términos FM y sus respectivos estimadores de la medida de efecto. Las agrupaciones se han efectuado según el valor del logaritmo de *odds ratio*, de manera que aquellas funciones moleculares que se encuentran como *UP* se hallan sobrerrepresentadas en mujeres (LOR > 0); y las representadas como *DOWN* se encuentran sobrerrepresentadas en hombres (LOR < 0). LI y LS hacen referencia a los límites inferior y superior del intervalo de confianza del LOR (al 95%). La última columna representa mediante símbolos el sexo en el que se encuentra sobrerrepresentada la función (♂ para hombres, ♀ para mujeres).

ID	NOMBRE	LI 95%	LOR	LS 95%	SEXO
GO:0005178	<i>integrin binding</i>	-0,469	-0,351	-0,233	♂
GO:0045296	<i>cadherin binding</i>	-0,365	-0,26	-0,154	♂
GO:0005096	<i>GTPase activator activity</i>	-0,306	-0,219	-0,131	♂
GO:0003735	<i>structural constituent of ribosome</i>	0,214	0,721	1,229	♀

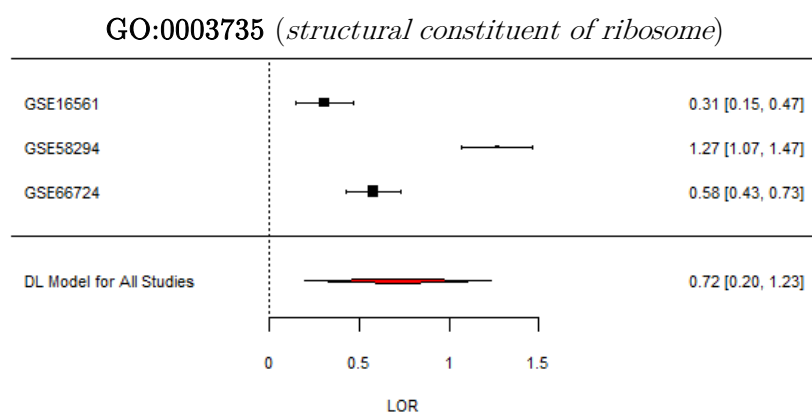


Figura 12. Gráfico de bosque para la función GO:0003735 (*structural constituent of ribosome*). En la parte izquierda se indican los estudios incluidos en el metaanálisis para esta función, mientras que a la derecha la estimación de la medida individual de cada uno con su respectivo intervalo de confianza al 95%. El eje discontinuo vertical indica la posición del LOR cuando su valor es 0.

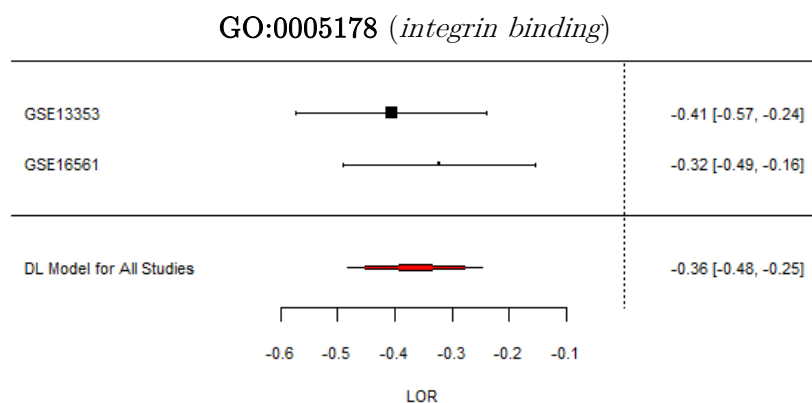


Figura 13. Gráfico de bosque para la función GO:0005178 (*integrin binding*). En la parte izquierda se indican los estudios incluidos en el metaanálisis para esta función, mientras que a la derecha la estimación de la medida individual de cada uno con su respectivo intervalo de confianza al 95%. El eje discontinuo vertical indica la posición del LOR cuando su valor es 0.

Con los resultados obtenidos respecto a las funciones moleculares relacionadas, se han hallado funciones sobrerrepresentadas en los hombres relacionadas con la adhesión celular o la actividad

4. RESULTADOS

GTPasa; mientras que en las mujeres se ha encontrado una función molecular que hace referencia al constituyente estructural del ribosoma.

Términos GO para componentes celulares

Se detectaron 14 componentes celulares significativos: 9 sobrerrepresentados en hombres y 5 en mujeres. En la siguiente tabla se detalla la información correspondiente a estos términos GO encontrados con el metaanálisis (Tabla 16). Por otro lado, tanto el LOR y el intervalo de confianza para cada estudio individual, y el global del metaanálisis, para los términos GO:0062023 y GO:0031464 pueden observarse en la Figura 14 y 15, respectivamente.

Tabla 16. Componentes celulares (CC) GO significativos en el metaanálisis. Se muestran los identificadores de los términos CC y sus respectivos estimadores de la medida de efecto. Las agrupaciones se han efectuado según el valor del logaritmo de *odds ratio*, de manera que aquellas funciones moleculares que se encuentran como *UP* se hallan sobrerrepresentadas en mujeres (LOR > 0); y las representadas como *DOWN* se encuentran sobrerrepresentadas en hombres (LOR < 0). LI y LS hacen referencia a los límites inferior y superior del intervalo de confianza del LOR (al 95%). La última columna representa mediante símbolos el sexo en el que se encuentra sobrerrepresentada la función (♂ para hombres, ♀ para mujeres).

ID	NOMBRE	LI 95%	LOR	LS 95%	SEXO
GO:0042405	<i>nuclear inclusion body</i>	-1,321	-0,925	-0,528	♂
GO:0071013	<i>catalytic step 2 spliceosome</i>	-0,524	-0,401	-0,277	♂
GO:0005681	<i>spliceosomal complex</i>	-0,518	-0,369	-0,22	♂
GO:0016605	<i>PML body</i>	-0,525	-0,351	-0,176	♂
GO:0005770	<i>late endosome</i>	-0,452	-0,334	-0,215	♂
GO:0005788	<i>endoplasmic reticulum lumen</i>	-0,388	-0,307	-0,226	♂
GO:0005769	<i>early endosome</i>	-0,438	-0,259	-0,08	♂
GO:0062023	<i>collagen-containing extracellular matrix</i>	-0,277	-0,217	-0,158	♂
GO:0031410	<i>cytoplasmic vesicle</i>	-0,282	-0,199	-0,117	♂
GO:0005759	<i>mitochondrial matrix</i>	0,114	0,187	0,26	♀
GO:0005743	<i>mitochondrial inner membrane</i>	0,197	0,293	0,389	♀
GO:0005762	<i>mitochondrial large ribosomal subunit</i>	0,359	0,508	0,657	♀
GO:0005763	<i>mitochondrial small ribosomal subunit</i>	0,375	0,586	0,796	♀
GO:0031464	<i>Cul4A-RING E3 ubiquitin ligase complex</i>	0,605	0,987	1,369	♀

En el caso de los componentes celulares, se hallaron términos sobrerrepresentados en hombres relacionados principalmente con vesículas citoplasmáticas, el núcleo, el retículo endoplasmático y la matriz extracelular; mientras que se obtuvieron términos sobrerrepresentados en mujeres relacionados con la estructura mitocondrial y con el complejo ubiquitina ligasa.

4. RESULTADOS

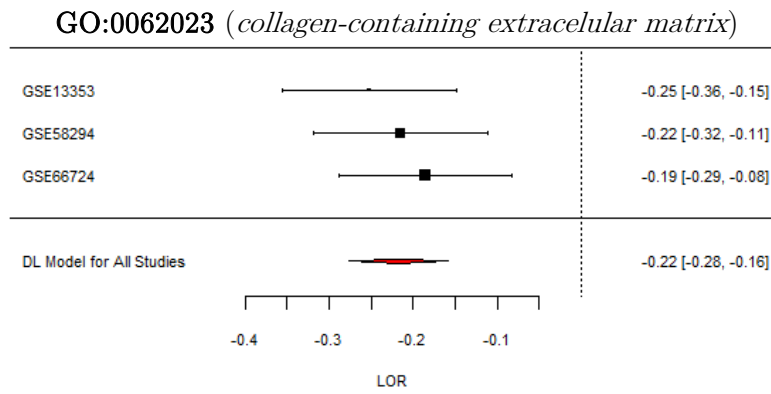


Figura 14. Gráfico de bosque para la función GO:0062023 (*collagen-containing extracellular matrix*). En la parte izquierda se indican los estudios incluidos en el metaanálisis para esta función, mientras que a la derecha la estimación de la medida individual de cada uno con su respectivo intervalo de confianza al 95%. El eje discontinuo vertical indica la posición del LOR cuando su valor es 0.

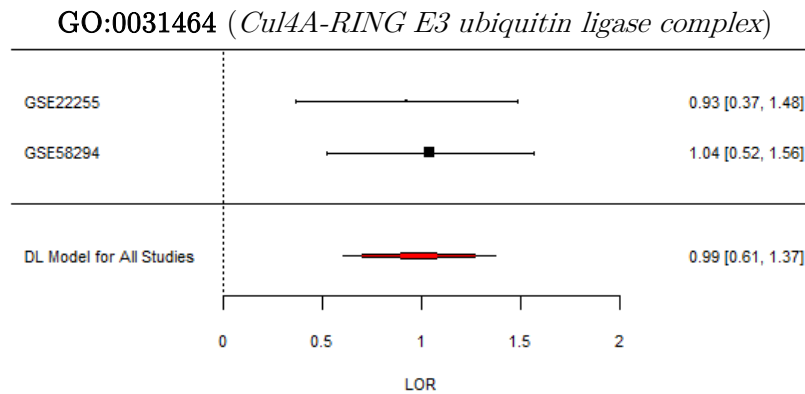


Figura 15. Gráfico de bosque para la función GO:0031464 (*Cul4A-RING E3 ubiquitin ligase complex*). En la parte izquierda se indican los estudios incluidos en el metaanálisis para esta función, mientras que a la derecha la estimación de la medida individual de cada uno con su respectivo intervalo de confianza al 95%. El eje discontinuo vertical indica la posición del LOR cuando su valor es 0.

Rutas metabólicas

Se identificaron 8 rutas significativas: 6 con una mayor activación en hombres y 2 en mujeres (Tabla 17).

4. RESULTADOS

Tabla 17. Rutas KEGG significativas en el metaanálisis. Se muestran los identificadores de las rutas KEGG y sus respectivos estimadores de la medida de efecto. Las agrupaciones se han efectuado según el valor del logaritmo de *odds ratio*, de manera que aquellas rutas KEGG que se encuentran como *UP* se hallan sobrerrepresentadas en mujeres ($LOR > 0$); y las representadas como *DOWN* se encuentran sobrerrepresentadas en hombres ($LOR < 0$). LI y LS hacen referencia a los límites inferior y superior del intervalo de confianza del LOR (al 95%).

ID	DESCRIPCIÓN	LI 95%	LOR	LS 95%	SEXO
hsa04660	<i>T cell receptor signaling pathway</i>	-0,532	-0,397	-0,262	♂
hsa04360	<i>Axon guidance</i>	-0,425	-0,303	-0,180	♂
hsa04144	<i>Endocytosis</i>	-0,435	-0,3	-0,165	♂
hsa04380	<i>Osteoclast differentiation</i>	-0,414	-0,29	-0,167	♂
hsa05200	<i>Pathways in cancer</i>	-0,282	-0,204	-0,127	♂
hsa04010	<i>MAPK signaling pathway</i>	-0,287	-0,201	-0,115	♂
hsa00190	<i>Oxidative phosphorylation</i>	0,285	0,415	0,545	♀
hsa05012	<i>Parkinson's disease</i>	0,285	0,416	0,547	♀

Una vez realizado el metaanálisis de rutas KEGG, se observó que las rutas que aparecían sobrerrepresentadas en hombres estaban relacionadas con la señalización del receptor de los linfocitos T, con la endocitosis o con el cáncer, entre otros; mientras que en las mujeres se observó una sobrerrepresentación de las rutas involucradas en la fosforilación oxidativa y en la enfermedad del Parkinson.

5. DISCUSIÓN

A pesar de encontrar un número considerable de estudios relacionados con el ACV, la mayoría de éstos no incluía información crucial como es el sexo de los individuos de donde procedían las muestras. La incorporación de la perspectiva del sexo es fundamental a la hora de llevar a cabo los estudios biomédicos, puesto que permite efectuar una investigación más acotada y específica, pudiendo así aplicar una terapia de precisión. En este trabajo, se ha optado por la elección de aquellos estudios relacionados con el ACV que incluyesen la descripción del sexo de sus individuos.

Sólo un grupo reducido de estudios presentaban la información del sexo en sus participantes y en algunos otros estudios fue posible recuperar esta información mediante métodos de imputación basados en los niveles de expresión de genes específicos relacionados con el sexo. Por ello, es necesario insistir en la importancia de compartir los datos con información completa y de calidad, tal como se detalla en los estándares FAIR (*Findability, Accessibility, Interoperability, Reusability*) (Wilkinson *et al.*, 2016), garantizando así la validez y la reproducibilidad de la investigación científica.

En el presente trabajo se han tenido en cuenta estos principios FAIR, reutilizando así datos de estudios anteriores y obteniendo resultados comunes entre estudios divergentes a nivel individual. Asimismo, todas las herramientas empleadas para el desarrollo de este trabajo están disponibles en el repositorio GitLab, facilitando de esta forma su accesibilidad y reutilización para otros proyectos.

Son escasos los estudios que han decidido realizar un abordaje por diferencias de sexo en el ACV (Li *et al.*, 2016; Xie *et al.*, 2020). En el caso del primero de los artículos, Li y colaboradores llevaron a cabo un análisis integrativo de los dos estudios que hallaron, identificando así genes diferencialmente expresados en 3 comparaciones diferentes (individuos controles frente a pacientes de ictus; hombres frente a mujeres, agrupados en controles o pacientes; y jóvenes frente a ancianos, agrupados en controles o pacientes). No obstante, en este análisis no se efectúa una comparación completa entre mujeres y hombres enfermos. Así pues, no se lleva a cabo el metaanálisis de los genes diferencialmente expresados.

En este trabajo, se presenta una estrategia más robusta por diversos motivos. A pesar de sólo encontrar 2 estudios que integraban la información del sexo, se ha optado por la imputación de dicha variable en estudios que no la incluían en los metadatos, pero sí en el artículo asociado. Comprobar que esta imputación ha resultado satisfactoria ha posibilitado la inclusión de un mayor número de estudios en este trabajo. Asimismo, se ha efectuado un metaanálisis de todos los estudios hallados, obteniendo funciones y rutas metabólicas de mayor significancia estadística que un análisis individualizado (Ahn y Kang, 2018).

En el metaanálisis funcional relacionado con los procesos biológicos de la GO (Tabla 14), únicamente se han identificado funciones sobrerrepresentadas en hombres. Dichas funciones están relacionadas con el desarrollo de eritrocitos (GO:0048821), donde una elevada amplitud de

5. DISCUSIÓN

distribución eritrocitaria puede estar asociado con la arteriosclerosis carotídea y embolismo cerebral, conduciendo así a un ACV isquémico (Ekici Günay y Günay, 2020; Feng *et al.*, 2017), por lo que un índice elevado de esta medida podría emplearse para anteponerse frente a estas condiciones y remediarlas antes de que se manifieste el ictus. Asimismo, se ha hallado la función de organización del citoesqueleto de actina (GO:0030036), la cual es crucial para el mantenimiento de la integridad de las células endoteliales (EC). En condiciones fisiológicas, las EC cerebrales se fusionan mediante uniones intercelulares, así como las *tight junctions* y las uniones adherentes (Wallez y Huber, 2008). La integridad de estas uniones endoteliales se disuelve cuando se produce una isquemia, aumentando así la permeabilidad de la barrera hematoencefálica. Dicha permeabilidad permite la entrada de leucocitos, citoquinas, quimiocinas y metaloproteinasas de matriz (MMP), exacerbando así el daño de la unidad neurovascular (del Zoppo, 2009; Dimitrijevic *et al.*, 2006). Es por ello por lo que es de elevada importancia restaurar la integridad y la organización del citoesqueleto para proteger el cerebro de un ACV isquémico. En cuanto a la degranulación plaquetaria (GO:0002576), existen evidencias que vinculan el efecto patológico plaquetario con eventos de isquemia cerebral (del Zoppo, 1998), donde la formación de trombos a partir de agregación de plaquetas, así como la activación de protrombina y plasminógeno, ha conducido a ACV isquémicos. Por otro lado, GO:0019221 y GO:0006954 están relacionados con la ruta de señalización de las citoquinas y la respuesta inflamatoria, respectivamente. A pesar de que muchos de los aspectos de la respuesta inflamatoria presentan beneficios para la restauración de la homeóstasis, existe un daño colateral que contribuye al daño producido por la isquemia (Anrather y Iadecola, 2016). La respuesta inflamatoria después de un accidente cerebrovascular es una forma importante de eliminar el tejido necrótico de la lesión cerebral, pero una respuesta inflamatoria excesiva también puede causar daño inflamatorio secundario. Aunque la inmunosupresión inducida por accidente cerebrovascular es un mecanismo de retroalimentación protectora necesario para que el cuerpo resista las lesiones inflamatorias, también es un factor de riesgo importante para los pacientes con accidente cerebrovascular, que podría reducir su defensa inmunológica, aumentar la susceptibilidad a las infecciones, conducir a complicaciones infecciosas graves y aumentar la mortalidad por accidente cerebrovascular. Por lo tanto, mantener el equilibrio entre proinflamación y antiinflamación es crucial para la recuperación de los pacientes con ACV isquémico (Xie *et al.*, 2020).

En cuanto a las funciones moleculares (Tabla 15), se obtuvieron términos sobrerrepresentados en hombres relacionados con el anclaje celular, tanto las integrinas (GO:0005178) como cadherinas (GO:0045296).

Las integrinas son proteínas transmembranales que, unidas a las proteínas de la matriz extracelular, se encargan del mantenimiento de la integridad de la barrera hematoencefálica, entre otras funciones (Hynes, 1992; Takada *et al.*, 2007). Después de un ACV isquémico, los niveles de expresión de diversas integrinas aumentan ya que se incrementa la expresión del factor de crecimiento vascular endotelial (Abumiya *et al.*, 1999). Por consiguiente, aumenta la expresión de proteínas degradadoras de la matriz extracelular (MEC), así como MMP-2 o MMP-9 (Edwards y Bix, 2019), iniciándose así la angiogénesis. Este fenómeno permite el suministro de oxígeno y nutrientes a la zona afectada, aunque una formación de vasos prolongada puede derivar en una permeabilidad crónica de la barrera hematoencefálica (Sandoval y Witt, 2008). Respecto a las cadherinas, se ha visto que, en situaciones de hipoxia, la expresión de cadherinas endoteliales

5. DISCUSIÓN

disminuye, resultando en un aumento de la permeabilidad de esta barrera y produciendo así un efecto patofisiológico que podía derivar en un edema vasogénico cerebral o en una inflamación severa. Análogamente, la cadherina vascular endotelial se halla altamente expresada, pudiendo reconstruir y restaurar la barrera hematoencefálica. (Orsenigo *et al.*, 2012). Es por ello que tanto las integrinas como las cadherinas podrían ser dianas a tener en cuenta para hacer frente a un ACV.

Por otro lado, en el caso de las mujeres se obtuvo una sobrerrepresentación del término GO:0003735, el cual hace referencia a los ARN ribosomales, moléculas que contribuyen a la integridad estructural del ribosoma. El equipo de Xie (2020) realizó un análisis de integración bioinformática de diferencias de sexo del ribosoma donde los genes que resultaron diferencialmente expresados se sometieron a un enriquecimiento funcional. Se encontró que las mujeres responden mejor a la inmunosupresión inducida por ACV, y que los genes relacionados con los ribosomas desempeñan funciones importantes en dicha inmunosupresión.

Respecto a los términos GO relacionados con los componentes celulares (Tabla 16), se obtuvo como resultado una sobrerrepresentación en hombres de los términos relacionados con la MEC (GO:0062023). Cuando se produce un ACV, numerosos cambios se producen en la MEC cerebral. Se ha visto que una hemorragia subaracnoidea resulta en una disrupción de la barrera hematoencefálica. Esta alteración provoca que el colágeno que compone la MEC se degrade a causa de la entrada de proteínas proteolíticas, así como las MMPs, resultando así en una mayor disrupción de la barrera hematoencefálica, inflamación secundaria, edema, entre otros efectos (Bingham, 1982; del Zoppo, 2009; Fukuda *et al.*, 2004; Schöller *et al.*, 2007). Es por ello por lo que se propone en este trabajo la inhibición de las MMPs para la reducción de los efectos del ictus.

Por otra parte, GO:0005759, GO:0005743, GO:0005762 y GO:0005763 eran algunos de los otros términos que se hallaron sobrerrepresentados en mujeres. Dichos términos hacen referencia a la integridad mitocondrial. La disfunción mitocondrial es la característica principal que se observa durante el inicio de la fisiopatología del accidente cerebrovascular. Por lo tanto, dirigirse a las vías moleculares involucradas en la patología del ictus, como la sobreproducción de especies reactivas de oxígeno (ROS), la sobrecarga de Ca^{2+} mitocondrial o la inflamación representa una estrategia prometedora para atenuar el daño y la muerte celulares (Carinci *et al.*, 2021). Entre los mecanismos de protección mitocondrial, la transferencia mitocondrial y la transmitofagia se han convertido en el centro de atención, ya que ambas estrategias están dirigidas a la reposición de mitocondrias sanas en las células dañadas. Esta mayor comprensión de la implicación de la mitocondria en el ictus destaca una nueva perspectiva sobre el desarrollo de enfoques terapéuticos para esta enfermedad, siendo una de las tendencias modernas de la rama neurológica de la medicina regenerativa el uso de células madre como fuentes sustitutas de mitocondrias funcionales para las células amenazadas por isquemia.

Por lo que respecta a las rutas KEGG (Tabla 17), las ruta relacionada con la señalización de los receptores de los linfocitos T (hsa04660) y con la guía axonal (hsa04360) se encontraron sobrerrepresentadas en hombres.

5. DISCUSIÓN

Un número creciente de investigadores ha descubierto que los linfocitos T están involucrados de manera vital en la aparición y progresión del ACV isquémico. Estas células no sólo promueven la aparición de la inflamación al infiltrarse en el cerebro lesionado en las primeras etapas después de la isquemia (Gill y Veltkamp, 2016), sino que también ejercen efectos sobre la reparación y mejora funcional en la etapa tardía después de este tipo de accidente (Liesz *et al.*, 2009). Este tipo de linfocitos desempeñan un papel indispensable a largo plazo después de un accidente isquémico a través de mecanismos como la remodelación y revascularización tisular (Stubbe *et al.*, 2013) y, por lo tanto, son un nuevo objetivo para el tratamiento clínico del ictus.

Por otro lado, la señalización de la guía axonal está involucrada principalmente en la formación de la red neuronal. Los axones están guiados por una variedad de factores de guía, así como netrinas o efrinas, involucradas en rutas de transducción de señales que convergen en la extensión, repulsión y atracción axonales, siendo también de gran relevancia la ruta de señalización de las proteínas quinasas activadas por mitógenos (MAPK) (hsa04010). En el contexto de esta enfermedad, se ha visto que la netrina-1 regula la angiogénesis cerebral, permitiendo el aumento de suministro de oxígeno y nutrientes en las zonas donde se ha producido el accidente. Así pues, los nuevos vasos sanguíneos actúan como estructuras para la migración de los neuroblastos a las zonas afectadas (Ding *et al.*, 2014).

En cuanto a las rutas sobrerrepresentadas en mujeres, destacan la fosforilación oxidativa (hsa00190) y la enfermedad de Parkinson (EP) (hsa05012).

La fosforilación oxidativa es la vía metabólica en la que las células utilizan enzimas para oxidar los nutrientes, liberando así la energía química almacenada dentro de los nutrientes para producir trifosfato de adenosina (Schmidt-Rohr, 2020). Un estudio anterior demostró que las variantes mitocondriales comunes influyen en el riesgo del ACV isquémico (Anderson *et al.*, 2011). El genoma mitocondrial es vital para el ensamblaje del sistema mitocondrial de fosforilación oxidativa (OXPHOS), pero la mayoría de las proteínas estructurales de OXPHOS están codificadas dentro del genoma autosómico. Este sistema mitocondrial consta de 5 complejos que son necesarios para mantener la homeostasis aeróbica y preservar el equilibrio de reducción / oxidación en el entorno celular. Múltiples trastornos raros son causados por mutaciones de genes OXPHOS, muchos de los cuales dan como resultado fenotipos neurodegenerativos o similares a accidentes cerebrovasculares, que incluyen convulsiones, infartos metabólicos y encefalomiopatías (DiMauro y Schon, 2008). Además, OXPHOS juega un papel en la respuesta neuronal y la recuperación después del estrés oxidativo (Nicholls, 2008).

Por otro lado, ha resultado sobrerrepresentada la EP en las mujeres. Vinculando esta ruta con el ictus, los índices de incidencia de EP y ACV aumentan con el envejecimiento (Beilina y Cookson, 2016). Así pues, la disfunción mitocondrial y la agregación anormal de ROS pueden atribuirse a la asociación positiva entre la EP y el accidente cerebrovascular: el daño oxidativo, simultáneamente, da como resultado la degeneración de las neuronas dopaminérgicas y la aterosclerosis (Schapira y Jenner, 2011). Además, los accidentes cerebrovasculares como el infarto y la hemorragia están asociados con el desarrollo de la EP. El ACV aterotrombótico, el tipo más común de ictus, a menudo refleja una aterosclerosis cerebral más grave. Las lesiones ateroscleróticas en el lecho cerebrovascular pueden alterar la permeabilidad endotelial de la barrera hematoencefálica, lo que permite una mayor exposición del parénquima cerebral a

5. DISCUSIÓN

moléculas circulantes por el torrente sanguíneo, así como oxidantes o citocinas, que aumentan el riesgo de EP.

A pesar de todas las diferencias encontradas en este trabajo, las mujeres se encuentran infrarrepresentadas en la investigaciones biomédicas (Kim *et al.*, 2010), habiendo gran cantidad de estudios que no incluyen la variable sexo en sus diseños experimentales. Como se ha visto en este trabajo, la incorporación del sexo en los estudios permite la determinación de biomarcadores que se pueden emplear en terapia personalizada. Para describir las diferencias de sexo en el ACV, se llevó a cabo en este trabajo un metaanálisis, una estrategia de estudio que aumenta el poder y la significancia estadísticos, mejorando la precisión de los resultados y respondiendo a preguntas que no se pueden contestar empleando análisis individuales.

6. CONCLUSIONES

6. CONCLUSIONES

1. El metaanálisis de estudios transcriptómicos es una estrategia robusta que permite combinar datos originados de diferentes estudios individuales y seleccionados a partir de una revisión sistemática estricta. Su aplicación potencia la reutilización de datos de otros estudios, cumpliendo con los criterios FAIR y generando información relevante en la caracterización de las enfermedades humanas.
2. Se ha identificado en mujeres una sobrerrepresentación de funciones y rutas relacionadas con la integridad del ribosoma y la mitocondria, y la fosforilación oxidativa, previamente descritos y asociados a la enfermedad.
3. En el caso de los hombres, se detectó una sobrerrepresentación de funciones y rutas vinculadas con factores relacionados con la homeostasis de la barrera hematoencefálica, con la respuesta inmune y con guía axonal, algunos de los cuales había sido asociados previamente con la enfermedad.
4. La estrategia aplicada en este trabajo ha permitido identificar y realizar una mejor caracterización de diferencias debidas al sexo en ACV, aproximándonos a la necesaria perspectiva de la medicina personalizada y de precisión en los estudios biomédicos.

7. BIBLIOGRAFÍA

7. BIBLIOGRAFÍA

ABUMIYA, T., LUCERO, J., HEO, J. H., TAGAYA, M., KOZIOL, J. A., COPELAND, B. R., Y DEL ZOPPO, G. J. (1999b). Activated Microvessels Express Vascular Endothelial Growth Factor and Integrin $\alpha v\beta 3$ During Focal Cerebral Ischemia. *Journal of Cerebral Blood Flow & Metabolism*, 19(9), 1038–1050. <https://doi.org/10.1097/00004647-199909000-00012>

AHN, E.; KANG, H. (2018). Introduction to systematic review and meta-analysis. *Korean Journal of Anesthesiology*, 71(2), 103-112. <https://doi.org/10.4097/kjae.2018.71.2.103>

ALLENDE, M.; MOLINA, E.; GURUCEAGA, E.; TAMAYO, I.; GONZÁLEZ-PORRAS, J. R.; GONZALEZ-LÓPEZ, T. J.; TOLEDO, E.; RABAL, O.; UGARTE, A.; ROLDÁN, V.; RIVERA, J.; OYARZABAL, J.; MONTES, R.; HERMIDA, J. (2016). Hsp70 protects from stroke in atrial fibrillation patients by preventing thrombosis without increased bleeding risk. *Cardiovascular Research*, 110(3), 309-318. <https://doi.org/10.1093/cvr/cvw049>

AMERICAN STROKE ASSOCIATION (2021). *The Known and Unknown of Stroke Infographic*, visto el 5 de abril de 2021. <https://www.stroke.org/en/professionals/stroke-resource-library/post-stroke-care/the-known-and-unknown-of-stroke-infographic>

ANDERSON, C. D., BIFFI, A., RAHMAN, R., ROSS, O. A., JAGIELLA, J. M., KISSELA, B., COLE, J. W., CORTELLINI, L., ROST, N. S., CHENG, Y. C., GREENBERG, S. M., DE BAKKER, P. I., BROWN, R. D., BROTT, T. G., MITCHELL, B. D., BRODERICK, J. P., WORRALL, B. B., FURIE, K. L., KITTNER, S. J., . . . ROSAND, J. (2010). Common mitochondrial sequence variants in ischemic stroke. *Annals of Neurology*, 69(3), 471–480. <https://doi.org/10.1002/ana.22108>

ANRATHER, J. y IADECOLA, C. (2016). Inflammation and Stroke: An Overview. *Neurotherapeutics*, 13(4), 661-670. <https://doi.org/10.1007/s13311-016-0483-x>

APPELROS, P.; STEGMAYR, B. y TERÉNT, A. (2009). A review on sex differences in stroke treatment and outcome: Sex differences in stroke treatment and outcome. *Acta Neurologica Scandinavica*, 121(6), 359-369. <https://doi.org/10.1111/j.1600-0404.2009.01258.x>

ATTARIAN, H.; BRANDES, J.; DAFER, R.; GERARD, E.; GIESSER, B. (2015). Sex Differences in the Study of Neurological Illnesses. *Behavioural Neurology*, 2015, 1–2. <https://doi.org/10.1155/2015/676531>

BARR, T. L.; CONLEY, Y.; DING, J.; DILLMAN, A.; WARACH, S.; SINGLETON, A.; MATARIN, M. (2010). Genomic biomarkers and cellular pathways of ischemic stroke by RNA gene expression profiling. *Neurology*, 75(11), 1009-1014. <https://doi.org/10.1212/WNL.0b013e3181f2b37f>

BARRETT, T.; WILHITE, S. E.; LEDOUX, P.; EVANGELISTA, C.; KIM, I. F.; TOMASHEVSKY, M.; MARSHALL, K. A.; PHILLIPPY, K. H.; SHERMAN, P. M.; HOLKO, M.; YEFANOV, A.; LEE, H.; ZHANG, N.; ROBERTSON, C. L.; SEROVA, N.; DAVIS, S.; SOBOLEVA, A. (2012). NCBI GEO: Archive for functional genomics data sets—update. *Nucleic Acids Research*, 41(D1), D991-D995. <https://doi.org/10.1093/nar/gks1193>

7. BIBLIOGRAFÍA

- BEILINA, A.; COOKSON, M. R. (2016). Genes associated with Parkinson's disease: Regulation of autophagy and beyond. *Journal of Neurochemistry*, 139, 91-107. <https://doi.org/10.1111/jnc.13266>
- BENJAMINI, Y.; HOCHBERG, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 57(1), 289-300.
- BINGHAM, H. G. (1982). Effect of plasminogen activator (urokinase), plasmin, and thrombin on glycoprotein and collagenous components of basement membrane. *Plastic and Reconstructive Surgery*, 69(6), 1042. <https://doi.org/10.1097/00006534-198206000-00105>
- BOEHME, A. K.; ESENWA, C.; ELKIND, M. S. V. (2017). Stroke Risk Factors, Genetics, and Prevention. *Circulation Research*, 120(3), 472-495. <https://doi.org/10.1161/CIRCRESAHA.116.308398>
- BUSHNELL, C.; HOWARD, V. J.; LISABETH, L.; CASO, V.; GALL, S.; KLEINDORFER, D.; CHATURVEDI, S.; MADSEN, T. E.; DEMEL, S. L.; LEE, S. J.; REEVES, M. (2018). Sex differences in the evaluation and treatment of acute ischaemic stroke. *The Lancet Neurology*, 17(7), 641-650. [https://doi.org/10.1016/S1474-4422\(18\)30201-1](https://doi.org/10.1016/S1474-4422(18)30201-1)
- CAMPBELL, B. C. V.; DE SILVA, D. A.; MACLEOD, M. R.; COUTTS, S. B.; SCHWAMM, L. H.; DAVIS, S. M.; DONNAN, G. A. (2019). Ischaemic stroke. *Nature Reviews Disease Primers*, 5(1), 70. <https://doi.org/10.1038/s41572-019-0118-8>
- CARCEL, C.; WOODWARD, M.; WANG, X.; BUSHNELL, C.; SANDSET, E. C. (2020). Sex matters in stroke: A review of recent evidence on the differences between women and men. *Frontiers in Neuroendocrinology*, 59, 100870. <https://doi.org/10.1016/j.yfrne.2020.100870>
- CARINCI, M.; VEZZANI, B.; PATERGNANI, S.; LUDEWIG, P.; LESSMANN, K.; MAGNUS, T.; CASETTA, I.; PUGLIATTI, M.; PINTON, P.; GIORGI, C. (2021). Different Roles of Mitochondria in Cell Death and Inflammation: Focusing on Mitochondrial Quality Control in Ischemic Stroke and Reperfusion. *Biomedicines*, 9(2), 169. <https://doi.org/10.3390/biomedicines9020169>
- CARLSON, M. (2019). *GO.db: A set of annotation maps describing the entire Gene Ontology*. R package version 3.8.2., <https://doi.org/doi:10.18129/B9.bioc.GO.db>
- CENTER OF DISEASE CONTROL (2021). *Stroke treatment*, visto el 27 de marzo de 2021. <https://www.cdc.gov/stroke/treatments.htm>
- CHENG, N. T.; KIM, A. S. (2015). Intravenous Thrombolysis for Acute Ischemic Stroke Within 3 Hours Versus Between 3 and 4.5 Hours of Symptom Onset. *The Neurohospitalist*, 5(3), 101-109. <https://doi.org/10.1177/1941874415583116>
- CRICK, F. Central Dogma of Molecular Biology. *Nature* 227, 561-563 (1970). <https://doi.org/10.1038/227561a0>
- DEL ZOPPO, G. J. (1998). The role of platelets in ischemic stroke. *Neurology*, 51(Issue 3, Supplement 3), S9-S14. https://doi.org/10.1212/WNL.51.3_Suppl_3.S9

7. BIBLIOGRAFÍA

- DEL ZOPPO, G. J. (2009). Inflammation and the neurovascular unit in the setting of focal cerebral ischemia. *Neuroscience*, 158(3), 972-982. <https://doi.org/10.1016/j.neuroscience.2008.08.028>
- DELLA-MORTE, D.; GUADAGNI, F.; PALMIROTTA, R.; TESTA, G.; CASO, V.; PACIARONI, M.; ABETE, P.; RENGO, F.; FERRONI, P.; SACCO, R. L.; RUNDEK, T. (2012). Genetics of ischemic stroke, stroke-related risk factors, stroke precursors and treatments. *Pharmacogenomics*, 13(5), 595-613. <https://doi.org/10.2217/pgs.12.14>
- DERSIMONIAN, R.; LAIRD, N. (1986). Meta-analysis in clinical trials. *Controlled Clinical Trials*, 7(3), 177-188. [https://doi.org/10.1016/0197-2456\(86\)90046-2](https://doi.org/10.1016/0197-2456(86)90046-2)
- DÍAZ-GUZMÁN, J.; BERMEJO-PAREJA, F.; BENITO-LEÓN, J.; VEGA, S.; GABRIEL, R.; MEDRANO, M. J. (2008). Prevalence of Stroke and Transient Ischemic Attack in Three Elderly Populations of Central Spain. *Neuroepidemiology*, 30(4), 247-253. <https://doi.org/10.1159/000135643>
- DÍAZ-GUZMÁN, J.; EGIDO, J. A.; GABRIEL-SÁNCHEZ, R.; BARBERÁ-COMES, G.; FUENTES-GIMENO, B.; FERNÁNDEZ-PÉREZ, C. (2012). Stroke and Transient Ischemic Attack Incidence Rate in Spain: The IBERICTUS Study. *Cerebrovascular Diseases*, 34(4), 272-281. <https://doi.org/10.1159/000342652>
- DICHGANS, M.; PULIT, S. L.; ROSAND, J. (2019). Stroke genetics: Discovery, biology, and clinical applications. *The Lancet Neurology*, 18(6), 587-599. [https://doi.org/10.1016/S1474-4422\(19\)30043-2](https://doi.org/10.1016/S1474-4422(19)30043-2)
- DIMAURO, S.; SCHON, E. A. (2008). Mitochondrial Disorders in the Nervous System. *Annual Review of Neuroscience*, 31(1), 91-123. <https://doi.org/10.1146/annurev.neuro.30.051606.094302>
- DIMITRIJEVIC, O. B.; STAMATOVIC, S. M.; KEEP, R. F.; ANDJELKOVIC, A. V. (2006). Effects of the Chemokine CCL2 on Blood-Brain Barrier Permeability during Ischemia-Reperfusion Injury. *Journal of Cerebral Blood Flow & Metabolism*, 26(6), 797-810. <https://doi.org/10.1038/sj.jcbFM.9600229>
- DING, Q.; LIAO, S. J.; YU, J. (2014). Axon guidance factor netrin-1 and its receptors regulate angiogenesis after cerebral ischemia. *Neuroscience Bulletin*, 30(4), 683-691. <https://doi.org/10.1007/s12264-013-1441-9>
- EDWARDS, D. N.; BIX, G. J. (2019). The Inflammatory Response After Ischemic Stroke: Targeting β 2 and β 1 Integrins. *Frontiers in Neuroscience*, 13, 540. <https://doi.org/10.3389/fnins.2019.00540>
- EKICI GÜNAY, N.; GÜNAY, N. (2020). The distribution indices of erythrocytes: Which one for acute ischemic stroke? *Turkish Journal of Biochemistry*, 45(1), 65-75. <https://doi.org/10.1515/tjb-2019-0103>
- FENG, G. H.; LI, H. P.; LI, Q. L.; FU, Y.; HUANG, R. B. (2017). Red blood cell distribution width and ischaemic stroke. *Stroke and Vascular Neurology*, 2(3), 172-175. <https://doi.org/10.1136/svn-2017-000071>
- FRANZIN, A.; SAMBO, F.; DI CAMILLO, B. (2016). bnstruct: An R package for Bayesian Network structure learning in the presence of missing data. *Bioinformatics*, btw807. <https://doi.org/10.1093/bioinformatics/btw807>
- FUKUDA, S.; FINI, C. A.; MABUCHI, T.; KOZIOL, J. A.; EGGLESTON, L. L.; DEL ZOPPO, G. J. (2004). Focal Cerebral Ischemia Induces Active Proteases That Degrade Microvascular Matrix. *Stroke*, 35(4), 998-1004. <https://doi.org/10.1161/01.STR.0000119383.76447.05>

7. BIBLIOGRAFÍA

- GARCÍA-GARCÍA, F. (2016) *Methods of functional enrichment analysis in genomic studies*. Tesis Doctoral del Departamento de Bioquímica y Biología Molecular. Programa de Doctorado de Biomedicina y Biotecnología. Universitat de València. 190 pp.
- GILL, D.; VELTKAMP, R. (2016). Dynamics of T cell responses after stroke. *Current Opinion in Pharmacology*, 26, 26-32. <https://doi.org/10.1016/j.coph.2015.09.009>
- HAYNES, W. A.; VALLANIA, F.; LIU, C.; BONGEN, E.; TOMCZAK, A.; ANDRES-TERRÈ, M.; LOFGREN, S.; TAM, A.; DEISSEROTH, C. A.; LI, M. D.; SWEENEY, T. E.; KHATRI, P. (2016). Empowering multi-cohort gene expression analysis to increase reproducibility. *Biocomputing 2017*. https://doi.org/10.1142/9789813207813_0015
- HYNES, R. O. (1992). Integrins: Versatility, modulation, and signaling in cell adhesion. *Cell*, 69(1), 11-25. [https://doi.org/10.1016/0092-8674\(92\)90115-S](https://doi.org/10.1016/0092-8674(92)90115-S)
- KIM, A. M.; TINGEN, C. M.; WOODRUFF, T. K. (2010). Sex bias in trials and treatment must end. *Nature*, 465(7299), 688-689. <https://doi.org/10.1038/465688a>
- KOLESNIKOV, N.; HASTINGS, E.; KEAYS, M.; MELNICHUK, O.; TANG, Y. A.; WILLIAMS, E.; DYLAG, M.; KURBATOVA, N.; BRANDIZI, M.; BURDETT, T.; MEGY, K.; PILICHEVA, E.; RUSTICI, G.; TIKHONOV, A.; PARKINSON, H.; PETRYSZAK, R.; SARKANS, U.; BRAZMA, A. (2015). ArrayExpress update—Simplifying data submissions. *Nucleic Acids Research*, 43(D1), D1113-D1116. <https://doi.org/10.1093/nar/gku1057>
- KRUG, T.; GABRIEL, J. P.; TAIPA, R.; FONSECA, B. V.; DOMINGUES-MONTANARI, S.; FERNANDEZ-CADENAS, I.; MANSO, H.; GOUVEIA, L. O.; SOBRAL, J.; ALBERGARIA, I.; GASPAR, G.; JIMÉNEZ-CONDE, J.; RABIONET, R.; FERRO, J. M.; MONTANER, J.; VICENTE, A. M.; SILVA, M. R.; MATOS, I.; LOPES, G.; OLIVEIRA, S. A. (2012). TTC7B Emerges as a Novel Risk Factor for Ischemic Stroke Through the Convergence of Several Genome-Wide Approaches. *Journal of Cerebral Blood Flow & Metabolism*, 32(6), 1061-1072. <https://doi.org/10.1038/jcbFM.2012.24>
- KURKI, M. I.; HÄKKINEN, S. K.; FRÖSEN, J.; TULAMO, R.; VON UND ZU FRAUNBERG, M.; WONG, G.; TROMP, G.; NIEMELÄ, M.; HERNESNIEMI, J.; JÄÄSKELÄINEN, J. E.; YLÄ-HERTTUALA, S. (2011). Upregulated Signaling Pathways in Ruptured Human Saccular Intracranial Aneurysm Wall: An Emerging Regulative Role of Toll-Like Receptor Signaling and Nuclear Factor- κ B, Hypoxia-Inducible Factor-1A, and ETS Transcription Factors. *Neurosurgery*, 68(6), 1667-1676. <https://doi.org/10.1227/NEU.0b013e318210f001>
- LI, W. X.; DAI, S. X.; WANG, Q.; GUO, Y. C.; HONG, Y.; ZHENG, J. J.; LIU, J. Q.; LIU, D.; LI, G. H.; HUANG, J. F. (2016). Integrated analysis of ischemic stroke datasets revealed sex and age difference in anti-stroke targets. *PeerJ*, 4, e2470. <https://doi.org/10.7717/peerj.2470>
- LIBERATI, A.; ALTMAN, D. G.; TETZLAFF, J.; MULROW, C.; GÖTZSCHE, P. C.; IOANNIDIS, J. P. A.; CLARKE, M.; DEVEREAUX, P. J.; KLEIJNEN, J.; MOHER, D. (2009). The PRISMA Statement for Reporting Systematic Reviews and Meta-Analyses of Studies That Evaluate Health Care Interventions: Explanation and Elaboration. *PLoS Medicine*, 6(7), e1000100. <https://doi.org/10.1371/journal.pmed.1000100>

7. BIBLIOGRAFÍA

LIESZ, A.; SURI-PAYER, E.; VELTKAMP, C.; DOERR, H.; SOMMER, C.; RIVEST, S.; GIESE, T.; VELTKAMP, R. (2009). Regulatory T cells are key cerebroprotective immunomodulators in acute experimental stroke. *Nature Medicine*, 15(2), 192-199. <https://doi.org/10.1038/nm.1927>

MAYO CLINIC (2021). *Accidente cerebrovascular - Síntomas y causas - Mayo Clinic*, visto el 15 de marzo de 2021. <https://www.mayoclinic.org/es-es/diseases-conditions/stroke/symptoms-causes/syc-20350113>

NICHOLLS, D. G. (2008). Oxidative Stress and Energy Crises in Neuronal Dysfunction. *Annals of the New York Academy of Sciences*, 1147(1), 53-60. <https://doi.org/10.1196/annals.1427.002>

ORSENIGO, F.; GIAMPIETRO, C.; FERRARI, A.; CORADA, M.; GALAUP, A.; SIGISMUND, S.; RISTAGNO, G.; MADDALUNO, L.; YOUNG KOH, G.; FRANCO, D.; KURTCUOGLU, V.; POULIKAKOS, D.; BALUK, P.; MCDONALD, D.; GRAZIA LAMPUGNANI, M.; DEJANA, E. (2012). Phosphorylation of VE-cadherin is modulated by haemodynamic forces and contributes to the regulation of vascular permeability in vivo. *Nature Communications*, 3(1), 1208. <https://doi.org/10.1038/ncomms2199>

R CORE TEAM (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>

RITZ, S. A.; ANTLE, D. M.; CÔTÉ, J.; DEROY, K.; FRALEIGH, N.; MESSING, K.; PARENT, L.; ST-PIERRE, J.; VAILLANCOURT, C.; MERGLER, D. (2014). First steps for integrating sex and gender considerations into basic experimental biomedical research. *The FASEB Journal*, 28(1), 4-13. <https://doi.org/10.1096/fj.13-233395>

RUIZ CANTERO, M. T. (2019). *Perspectiva de género en medicina*.

RUSSO, M.W. (2007) How to Review a Meta-analysis. *Gastroenterol Hepatol (NY)*, 3(8):637-42. PMID: 21960873

SANDOVAL, K. E.; WITT, K. A. (2008). Blood-brain barrier tight junction permeability and ischemic stroke. *Neurobiology of Disease*, 32(2), 200-219. <https://doi.org/10.1016/j.nbd.2008.08.005>

SCHAPIRA, A. H.; JENNER, P. (2011). Etiology and pathogenesis of Parkinson's disease: Etiology and Pathogenesis. *Movement Disorders*, 26(6), 1049-1055. <https://doi.org/10.1002/mds.23732>

SCHMIDT-ROHR, K. (2020). Oxygen Is the High-Energy Molecule Powering Complex Multicellular Life: Fundamental Corrections to Traditional Bioenergetics. *ACS Omega*, 5(5), 2221-2233. <https://doi.org/10.1021/acsomega.9b03352>

SCHÖLLER, K.; TRINKL, A.; KLOPOTOWSKI, M.; THAL, S. C.; PLESNILA, N.; TRABOLD, R.; HAMANN, G. F.; SCHMID-ELSAESSER, R.; ZAUSINGER, S. (2007). Characterization of microvascular basal lamina damage and blood-brain barrier dysfunction following subarachnoid hemorrhage in rats. *Brain Research*, 1142, 237-246. <https://doi.org/10.1016/j.brainres.2007.01.034>

SHORTEN, A.; SHORTEN, B. (2013). What is meta-analysis? *Evidence Based Nursing*, 16(1), 3-4. <https://doi.org/10.1136/eb-2012-101118>

7. BIBLIOGRAFÍA

- SOHRABJI, F.; OKOREEH, A.; PANTA, A. (2019). Sex hormones and stroke: Beyond estrogens. *Hormones and Behavior*, 111, 87–95. <https://doi.org/10.1016/j.yhbeh.2018.10.010>
- STAMOVA, B.; JICKLING, G. C.; ANDER, B. P.; ZHAN, X.; LIU, D.; TURNER, R.; HO, C.; KHOURY, J. C.; BUSHNELL, C.; PANCIOLI, A.; JAUCH, E. C.; BRODERICK, J. P.; SHARP, F. R. (2014). Gene Expression in Peripheral Immune Cells following Cardioembolic Stroke Is Sexually Dimorphic. *PLoS ONE*, 9(7), e102550. <https://doi.org/10.1371/journal.pone.0102550>
- STUBBE, T.; EBNER, F.; RICHTER, D.; ENGEL, O. R.; KLEHMET, J.; ROYL, G.; MEISEL, A.; NITSCH, R.; MEISEL, C.; BRANDT, C. (2013). Regulatory T Cells Accumulate and Proliferate in the Ischemic Hemisphere for up to 30 Days after MCAO. *Journal of Cerebral Blood Flow & Metabolism*, 33(1), 37-47. <https://doi.org/10.1038/jcbFM.2012.128>
- SUBRAMANIAN, A.; TAMAYO, P.; MOOTHA, V. K.; MUKHERJEE, S.; EBERT, B. L.; GILLETTE, M. A.; PAULOVICH, A.; POMEROY, S. L.; GOLUB, T. R.; LANDER, E. S.; MESIROV, J. P. (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences*, 102(43), 15545-15550. <https://doi.org/10.1073/pnas.0506580102>
- TAKADA, Y., YE, X.; SIMON, S. (2007). The integrins. *Genome Biology*, 8(5), 215. <https://doi.org/10.1186/gb-2007-8-5-215>
- THE LANCET NEUROLOGY. (2019). A spotlight on sex differences in neurological disorders. *The Lancet Neurology*, 18(4), 319. [https://doi.org/10.1016/S1474-4422\(19\)30001-8](https://doi.org/10.1016/S1474-4422(19)30001-8)
- BOIX, R.; DEL BARRIO, J. L.; SAZ, P.; REÑÉ, R.; MANUBENS, J. M.; LOBO, A.; GASCÓN, J.; DE ARCE, A.; DÍAZ-GUZMÁN, J.; BERGARECHE, A.; BERMEJO-PAREJA, F.; DE PEDRO-CUESTA, J. (2006). Stroke prevalence among the Spanish elderly: an analysis based on screening surveys. *BMC Neurology*, 6(1). <https://doi.org/10.1186/1471-2377-6-36>
- TIERNEY, M. C.; CURTIS, A. F.; CHERTKOW, H.; RYLETT, R. J. (2017). Integrating sex and gender into neurodegeneration research: A six-component strategy. *Alzheimer's & Dementia: Translational Research & Clinical Interventions*, 3(4), 660-667. <https://doi.org/10.1016/j.trci.2017.10.006>
- WALLEZ, Y., y HUBER, P. (2008). Endothelial adherens and tight junctions in vascular homeostasis, inflammation and angiogenesis. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1778(3), 794–809. <https://doi.org/10.1016/j.bbamem.2007.09.003>
- WILKINSON, M. D.; DUMONTIER, M.; AALBERSBERG, I. J.; APPLETON, G.; AXTON, M.; BAAK, A.; BLOMBERG, N.; BOITEN, J. W.; DA SILVA SANTOS, L. B.; BOURNE, P. E.; BOUWMAN, J.; BROOKES, A. J.; CLARK, T.; CROSAS, M.; DILLO, I.; DUMON, O.; EDMUNDS, S.; EVELO, C. T.; FINKERS, R.; MONS, B. (2016). *The FAIR Guiding Principles for scientific data management and stewardship*. *Scientific Data*, 3(1). <https://doi.org/10.1038/sdata.2016.18>
- XIE, J. Q.; LU, Y. P.; SUN, H. L.; GAO, L. N.; SONG, P. P.; FENG, Z. J.; YOU, C. G. (2020). Sex Difference of Ribosome in Stroke-Induced Peripheral Immunosuppression by Integrated Bioinformatics Analysis. *BioMed Research International*, 2020, 1–15. <https://doi.org/10.1155/2020/3650935>

7. BIBLIOGRAFÍA

YADAV, D.; TANVEER, A.; MALVIYA, N.; YADAV, S. (2018). Overview and Principles of Bioengineering. *Omics Technologies and Bio-Engineering*, 3-23. <https://doi.org/10.1016/b978-0-12-804659-3.00001-4>

ZAGNI, E.; SIMONI, L.; COLOMBO, D. (2016). Sex and Gender Differences in Central Nervous System-Related Disorders. *Neuroscience Journal*, 2016, 1-13. <https://doi.org/10.1155/2016/2827090>

8. ANEXOS

8.1. Anexo de figuras.

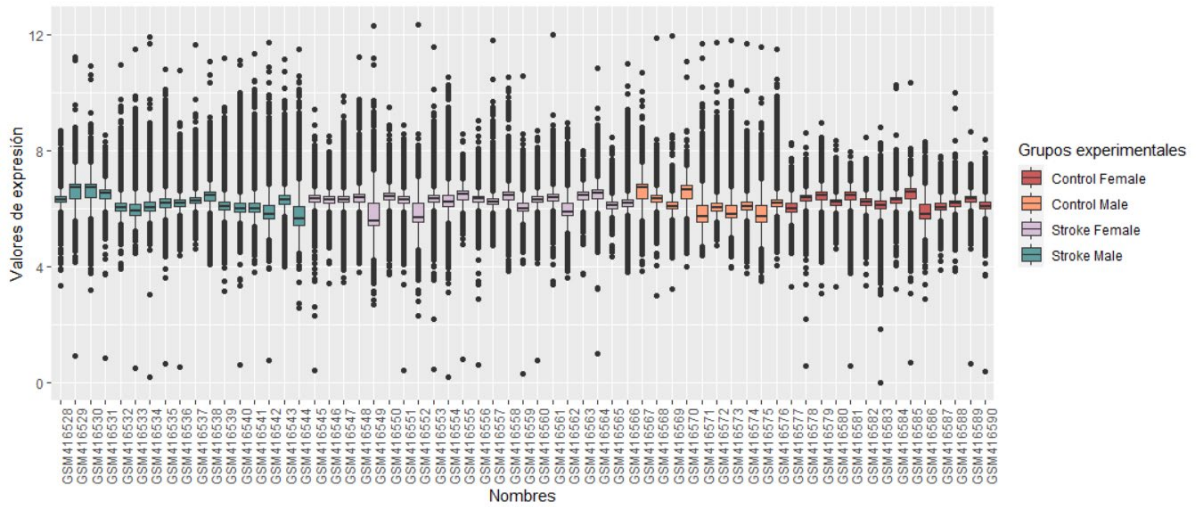


Figura A.1. Diagrama de cajas de los valores de expresión de las muestras empleadas en el estudio GSE16561. Cada color hace referencia a un grupo experimental distinto.

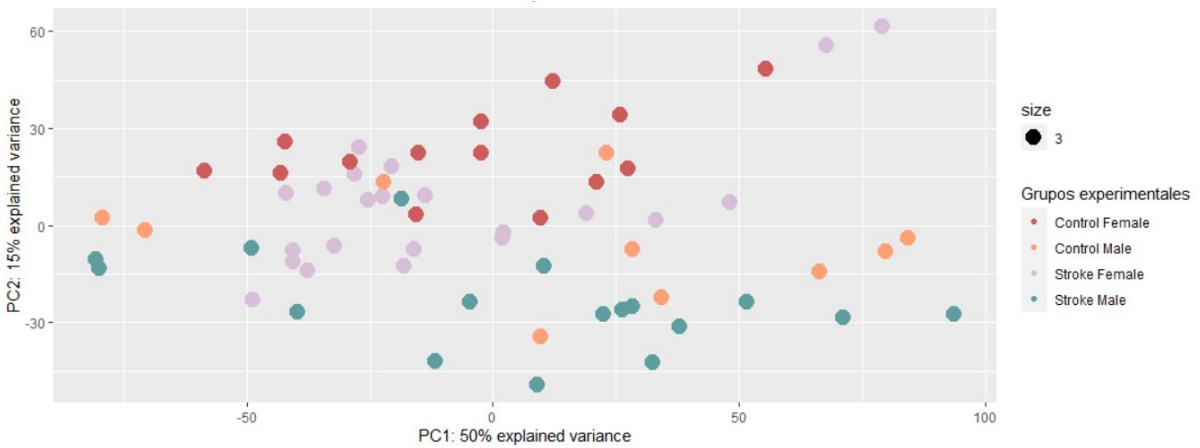


Figura A.2. Análisis de componentes principales (PCA) de las muestras del estudio GSE16561. Cada color hace referencia a un grupo experimental distinto.

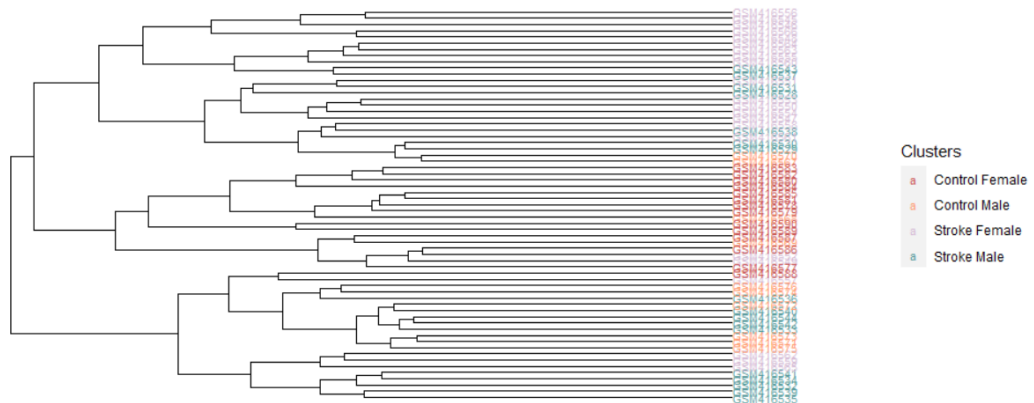


Figura A.3. Análisis de agrupamiento de las muestras del estudio GSE16561. Cada color hace referencia a un grupo experimental distinto.

8. ANEXOS

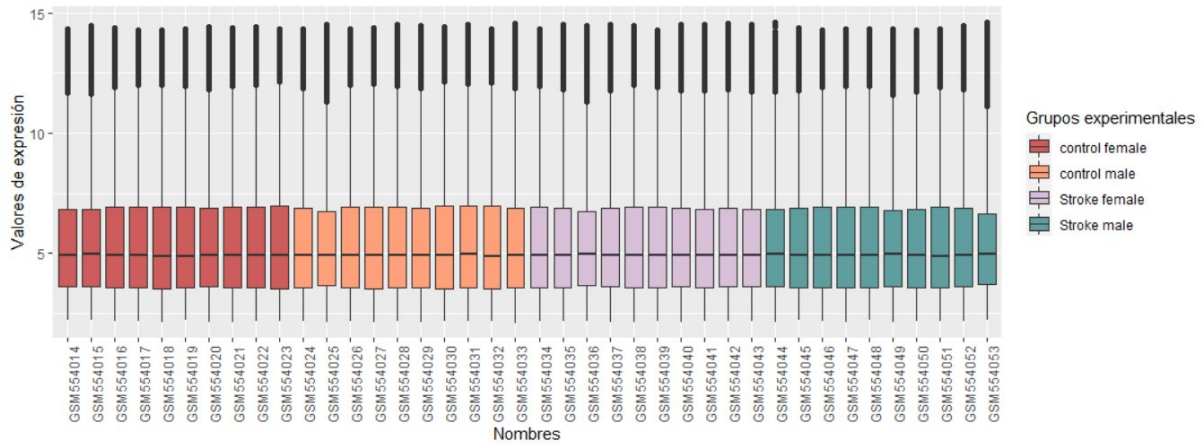


Figura A.4. Diagrama de cajas de los valores de expresión de las muestras empleadas en el estudio GSE22255. Cada color hace referencia a un grupo experimental distinto.

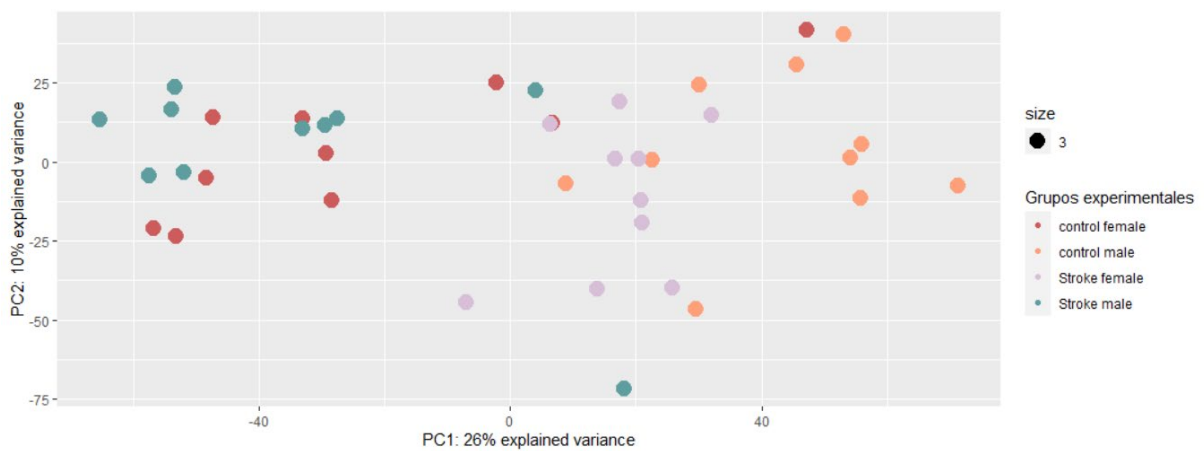


Figura A.5. Análisis de componentes principales (PCA) de las muestras del estudio GSE22255. Cada color hace referencia a un grupo experimental distinto.

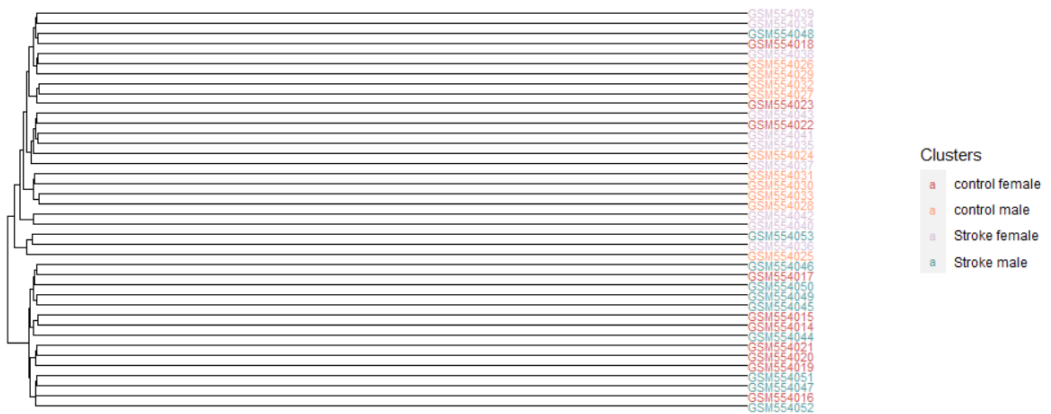


Figura A.6. Análisis de agrupamiento de las muestras del estudio GSE22255. Cada color hace referencia a un grupo experimental distinto.

8. ANEXOS

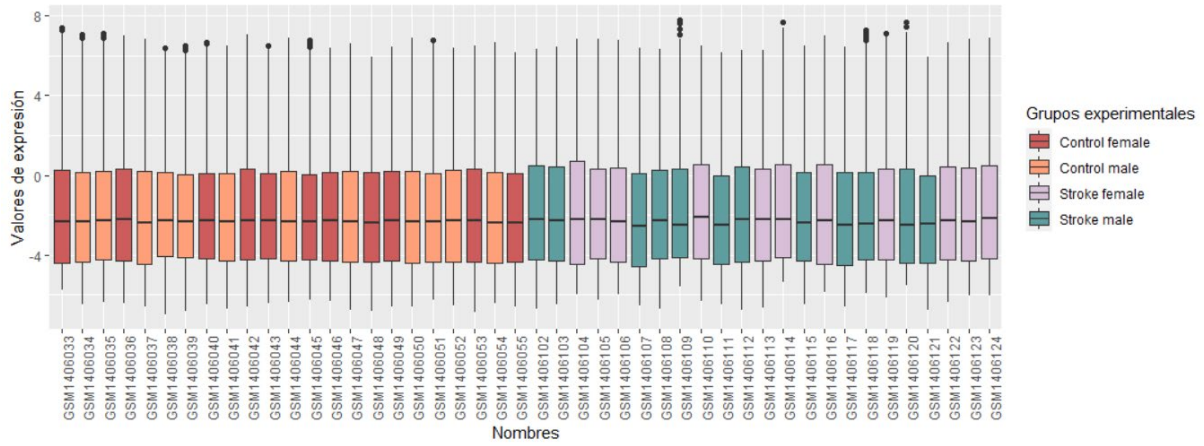


Figura A.7. Diagrama de cajas de los valores de expresión de las muestras empleadas en el estudio GSE58294. Cada color hace referencia a un grupo experimental distinto.

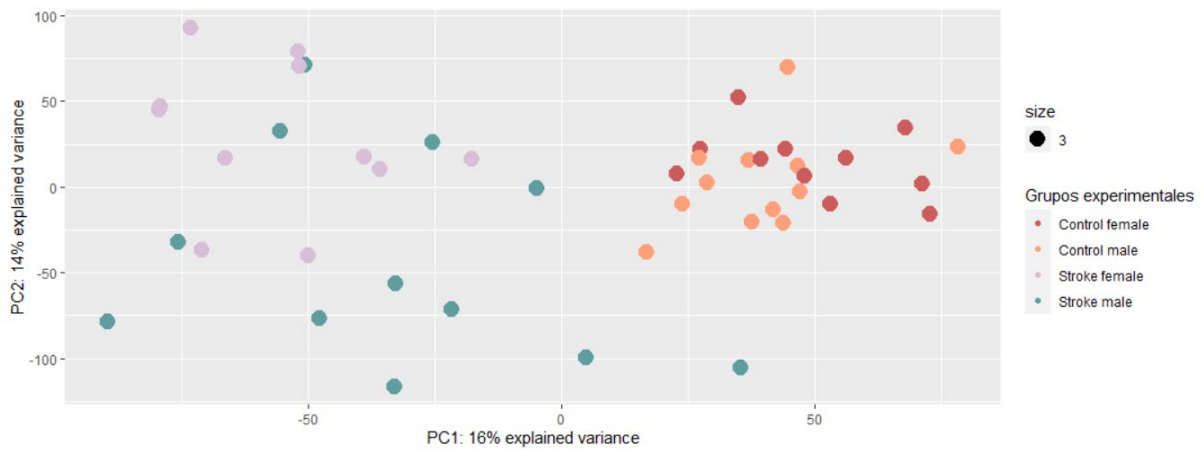


Figura A.8. Análisis de componentes principales (PCA) de las muestras del estudio GSE58294. Cada color hace referencia a un grupo experimental distinto.

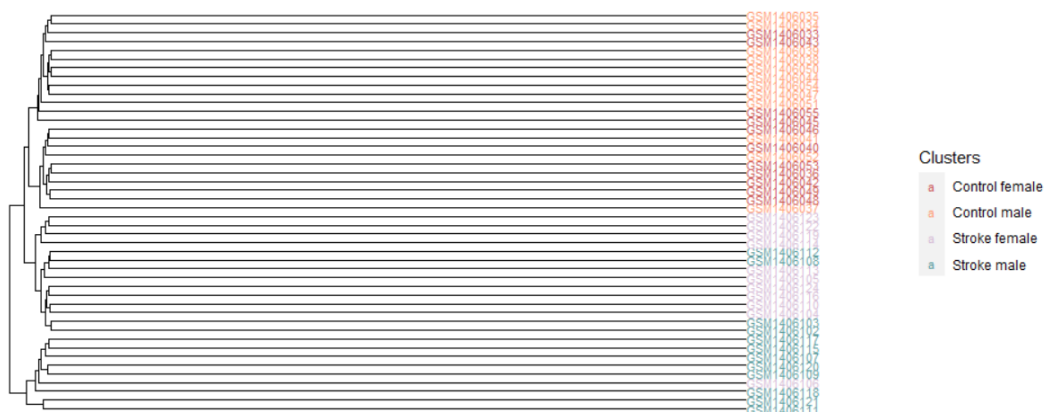


Figura A.9. Análisis de agrupamiento de las muestras del estudio GSE58294. Cada color hace referencia a un grupo experimental distinto.

8. ANEXOS

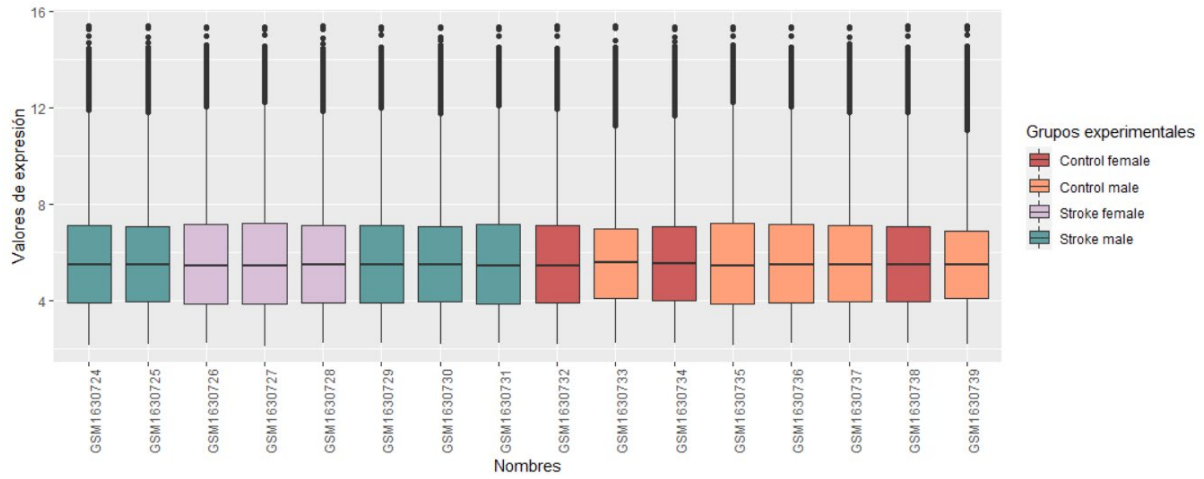


Figura A.10. Diagrama de cajas de los valores de expresión de las muestras empleadas en el estudio GSE66724. Cada color hace referencia a un grupo experimental distinto.

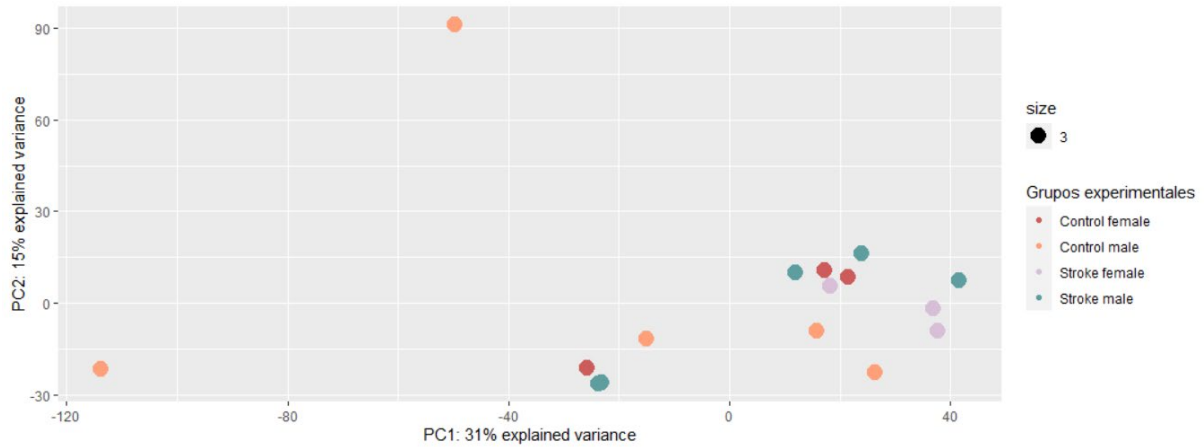


Figura A.11. Análisis de componentes principales (PCA) de las muestras del estudio GSE66724. Cada color hace referencia a un grupo experimental distinto.



Figura A.12. Análisis de agrupamiento de las muestras del estudio GSE66724. Cada color hace referencia a un grupo experimental distinto.